

W. Eric Wong  
Tingshao Zhu *Editors*

# Computer Engineering and Networking

Proceedings of the 2013 International  
Conference on Computer Engineering  
and Network (CENet2013)





W. Eric Wong • Tingshao Zhu  
Editors

# Computer Engineering and Networking

Proceedings of the 2013 International  
Conference on Computer Engineering  
and Network (CENet2013)

 Springer

*Editors*

W. Eric Wong  
University of Texas at Dallas  
Richardson, Texas  
USA

Tingshao Zhu  
Chinese Academy of Sciences  
Beijing, China, People's Republic

ISSN 1876-1100

ISBN 978-3-319-01765-5

DOI 10.1007/978-3-319-01766-2

Springer Cham Heidelberg New York Dordrecht London

ISSN 1876-1119 (electronic)

ISBN 978-3-319-01766-2 (eBook)

Library of Congress Control Number: 2014931403

© Springer International Publishing Switzerland 2014

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed. Exempted from this legal reservation are brief excerpts in connection with reviews or scholarly analysis or material supplied specifically for the purpose of being entered and executed on a computer system, for exclusive use by the purchaser of the work. Duplication of this publication or parts thereof is permitted only under the provisions of the Copyright Law of the Publisher's location, in its current version, and permission for use must always be obtained from Springer. Permissions for use may be obtained through RightsLink at the Copyright Clearance Center. Violations are liable to prosecution under the respective Copyright Law.

The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.

While the advice and information in this book are believed to be true and accurate at the date of publication, neither the authors nor the editors nor the publisher can accept any legal responsibility for any errors or omissions that may be made. The publisher makes no warranty, express or implied, with respect to the material contained herein.

Printed on acid-free paper

Springer is part of Springer Science+Business Media ([www.springer.com](http://www.springer.com))

# Preface

This book is a collection of the papers accepted by CENet 2013—the third International Conference on *Computer Engineering and Network* (CENet), which was held from 20 to 21 July, 2013 in Shanghai, China. It has two volumes and three parts in each. Part I focuses on Algorithm Design with 29 papers over 232 pages; Part II emphasizes Data Processing containing 184 pages divided among 22 papers; Part III Pattern Recognition includes 29 papers in 234 pages; Part IV has 22 papers and 187 pages devoted to one of the most exciting technologies currently surging in popularity—Cloud Computing; Part V covers recent advances in Embedded Systems with 28 papers in 228 pages; and finally Part VI has 28 papers spanning 234 pages dedicated to Network Optimization.

Each part can be used as an excellent reference by industry practitioners, university faculty, and undergraduate as well as graduate students who need to build a knowledge base of the most current advances and state of practice in the topics covered by this book. This will enable them to produce, maintain, and manage systems with high levels of trustworthiness and complexity that provide critical services in a variety of applications.

Thanks go to the authors for their hard work and dedication as well as the reviewers for ensuring the selection of only the highest quality papers; their efforts made this book possible. Invaluable assistance with the publication was also provided by the editorial staff at Springer, especially Mr. Brett Kurzman and Miss Rebecca Hytowitz.

Richardson, Texas, USA  
Beijing, China

W. Eric Wong  
Tingshao Zhu



# Contents

## Volume 1

### Part I Algorithm Design

<b>1 Simulation Algorithm of Adaptive Scheduling in Missile-Borne Phased Array Radar . . . . .</b>	<b>3</b>
Qizhong Li, Shanshan Sun, and Jianfei Zhao	
<b>2 A Second-Order Algorithm for Curve Parallel Projection on Parametric Surfaces . . . . .</b>	<b>9</b>
Xiongbing Fang and Hai-Yin Xu	
<b>3 Computation Method of Processing Time Based on BP Neural Network and Genetic Algorithm . . . . .</b>	<b>21</b>
Danchen Zhou and Chao Guo	
<b>4 Integral Sliding Mode Controller for an Uncertain Network Control System with Delay . . . . .</b>	<b>31</b>
Zhenbin Gao	
<b>5 Synthesis of Linear Antenna Array Using Genetic Algorithm to Control Side Lobe Level . . . . .</b>	<b>39</b>
Zhigang Zhang, Ting Li, Feng Yuan, and Li Yin	
<b>6 Wavelet Analysis Combined with Artificial Neural Network for Predicting Protein-Protein Interactions . . . . .</b>	<b>47</b>
Juanjuan Li, Yuehui Chen, and Fenglin Wang	
<b>7 Application Analysis of Slot Allocation Algorithm for Link16 . . . . .</b>	<b>55</b>
Hui Zeng, Qiang Chen, Xiaoqiang Li, and Jianguo Shen	
<b>8 An Improved Cluster Head Algorithm for Wireless Sensor Network . . . . .</b>	<b>65</b>
Feng Yu, Wei Liu, and Gang Li	



**9 An Ant Colony System for Dynamic Voltage Scaling Problem in Heterogeneous System . . . . . 73**  
 Yan Kang, Ying Lin, Yifan Zhang, and He Lu

**10 An Improved Ant Colony System for Task Scheduling Problem in Heterogeneous Distributed System . . . . . 83**  
 Yan Kang, Yifan Zhang, Ying Lin, and He Lu

**11 Optimization of Green Agri-Food Supply Chain Network Using Particle Swarm Optimization Algorithm . . . . . 91**  
 Qian Tao, Zhexue Huang, Chunqin Gu, and Chenxin Zhang

**12 A New Model for Short-Term Power System Load Forecasting Using Wavelet Transform Fuzzy RBF Neural Network . . . . . 99**  
 Jingduan Dong, Changhao Xia, and Wei Zhang

**13 Energy-Effective Frequency-Based Adaptive Sampling Algorithm for Clustered Wireless Sensor Network . . . . . 107**  
 Meiyang Zhang, Wenyu Cai, Liping Zhou, and Jilai Liu

**14 An Indoor Three-Dimensional Positioning Algorithm Based on Difference Received Signal Strength in WiFi . . . . . 115**  
 Yibo Li and Xiting Liu

**15 The Universal Approximation Capability of Double Flexible Approximate Identity Neural Networks . . . . . 125**  
 Saeed Panahian Fard and Zarita Zainuddin

**16 A Novel and Real-Time Hand Tracking Algorithm for Gesture Manipulation . . . . . 135**  
 Zhiqin Zhang

**17 A Transforming Quantum-Inspired Genetic Algorithm for Optimization of Green Agricultural Products Supply Chain Network . . . . . 145**  
 Chunqin Gu and Qian Tao

**18 A Shortest Path Algorithm Suitable for Navigation Software . . . . . 153**  
 Peng Luo, Qizhi Qiu, Wenyan Zhou, and Pei Fang

**19 An Energy-Balanced Clustering Routing Algorithm for Wireless Sensor Networks . . . . . 163**  
 Mingqiang Chen and Xianhai Tan

**20 Simulation and Analysis of Binary Frequency Shift Keying Noise Cancel Adaptive Filter Based on Least Mean Square Error Algorithm . . . . . 171**  
 Zhongping Chen and Jinding Gao

<b>21</b>	<b>Density-Sensitive Semi-supervised Affinity Propagation Clustering</b> . . . . .	177
	Kunlun Li, Qi Meng, Shangzong Luo, Hexin Li, and Qian Wang	
<b>22</b>	<b>The Implementation of a Hybrid Particle Swarm Optimization Algorithm Based on Three-Level Parallel Model</b> . . . . .	185
	Yi Xiao and Yu Liu	
<b>23</b>	<b>Optimization of Inverse Planning Based on an Improved Non-dominated Neighbor-Based Selection in Intensity Modulated Radiation Therapy</b> . . . . .	193
	Xiao Zhang, Guoli Li, and Zhizhong Li	
<b>24</b>	<b>A Recommendation System for Paper Submission Based on Vertical Search Engine</b> . . . . .	201
	Zhen Xu, Yi Yang, Fei Wang, Jiao Xu, Zhong Li, Fuqiang Mu, and Lian Li	
<b>25</b>	<b>Analysis and Improvement of SPRINT Algorithm Based on Hadoop</b> . . . . .	209
	Shanshan Fei, Qiaoyan Wen, and Zhengping Jin	
<b>26</b>	<b>Prediction Model for Trend of Web Sentiment Using Extension Neural Network and Nonparametric Auto-regression Method</b> . . . . .	219
	Haitao Zhang, Binjun Wang, and Guangxuan Chen	
<b>27</b>	<b>K-Optimal Chaos Ant Colony Algorithm and Its Application on Dynamic Route Guidance System</b> . . . . .	227
	Hai Yang	
<b>28</b>	<b>A Certainty-Based Active Learning Framework of Meeting Speech Summarization</b> . . . . .	235
	Jian Zhang and Huaqiang Yuan	
<b>29</b>	<b>Application of Improved BP Neural Network in the Frequency Identification of Piano Tone</b> . . . . .	243
	Xu Chen and Jun Tang	
 <b>Part II Data Processing</b>		
<b>30</b>	<b>Implicit Factoring with Shared Middle Discrete Bits</b> . . . . .	255
	Meng Shi, Xianghui Liu, and Wenbao Han	
<b>31</b>	<b>Loading Data into HBase</b> . . . . .	265
	Juan Yang and Xiaopu Feng	
<b>32</b>	<b>Incomplete Decision-Theoretic Rough Set Model Based on Improved Complete Tolerance Relation</b> . . . . .	273
	Xia Wang	

<b>33</b>	<b>A New Association Rule Mining Algorithm Based on Compression Matrix</b> . . . . .	281
	Sihui Shu	
<b>34</b>	<b>Retraction: Decoupling Interrupts from the Internet in Markov Models</b> . . . . .	291
	Jinwen Ma, Jingchun Zhang, and Jinrong Guo	
<b>35</b>	<b>Parallel Feature Selection Based on MapReduce</b> . . . . .	299
	Zhanquan Sun	
<b>36</b>	<b>Initial State Modeling of Interlocking System Using Maude</b> . . . . .	307
	Rui Ma, Zhongwei Xu, Zuxi Chen, and Shuqing Zhang	
<b>37</b>	<b>Semi-supervised Learning Using Nonnegative Matrix Factorization and Harmonic Functions</b> . . . . .	321
	Lin Li, Zhenyu Zhao, Chenping Hou, and Yi Wu	
<b>38</b>	<b>Exploring Data Communication at System Level Through Reverse Engineering: A Case Study on USB Device Driver</b> . . . . .	329
	Leela Sedaghat, Brad Duerling, Xiaoxi Huang, and Ziyang Tang	
<b>39</b>	<b>Using Spatial Analysis to Identify Tuberculosis Transmission and Surveillance</b> . . . . .	337
	Jinrong Bai, Guozhong Zou, Shiguang Mu, and Yu Ma	
<b>40</b>	<b>Construction Method of Exception Control Flow Graph for Business Process Execution Language Process</b> . . . . .	345
	Caoqing Jiang, Shi Ying, Shanming Hu, and Hua Guan	
<b>41</b>	<b>P300 Detection in Electroencephalographic Signals for Brain-Computer Interface Systems: A Neural Networks Approach</b> . . . . .	355
	Seyed Aliakbar Mousavi, Muhammad Rafie Hj. Mohd. Arshad, Hasimah Hj. Mohamed, Putra Sumari, and Saeed Panahian Fard	
<b>42</b>	<b>Web Content Extraction Technology</b> . . . . .	365
	Zhenyu Jiao, Xiaoben Yan, Jinjin Sun, Yuchen Wang, and Jiangbin Chen	
<b>43</b>	<b>A New Data-Intensive Parallel Processing Framework for Spatial Data</b> . . . . .	375
	Dong Zhao, Yang Gu, and Zhenchun Huang	
<b>44</b>	<b>The Approach of Graphical User Interface Testing Guided by Bayesian Model</b> . . . . .	385
	Zhifang Yang, Zhongxing Yu, and Chenggang Bai	
<b>45</b>	<b>A Model for Reverse Logistics with Collection Sites Based on Heuristic Algorithm</b> . . . . .	395
	Xiaoqing Geng and Yu Wang	

**46 The Storage of Wind Turbine Mass Data Based on MongoDB** . . . . . 403  
 Qile Wang, Zhu Shen, Long Ma, and Shi Yin

**47 Improvement of Extraction Method of Correlation Time Delay Based on Connected-Element Interferometry** . . . . . 411  
 Fei Wang, Zhenfei Wang, Dun Li, and Bingjie Yang

**48 Modeling and Evaluation of the Performance of Parallel/Distributed File System** . . . . . 421  
 Tiezhu Zhao, Xin Ao, and Huaqiang Yuan

**49 CoCell: A Low-Diameter, High-Performance Data Center Network Architecture** . . . . . 429  
 Peng Wang, Huaxi Gu, Yan Zhao, and Xiaoshan Yu

**50 Simulation Investigation of Counterwork Between Anti-radiation Missile and Active Decoy System** . . . . . 437  
 Huaqiang Hu and Dandan Wen

**51 Simulation Jamming Technique on Binary Phase-Coded Pulse Compression Radar** . . . . . 445  
 Yulin Yang and Lijuan Qiu

**Part III Pattern Recognition**

**52 Personalized Information Service Recommendation System Based on Clustering and Classification** . . . . . 455  
 Yu Wang

**53 Palmprint Recognition Based on Subclass Discriminant Analysis** . . . . . 465  
 Pengfei Yu, Haiyan Li, Hao Zhou, and Dan Xu

**54 A Process Quality Monitoring Approach of Automatic Aircraft Component Docking** . . . . . 473  
 Guowei Yang, Chengjing Zhang, and Xiaofeng Zhang

**55 Overhead Transmission Lines Sag Measurement Based on Image Processing** . . . . . 481  
 Wengang Cheng and Long Chen

**56 Chinese Domain Ontology Learning Based on Semantic Dependency and Formal Concept Analysis** . . . . . 489  
 Lixin Hou, Shanhong Zheng, Haitao He, and Xinyi Peng

**57 Text Classification Algorithm Based on Rough Set** . . . . . 499  
 Zhiyong Hong

**58 Robust Fragment-Based Tracking with Online Selection of Discriminative Features** . . . . . 507  
 Yongqiang Huang and Long Zhao

**59 Extraction Method of Gait Feature Based on Human Centroid Trajectory . . . . . 515**  
Xin Chen and Tianqi Yang

**60 An Algorithm for Bayesian Network Structure Learning Based on Simulated Annealing with Adaptive Selection Operator . . . . . 525**  
Ao Lin, Bing Xiao, and Yi Zhu

**61 Static Image Segmentation Using Polar Space Transformation Technique . . . . . 533**  
Xuan Luo, Tiancai Liang, and Weifeng Wang

**62 Image Restoration via Nonlocal *P*-Laplace Regularization . . . . . 541**  
Chen Yao, Lijuan Hong, and Yunfei Cheng

**63 Analysis and Application of Computer Technology on Architectural Space Lighting Visual Design . . . . . 549**  
Yiwen Cao

**64 Improving Online Gesture Recognition with Warping LCSS by Multi-Sensor Fusion . . . . . 559**  
Chao Chen and Haibin Shen

**65 The Lane Mark Identifying and Tracking in Intense Illumination . . . . . 567**  
Yanyun Xing, Bo Yu, and Fangqun Yang

**66 Classification Modeling of Multi-Featured Remote Sensing Images Based on Sparse Representation . . . . . 577**  
Xiaoting Hao, Chunmei Zhang, Jing Bai, Mo Dai, Wenxing Bao, and Wei Feng

**67 A Parallel and Convergent Support Vector Machine Based on MapReduce . . . . . 585**  
Yingying Ma, Liming Wang, and Longpu Li

**68 Vehicle Classification Based on Hierarchical Support Vector Machine . . . . . 593**  
Mengwan Jiang and Haoliang Li

**69 Image Splicing Detection Based on Machine Learning Algorithm . . . . . 601**  
Yan Xiao

**70 A Lane Detection Algorithm Based on Hyperbola Model . . . . . 609**  
Chaobo Chen, Bofeng Zhang, and Song Gao

**71 Comparisons and Analyses of Image Softproofing Under Different Profile Rendering Intent . . . . . 617**  
Qingxue Yu, Yunhui Luo, Maohai Lin, and Quantao Liu

**72 An Improved Dense Matching Algorithm for Face Based on Region Growing . . . . . 625**  
Xin Xia and Shaoyan Gai

**73 An Improved Feature Selection Method for Chinese Short Texts Clustering Based on HowNet . . . . . 635**  
Xin Chen, Yuqing Zhang, Long Cao, and Donghui Li

**74 Internet Worm Detection and Classification Based on Support Vector Machine . . . . . 643**  
Huihui Liang, Min Li, and Jiwen Chai

**75 Real-Time Fall Detection Based on Global Orientation and Human Shape . . . . . 653**  
Shuangcheng Wang, Yepeng Guan, and Ruiyue Xu

**76 The Classification of Synthetic Aperture Radar Oil Spill Images Based on the Texture Features and Deep Belief Network . . . . . 661**  
Xixi Huang and Xiaofeng Wang

**77 The Ground Objects Identification for Digital Remote Sensing Image Based on the BP Neural Network . . . . . 671**  
Shengkui Cao, Guangchao Cao, Kelong Chen, Chengyong Wu, Tao Zhang, and Jie Yuan

**78 Detection of Image Forgery Based on Improved PCA-SIFT . . . . . 679**  
Kunlun Li, Hexin Li, Bo Yang, Qi Meng, and Shangzong Luo

**79 A Thinning Model for Handwriting-Like Image Skeleton . . . . . 687**  
Shijiao Zhu, Jun Yang, and Xue-fang Zhu

**80 Discrimination of the White Wine Based on Sparse Principal Component Analysis and Support Vector Machine . . . . . 695**  
Rong Wang, Wu Zeng, and Jiao Ming

**Volume 2**

**Part IV Cloud Computing**

**81 Design of Mobile Electronic Payment System . . . . . 705**  
Ting Huang

**82 Power Saving-Based Radio Resource Scheduling in Long-Term Evolution Advanced Network . . . . . 713**  
Yen-Yin Chu, I-Hsuan Peng, Yen-Wen Chen, Chi-Fu Yi, and Addison Y.S. Su

**83 Dispatching and Management Model Based on Safe Performance Interface for Improving Cloud Efficiency . . . . . 723**  
Bin Chen, Zhijian Wang, and Yu Wang

**84 A Proposed Methodology for an E-Health Monitoring System Based on a Fault-Tolerant Smart Mobile . . . . . 731**  
Ahmed Alahmadi and Ben Soh

**85 Design and Application of Indoor Geographical Information System . . . . . 739**  
Yongfeng Suo, Tianhe Chi, and Tianyue Liu

**86 Constructing Cloud Computing Infrastructure Platform of the Digital Library Base on Virtualization Technology . . . . . 747**  
Tingbo Fu, Jinsheng Yang, Yu Gao, and Guang Yu

**87 A New Single Sign-on Solution in Cloud . . . . . 755**  
Guangxuan Chen, Yanhui Du, Panke Qin, Lei Zhang, and Jin Du

**88 A Collaborative Load Control Scheme for Hierarchical Mobile IPv6 Network . . . . . 763**  
Yi Yang, QingShan Man, and PingLiang Rui

**89 A High Efficient Selective Content Encryption Method Suitable for Satellite Communication System . . . . . 775**  
Yanyan Xu, Bo Yang, Zhengquan Xu, and Tengyue Mao

**90 Network Design of a Low-Power Parking Guidance System . . . . . 783**  
Ming Xia, Yabo Dong, Qingzhang Chen, Kai Wang, and Rongjie Wu

**91 Strategy of Domain and Cross-Domain Access Control Based on Trust in Cloud Computing Environment . . . . . 791**  
Bo Li, Ming Tian, Yongsheng Zhang, and Shenjuan Lv

**92 Detecting Unhealthy Cloud System Status . . . . . 799**  
Zhidong Chen, Buyang Cao, and Yuanyuan Liu

**93 Scoring System of Simulation Training Platform Based on Expert System . . . . . 809**  
Wei Nie, Ying Wu, and Dabin Hu

**94 Analysis of Distributed File Systems on Virtualized Cloud Computing Environment . . . . . 817**  
Tiezhu Zhao, Zusheng Zhang, and Huaqiang Yuan

**95 A Decision Support System with Dynamic Probability Adjustment for Fault Diagnosis in Critical Systems . . . . . 825**  
Qiang Chen and Yun Xue

<b>96</b>	<b>Design and Implementation of an SD Interface to Multiple-Target Interface Bridge . . . . .</b>	835
	Guoyong Li, Leibo Liu, Shouyi Yin, Dajiang Liu, and Shaojun Wei	
<b>97</b>	<b>Cloud Storage Management Technology for Small File Based on Two-Dimensional Packing Algorithm . . . . .</b>	847
	Zhiyun Zheng, Shaofeng Zhao, Xingjin Zhang, Zhenfei Wang, and Liping Lu	
<b>98</b>	<b>Advertising Media Selection and Delivery Decision-Making Using Influence Diagram . . . . .</b>	855
	Xiaoxuan Hu and Fan Jiang	
<b>99</b>	<b>The Application of Trusted Computing Technology in the Cloud Security . . . . .</b>	865
	Bo Li, Shenjuan Lv, Yongsheng Zhang, and Ming Tian	
<b>100</b>	<b>The Application Level of E-commerce in Enterprises in China . . . . .</b>	873
	Yinghan H. Tang	
<b>101</b>	<b>Toward a Trinity Model of Digital Education Resources Construction and Management . . . . .</b>	883
	Yong Huang and Qingchun Hu	
<b>102</b>	<b>Geographic Information System in the Cloud Computing Environment . . . . .</b>	893
	Yichun Peng and Yunpeng Wang	
<b>Part V Embedded Systems</b>		
<b>103</b>	<b>Memory Controller Design Based on Quadruple Modular Redundant Architecture . . . . .</b>	905
	Yuanyuan Cui, Wei Li, and Xunying Zhang	
<b>104</b>	<b>Computer Power Management System Based on the Face Detection . . . . .</b>	913
	Li Xie, Yong He, Yanfang Tian, and Tinghong Yang	
<b>105</b>	<b>Twist Rotation Deformation of Titanium Sheet Metal in Laser Curve Bending Based on Finite Element Analysis . . . . .</b>	921
	Peng Zhang, Qian Su, and Dong Luan	
<b>106</b>	<b>Voltage Transient Stability Analysis by Changing the Control Modes of the Wind Generator . . . . .</b>	929
	Yu Shao, Feng Shi, and Xiang Li	



**107 The Generator Stator Fault Analysis Based on the Multi-loop Theory . . . . . 939**  
Yu Shao, Feng Shi, and Xiang Li

**108 An Improved Edge Flag Algorithm Suitable for Hardware Implementation . . . . . 947**  
Lixiang Wang and Tiejun Xiao

**109 A Handheld Controller with Embedded Real-Time Video Transmission Based on TCP/IP Protocol . . . . . 955**  
Mingjie Dong, Wusheng Chou, and Yihan Liu

**110 Evaluating the Energy Consumption of InfiniBand Switch Based on Time Series . . . . . 963**  
Huifeng Wang, Zhanhuai Li, Xiaonan Zhao, Qinlu He, and Jian Sun

**111 Real-Time Filtering Method Based on Neuron Filtering Mechanism and Its Application on Robot Speed Signals . . . . . 971**  
Wa Gao, Fusheng Zha, Baoyu Song, Mantian Li, Pengfei Wang, Zhenyu Jiang, and Wei Guo

**112 Multiple-View Spectral Embedded Clustering Using a Co-training Approach . . . . . 979**  
Hong Tao, Chenping Hou, and Dongyun Yi

**113 Feedback Earliest Deadline First Exploiting Hardware Assisted Voltage Scaling . . . . . 989**  
Chuansheng Wu

**114 Design and Realization of General Interface Based on Object Linking and Embedding for Process Control . . . . . 997**  
Jiguang Liu, Jianbing Wu, and Zhiguo He

**115 A Stateful and Stateless IPv4/IPv6 Translator Based on Embedded System . . . . . 1007**  
Yanlin Yin and Dalin Jiang

**116 A Novel Collaborative Filtering Approach by Using Tags and Field Authorities . . . . . 1017**  
Zhi Xue, Yaoxue Zhang, Yuezhi Zhou, and Wei Hu

**117 Characteristics of Impedance for Plasma Antenna . . . . . 1027**  
Bo Yin and Feng Yang

**118 A Low-Voltage 5.8-GHz Complementary Metal Oxide Semiconductor Transceiver Front-End Chip Design for Dedicated Short-Range Communication Application . . . . . 1035**  
Jhin-Fang Huang, Jiun-Yu Wen, and Yong-Jhen Jiangn

**119 A 5.8-GHz Frequency Synthesizer with Dynamic Current-Matching Charge Pump Linearization Technique and an Average Varactor Circuit . . . . . 1045**  
 Jhin-Fang Huang, Jia-Lun Yang, and Kuo-Lung Chen

**120 Full-Wave Design of Wireless Charging System for Electronic Vehicle . . . . . 1055**  
 Yongxiang Liu, Yi Ren, and Yi Wang

**121 A Hierarchical Local-Interconnection Structure for Reconfigurable Processing Unit . . . . . 1063**  
 Yujia Zou, Leibo Liu, Shouyi Yin, Min Zhu, and Shaojun Wei

**122 High Impedance Fault Location in Distribution System Based on Nonlinear Frequency Analysis . . . . . 1073**  
 Jinqian Zhai, Di Su, Wenjian Li, Feng Li, and Guohong Zhang

**123 Early Fault Detection of Distribution Network Based on High-Frequency Component of Residual Current . . . . . 1083**  
 Jinqian Zhai, Di Su, Wenjian Li, Feng Li, and Guohong Zhang

**124 A Complementary Metal Oxide Semiconductor D/A Converter with R-2R Ladder Based on T-Type Weighted Current Network . . . . . 1091**  
 Junshen Jiao

**125 Detecting Repackaged Android Applications . . . . . 1099**  
 Zhongyuan Qin, Zhongyun Yang, Yuxing Di, Qunfang Zhang, Xinshuai Zhang, and Zhiwei Zhang

**126 Design of Wireless Local Area Network Security Program Based on Near Field Communication Technology . . . . . 1109**  
 Pengfei Hu and Leizhen Wang

**127 A Mechanism of Transforming Architecture Analysis and Design Language into Modelica . . . . . 1117**  
 Shuguang Feng and Lichen Zhang

**128 Aspect-Oriented QoS Modeling of Cyber-Physical Systems by the Extension of Architecture Analysis and Design Language . . . . . 1125**  
 Lichen Zhang and Shuguang Feng

**129 Using RC4-BHF to Construct One-way Hash Chains . . . . . 1133**  
 Qian Yu and Chang N. Zhang

**130 Leakage Power Reduction of Instruction Cache Based on Tag Prediction and Drowsy Cache . . . . . 1143**  
 Wei Li and Jianqing Xiao

**Part VI Network Optimization**

**131 The Human Role Model of Cyber Counterwork . . . . . 1155**  
Fang Zhou

**132 A Service Channel Assignment Scheme for IEEE 802.11p  
Vehicular Ad Hoc Network . . . . . 1165**  
Yao Zhang, Licai Yang, Haiqing Liu, and Lei Wu

**133 An Exception Handling Framework for Web Service . . . . . 1173**  
Hua Guan, Shi Ying, and Caoqing Jiang

**134 Resource Congestion Based on SDH Network Static  
Resource Allocation . . . . . 1181**  
Fuyong Liu, Jianghe Yao, Gang Wu, and Huanhuan Wu

**135 Multilayered Reinforcement Learning Approach for Radio  
Resource Management . . . . . 1191**  
Kevin Collados, Juan-Luis Gorricho, Joan Serrat, and Hu Zheng

**136 A Network Access Security Scheme for Virtual Machine . . . . . 1201**  
Mingkun Xu, Wenyuan Dong, and Cheng Shuo

**137 Light Protocols in Chain Network . . . . . 1209**  
Ying Wang, Yifang Chen, and Lenan Wu

**138 Research and Implementation of a Peripheral Environment  
Simulation Tool with Domain-Specific Languages . . . . . 1217**  
Maodi Zhang, Zili Wang, Ping Xu, and Yi Li

**139 Probability Model for Information Dissemination  
on Complex Networks . . . . . 1225**  
Juan Li and Xueguang Zhou

**140 Verification of UML Sequence Diagrams in Coq . . . . . 1233**  
Liang Dou, Lunjin Lu, Ying Zuo, and Zongyuan Yang

**141 Quantitative Verification of the Bounded  
Retransmission Protocol . . . . . 1245**  
Xu Guo, Ming Xu, and Zongyuan Yang

**142 A Cluster-Based and Range-Free Multidimensional  
Scaling-MAP Localization Scheme in WSN . . . . . 1253**  
Ke Xu, Yuhua Liu, Cui Xu, and Kaihua Xu

**143 A Resource Information Organization Method Based  
on Node Encoding for Resource Discovering . . . . . 1263**  
Zhuang Miao, Qianqian Zhang, Songqing Wang, Yang Li,  
Weiguang Xu, and Jiang Xiao

**144 The Implementation of Electronic Product Code System Based on Internet of Things Applications for Trade Enterprises . . . . . 1271**  
 Huiqun Zhao and Biao Shi

**145 The Characteristic and Verification of Length of Vertex-Degree Sequence in Scale-Free Network . . . . . 1281**  
 Yanxia Liu, Wenjun Xiao, and Jianqing Xi

**146 A Preemptive Model for Asynchronous Persistent Carrier Sense Multiple Access . . . . . 1289**  
 Lin Gao and Zhijun Wu

**147 Extended Petri Net-Based Advanced Persistent Threat Analysis Model . . . . . 1297**  
 Wentao Zhao, Pengfei Wang, and Fan Zhang

**148 Energy-Efficient Routing Protocol Based on Probability of Wireless Sensor Network . . . . . 1307**  
 Kaiguo Qian

**149 A Dynamic Routing Protocols Switching Scheme in Wireless Sensor Networks . . . . . 1315**  
 Zusheng Zhang, Tiezhu Zhao, and Huaqiang Yuan

**150 Incipient Fault Diagnosis in the Distribution Network Based on S-Transform and Polarity of Magnitude Difference . . . . . 1323**  
 Jinqian Zhai and Xin Chen

**151 Network Communication Forming Coalition S4n-Knowledge Model Case . . . . . 1331**  
 Takashi Matsuhisa

**152 An Optimization Model of the Layout of Public Bike Rental Stations Based on B+R Mode . . . . . 1341**  
 Liu He, Xuhong Li, and Dawei Chen

**153 Modeling of Train Control Systems Using Formal Techniques . . . . . 1349**  
 Bingqing Xu and Lichen Zhang

**154 A Clock-Based Specification of Cyber-Physical Systems . . . . . 1357**  
 Bingqing Xu and Lichen Zhang

**155 Polymorphic Worm Detection Using Position-Relation Signature . . . . . 1365**  
 Huihui Liang, Jiwen Chai, and Yong Tang

**156 Application of the Wavelet-ANFIS Model . . . . . 1373**  
 Rijun Zhang, Caishui Hou, Hui Lin, Meiyang Zhuo, Meixin Zhang, Zhongsheng Li, Liwu Sun, and Fengqin Lin

**157 Visualization of Clustered Network Graphs Based on Constrained Optimization Partition Layout . . . . . 1381**  
Fang Huang, Wenjie Xiao, and Hao Zhang

**158 An Ultra-Wideband Cooperative Communication Method Based on Transmitted Cooperative Reference . . . . . 1395**  
Tiefeng Li, Ou Li, and Zewen Zhou

**Author Index for Volume 1 . . . . . 1407**

**Subject Index for Volume 1 . . . . . 1411**

**Author Index for Volume 2 . . . . . 1417**

**Subject Index for Volume 2 . . . . . 1421**

# CENet 2013 Committee

## Advisory Chairs

Aniruddha Bhattacharjya     Amrita University, India

## Program Chairs

C. E. Tapie Rohm     California State University San Bernardino, USA  
W. Eric Wong     University of Texas at Dallas, USA  
Hong Jiang     Hohai University, China  
Jin Wang     Nanjing University of Information Science & Technology, China  
Tingshao Zhu     Chinese Academy of Sciences, China

## Program Committee

Abdaloussein Rezaei     ACECR (Academic Center for Education, Culture and Research) and Semnan University, Iran  
Akram Rashid     Air University, Islamabad, Pakistan  
Amit Joshi     Sardar Vallabhbhai National Institute of Technology, India  
Chen Hong     Beijing University of Aeronautics and Astronautics, China  
Fatimah De'nan     Universiti Sains Malaysia, Malaysia  
Feilong Liu     Hunan Institute of Science and Technology, China  
Feng Xu     Hohai University, China  
Fengjun Shang     Chongqing University of Posts and Telecommunications, China  
Gyanendra Prasad Joshi     Yeungnam University, Korea  
Hsieh Tzung-Yu     MingDao Univesity, Taiwan, China  
Jesús C. Hernandez     University of Jaen, Spain  
Jiandong Sun     Zhejiang University, China  
Qian Yu     University of Regina, Canada

RADJEF Mohammed Said	University of Bejaia, Algeria
Rui Chen	Xi'an University of Electronic Science and Technology, China
S. EL-Rabaie	Faculty of Electronic Engineering, 32952 Menouf, Egypt
Stephan Chalup	Heidelberg University, Germany
Sunil Kumar Khatri	Amity University, Uttar Pradesh, India
Tao Zhou	North China Electric Power University, China
Vong Chi-Man	University of Macau, Macau, China
Xianzhong Yi	Yangtze University, China
Yan Pei	Kyushu University, Japan
Yuan Gao	Tsinghua University, China
Yuduo Wang	Beijing Information Science and Technology University, China
Zhongtian Jia	University of Jinan, China

**Part I**  
**Algorithm Design**



# Chapter 1

## Simulation Algorithm of Adaptive Scheduling in Missile-Borne Phased Array Radar

Qizhong Li, Shanshan Sun, and Jianfei Zhao

**Abstract** In contrast to conventional radars, phased array radars have the capability to switch the direction of the radar beam very quickly without inertia. The measurements from phased array radar can contribute to many application fields such as data and intelligence process and radar performance evaluation. However, it often costs more than we can bear to obtain phased array radar measurements. It is necessary to model and simulate the phased array radar, especially for missile-borne phased array radar. This chapter lays a strong emphasis on the search and simulation technology of missile-borne phased array radar. According to operational theory of phased array radar, this chapter focuses on the functional modeling and simulation techniques. It contains three parts: beam arrangement of phased array radar in sin coordinate, parameter optimization of missile-borne phased array radar, and a function simulation model of missile-borne phased array radar.

### 1.1 Introduction

In ballistic missile defense system, the beam width of tracking radar is usually very narrow, and the dwell time is longer, so the implementation of routine large airspace search is not realistic. Only under the guide information of early warning system, the radar can intercept the target in small space [1]. The flight time of ballistic missile is usually short, and the lethality is enormous, which requires that the tracking radar can intercept the target as early as possible. Therefore, the missile-borne radar search strategy, especially on the wave search order, is very important.

Based on the previous research, considering the constraints of radar, we complete the simulation of adaptive scheduling method and make the constraints

---

Q. Li (✉) • S. Sun • J. Zhao  
North China Institute of Science and Technology, Beijing 101601, China  
e-mail: [li\\_qz@126.com.cn](mailto:li_qz@126.com.cn)

modeling for recursive form, through adjusting the scheduling of periodic inspection the constraint conditions [2]. This chapter presents the realization method and process of the algorithm in detail, through the scenario simulation, and verifies the effectiveness and practicability of the method.

## 1.2 Scheduling Model

### 1.2.1 Single Beam Search Simulation

In the distance occlusion problem under a single frequency, pulse Doppler waveform exists, which is mainly due to pulse Doppler radar seeker using the same antenna for transmitting and receiving. Transmitting and receiving is timeshare; when the transmitter sends electromagnetic wave, the receiver is in a closed state, which can avoid the leakage of the transmitter signal to the receiving system and burn the frequency receiver [3]. The high-frequency receiver uses gating switch to guarantee isolation in certain, if the target echo arrival seeker at the time of the transmitter sends pulse. Because the receiving system is in the closed state, the return signal cannot be received by target seeker in the “shelter area.” And while the missile is close to the target, the delay time of the target echo pulse with respect to the transmission pulse is fluxing. Thus, the blocking periodic phenomenon appears, which has entered the “transparent area,” “translucent area,” “shelter area,” “translucent area,” and then the “transparent area.” In the “translucent area,” the return signal is not affected; in the “shelter area,” it cannot receive the return signal completely; and in the “translucent area,” the width of echo pulse which is received by seeker gradually increases to the width of transmission pulse or from the width of transmitted pulse which decreases to zero [4].

We assume that the occlusion period is  $T_{zr}$ , width of “shelter area” is  $\tau_z$ , width of “transparent area” is  $\tau_m$ , and width of “translucent area” is  $\tau_{bm}$ ; they were calculated by the following equation:

$$T_{zr} = \frac{c}{2v_r} T_r \quad (1.1)$$

$$\tau_z = \frac{c}{2v_r} (T_r - \tau_r - \tau_l) \quad (1.2)$$

$$\tau_m = \frac{c}{2v_r} (\tau_r - \tau_l) \quad (1.3)$$

$$\tau_{bm} = \frac{c}{2v_r} \tau_l \quad (1.4)$$

In formula (1.1) to (1.4),  $C$  is the speed of light,  $v_r$  is the missile-target relative velocity,  $T_r$  is pulse width,  $\tau_r$  is received pulse width, and  $\tau_l$  is transmitted pulse width.

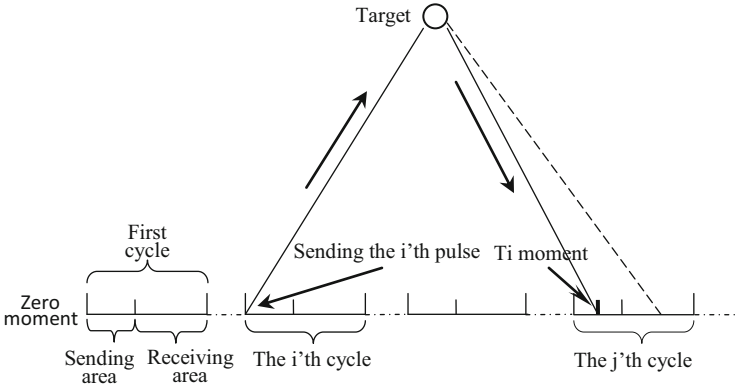


Fig. 1.1 The radar “eclipse” phenomenon mathematical model diagram

### 1.2.2 Establish Mathematical Model

According to the basic principle of the radar “eclipse” phenomenon, we can start from a simple mathematical model as shown in Fig. 1.1 and establish the idealized mathematical model of “eclipse” phenomenon step by step.

Assume that the target can be detected by radar, and the distance to the target is  $D$ , the radar frequency is  $f$ , the speed of light is  $C$ , pulse width of radar is  $T_r$ , the return moments of the  $i$  beam radar is  $T_i$ , and  $V_1, V_2 \dots V_i$  is the speed of the target in different intervals.

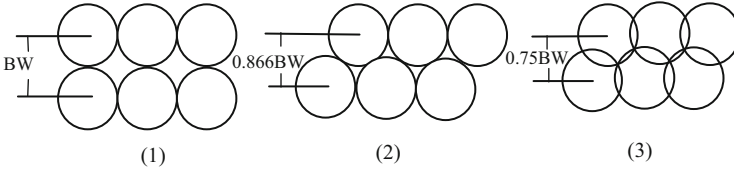
$$T_i = (i - 1) \times T_r + \frac{2(D - (V_1 + V_2 \dots + V_{i-1})) \times T_r}{C + v_i} \quad (1.5)$$

## 1.3 Multi-beam Search Simulation Modeling and Algorithm

### 1.3.1 Beam Position Arrangement Method

The antenna of phased array radar is fixed while scanning. When the scanning angle deviates from the normal direction, the beam will change. So the beam arrangement was usually complete in the sinusoidal spherical coordinates.

First step: determine the scanning space which was required in the radar station coordinates and change the space into the front spherical coordinates through coordinate transformation.



**Fig. 1.2** Beam accumulation method

**Table 1.1** Three waves of a style of the coverage and overlap rate

Arrangement	The wave number	Fraction of coverage (%)	Coincidence rate (%)
Vertical and horizontal arrangement wave	645	86.4	0
Staggered wave	783	98.7	0
Crisscross wave	920	100	3.56

Second step: change the space of the front spherical coordinates into front sine coordinate system and complete the beam arrangement in the sine coordinate system.

Third step: after the completion of beam arrangement in sine coordinate system, we can get the beam distribution in radar station coordinates by coordinate transformation.

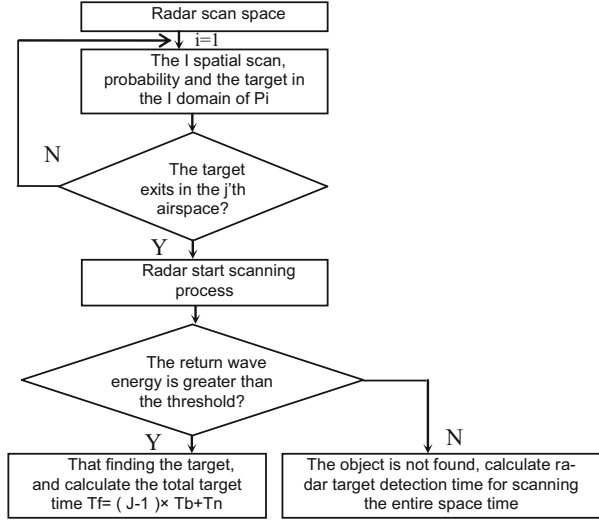
The optimal beam position arrangement aims at making full use of the prior information and the phased array radar system resources and makes the average time of finding target as short as possible. So we can think it is an optimization problem under constraints [5], as shown in Fig. 1.2 (Table 1.1).

From the table above, we know that three kinds of beam position arrangement each have advantages. The first arrangement needs the minimum number of wave, but the coverage rate is only 86.4 %. When the resource is limited or in the small target distribution density area, we can use it. Although the coverage rate of the third arrangement can reach 100 %, it is easy to cause the redundancy detection, which limits its application. In the second arrangement, the wave number and the coverage rate achieve a better balance, so it is a common beam position arrangement style. To make a long story short, the specific choice of what kind of arrangement style cannot lump together; we should think about the actual background and phased array radar system resources to make a reasonable choice.

### 1.3.2 Mathematical Model

According to the phased array radar beam position arrangement theory and calculation, we use the above three kinds of beam position arrangement. If the number of beam is  $N$  in the first forms, according to the theory, we need  $1.15N$  beam in the second forms and  $1.54N$  beam for the third forms.

**Fig. 1.3** Radar scanning airspace flow chart



Assuming a certain air-to-air missile radar scan area for  $[10^0,10^0]$ ,  $[10^0,10^0]$ , radar beam width is  $2^0$ , then the first form needs 50 beams in the area, the second form needs 58 beams, and the third form needs 77 beams.

So we can design a mode which has three different scanning forms. Assume that the target must be in the airspace that the radar will scan; to find the target or not, we will use the basic model of radar scanning to determine by target's frequency, radar's duty ratio, radar return wave threshold, and other factors. At the same time, we assume that the probability of the target appears in a wave of the space is different.

To take the first mode, for example, in the 50 wave, we will begin numbering from the largest probability of targets appearing as no. 1, 2 . . . N and the probability as P1, P2 . . . Pn (n is the wave number). We design an algorithm of target emerging randomly by the probability and make probability of the target emerging in the wave I as Pi. We set radar to start scanning from the maximum probability of waves until it finds the target. If it has scanned j wave from the start of radar scanning, the time the target is found is  $T_f = (J-1) \times T_b + T_n$ , where  $T_b$  is the time of scanning a wave and  $T_n$  is the time of finding the target in the jth wave, as shown in Fig. 1.3.

The simulation program uses C/C++ language to compile the simulation model of the function such as source code, establishes user interface in the VC++6.0 environment, tests results under various conditions through the simulation calculation, and uses the program for calibration and correction of simulation model [6]. We select Access as the background data management and ODBC for accessing dynamic target echo simulation database.

## 1.4 Conclusion

Using the above method, we can do the simulation for each beam position arrangement. For each method, we tested the time of radar target detection as  $T_f$  for three groups of 100 times, and then we calculated the average time as  $T_a$ . We can analyze the data to find the shortest average time which we take as the optimal model. Of course it should consider radar's frequency, target distance, duty ratio, and other factors. We need to complete simulation and calculation under the conditions of different target distance, frequency, and duty ratio and analyze large amounts of data to derive the optimal searching adaptive algorithm that can automatically adjust the radar's frequency and duty ratio in order to find the target at the shortest time under the changing target distance.

**Acknowledgments** A lot of thanks to the fundamental research funds for the Central Universities Fund (DX2013B01).

## References

1. Skolnik, M. I. (1990). *Radar handbook* (2nd ed., pp. 231–233). New York: McGraw-Hill.
2. Horne, G., & Meyer, T. (2004). Data farming: Discovering surprise. In *Proceedings of the 2004 Winter Simulation Conference* (pp. 807–813). Huntsville, AL: Society for Computer Simulation International.
3. Liao, S. Y., Lu, H. W., Chen, J., et al. (2006). Research on conceptual framework for agent-based modeling and simulation. *Journal of System Simulation (S1004-731X)*, 18(S2), 616–620 (In Chinese).
4. Wang, X., Li, D., Li, B., & Wang, X. S. (2010). Coherent video modeling and simulation method of phased array radar. *Journal of System Simulation (S1004-731X)*, 27(3), 741–747 (In Chinese).
5. Yang, L. B., Wang, X. S., Dan, M., & Xiao, S. P. (2004). Simulation and evaluation of phased array radar's tactical ballistic missile defense capability. *Journal of System Simulation (S1004-731X)*, 16(7), 1417–1426 (In Chinese).
6. Brandstein, A., & Horne, G. (2008). *Maneuver warfare science 2001 [EB/OL]*. Retrieved from <http://www.projectalbert.org/files/MWS2001On-line.pdf>

# Chapter 2

## A Second-Order Algorithm for Curve Parallel Projection on Parametric Surfaces

Xiongbing Fang and Hai-Yin Xu

**Abstract** A second-order algorithm is presented to calculate the parallel projection of a parametric curve onto a parametric surface in this chapter. The essence of our approach is to transform the problem of computing parallel projection curve on the parametric surface into that of computing parametric projection curve in the two-dimensional parametric domain of the surface. First- and second-order differential geometric characteristics of the parametric projection curve in the parametric domain of the surface are firstly analyzed. A marching method based on second-order Taylor Approximation is formulated to calculate the parametric projection curve. A first-order correction technique is developed to depress the error caused by the truncated higher order terms in the marching method. Several examples are finally implemented to demonstrate the effectiveness of the proposed scheme. Experimental results indicate that both the computational efficiency and accuracy of the presented method have dominant performance as compared with the first-order differential equation method.

### 2.1 Introduction

Curves on a surface have a wide range of applications in the fields of Computer Graphics, Computer-Aided Geometric Design, Computer Animation, CNC, etc. For instance, curves on a surface can be used for surface trimming [1], surface blending [2], NC tool path generation [3, 4], and so on. According to the designing manner, curves on a surface can be the intersection curve of two surfaces [4],

---

X. Fang (✉)  
China Ship Development and Design Center, Wuhan 430064, China  
e-mail: [fangxb2013@sina.cn](mailto:fangxb2013@sina.cn)

H.-Y. Xu  
School of Computer Science and Technology, Huazhong University of Science and Technology, Wuhan 430074, China

the offset of a given curve on a surface [3], the projection curve of a spatial curve onto a surface [5–9], the image of a curve in the parametric domain of a parametric surface [1, 2], or the fitting curve of a sequence of points lying on a surface [10].

In this chapter, we focus on computation of parallel projection of a parametric curve onto a parametric surface. Presently, there are two ways to do this problem. One is the first-order differential equation method [6] and the other is discrete method.

For the problem of calculating the parallel projection of parametric curves onto parametric surfaces, Wang, et al. transformed the condition of parallel projection into a system of differential equations and then formulated the problem as a first-order initial value problem. Numerical methods such as Runge–Kutta and Adams–Bashforth can be utilized to solve the initial value problem to generate a sequence of points. Under this transformation, difficulties lie in the choice of an accurate initial value and the stability of the adopted numerical method. For the discrete method, the parametric curve should first be discretized into a series of points. Parallel projections of these separated points can be calculated through the technique of intersection of a line with a surface. Usually, the computational efficiency of the discrete method depends on that of the intersection algorithm. Though the current projected point can be taken as the initial value of the next iteration, efficiency of the discrete method is generally slow as it does not fully utilize the differential geometric properties of both the parametric curve and the projection curve. The projection curve can be constructed by fitting the projected points generated by the two aforementioned types of approaches.

A second-order algorithm is put forward for tracing the parallel projection of a parametric curve onto a parametric surface. Experimental results show that the proposed scheme has dominant performance in both efficiency and computational accuracy as compared with Wang’s approach [6]. The rest of the chapter is organized as follows. An overview for our approach is presented in the next section. A second-order technique with error adjustment for tracing the projection curve is given in Sect. 2.3. Implementation and experimental results of our approach are carried out in Sect. 2.4 and we conclude the chapter in Sect. 2.5.

## 2.2 Overview

A 3D parametric curve  $\mathbf{p}(t)$  and a parametric surface  $\mathcal{S}(u, w)$  given in Fig. 2.1a are represented as

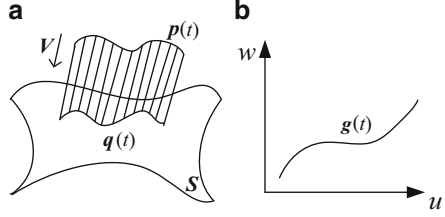
$$\mathbf{p}(t) = [x(t) \ y(t) \ z(t)] \quad (2.1)$$

and

$$\mathcal{S}(u, w) = [x(u, w) \ y(u, w) \ z(u, w)] \quad (2.2)$$



**Fig. 2.1** Parallel projection of 3D curve onto surface



respectively. Suppose  $p$  is a point on the curve  $p(t)$  and the  $q$  is the corresponding point generated by projecting the point  $p$  along with the direction  $V$  onto the surface  $S$ . While the point  $p$  moving along the curve  $p(t)$ , its parallel projection point  $q$  also moves along a curve on the surface, which is termed as the parallel projection curve of  $p(t)$  along with the direction  $V$  onto the surface  $S$ . According to the definition of parallel projection, one has

$$(q(t) - p(t)) \times V = 0 \quad (2.3)$$

where “ $\times$ ” denotes the cross product of two vectors. As the point lies on the parametric surface  $S(u, w)$ , the curve  $q(t)$  can be represented as  $q(t) = [x(u(t), w(t)), y(u(t), w(t)), z(u(t), w(t))]$ .

From the above analysis, there is a curve  $g(t) = [u(t), w(t)]$  in the parametric domain of the surface (please see Fig. 2.1b), which has a one-to-one corresponding relationship with the parallel projection curve on the surface. Thus a one-to-one corresponding relationship exists between the 3D curve  $p(t)$  and the 2D curve  $g(t)$ . For the convenience of description, we call  $g(t)$  as parametric projection curve in the remainder.

In this chapter, we transform the problem of computing parallel projection curve on parametric surface into the one of tracing parametric projection curve in 2D  $u$ – $w$  parametric domain. In Sect. 2.3.1, the first- and second-order differential quantities of the parametric projection curve are analyzed. A second-order iteration method for marching a series of points on the parametric projection curve based on Taylor Approximation is established and two methods for choosing the iterative step size are developed in Sect. 2.3.2. Considering the error caused by truncated higher order terms, a simple and efficient way to decrease the error is put forward in Sect. 2.3.3.

## 2.3 Parallel Projection Marching

### 2.3.1 The First- and Second-Order Differential Quantities of the Parametric Projection Curve

In order to calculate the first-order differential quantities, i.e.,  $u_t$  and  $w_t$ , of the parameters  $u$  and  $w$  with respect to the parameter  $t$  of the 3D curve, differentiating Eq. (2.3) with respect to  $t$  produces

$$(\mathbf{S}_u \times \mathbf{V})u_t + (\mathbf{S}_w \times \mathbf{V})w_t = \mathbf{p}_t \times \mathbf{V} \quad (2.4)$$

Taking the cross product of the  $\mathbf{S}_w$  and both sides of the Eq. (2.4), one has

$$([\mathbf{S}_u \times \mathbf{V}] \cdot \mathbf{S}_w)u_t = [\mathbf{p}_t \times \mathbf{V}] \cdot \mathbf{S}_w \quad (2.5)$$

Substituting  $[\mathbf{S}_u \times \mathbf{V}] \cdot \mathbf{S}_w = -(\mathbf{S}_u \times \mathbf{S}_w) \cdot \mathbf{V}$  into Eq. (2.5), we have

$$u_t = \frac{-\mathbf{L} \cdot \mathbf{S}_w}{[\mathbf{S}_u \times \mathbf{S}_w] \cdot \mathbf{V}} \quad (2.6)$$

where  $\mathbf{L} = \mathbf{p}_t \times \mathbf{V}$ . Similarly, dot-multiplying Eq. (2.4) by  $\mathbf{S}_u$  gives

$$w_t = \frac{\mathbf{L} \cdot \mathbf{S}_u}{[\mathbf{S}_u \times \mathbf{S}_w] \cdot \mathbf{V}} \quad (2.7)$$

One can continue to differentiate Eq. (2.3) with respect to the parameter  $t$

$$\left(\frac{d\mathbf{S}_u}{dt} \times \mathbf{V}\right)u_t + (\mathbf{S}_u \times \mathbf{V})u_{tt} + \left(\frac{d\mathbf{S}_w}{dt} \times \mathbf{V}\right)w_t + (\mathbf{S}_w \times \mathbf{V})w_{tt} = \mathbf{p}_{tt} \times \mathbf{V} \quad (2.8)$$

Since  $\frac{d\mathbf{S}_u}{dt} = \mathbf{S}_{uu}u_t + \mathbf{S}_{uw}w_t$  and  $\frac{d\mathbf{S}_w}{dt} = \mathbf{S}_{uw}u_t + \mathbf{S}_{ww}w_t$ , Eq. (2.8) can be rewritten as

$$\begin{aligned} & \left[ \mathbf{S}_{uu}(u_t)^2 + 2\mathbf{S}_{uw}u_t w_t + \mathbf{S}_{ww}(w_t)^2 \right] \times \mathbf{V} + (\mathbf{S}_u \times \mathbf{V})u_{tt} + (\mathbf{S}_w \times \mathbf{V})w_{tt} \\ & = \mathbf{p}_{tt} \times \mathbf{V} \end{aligned} \quad (2.9)$$

Dot-multiplying Eq. (2.9) by  $\mathbf{S}_w$  and  $\mathbf{S}_u$ , respectively, gives

$$\begin{cases} u_{tt} = \frac{-\mathbf{J} \cdot \mathbf{S}_w}{(\mathbf{S}_u \times \mathbf{S}_w) \cdot \mathbf{V}} \\ w_{tt} = \frac{\mathbf{J} \cdot \mathbf{S}_u}{(\mathbf{S}_u \times \mathbf{S}_w) \cdot \mathbf{V}} \end{cases} \quad (2.10)$$

where  $\mathbf{J} = [\mathbf{p}_{tt} - \mathbf{S}_{uu}(u_t)^2 - 2\mathbf{S}_{uw}u_t w_t - \mathbf{S}_{ww}(w_t)^2] \times \mathbf{V}$ , and  $u_t, w_t$  are computed by Eqs. (2.6) and (2.7), respectively.

### 2.3.2 Parametric Projection Curve Marching

In this section, we will use a second-order Taylor Approximation method to trace the parametric projection curve  $\mathbf{g}(t)$  to generate a series of points.

Let  $\mathbf{g}_i = (u_i, w_i)$  and  $\mathbf{g}_{i+1} = (u_{i+1}, w_{i+1})$  be the current and the next parametric projected point in the  $u$ - $w$  domain respectively. Then the position of point  $\mathbf{g}_{i+1}$  can be iteratively calculated through the current projected point and the first- and second-order differential quantities  $u_t, w_t, u_{tt}, w_{tt}$  at the point as follows

$$[u_{i+1}, w_{i+1}] = [u_i, w_i] + [u_t, w_t][k + 1/2]u_{tt}, w_{tt}k^2 \quad (2.11)$$

where  $k$  is a constant iteration step size.

If  $k = \Delta t$  is constant, one can utilize the Eq. (2.11) to get a sequence of points  $(u_i, w_i)$  on the parametric projection curve  $\mathbf{g}(t)$ , where  $i = 1, 2, \dots$ . However, the distribution of the projected points may not be well-proportioned. In order to fully use the differential geometric characteristics of  $\mathbf{g}(t)$  and  $\mathbf{p}(t)$ , we will give two ways to choose the iteration step size.

One can consider marching along the parametric projection curve based on a constant step size  $v_p$  along the 3D parametric curve. Suppose  $r$  is the arc-length parameter of the 3D parametric curve, one has

$$dr = \|\dot{x}_t \ \dot{y}_t \ \dot{z}_t\|dt = \|\mathbf{p}_t\|dt \quad (2.12)$$

Marching along the 3D parametric curve uses

$$t_{i+1} = t_i + k \text{ with } k = \frac{1}{\|\mathbf{p}_t\|}v_p \quad (2.13)$$

in addition to using Eqs. (2.6), (2.7), (2.10) and (2.11).

For marching along the parametric projection curve based on constant step size  $v_q$  along the projection curve, one can write

$$ds = \|\mathbf{q}_t\|dt \quad (2.14)$$

where  $s$  is the arc-length parameter of the parallel projection curve and  $\mathbf{q}_t = \mathbf{S}_u u_t + \mathbf{S}_w w_t$ . Marching along the parallel projection curve uses

$$t_{i+1} = t_i + k \text{ with } k = \frac{1}{\|\mathbf{S}_u u_t + \mathbf{S}_w w_t\|}v_q \quad (2.15)$$

in addition to using Eqs. (2.6), (2.7), (2.10) and (2.11).

### 2.3.3 Error Adjustment

Considering the truncated higher order terms, the position of point computed through Eq. (2.11) may depart from the parametric projection curve  $\mathbf{g}(t)$  in  $u$ - $w$  domain. Hence, a first-order adjustment technique is presented to reduce the error.

Suppose  $\hat{\mathbf{g}}_{i+1} = (\hat{u}_{i+1}, \hat{w}_{i+1})$  be the parametric projection point computed by Eq. (2.11)

$$[\hat{u}_{i+1}, \hat{w}_{i+1}] = [u_i, w_i] + [u_t, w_t]k + 1/2[u_{tt}, w_{tt}]k^2$$

As the above iteration formula omits higher order terms, the image point of  $\hat{\mathbf{g}}_{i+1}$  may deviate from the parallel projection curve on the parametric surface  $\mathcal{S}$ , i.e.,

$$(\hat{\mathbf{q}}_{i+1} - \mathbf{p}_{i+1}) \times \mathbf{V} = \boldsymbol{\varepsilon} \neq \mathbf{0} \quad (2.16)$$

where  $\hat{\mathbf{q}}_{i+1} = \mathcal{S}(\hat{u}_{i+1}, \hat{w}_{i+1})$ ,  $\mathbf{p}_{i+1} = \mathbf{p}(t_{i+1})$ .

Let  $[\Delta u, \Delta w]$  be the correction vector, the corrected parametric projection point and parallel projection point be  $\mathbf{g}_{i+1} = (u_{i+1}, w_{i+1}) = (\hat{u}_{i+1} + \Delta u, \hat{w}_{i+1} + \Delta w)$  and  $\mathbf{q}_{i+1} = \mathcal{S}(u_{i+1}, w_{i+1})$ , respectively. According to the definition of parallel projection, one has

$$(\mathbf{q}_{i+1} - \mathbf{p}_{i+1}) \times \mathbf{V} = \mathbf{0} \quad (2.17)$$

In order to calculate the error adjustment vector  $[\Delta u, \Delta w]$ , we use a first-order Taylor formula to expand the point  $\mathbf{q}_{i+1}$  at  $\hat{\mathbf{g}}_{i+1}$ , i.e.,  $\mathbf{q}_{i+1} = \mathcal{S}(\hat{u}_{i+1}, \hat{w}_{i+1}) + \mathbf{S}_u \Delta u + \mathbf{S}_w \Delta w$ . Substituting  $\mathbf{q}_{i+1}$  into Eq. (2.17) produces

$$(\mathbf{S}_u \Delta u + \mathbf{S}_w \Delta w) \times \mathbf{V} = -\boldsymbol{\varepsilon} \quad (2.18)$$

Dot-multiplying Eq. (2.18) by  $\mathbf{S}_w$  and  $\mathbf{S}_u$  respectively, gives

$$\begin{cases} \Delta u = \frac{\boldsymbol{\varepsilon} \cdot \mathbf{S}_w}{(\mathbf{S}_u \times \mathbf{S}_w) \cdot \mathbf{V}} \\ \Delta w = \frac{-\boldsymbol{\varepsilon} \cdot \mathbf{S}_u}{(\mathbf{S}_u \times \mathbf{S}_w) \cdot \mathbf{V}} \end{cases} \quad (2.19)$$

Note that all variables in Eq. (2.19) are evaluated at point  $\hat{\mathbf{g}}_{i+1}$  and  $t_{i+1}$ . When the deviation  $\|\boldsymbol{\varepsilon}\|$  calculated by Eq. (2.16) is greater than a given threshold  $\alpha$ , i.e.,  $\|\boldsymbol{\varepsilon}\| > \alpha$ , one can adopt the error adjustment vector calculated by Eq. (2.19) to adjust the parametric projection point  $\hat{\mathbf{g}}_{i+1}$ . In the following, the quantities  $\|\boldsymbol{\varepsilon}\|$  and  $\alpha$  are called as projection error and projection error threshold, respectively.

Suppose  $\mathbf{g}_i$  and  $\mathbf{g}_{i+1}$  are two adjacent points after error correction lying on the parametric projection curve  $\mathbf{g}(t) = [u(t) \ w(t)]$ . One can use a line segment to connect the two points  $\mathbf{g}_i$  and  $\mathbf{g}_{i+1}$  as follows:

$$\begin{cases} u = t \quad (u_i \leq t \leq u_{i+1}) \\ w = w_i + l(t - u_i) \end{cases} \quad (2.20)$$

where  $l = (w_{i+1} - w_i)/(u_{i+1} - u_i)$ . A  $G^0$  continuous parallel projection curve is acquired by substituting Eq. (2.20) into Eq. (2.2). One can also use the  $G^1$  approximation method [10] to fit the sequence of points  $S(u_i, w_i)$ ,  $i = 1, 2, \dots$ , to get a parallel projection curve with  $G^1$  or higher continuity.

## 2.4 Demonstrations

Our parallel projection algorithm framework is outlined in Sect. 2.4.1 and a number of examples are given to demonstrate the validity of our proposed method in Sect. 2.4.2. The examples make use of cubic NURBS curves and bicubic NURBS surfaces. In this context, these are viewed as parametric curves and surfaces, respectively.

### 2.4.1 Outline of Our Parallel Projection Algorithm

Given the control points and knot vectors for a NURBS curve and a NURBS surface, we need to compute a series of parallel projection points on the surface.

---

#### Algorithm 1. Parallel projection of parametric curves onto parametric surfaces

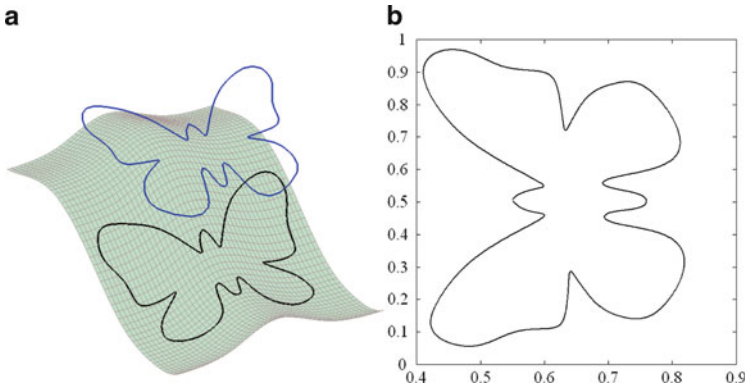
---

**Input:** Parametric curve  $p(t)$ , parametric surface  $S(u, w)$ , the initial projection point  $p_0=p(t_0)$  and its parallel projection point  $q_0=S(u_0, w_0)$  on surface  $S$ , projection error threshold  $\alpha$

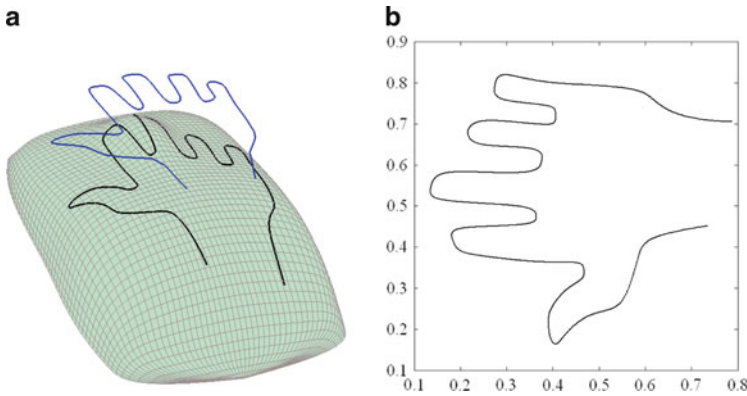
**Output:** A series of parallel projection points  $q_i=S(u_i, w_i)$  ( $i=1,2,\dots$ ) on surface  $S$ .

**Algorithm Description:**

1.  $t_i = t_0$ ,  $(u_{i+1}, w_{i+1}) = (u_0, w_0)$ ,  $p_i = p_0$ ,  $q_i = q_0$ ;
  2. **do** {
  3.   Compute step size  $k$ ,  $t_{i+1} = t_i + k$  and  $p(t_{i+1})$ ;
  4.   Compute  $(u_{i+1}, w_{i+1})$  and point  $q_{i+1} = S(u_{i+1}, w_{i+1})$ ;
  5.   Compute the projection error  $\varepsilon$ ;
  6.   **while**(  $\|\varepsilon\| > \alpha$  ) **do** {
  7.     Compute the adjustment vector  $[\Delta u, \Delta w]$  with equation (19);
  8.     Compute  $(\hat{u}_{i+1}, \hat{w}_{i+1}) = (u_{i+1}, w_{i+1}) + (\Delta u, \Delta w)$  and  $\hat{q}_{i+1} = S(\hat{u}_{i+1}, \hat{w}_{i+1})$ ;
  9.     Renew  $(u_{i+1}, w_{i+1})$ :  $(u_{i+1}, w_{i+1}) = (\hat{u}_{i+1}, \hat{w}_{i+1})$  and  $q_{i+1} = \hat{q}_{i+1}$ ;
  10.    Recalculate the projection error  $\varepsilon$  with new parameters  $(u_{i+1}, w_{i+1})$  and  $q_{i+1}$ ;
  11.   } **end while**
  12. Renew  $t$ :  $t_i = t_{i+1}$ ,  $(u_i, w_i) = (u_{i+1}, w_{i+1})$  and  $q_i = q_{i+1}$ ;
  13. Output the parallel projection point  $q_i$ ;
  14. } **while**( $p(t_i)$  is at the end of the 3D parametric curve  $p(t)$ )
-



**Fig. 2.2** Parallel projection example 1. (a) A butterfly curve onto an undee surface, and (b) parametric projection curve in  $u$ - $w$  domain

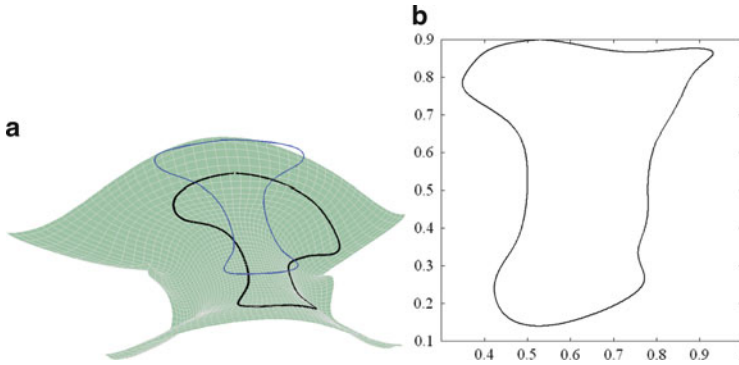


**Fig. 2.3** Parallel projection example 2. (a) A hand-form curve onto a mouse surface, and (b) parametric projection curve in  $u$ - $w$  domain

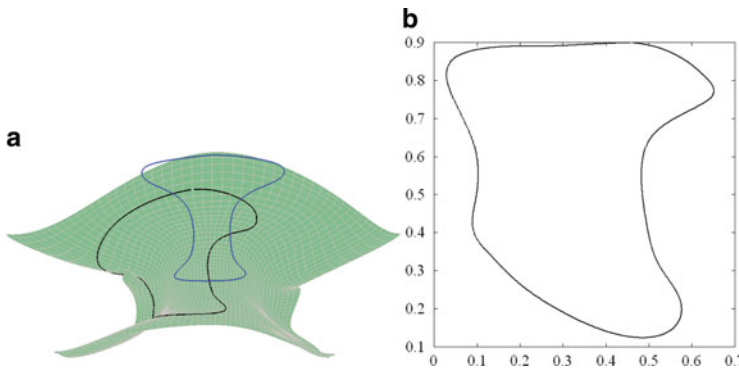
### 2.4.2 Examples

In this section we firstly give four examples (please see Figs. 2.2, 2.3, 2.4, and 2.5) obtained by the **Algorithm 1**. Further detailed comparisons of our scheme with Wang's first-order differential equation method [6] are carried out to testify the computational efficiency and accuracy of our methods (please see Tables 2.1, 2.2, 2.3, and 2.4). All the examples are implemented with Matlab, and run on a PC with 2.80 GHz CPU and 1 GB memory.

Figure 2.2a shows an example of projecting a closed butterfly curve onto an undee surface. The corresponding parametric projection curve in the  $u$ - $w$  parametric domain is shown in Fig. 2.2b. The second example shows the projection of a small hand-form curve onto a mouse surface. The hand-form curve and its parallel



**Fig. 2.4** Parallel projection example 3. (a) A closed NURBS curve onto a ridge surface, and (b) parametric projection curve in  $u$ - $w$  domain



**Fig. 2.5** Parallel projection example 4. (a) A closed NURBS curve onto a ridge surface, and (b) parametric projection curve in  $u$ - $w$  domain

projection curve on the NURBS surface is shown in Fig. 2.3a. Figure 2.3b denotes the parametric projection curve in the  $u$ - $w$  parametric domain.

Figures 2.3a and 2.4a are examples of projecting a same closed NURBS curve onto a same complex ridge NURBS surface along two different directions. The corresponding parametric projection curves in the  $u$ - $w$  parametric domain are shown in Figs. 2.3b and 2.4b.

For the parallel projection examples 1 and 2, we compared our method with Wang's first-order differential equation method [6] focusing on computational accuracy and efficiency (please see Tables 2.1, 2.2, 2.3, and 2.4). For the sake of fairness, the step size is specified by constant parametric increment  $\Delta t$  and the tolerance  $\alpha$  of the parallel projection error  $\epsilon$  is set as constant ( $\alpha$  is set as  $1.0e-009$  for all the comparisons). Moreover, the initial values for the two methods are the same. As the parallel projection error of Wang's method may overrun the error threshold  $\alpha$ , it is solved with the classical fourth-order Runge-Kutta method and our error adjustment technique deduced in Sect. 2.3.3.

**Table 2.1** Comparison of accuracy: projection of a butterfly curve onto an undee surface

Step size, $\Delta t$	APE of Wang's method	APE of our method	
	(no error adjustment)	(no error adjustment)	(error adjustment)
0.05	2.2598e-009	1.3133e-004	1.8662e-012
0.04	4.7598e-010	8.3957e-005	4.8905e-013
0.03	2.0282e-007	4.9389e-005	8.6769e-014
0.02	6.7689e-011	2.0942e-005	7.6519e-015

**Table 2.2** Comparison of accuracy: projection of a hand-form curve onto a mouse surface

Step size, $\Delta t$	APE of Wang's method	APE of our method	
	(no error adjustment)	(no error adjustment)	(error adjustment)
0.05	2.8146e-004	4.1585e-002	7.7697e-011
0.04	1.2407e-004	2.4439e-002	2.6262e-011
0.03	4.1559e-005	1.2371e-002	5.4997e-012
0.02	8.4125e-006	4.8313e-003	1.4454e-012
0.01	5.0854e-007	1.0423e-003	4.5111e-013

**Table 2.3** Comparison of efficiency: projection of a butterfly curve onto an undee surface

Step size, $\Delta t$	Wang's method (error adjustment)		Our method (error adjustment)	
	APE	CPU time (s)	APE	CPU time (s)
0.05	3.0445e-010	30.1094	1.8662e-012	27.2031
0.04	2.1330e-010	37.2500	4.8905e-013	33.0781
0.02	6.7689e-011	73.0781	7.6519e-015	66.1406

For comparison of accuracy of the two methods, the first-order method is realized just by using the classical four-order Runge–Kutta method, while our method is implemented in two ways that one is carried out with error adjustment and the other is without correction. From the results in Tables 2.1 and 2.2 (Note that the symbol “APE” in Tables 2.1, 2.2, 2.3, and 2.4 denotes average projection error), we can see that our method is much lower than Wang's in precision when implemented without error correction. After the error adjustment, the accuracy of our method is much finer than Wang's. From the results of large numbers of experiments, we found that the CPU time of our method with error adjustment is less than that of Wang's as projecting the same number of points onto a surface, while the precision of ours still has predominant performance.

Tables 2.3 and 2.4 are the comparison results of accuracy of our method with Wang's, which are gained under the conditions of same initial values, iteration step size, and error threshold. Both of the two methods adopted the error correction technique given in Sect. 2.3.3. From the column of CPU time in Tables 2.3 and 2.4, we can see that the efficiency of our method is about 1.1 times of Wang's. On the



**Table 2.4** Comparison of efficiency: projection of a hand-form curve onto a mouse surface

Step size, $\Delta t$	Wang's method (error adjustment)		Our method (error adjustment)	
	APE	CPU time (s)	APE	CPU time (s)
0.05	4.5873e-010	30.5469	7.7697e-011	26.9219
0.03	1.7461e-010	49.7031	5.4997e-012	43.7813
0.01	3.9048e-010	145.3438	4.5111e-013	131.0313

other hand, from the comparisons of Tables 2.1, 2.2 and Tables 2.3, 2.4, one can also find that our first-order error adjustment approach could efficiently improve the computational precision of the Wang's method.

## 2.5 Conclusion

A second-order algorithm based on Taylor Approximation is proposed to compute the parallel projection of a parametric curve onto a parametric surface. Several examples are presented to demonstrate the effectiveness of the presented approach. Experimental results and comparisons indicate that the computational efficiency of our method is about 1.1 times of that of the first-order differential equation method and our method has superior performance in the computational accuracy.

## References

1. Sederberg, T. W., Finnigan, G. T., Li, X., Lin, H. W., & Ipson, H. (2008). Watertight trimmed NURBS. *ACM Transaction on Graphics*, 27(3), 79:1–79:8.
2. Chuang, J. H., Lin, C. H., & Hwang, W. C. (1995). Variable-radius blending of parametric surfaces. *The Visual Computer*, 11(10), 513–525.
3. Tam, H.-Y., Law, H. W., & Xu, H.-Y. (2004). A geometric approach to the offsetting of profiles on the three-dimensional surfaces. *Computer-Aided Design*, 36(10), 887–902.
4. Xu, H.-Y., Tam, H.-Y., Fang, X., & Hu, L. (2009). Quart-parametric interpolations for intersecting paths. *Computer-Aided Design*, 41(6), 432–440.
5. Pegna, J., & Wolter, F. E. (1996). Surface curve design by orthogonal projection of space curves onto free-form surface. *Journal of Mechanical Design*, 118(1), 45–52.
6. Wang, X., Wei, W., & Zhang, W.-Z. (2010). Projecting curves onto free-form surfaces. *International Journal of Computer Applications in Technology*, 37(2), 153–159.
7. Wang, X., An, L. L., Zhou L. S., & Zhang, L. (2010). Constructing  $G^2$  continuous curve on freeform surface with normal projection. *Chinese Journal of Aeronautics*, 23(1), 137–144.
8. Xu, H.-Y., Fang, X., Hu, L., Xiao F., & Li, D. (2010). An algorithm for curve orthogonal projections onto implicit surfaces. *Journal of Computer-Aided Design and Computer Graphics*, 22(12), 2103–2110.
9. Xu, H.-Y., Fang, X., Tam, H.-Y., Wu, X., & Hu, L. (2012). A second-order algorithm for curve orthogonal projection onto parametric surface. *International Journal of Computer Mathematics*, 89(1), 98–111.
10. Yang, Y.-J., Zeng, W., Yang, C.-L., Meng, X.-X., Yong J.-H., & Deng, B. (2012).  $G^1$  continuous approximate curves on NURBS surfaces. *Computer-Aided Design*, 44(9), 824–834.

# Chapter 3

## Computation Method of Processing Time Based on BP Neural Network and Genetic Algorithm

Danchen Zhou and Chao Guo

**Abstract** Looking-up standard processing time table is a commonly used and important determination method of processing time. However, the large error in nonstandard nodes brings adverse effect on its accuracy. In view of the problem, a computation method of processing time based on back propagation neural network (BPNN) and genetic algorithm (GA) is proposed. Several key technologies of BPNN based on Matlab, including computation of the number of neurons in hidden layer, determination of training algorithm, and affecting factors of generalization ability, are researched in depth. In order to improve the training efficiency of BPNN, GA is used to optimize its connection weights and thresholds. The encoding method, selection operation, crossover, and mutation operation of GA are discussed in detail. The higher computation precision and faster operation speed of the proposed method is demonstrated through application cases.

### 3.1 Introduction

Processing time is the time consumption of per unit product or per unit work under the condition of certain production technology and organization manner. The main methods for determining processing time include experience estimation, probability estimation, statistic analysis, analogism comparison, and standard looking-up. On the basis of systematical and whole-set processing time standards, the method of standard looking-up obtains processing time of job, operation, or step in terms of decomposition of work elements and confirmation of one-to-one corresponding factor. As the criterion for determining processing time, processing time standards

---

D. Zhou (✉) • C. Guo  
Institute of Machinery Manufacturing Technology, China Academy of Engineering Physics,  
Mianyang 621900, Sichuan, China  
e-mail: [zdc69@163.com](mailto:zdc69@163.com)

**Table 3.1** An example of standard processing time table (min)

Width (B/mm)	Depth (t/mm)	Length (L/mm)								
		20	30	43	64	94	138	204	300	426
14	3	2.0	2.1	2.3	2.5	2.9	3.4	4.2	5.3	6.8
22	5	2.1	2.2	2.4	2.7	3.2	3.8	4.8	6.2	8.1
30	7	2.2	2.4	2.7	3.1	3.7	4.5	5.8	7.6	10.0
38	9	2.4	2.7	3.0	3.5	4.3	4.4	7.1	9.5	12.7
50	11	2.8	3.2	3.7	4.5	5.6	7.3	9.8	13.4	18.2
Working condition		Cutting tool: high-speed steel end mill								

generally are established based on multiple measurements by taking processing factors (surface roughness, cutting tool, length, width, depth, height, diameter, radius, volume, area, modulus, etc.) into full consideration under the premise of the same process conditions, such as machine, working environment, and personnel skill. Processing time standard is usually represented into the tabular form, which is called standard processing time table (SPTT). An example of SPTT with three factors (length, width, and depth) is shown in Table 3.1.

As for factors in standard nodes, processing time value can be obtained directly from SPTT. However, as for factors in nonstandard nodes, processing time value can only be obtained according to the principle of taking value of the closest to a standard node. Obviously, it is difficult to satisfy customers' requirements due to a large error. Aiming at the problem, for SPTT with only one factor, data fitting method can be used to transform SPTT into mathematical model. Unfortunately, the great majority of SPTT have more than one factor, each of whose influencing regularity on processing time is very difficult to be mastered at present. As a result, it is almost impossible to implement data fitting under the condition of indefinite curve equation.

Nowadays, artificial neural network (ANN) has been widely applied in function approach, data fitting, and structural optimization. Compared with general data fitting method, it has greater advantage and higher precision in training and computing data with no distinct regularity and multiple parameters. As a kind of multilayer feed-forward neural network based on error back propagation algorithm, back propagation neural network (BPNN) is the most popular neural network learning algorithm [1–3]. However, BPNN has some disadvantages such as over long training time and convergence difficulties, thanks to easily getting stuck in local minimum caused by randomness of connection weights and thresholds [4, 5]. Thus, this paper attempts to find an accurate and fast computation method of processing time based on the combination of BPNN and genetic algorithm (GA) according to the characteristics of SPTT.

## 3.2 GA–BPNN Algorithm

GA–BPNN algorithm is the combination of BPNN and GA. Based on the evolutionary ideas of natural selection and genetics, GA is a stochastic parallel search algorithm with high robust performance and strong global search ability [6–8]. In order to improve training efficiency and convergence rate of BPNN by overcoming the randomness of its initialization, GA is used to optimize the connection weights and thresholds of BPNN.

The main ideas of GA–BPNN algorithm are as follows. Firstly, population size and chromosome length of GA are determined in accordance with the structure of BPNN. Secondly, chromosome with the maximal fitness value is decoded as initial connection weights and thresholds of BPNN by means of the selection, crossover, and mutation operation of GA. Eventually, the satisfactory model of BPNN is determined through repeated training, by which the processing time can be accurately computed.

According to Kolmogorov theorem, a three-layer feed-forward neural network can exactly approach sample sets with regularity [5]. Accordingly, BPNN involved in this paper belongs to three-layer neural network, including input layer, hidden layer, and output layer, and research works are based on neural network toolbox and genetic algorithm toolbox in Matlab. In addition, all sample sets (SPTT) used in this paper are from processing time of petroleum machinery (SY/T 5179–93) in Chinese oil and natural gas industry standards.

## 3.3 BPNN-Based Computation Method of Processing Time

### 3.3.1 *Computation of the Number of Neurons in Hidden Layer*

For three-layer neural network, the number of neurons in input layer and output layer neurons is determined by sample sets. As for SPTT, the former is the number of factors in a table, and the latter is always one (processing time value). Hence, as long as the number of neurons in hidden layer has been decided, the structure of BPNN will be fixed. However, there is no theoretical guideline for its determination at present [8]. Consequently, some empirical formulas are used to determine its approximate range at first, and then, the model of BPNN meeting the error requirement is found out by gradual adjustment of the number of neurons in hidden layer and repeated training. Five tested empirical formulas for the number of neurons in hidden layer are listed as below.

$$h = \sqrt{0.43mn + 2.54m + 0.77n + 0.35 + 0.12n^2} + 0.51 \quad (3.1)$$

$$h = 2m + 1 \quad (3.2)$$

$$h = \sqrt{m + n} + a \quad (3.3)$$

$$h = \frac{m + n + 0.5n(m^2 + m)}{2} + \sqrt{p} + b \quad (3.4)$$

$$h = \text{round}\left(\frac{n(p-1)}{m+n+1}\right) \quad (3.5)$$

where  $h$  is the number of neurons in hidden layer,  $m$  is the number of neurons in input layer,  $n$  is the number of neurons in output layer (here,  $n = 1$ ),  $a$  is the integer between 0 and 10,  $p$  is the quantity of samples,  $b$  is 1 or  $-1$ , and  $\text{round}(\ )$  is integral function.

The experimental results show that, for five empirical formulas, Eq. (3.3) has the highest predicting accuracy for SPTT. However, because value obtained by Eq. (3.3) is only an estimation range, it will spend more training times to determine the most appropriate structure of BPNN. Therefore, Eq. (3.2) is selected as preferred formula because its accuracy is secondary to Eq. (3.3) and the value is definite. At the same time, if the structure of BPNN determined by Eq. (3.2) does not converge, or has poor generalization ability in the process of training, Eq. (3.3) will be selected.

### 3.3.2 Determination of Training Algorithm

Multiple training algorithms are provided by neural network toolbox in Matlab [9]. For selecting the appropriate training algorithms, an experiment scheme for evaluating training performance of every algorithm oriented to SPTT is designed. It is divided into two parts. The first part is to examine the training epochs (maximum is 5,000) and time of every algorithm under the condition of reaching specified convergence error (0.00001). The second part is to examine minimal convergence error and time of every algorithm under the condition of setting fixed training epochs (10,000). The comparisons of two parts are shown in Tables 3.2 and 3.3, respectively.

It can be seen that `trainlm` algorithm is the most excellent in training speed and convergence error among these algorithms. However, it is also discovered during the experiment that, when there are a large number of training samples, `trainlm` algorithm will take a large amount of resources of computer. By comprehensive consideration, `trainlm`, `trainidx`, and `trainrp` algorithms are selected as training algorithm of BPNN.

**Table 3.2** Comparison of training performance for different algorithms through first part of experiment

Algorithm name	Training epochs	Whether reaching specified convergence error	Time (s)
Traingd	5,000	No	29
Traingdm	5,000	No	31
Traingda	5,000	No	31.2
Traingdx	3,776	Yes	23.6
Trainrp	549	Yes	3.9
Trainlm	8	Yes	0.9
Trainbr	64	Yes	1.6

**Table 3.3** Comparison of training performance for different algorithms through the second part of experiment

Algorithm name	Training epochs	Minimal convergence error	Time (s)
Traingd	10,000	3.5E-3	56.5
Traingdm	10,000	4.6E-3	59.1
Traingda	10,000	2.7E-4	65.4
Traingdx	10,000	9.1E-5	59.6
Trainrp	10,000	1.6E-5	60.4
Trainlm	1,019	3.2E-28	10.9
Trainbr	6,965	5.6E-20	96

### 3.3.3 Affecting Factors of Generalization Ability

#### 3.3.3.1 Quality of Samples

Quality of samples denotes whether data in sample sets are accurate and their distributions are uniform. In other words, bad quality of samples will directly lead to poor generalization ability. Because all sample sets used in this paper are from SPTT, in which data have been verified by practice and have better regularity, it can be believed that quality of samples has less effect on generalization ability.

#### 3.3.3.2 Structure of BPNN

The structure of BPNN has greater impact on generalization ability, which is mainly reflected in the number of neurons in hidden layer. Namely, the appropriate number of neurons in hidden layer will make BPNN reach the specified convergence error with the least training epochs and meanwhile easily meet generalization error requirement. Through extensive experiments for SPTT, relations of the number of neurons in hidden layer, convergence error and generalization error are studied. An experimental result is shown in Fig. 3.1.

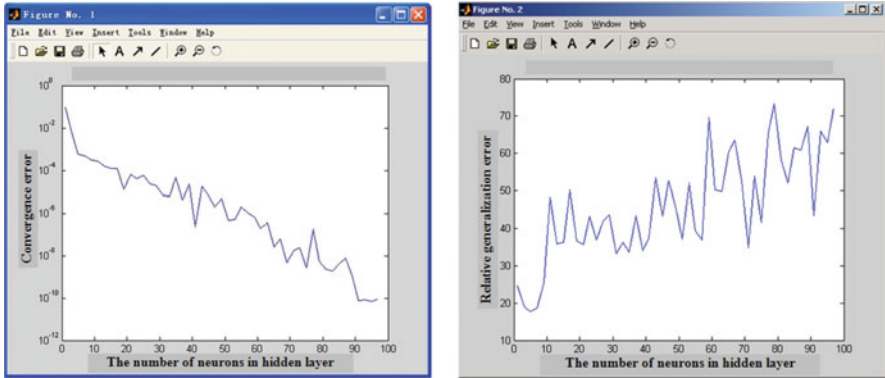
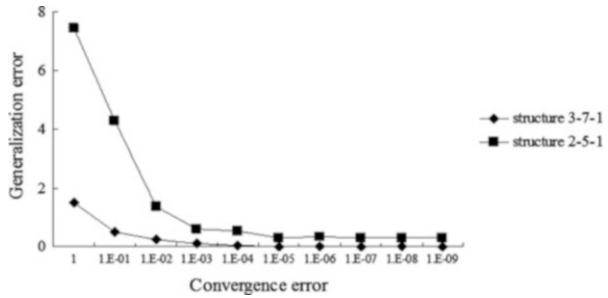


Fig. 3.1 Experimental result for structure of BPNN

Fig. 3.2 Relation between generalization error and convergence error



It can be seen from Fig. 3.1 that with the increase of number of neurons in hidden layer, the tendency is gradual reduction of convergence error and gradual increase of generalization error, but in the beginning, both of them reduce simultaneously.

### 3.3.3.3 Convergence Error

Experimental result of relation between generalization error and convergence error for two different structures of BPNN is shown in Fig. 3.2.

It can be seen from Fig. 3.2 that for a fixed structure of BPNN, generalization error also reduces correspondingly with the reduction of convergence error, which is maybe because the hidden regularity in training samples is learned in depth step-by-step. Due to higher correctness of the regularity guaranteed by higher reliability of data in SPTT, predicting results are more and more accurate.

### 3.4 GA-Based Optimization for BPNN

#### 3.4.1 Chromosome Encoding Method

For GA-based optimization for BPNN, binary encoding, real encoding, and Gray code encoding are dominant chromosome encoding methods. To test the performance of three encoding methods, a comparative experiment is conducted. Therefore, three SPTTs with 216, 60, and 72 data are selected as sample sets. The structures of BPNN are 4-9-1, 3-7-1, and 2-5-1, respectively. In the experiment, their fitness function is defined as reciprocal of the mean squared error (MSE), which is expressed in Eq. (3.6).

$$fitness = \frac{1}{1 + \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - d_i)^2}}, \quad (3.6)$$

where  $N$  represents the number of training samples and  $y_i$  and  $d_i$  represent the computing and expected processing time value, respectively. The time consumption and maximum of fitness values of three encoding methods are listed in Table 3.4.

It can be found from experimental result that three encoding methods have the approximate maximum of fitness value, but real encoding method consumes the least running time of program. Therefore, the real encoding is chosen as encoding method.

#### 3.4.2 Selection Operation

An improved ranking selection method is adopted in this paper. The steps are as below.

Step 1: Rank all chromosomes of initial parent population in descending order according to their own fitness values.

**Table 3.4** Performance comparison among three encoding methods

Comparison item	Sample set	Binary encoding	Real encoding	Gray code encoding
Time consumption (s)	SPTT1	809.940	718.000	3,562.4
	SPTT2	359.391	318.015	1,725.6
	SPTT3	329.830	299.580	1,124.7
Maximum of fitness values	SPTT1	0.998 65	0.998 75	0.994 92
	SPTT2	0.998 54	0.999 66	0.866 17
	SPTT3	0.999 74	0.999 50	0.942 84



- Step 2: Choose the top N chromosomes as initial child population for crossover and mutation operation.
- Step 3: Generate new child population through carrying out crossover and mutation operation for initial child population.
- Step 4: Generate new parent population through mixing new child population into initial parent population.
- Step 5: Rank all chromosomes of new parent population in descending order according to their own fitness values.
- Step 6: Choose the top N chromosomes as next-generation population.

### 3.4.3 Crossover and Mutation Operation

In this paper, single-point crossover method is adopted due to its low possibility of destroying structure of chromosome in population. At the same time, uniform mutation method is adopted because, on the one hand, it is suitable for real encoding, and on the other hand, it accords with the actual states of nature for the reason that the number in mutation range is randomly generated with the same possibility.

## 3.5 Application Cases

To verify the validity of the proposed method, two standard processing time tables with three factors, which have 65 and 60 data, respectively, are selected as sample sets. The structure of BPNN is 3-10-1 and 3-7-1, respectively. Parameters setting of BPNN and GA are shown in Table 3.5. Comparison between computing values and actual values of testing samples is shown in Table 3.6. The convergence curves of training error for two SPTTs are shown in Fig. 3.3.

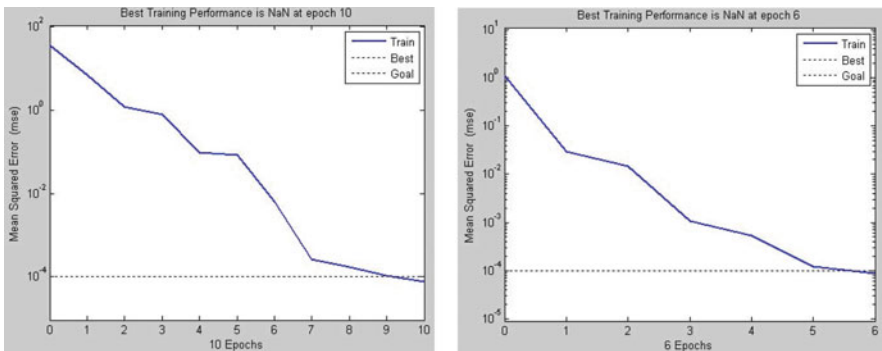
It can be calculated from Table 3.6 that average relative errors are 0.376 % and 0.620 % and the MSE are 0.016 and 0.048, respectively. The result proves the

**Table 3.5** Parameters setting of BPNN and GA

BPNN		GA	
Parameter	Value	Parameter	Value
Required generalization error	0.05	Population size	50
Training epochs	2,000	Maximum of generation	500
Specified convergence error	0.0001	Generation gap	0.7
Learning rate	0.7	Crossover probability	0.75
Momentum	0.5	Mutation probability	0.1
Training algorithm	Trainidx	Maximum and minimum of encoding	(2, -2)

**Table 3.6** Comparison between computing values and actual values of testing samples

Sample	SPTT1			SPTT2		
	Actual value	Computing value	Relative error	Actual value	Computing value	Relative error
1	3.34	3.3313	0.26	3.5	3.4945	0.16
2	3.39	3.3758	0.42	3.7	3.6882	0.32
3	3.43	3.4229	0.21	3.8	3.8675	1.78
4	3.44	3.4500	0.29	3.9	3.8520	1.23
5	3.45	3.4358	0.41	4.0	3.9800	0.50
6	3.48	3.4690	0.32	4.1	4.1051	0.12
7	3.52	3.5154	0.13	4.2	4.1863	0.33
8	3.54	3.5570	0.48	4.4	4.3905	0.22
9	3.63	3.6439	0.38	4.5	4.4620	0.84
10	3.64	3.6552	0.42	4.7	4.7640	0.98
11	3.67	3.6943	0.66	4.8	4.8068	0.14
12	3.74	3.7472	0.19	5.2	5.2194	0.37
13	3.83	3.8357	0.15	5.3	5.2637	0.68
14	3.93	3.9230	0.18	5.8	5.7651	0.6
15	4.03	4.0112	0.47	6.5	6.5149	0.23
16	4.13	4.1023	0.67	6.8	6.6884	1.64
17	4.28	4.2483	0.74	7.4	7.3693	0.42
18	4.38	4.3542	0.59	7.5	7.5163	0.22



**Fig. 3.3** Convergence curves of training error for SPTT1 and SPTT2

models of GA-BPNN have better generalization ability. Meanwhile, Fig. 3.3 shows that the GA-BPNN models can converge to the specified accuracy with much fewer training epochs. Therefore, application cases demonstrate that the computation method proposed in this paper has higher computation precision and faster operation speed and is capable of accurately determining processing time.

### 3.6 Conclusion

In view of the characteristics of SPTT, a computation method of processing time based on BPNN and GA is proposed, and depending on the optimization of GA for connection weights and thresholds of BPNN, training efficiency and convergence rate are obviously improved because the probability of getting stuck in local minimum is greatly decreased. Application cases have proved its higher computation precision and faster operation speed by means of in-depth studies on some key technologies.

### References

1. Ding, S., Su, C., Yu, J. (2011). An optimizing BP neural network algorithm based on genetic algorithm. *Artificial Intelligence Review*, 36(2), 153–162.
2. Wang, H., & Liu, M. (2012). Design of robotic visual servo control based on neural network and genetic algorithm. *International Journal of Automation and Computing*, 9(1), 24–29.
3. Wang, Z., Zhao, Z., Wang, L., Wang, K. (2011). Study on the law of short fatigue crack using genetic algorithm-BP neural networks. *Lecture Notes in Computer Science*, 6677(1), 586–593.
4. Long, J., Lan, F., Chen, J., Yu, P. (2009). Mechanical properties prediction of the mechanical clinching joints based on genetic algorithm and BP neural network. *Chinese Journal of Mechanical Engineering*, 22(1), 36–40.
5. Zhang, J., Xu, C., Yi, M., Fang, B. (2012). Design of nano-micro-composite ceramic tool and die material with back propagation neural network and genetic algorithm. *Journal of Materials Engineering and Performance*, 21(4), 463–470.
6. Lin, J. (2012). A systematic estimation model for fraction nonconforming of a wafer in semiconductor manufacturing research. *Applied Soft Computing Journal*, 12(6), 1733–1740.
7. Zemin, F., & Mo, J. (2011). Springback prediction of high-strength sheet metal under air bending forming and tool design based on GA-BPNN. *The International Journal of Advanced Manufacturing Technology*, 53(5–8), 473–483.
8. Xu, M., Jin, B., Yu, Y., Shen, H., Li, W. (2010). Using artificial neural networks for energy regulation based variable-speed electrohydraulic drive. *Chinese Journal of Mechanical Engineering*, 23(3), 327–335.
9. Zain, A. M., Haron, H., Sharif, S. (2010). Prediction of surface roughness in the end milling machining using Artificial Neural Network. *Expert Systems with Applications*, 37(2), 1755–1768.

# Chapter 4

## Integral Sliding Mode Controller for an Uncertain Network Control System with Delay

Zhenbin Gao

**Abstract** Integral sliding mode control is formulated with respect to an uncertain continuous network control system with the state delay. The parameter uncertainty is assumed to be the norm-bounded and satisfy the sliding mode matching requirements. The switching function is presented which include the integral term of the state feedback gain and the sliding mode compensator. The sliding mode controller is designed which is divided into the equivalent controller and switching controller, so the reachability of the sliding surface is ensured. A sufficient condition is derived by means of linear matrix inequality such that the asymptotical stability of the closed-loop system is guaranteed. The validity and feasibility of the proposed approach is investigated via the corresponding numerical simulation.

### 4.1 Introduction

A networked control system (NCS) is a feedback control system with network channels used for the communications between spatially distributed system components, such as sensors, actuators, and controllers. Since the signals are transmitted over a communication network of limited bandwidth, there exist network-induced delays, which may deteriorate the performance and stability of the closed-loop control system [1, 2].

Owing to the fact that the conventional control methods do not take into account the uncertainties of the network environment, the study of new control strategies dealing with this problem is of practical importance.

Sliding mode control (SMC) has been successfully used in controlling many uncertain systems; it is preferred because it's robust character and superior performance [3, 4]. During the last decades, numerous contributions to SMC theory have

---

Z. Gao (✉)

Department of Applied Mathematics, School of Statistic, Xi'an University of Finance and Economics, Xi'an 710100, China  
e-mail: [gaozb2700@sina.com](mailto:gaozb2700@sina.com)

been made. The switching schemes, putting the differential equations into canonical forms, and generating simple SMC strategies are considered in detail. The application of SMC scheme to robotic manipulator is studied, and the quality of the scheme is discussed from the point of robustness [5]. The performance of SMC scheme is proven to be satisfactory in the face of external disturbances and uncertainties in the system model representation.

A new sliding mode, called integral sliding mode, is proposed [6]. Having this feature, integral sliding mode control has been widely applied to various systems [7, 8].

In this paper, we design a novel integral switching function to tackle a continuous NCS which is comprised of the parameter uncertainties and the state delay. The sliding mode controller is achieved and the reachability of SMC is analyzed. The example is given to verify the feasibility and effectiveness.

## 4.2 Problem Formulation

In an NCS, the controlled plant with parameter uncertainty is a continuous system that can be described as follows:

$$\begin{aligned} \dot{x}(t) &= (A + \Delta A)x(t) + (A_d + \Delta A_d)x(t-d) + (B + \Delta B)u(t) \\ x(t) &= \phi(t), t \in [-d, 0] \end{aligned} \quad (4.1)$$

where  $x(t) \in R^n$  is the state vector;  $u(t) \in R^m$  is the control input;  $d > 0$  is the delay constant;  $\phi(t)$  is the original state vector defined in  $[-d, 0]$ ;  $A, A_d, B$  are matrices of the appropriate dimension; and  $\Delta A, \Delta A_d, \Delta B$  are uncertain matrices, which are satisfied with the condition of the norm-bounded, that is,

$$[\Delta A, \Delta A_d, \Delta B] = DF[E_1, E_d, E_2]$$

where  $D, E_1, E_d, E_2$  are the known matrices of the appropriate dimension and  $F$  is the unknown matrix which satisfies  $F^T F \leq I$ .

The sliding mode matching requirements are satisfied with

$$[\Delta A \quad \Delta A_d \quad \Delta B] = B[\Delta \tilde{A} \quad \Delta \tilde{A}_d \quad \Delta \tilde{B}]$$

Then, the system (4.1) can be formed as follows:

$$\dot{x}(t) = Ax(t) + A_d x(t-d) + B[u(t) + W(t)] \quad (4.2)$$

where  $W(t) = B^+[\Delta Ax(t) + \Delta A_d x(t-d) + \Delta Bu(t)]$  and  $B^+$  is the pseudo inverse, namely,  $B^+ = (B^T B)^{-1} B^T$ .

### 4.3 Main Results

The main idea is to find a sliding mode controller for the abovementioned systems (4.1) and (4.2); the corresponding normal system is

$$\dot{x}(t) = Ax(t) + A_d x(t-d) + Bu(t) \quad (4.3)$$

Define the integral switching function as

$$s(t) = C \left[ x(t) - \int_0^t (A + BK)x(t) dt \right] + T \quad (4.4)$$

where the state feedback gain  $K$  is unknown,  $C$  is the matrix with the appropriate dimension and satisfy  $CB > 0$ , and  $T$  is the sliding mode compensator which satisfies

$$\dot{T} = -CA_d x(t-d) \quad (4.5)$$

During sliding surface  $s(t) = \dot{s}(t) = 0$ , from Eqs. (4.3), (4.4), and (4.5), we have

$$\dot{s}(t) = C\dot{x}(t) - C(A + BK)x(t) - CA_d x(t-d) = CBu(t) - CBKx(t) = 0 \quad (4.6)$$

Then, the equivalent controller is

$$u_{eq}(t) = Kx(t) \quad (4.7)$$

#### 4.3.1 SMC Design

For the system defined by Eq. (4.2), the SMC law is written as

$$u(t) = u_{eq}(t) + u_N(t) \quad (4.8)$$

where  $u_N(t)$  is the switching controller which is used to overcome the system uncertainties; the form is selected as

$$u_N(t) = -\gamma s(t) - f \operatorname{sgn}(s(t)) \quad (4.9)$$

where  $f \geq |W(t)|$ ,  $\gamma > 0$  and  $\operatorname{sgn}(\bullet)$  is the sign function.

**Theorem 1** Consider the system (4.2), design the switching function as Eq. (4.4), if the SMC law is given by Eq. (4.8), then, the system (4.2) can be asymptotically stable.

*Proof* Suppose Lyapunov function is  $V(t) = \frac{1}{2}s^2(t)$ ,  $s(t)$  is described as Eq. (4.4), then

$$\begin{aligned}
\dot{V}(t) &= s\dot{s} = s[C\dot{x} - C(A + BK)x - CA_d x(t-d)] \\
&= s[CB(u(t) + W(t)) - CBKx(t)] \\
&= s[CB(u_{eq}(t) + u_N(t) + W(t)) - CBKx(t)] \\
&= s[CB(-\gamma s - f \operatorname{sgn}(s) + W(t))] = -CB[\gamma|s|^2 + f|s| - sW] \leq 0
\end{aligned}$$

Then, the SMC law is considered to guarantee the reachability of the sliding surface. Thus, the proof is complete.

### 4.3.2 Stability of the Sliding Surface

**Lemma** Given a scalar, if there exist matrices and satisfying, where, is an identity matrix, then the following matrix inequality holds:

$$DFE + E^T F^T D^T \leq \varepsilon DD^T + \varepsilon^{-1} E^T E$$

Put the equivalent controller Eq. (4.7) into system (4.1), we have the sliding mode dynamics equation as follows:

$$\dot{x}(t) = [(A + \Delta A) + (B + \Delta B)K]x(t) + (A_d + \Delta A_d)x(t-d) \quad (4.10)$$

**Theorem 2** For system (4.1), given constant  $\varepsilon > 0$ , select switching function (4.4). If exist symmetric definite matrices  $X, V$  and matrix  $W$ , satisfy linear matrix inequality (LMI) as follows:

$$\begin{bmatrix}
\Theta & A_d V & (E_1 X + E_2 W)^T & X \\
* & -V & V E_d^T & 0 \\
* & * & -\varepsilon I & 0 \\
* & * & * & -V
\end{bmatrix} < 0 \quad (4.11)$$

where  $\Theta = AX + XA^T + BW + W^T B^T + \varepsilon DD^T$ . The asterisk denotes the transpose of the corresponding block above the main diagonal, and  $I$  denote the identity matrix of appropriate dimension. Then, the equivalent controller is  $u_{eq}(t) = WX^{-1}x(t)$ , and the sliding mode dynamics system (4.10) is globally asymptotically stable.

*Proof* Suppose the Lyapunov function is

$$V(t) = x^T(t)Px(t) + \int_{t-d}^t x^T(\tau)Rx(\tau)d\tau$$

where  $P, R$  are symmetry definite matrices. Calculate  $\dot{V}(t)$  as follows:

$$\begin{aligned}
\dot{V}(t) &= 2x^T(t)P\dot{x}(t) + x^T(t)Rx(t) - x^T(t-d)Rx(t-d) \\
&= 2x^T(t)P[(A + DFE_1) + (B + DFE_2)K]x(t) + 2x^T(t)P(A_d + DFE_d)x(t-d) \\
&\quad + x^T(t)Rx(t) - x^T(t-d)Rx(t-d) \\
&= x^T(t)(PA + A^TP + PBK + K^TB^TP + R)x(t) + x^T(t)(PA_d + A_d^TP)x(t-d) \\
&\quad - x^T(t-d)Rx(t-d) + x^T(t)(PDFE_1 + PDFE_2K + E_1^TF^TD^TP + K^TE_2^TF^TD^TP)x(t) \\
&\quad + x^T(t)PDFE_dx(t-d) + x^T(t-d)E_d^TF^TD^TPx(t)
\end{aligned}$$

Using Lemma, we get

$$\begin{aligned}
&x^T(t)(PDFE_1 + PDFE_2K + E_1^TF^TD^TP + K^TE_2^TF^TD^TP)x(t) \\
&\leq x^T(t)\left[\varepsilon PDD^TP + \varepsilon^{-1}(E_1 + E_2K)^T(E_1 + E_2K)\right]x(t) \\
&\quad x^T(t)PDFE_dx(t-d) + x^T(t-d)E_d^TF^TD^TPx(t) \\
&\leq x^T(t)\varepsilon PDD^TPx(t) + \varepsilon^{-1}x^T(t-d)E_d^TE_dx(t-d)
\end{aligned}$$

Consequently, we obtain

$$\dot{V}(t) \leq \xi^T(t)\Phi\xi(t)$$

where  $\xi(t) = [x(t) \quad x(t-d)]$ , and

$$\Phi = \begin{bmatrix} Z & PA_d \\ A_d^TP & -R + \varepsilon^{-1}E_d^TE_d \end{bmatrix} \quad (4.12)$$

and  $Z = PA + A^TP + PBK + K^TB^TP + R + \varepsilon PDD^TP + \varepsilon^{-1}(E_1 + E_2K)^T(E_1 + E_2K)$

Thus,  $\dot{V}(t) < 0$  is satisfied if  $\Phi < 0$ . From Eq. (4.12) pre-multiply and post-multiply the matrix  $\text{diag}\{P^{-1}, P^{-1}\}$ ; the nest LMI can be derived:

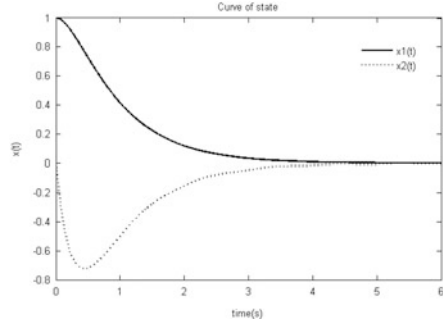
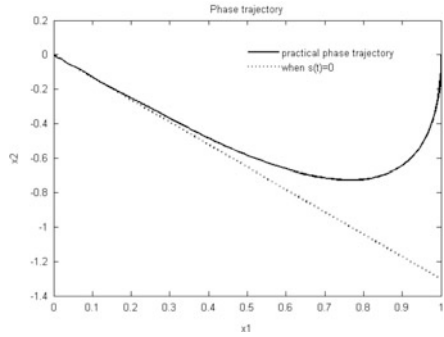
$$\begin{bmatrix} \Sigma & A_d & X(E_1 + E_2K)^T & X \\ * & -R & E_d^T & 0 \\ * & * & -\varepsilon I & 0 \\ * & * & * & -R^{-1} \end{bmatrix} < 0$$

where  $\Sigma = AX + XA^T + BKX + XK^TB^T + \varepsilon DD^T$ . Let  $W = KX$  and pre-multiply and post-multiply the  $\text{diag}\{I, R^{-1}, I, I\}$ , and set  $V = R^{-1}$ , we have the LMI (4.11). Thus, the proof is complete.

## 4.4 Experiment Studying

For the system (4.1), we have the system parameters as follows:



**Fig. 4.1** The curve of state**Fig. 4.2** The phase trajectory

$$A = \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 5 \end{bmatrix}, A_d = \begin{bmatrix} 0 & 0 \\ 0.1 & 0.1 \end{bmatrix}, C = [1.3 \quad 1], D = \begin{bmatrix} 0.5 \\ 0.1 \end{bmatrix}, E_1 = [0 \quad 0.1], \\ E_d = [0 \quad 0.1], E_2 = 0.1$$

the delay constant  $d = 3$ ; original state is  $X_0 = [1 \quad 0]^T$ .

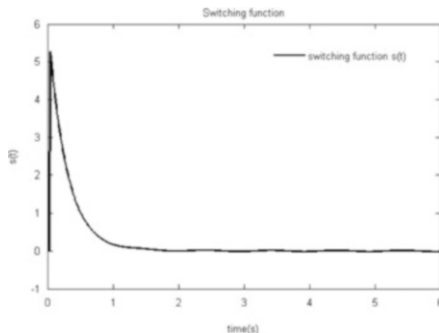
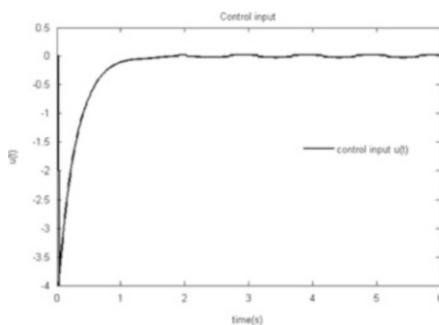
From the Eqs. (4.8) and (4.9), we have the SMC law  $u(t)$ . Set  $f = 0.01$ ,  $\gamma = 0.6$  and the perturbation  $W(t) = 0.005 \sin(2\pi t)$ . According to Theorem 2, using Matlab LMI toolbox, we have

$$X = \begin{bmatrix} 0.4002 & -0.2891 \\ -0.2891 & 0.3968 \end{bmatrix}, V = \begin{bmatrix} 1.1928 & -0.0078 \\ -0.0078 & 1.1850 \end{bmatrix}, \\ W = [-0.1470 \quad -0.0418]$$

then, the equivalent controller gain is  $K = WX^{-1} = -[0.9358 \quad 0.7870]$ .

The simulation results are shown in Figs. 4.1, 4.2, 4.3, and 4.4. The evolutions of two state variables  $x_1(t)$ ,  $x_2(t)$  are depicted in Fig. 4.1.

The motion in the phase plane is illustrated in Fig. 4.2. It shows that after a fast reaching mode, a sliding mode is maintained on the sliding surface  $s(t) = 0$  by the

**Fig. 4.3** Switching function**Fig. 4.4** Control input

suitable control input. It can efficiently eliminate the chatter due to apply the integral sliding mode controller.

The evolution of the switching function is depicted in Fig. 4.3. In Fig. 4.4, it is seen that the control signal slightly vibrates at first time and then tends to zero in the finite time.

## 4.5 Conclusion

In this paper, the proposed integral sliding mode controller is applied for an uncertain continuous NCS with the state delay. The parameter uncertainties satisfy the norm-bounded, and the sliding modes comply with the matching requirements. The main contribution of this work is to design the new integral switching function which includes the state feedback control gain and the sliding mode compensator. Furthermore, according to Lyapunov stability condition and in basis of the LMI convex optimal technology, a sufficient condition is derived so that the asymptotical stability of the closed-loop system is guaranteed. The numerical simulation is provided to verify the correctness and superiority of the approach.

**Acknowledgements** This paper is supported by Science Research Term of the Education Department of Shaanxi province of China (No. 12JK0548).

## References

1. Dai, S.L., Lin, H., Ge, S.S. (2010). Scheduling-and-control code-sign for a collection of networked control systems with uncertain delays. *IEEE Transactions on Control Systems Technology*, 18(1), 66–78.
2. Gao, H.J., Meng, X.Y., Chen, T.W. (2008). Stabilization of networked control system with a new delay characterization. *IEEE Transactions on Automatic Control*, 53(9), 2142–2148.
3. Utkin, V.I. (1992). *Sliding modes in control optimization* (pp. 145–154). New York: Springer.
4. Efe, M.O., Kaynak, O., Yu, X.H. (2000). Sliding mode control of a three degrees of freedom anthropoid robot by driving the controller parameters to an equivalent regime. *Transactions of the ASME Journal of Dynamic Systems, Measurement and Control*, 122(4), 632–640.
5. Gao, W.B., Hung, J.C. (1993). Variable structure control of nonlinear systems: A new approach. *IEEE Transactions on Industrial Electronics*, 40(1), 45–55.
6. Utkin, V.I., Shi, J. (1996). Integral sliding mode in systems operating under uncertainty conditions. In *Proceedings of the 35th IEEE Conference Decision Control* (pp. 4591–4596). New York: IEEE Press.
7. Chou, C., & Cheng, C. (2003). A decentralized model reference adaptive variable structure controller for large-scale time-varying delay systems. *IEEE Transactions on Automatic Control*, 48(7), 1213–1217.
8. Poznyak, A., Fridman, L., & Bejarano, F.J. (2004). Mini-max integral sliding mode control for multimodel linear uncertain systems. *IEEE Transactions on Automatic Control*, 49(1), 97–102.

# Chapter 5

## Synthesis of Linear Antenna Array Using Genetic Algorithm to Control Side Lobe Level

Zhigang Zhang, Ting Li, Feng Yuan, and Li Yin

**Abstract** In array pattern synthesis, it is often designed to achieve the desired radiation pattern. In this paper, real-coded genetic algorithm (RGA) optimization method is presented to optimize the value of weights of each antenna element to minimize side lobe level of the uniform spaced linear array geometries with a certain main beam width. The optimization program is done by using MATLAB. It compared with the conventional analytical methods such as Chebyshev and Taylor through radiation patterns with different number of elements and intervals of each element. The simulation results show that the optimization results are of little difference when  $d \geq \lambda/2$ , but GA can get a more optimal result when  $d < \lambda/2$ . The application of genetic algorithm for pattern synthesis is found to be useful.

### 5.1 Introduction

Synthesis of array antennas is very important to get the desired pattern, and how to achieve low side lobe in the condition of a fixed main beam width has been considered since a long period [1]. Conventional analytical methods such as Chebyshev and Taylor are widely used for uniformly spaced linear arrays with isotropic elements [2, 3]. However, these methods cannot synthesize antenna array with complicated geometry layout. In recent years, numerical approach has become more popular in synthesis of antenna arrays such as Powell's method, memetic algorithm (MA), tabu search (TS), particle swarm optimization (PSO), and genetic algorithm (GA) [4–6].

---

Z. Zhang • T. Li (✉) • L. Yin  
Electronic Engineering College, University of Naval Engineering,  
Wuhan 430033, Hubei, China  
e-mail: [liting89720@126.com](mailto:liting89720@126.com)

F. Yuan  
Department of Information, Command of East China Sea Fleet,  
Ningbo 315122, Zhejiang, China

As a general optimization algorithm, genetic algorithm is simple in coding and genetic operations. Optimization is not bounded by other constraint. Because of its global search ability, GA does not have to rely on good initial values to obtain the optimal solution. J. Michael Johnson and Yahya Rahmat-Samii applied genetic algorithm to antenna synthesis firstly, illustrated the basic theory of array synthesis by using genetic algorithm, and have proved the probability of synthesized one and two dimensions of uniform and nonuniform antenna array by genetic algorithm [7]. Randy L. Haupt used genetic algorithm to optimize the 200 symmetric linear array and plane array, respectively, to obtain the minimum side lobe level, got less than 20 db side lobe level, and provided a general method of synthesized array by genetic algorithm [8]. R. L. Haupt presented a genetic algorithm which could deal with real number coding and binary coded simultaneously [9]. V. R. Lakshmi and G. S. N Raju compared GA with synthesis of Taylor by different scan angles.

In this paper, it is assumed that the array is uniform, where all the antenna elements are identical and equally spaced. The design criterion here considered is to minimize the side lobe level with a fixed main beam width. In these conditions, GA compared with the methods of Chebyshev and Taylor through radiation patterns with different number of elements and intervals of each element.

## 5.2 Uniform Linear Antenna Array

In linear antenna array, all the antenna elements of mutually uncoupled isotropic radiator are arranged along the x-axis with equal spacing between them,  $N$  is the total number of elements in the antenna array, and  $d$  is the distance between two consecutive elements [10, 11].  $A_m$  is the amplitude of each element, A symmetric linear array with even number is shown in Fig. 5.1.

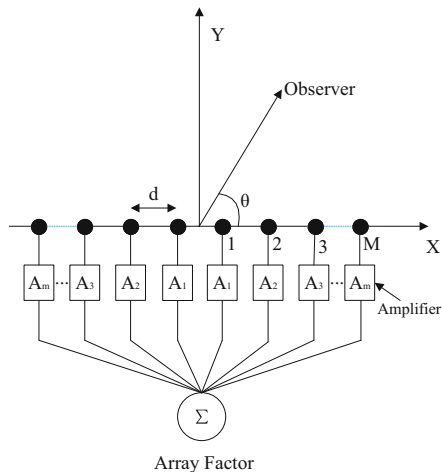
The far-field radiation pattern in the XY—plane of the array in the free space—can be obtained with the array factor (AF) as given in Eq. (5.1):

$$AF(\theta) = 2 \sum_{n=1}^M A_n \cos(kz_n \cos \theta + \alpha_n) \quad (5.1)$$

where  $A_n$  and  $\alpha_n$  are excitation magnitude and phase, respectively, the wave number of the carrier signal is  $k = 2\pi/\lambda$ ,  $\theta$  is the angle from broadside, and  $z_n$  is the distance between  $n$ th element with origin (the array center) as given in Eq. (5.2):

$$z_n = \begin{cases} \pm(n-1)d, n = 1, 2, \dots, M+1 (N = 2M+1) \\ \pm \frac{(2n-1)}{2}d, n = 1, 2, \dots, M (N = 2M) \end{cases} \quad (5.2)$$

**Fig. 5.1** Geometry of 2M elements array



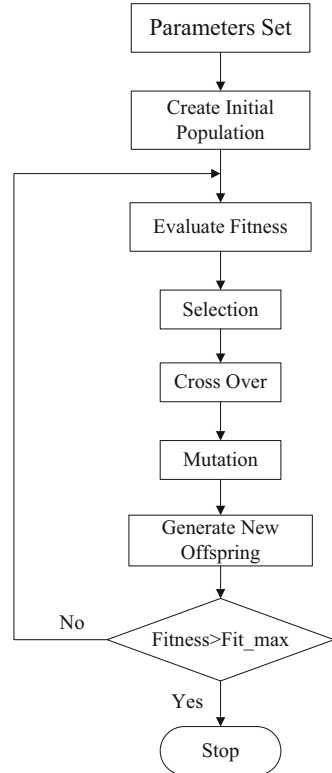
In this paper, only amplitude is considered in pattern synthesis, uniform the phase of excitation ( $\alpha_n = 0$ ), and formula (5.1) can be translated as given in Eq. (5.3):

$$AF'(\theta) = \begin{cases} 2 \sum_{n=1}^M A_n \cos \left( \frac{2\pi}{\lambda} (n-1) d \cos \theta \right), (N = 2M + 1) \\ 2 \sum_{n=1}^M A_n \cos \left( \frac{\pi}{\lambda} (2n-1) d \cos \theta \right), (N = 2M) \end{cases} \quad (5.3)$$

### 5.3 Optimizing by GA

Genetic algorithm is a stochastic global search algorithm based on the natural selection and genetic. GA is different with the traditional optimum algorithms, which are used to generate a deterministic test solution sequence merely based on the gradient calculation of evaluation function [12, 13]. Genetic algorithm searched the optimal solution through imitating the natural evolutionary process. It provides a generic framework for solving complex system optimization problems. It is not dependent on the problem of specific fields with a strong robustness. Genetic algorithm has a strong vitality in nonlinear numerical optimization. It is very useful for many electromagnetic optimization problems and adapted to beams forming of antenna array.

Fig. 5.2 Flowchart of GA



GA produced the first generation randomly and calculated the fitness value of each individual based on the fitness function. Fitness criterion suggested the advantages and disadvantages of each individual. The fittest individual has the biggest fitness value. The best will be survived to the next generation, and the worst will be eliminated directly. The rest are selected to produce new generation by recombination. Recombination consists of two genetic operators named crossover and mutation. And it will be recycled for a certain number of generations to get the optimal solution. The flowchart is shown in Fig. 5.2.

The important operators of GA can be summarized as follows:

1. Selection: The purpose of selection is to get good individuals from present population and then make them have chances to generate descendants as father generation. The selection probability is dependent on the fitness value.
2. Crossover: The new individuals can be generated through the operator of crossover and combined the characters of two individuals of their father generation. The operator happened based on the crossover rate.
3. Mutation: The change in a single individual depended on the mutation rate. The operator provides a chance to generate new individuals.

In amplitude-only synthesis, the main concern is to optimize the amplitude levels to produce the lowest side lobe level. The objective function associated with this array is the maximum side lobe level of its radiation field pattern to be minimized with a certain main beam width. The general form of the objective function is given by Eq. (5.4)

$$\text{Objective} = 20 \log_{10} \left( \frac{\max |AF'(\theta)|}{\max |AF'(\theta_0)|} \right) \quad (5.4)$$

where  $AF'(\theta)$  is array factor as given in Eq. (5.3), and  $0 \leq \theta_0 \leq \pi$ ,  $\theta$  belong a certain bands. Fitness function is ordered the values of objective function of every generation from small to big by nonlinear sequence. The fitness value decreased follow with the increasement of objective value. The range of fitness value is from 0 to 1.

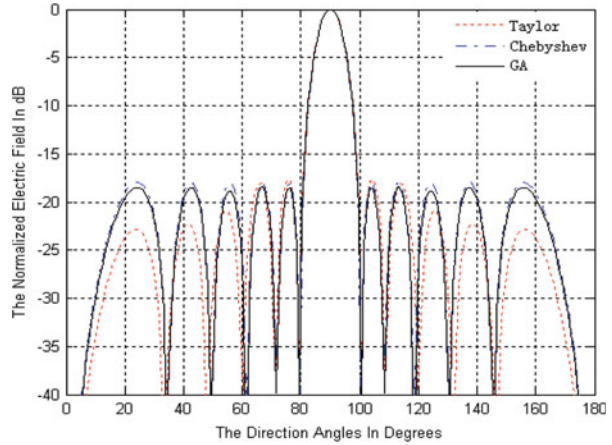
## 5.4 Experimental Results

Here we have used genetic algorithm for linear antenna array design. The results have been compared with that of Chebyshev optimization and Taylor optimization. The antenna model consists of certain number elements and equally spaced along the x-axis. Voltage sources are at the center segment of each element, and the amplitude of the voltage level is the antenna element weight. Only amplitude distribution is changed to find the optimum result; the array geometry and elements remain constant. GA with population size of 40, generation gap of 0.9, crossover rate of 0.8, and mutation rate of 0.2 runs for a total of 100 generations by using MATLAB. The optimal results will be founded each iteration. The fitness function is minimized the side lobe level of antenna pattern. Figure 5.3 shows that the antenna array with  $N = 12$  elements and interval of each element is  $d = \lambda/2$  has been designed for minimum side lobe level in bands  $[0, 80]$  and  $[100, 180]$ . The maximum of SLL calculated by GA is  $-18.49$  dB, approximated with the method of Chebyshev and Taylor. Figure 5.4 shows the maximum side lobe level value of every generation. The evolutionary tendency suggested that the optimal solution tends to be converged with the increase of generation and there is no premature convergence.

The synthesis of Chebyshev and Taylor can achieve the minimum side lobe level in a given main beam width based on the classical array theory. However, the optimal is conditional. It required array element spaced  $d \geq \lambda/2$ . When the array element spaced  $d < \lambda/2$ , Chebyshev and Taylor were not the optimal, but genetic algorithm was still valid. Figure 5.5 showed the radiation pattern with array element spaced  $d = 3\lambda/4$  and the minimum side lobe level in bands  $[0, 82.5]$  and  $[97.5, 180]$ .



**Fig. 5.3** Normalized patterns with  $d = \lambda/2$



**Fig. 5.4** Convergence of side lobe level of every generation

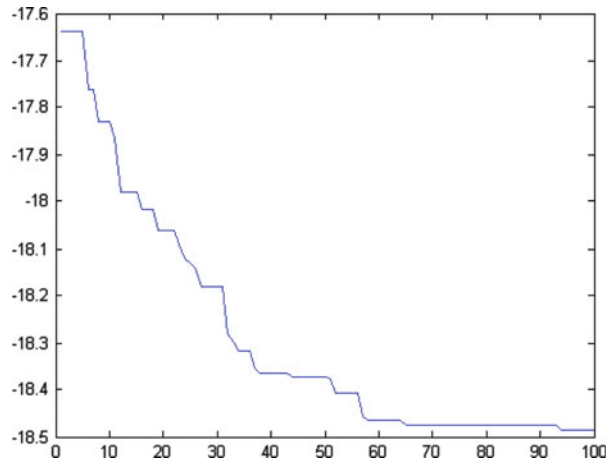
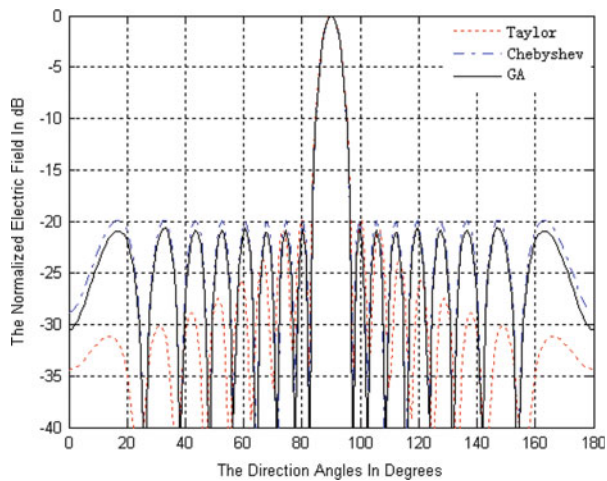


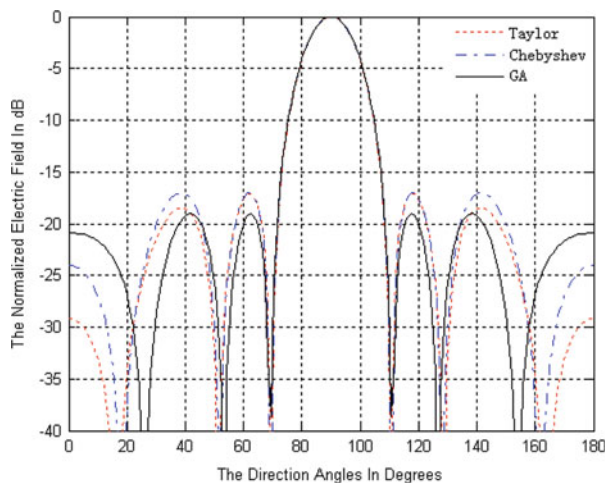
Figure 5.6 showed the radiation pattern with array element spaced  $d = \lambda/4$  and the minimum side lobe level in bands  $[0, 70]$  and  $[110, 180]$ . The maximum side lobe level calculated by GA is  $-19$  dB, which is lower about 3 dB than the result by Chebyshev and Taylor.

Figure 5.7 showed the radiation pattern with 24 elements spaced  $d = \lambda/4$  and the minimum side lobe level in bands  $[0, 80]$  and  $[100, 180]$ . The maximum side lobe level calculated by GA is  $-19.5$  dB, which is lower about 2 dB than the result by Chebyshev and Taylor.

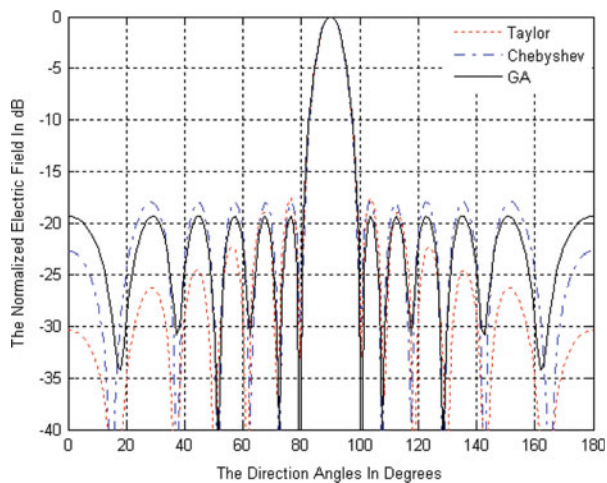
**Fig. 5.5** Normalized patterns with  $d = 3\lambda/4$



**Fig. 5.6** Normalized patterns with  $d = \lambda/4$



**Fig. 5.7** Normalized patterns with  $d = \lambda/4$  for 24 elements array



## 5.5 Conclusion

From the results, it is clear that genetic algorithm performs better than the method of Chebyshev and Taylor synthesis. As calculated in the previous section, the three methods performed almost the same when  $d \geq \lambda/2$ . However, GA can achieve about 3 dB lower than Chebyshev and Taylor in radiation pattern with the same main beam width when  $d < \lambda/2$  for 12 elements array. The application of genetic algorithm for pattern synthesis is found to be useful.

## References

1. Rajya Lakshmi, V., & Raju, G. S. N. (2011). Synthesis of linear antenna arrays using array designer and genetic algorithm. *IJAEST*, 9(1), 44–48.
2. Raju, G. S. N. (2005). *Antennas and wave propagation* (pp. 132–139). India: Pearson Education.
3. Elliot, R. S. (1981). *Antenna theory and design* (pp. 49–54). New Jersey: Prentice-Hall.
4. Jiao, Y. C., Wei, W. Y., Huang, L. W., & Wu, H. S. (1993). A new low side lobe pattern synthesis technique for conformal arrays. *Antenna Propagation*, 41(8), 824–831.
5. Er, M. H., Sim, S. L., & Koh, S. N. (1993). Application of constrained optimization techniques to array pattern synthesis. *Signal Processing*, 34(5), 323–334.
6. Chen, T. B., Chen, Y. B., Jiao, Y. C., & Zhang, E. S. (2005). Synthesis of antenna array using particle swarm optimization. In *Microwave conference proceedings 2005. Asia-Pacific conference proceedings, China* (pp. 4–9).
7. Johnson, J. M., & Rahmat-Samii, Y. (1994). Genetic algorithm optimization and its application to antenna design. In *Antennas and propagation society international symposium, 1994, USA, AP-S. Digest Volume 1* (pp. 326–329).
8. Haupt, R. L. (1994). Thinned arrays using genetic algorithms. *Antennas and Propagation*, 12(7), 993–999.
9. Haupt, R. L. (2007). Antenna design with a mixed integer genetic algorithm. *Antennas and Propagation*, 55(5), 577–582.
10. Shrivastava, S., & Cecil, K. (2012). Performance analysis of linear antenna array using genetic algorithm. *IJEIT*, 2(5), 84–88.
11. Laseetha, T. S. J., & Sukanesh, R. (2011). Synthesis of linear antenna array using genetic algorithm to maximize side lobe level reduction. *International Journal of Computer Applications*, 20(7), 27–33.
12. Mandal, D., & Chandra, A. (2010). Side lobe reduction of a concentric circular antenna array using genetic algorithm. *Serbian Journal of Electrical Engineering*, 7(2), 214–218.
13. Johnson, J. M., & Samii, Y. R. (1997). Genetic algorithm in engineering electromagnetic. *Antennas and Propagation*, 39(4), 7–21.

# Chapter 6

## Wavelet Analysis Combined with Artificial Neural Network for Predicting Protein–Protein Interactions

Juanjuan Li, Yuehui Chen, and Fenglin Wang

**Abstract** In order to solve the prediction problem of interaction between proteins, we use a wavelet coefficient combined with artificial neural network method, improving the prediction accuracy of the problem of protein–protein interactions. By introducing the Biorthogonal Wavelet 3.3 coefficients as the feature extraction method and the three-layer feedforward neural network as a classifier, we solve the problem of protein interaction effectively. Using the Human dataset verifies the validity of this method. Through testing the Human dataset, using Biorthogonal Wavelet 3.3 coefficient combined with the three-layer feedforward neural network, solve the prediction problem of protein interactions with well results. This combination of wavelet coefficients and the three-layer feedforward neural network to predict protein interaction problem is an effective method. At the same time, compared with other prediction methods, this method performs at least 4 % higher accuracy than the better accuracy of auto-covariance (11) combined with PNN on the same dataset.

### 6.1 Introduction

Proteins are major component of organism and play a very important role in organism. Proteins play biological function through the interactions in organism. The study of interaction between proteins starts from biological experiment, and commonly biological experimental methods have the following: yeast two-hybrid (Y2H) [1], a biological method aiming at the yeast; mass spectrometry protein complex identification (MS-PCI) [2], which is from the protein molecular and atomic micro, based on which forecast is carried on; and protein chip technology [3], a technology that solidifies some known proteins and chips are used to predict

---

J. Li (✉) • Y. Chen • F. Wang  
Computational Intelligence Laboratory, University of Jinan, Jinan 201305, China  
e-mail: [juanjuanli@126.com](mailto:juanjuanli@126.com)

the interactions of solidified proteins. Using methods of biological experiment for studying the interactions between protein prediction problems has its own advantages. Biological experiments operate simply and can get intuitive and reliable results. But with the development and use of high-throughput biological experiments, a large number of biological data are produced. Using biological experiment for high throughput data is a very large project and also time consuming. So, intelligent calculation methods are introduced to solve biological problems. Intelligent calculation method is a technology combining computer technology with biology. It is generated through the rapid development of computer and can be a very good solution in solving biological problems of high throughput calculation [4].

At present, many intelligent computation methods are used for the study of protein–protein interaction (PPI) problem. Among them, typical feature extraction methods have amino acid composition, pseudo-amino acid composition, and physicochemical properties. Classification methods have kernel-nearest neighbor, support vector machine [5], and probabilistic neural network. In this chapter, we proposed the resonant recognition model combined with artificial neural network to predict interaction between proteins. Experiment results show that our proposed method performs at least 4 % higher accuracy than the best of other related works.

## 6.2 Technical Introduction

The term wavelet means small wave; the so-called small means it is capable of decaying; and wave refers to the volatility, the amplitude of positive and negative form of shocks. Compared with the Fourier transform, the wavelet transform is a time frequency analysis (spatial) localization, signals (functions) multi-scale refinement through telescopic translation operations, and ultimately achieves frequency time segments and low frequency subdivision. The wavelet transform can automatically adapt to the time-frequency signal analysis requirements, which can focus on any signal details, solve the difficult problem of Fourier transform, and become a major breakthrough in the scientific methods of Fourier transform since [6, 7].

Neural network is an algorithm of the simulation to process human neurons. The input data are combined to a single output node through the nonlinear transformation. The typical structure of the neural network has three layers, and it has strong robustness and fault tolerance; at the same time, it can learn and adapt uncertain systems. The neural network needs to set the number of neurons of the input layer, hidden layer, and output layer. The neural network classification process has two steps: first, training the neural network parameters by continuous input training data; and second, inputting test data to get the results of the test data used in the optimization of neural network. Using artificial neural network (ANN) classification has the advantages of speed and potential super speed and good fault tolerance ability. It is suitable for solving some problem. Compared to some clustering

ideological classification, for example, support vector machine (SVM) and nuclear nearest neighbor, it has better classification results [8, 9].

### 6.3 Research Programs

The dataset used in this study is Human dataset with more difficulty. Human biological structure is very complex, and the human body to maintain its own physiological function needs a lot of proteins to participate, so the interaction between the human proteins is more complex, and then the prediction of this dataset has some difficulty. The human data contain 914 positive protein pairs and 941 negative protein pairs for a total of 1,882 protein pairs. This dataset is used in this study to verify the effectiveness of the method we chose [10].

Feature extraction. The following are the three main steps to extraction feature of the protein interaction sequence pairs and finally into a one-dimensional vector of the neural network input:

1. Assign the two test proteins to respective IC value, which will be converted to two numerical sequences. The IC values are shown in Table 6.1.
2. Transform the two protein sequences using the discrete wavelet transform of DWT, using Biorthogonal Wavelet 3.3 (see Fig. 6.1). The discrete wavelet transform is performed at three levels (see Fig. 6.2). The original numerical signal is decomposed into A3 (approximation at level 3), D3 (detail at level 3), D2 (detail at level 2), and D1 (detail at level 1) signals of four scales. D3, D2, and D1 contain details of the signal in the original signal at different levels, but A3 only retained a few low-frequency signals [11].
3. Wavelet coefficients are expressed by  $CD_n^{K,p}$ , and wavelet coefficients were normalized; see Eq. (6.1)

$$NOM_n^{K,p} = \frac{|CD_n^{K,p}|}{\max_n (|CD_n^{K,p}|)} \quad (6.1)$$

where  $n = 1, 2, \dots, N$ ;  $p = 1, 2$ ;  $K = A3, D3, D2, D1$ .

Among them, P is protein serial number.

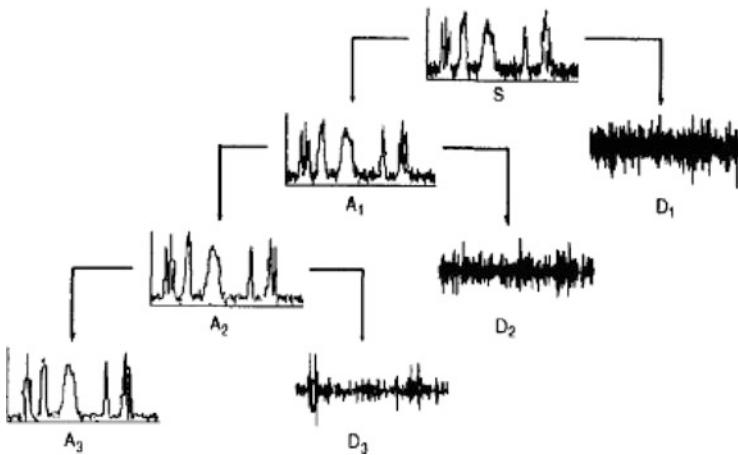
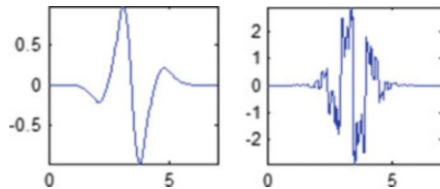
So we can get the wavelet coefficient, using the wavelet coefficients as the feature for protein–protein prediction. But because the sequence length is too long, the result in numerical sequences of wavelet coefficients is also very long; as such, the input of the classifier is not reasonable, so we choose 1 of the 20 largest values as the feature representation input to the classifier prediction. But we save these maximum values with their original order. Therefore, our neural network input is 20D.

Classify. Using artificial neural network to classify PPIs has two steps. Firstly, we need to use the training set to train parameters of neural network. Secondly, we need to use the test set to test the trained classifier. We choose particle swarm

**Table 6.1** IC value of the amino acid

Amino acid	L	A	N	C	I	P	W	T
IC value	2.40	2.30	2.20	1.96	2.40	2.00	2.37	2.09
Amino acid	Y	H	Q	F	H	K	M	R
IC value	2.20	2.30	2.06	1.98	2.46	2.20	2.17	1.82
Amino acid	V	E	S	D				
IC value	2.35	2.30	2.10	1.88				

**Fig. 6.1** Biorthogonal wavelet 3.3

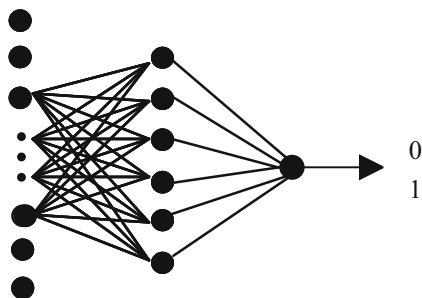


**Fig. 6.2** Discrete wavelet transform three levels

algorithm as parameter optimization algorithm to optimize parameters of neural network. It is a random search algorithm developed by simulating the foraging behavior of bird flocks based on group cooperation. It is generally considered swarm intelligence. Particle swarm optimization parameters mainly consist of two steps, one is the change of speed, and other is the change of position, using formulas (6.2) and (6.3):

$$V_{id} = w * V_{id} + c_1 * rand() * (P_{id} - X_{id}) + c_2 * rand() * (P_{gd} - X_{id}) \quad (6.2)$$

**Fig. 6.3** The model of the artificial neural network



$$X_{id} = X_{id} + V_{id} \quad (6.3)$$

Through the feature dimension obtained, we design the neural network structure as shown in Fig. 6.3. The input dimension is 20, and after trying, hidden layer is set to 6. Because this is a two-classification problem, the output is one dimensional.

## 6.4 Analysis of Results

This chapter uses 10-fold cross-validation test to verify this method performance. The 10-fold cross-validation divides the dataset into 10 groups, each with nine datasets to do the training set and the rest of the group to do the test set, turns to a total of 10 times, and then takes the average value as a result of the test. Each test set has 94 positive samples and 94 negative samples, a total of 188; each train set has 867 positive samples and 867 negative samples, a total of 1,694. Classification index using the total classification accuracy (Acc) and sensitivity (S) and precision (P) [12] is defined as follows:

$$Acc = \frac{PT + NT}{PT + PF + NT + NF} \quad (6.4)$$

$$S = \frac{PT}{PT + NF} \quad (6.5)$$

$$P = \frac{PT}{PT + PF} \quad (6.6)$$

in which PT is the correct number of positive samples, PF is the error number of positive samples, NT is the correct number of negative samples, and NF is error number of negative samples. The positive samples here refer to the interaction of the protein pairs, while the negative samples refer to the non-interaction protein pairs (Table 6.2).

The chapter is experimented many times to verify the selected methods' effectiveness; the following tables show the results of ten pairs of train set and test set and the best result of the integrated methods (Table 6.3).



**Table 6.2** Best result of the experiment

Dataset	P	S	Acc
Human	78.723	84.09	81.9149

**Table 6.3** Best results of 10-fold cross-validation

P	P1	P 2	P 3	P 4	P 5	
R	80.8511	80.8511	76.0638	77.6596	81.9149	
P	P 6	P 7	P 8	P 9	P 10	Best
R	80.8511	78.1915	80.8511	79.7872	78.1915	81.9149

**Table 6.4** Results of different methods used

Data	Classifier	Feature	Code	Result
Human	SVM	KMC + KNN + BIO [10]	Link	73.10
Human	KNN	ACC [9]	Summation	73.90
Human	PNN	AC (11)	Link	78.37
Human	ANN	Wavelet	Summation	81.9149

List the results of the different methods, where AC (11) presents calculating AC of 11 physicochemical properties (Table 6.4).

Through this table, we can achieve a conclusion: the wavelet coefficients combined with artificial neural network perform at least 4 % higher accuracy than the AC (11) combined with PNN on the same dataset.

## 6.5 Conclusion

This chapter uses ANN as a classifier to predict protein interactions. Through experiment of a variety of physicochemical properties of the protein, we find that the wavelet analysis combined with ANN reached good results. This feature extraction method combined with ANN achieves more high precision. It can obviously be seen that the accuracy is improved. Through confirmation, this method we used is effective.

**Acknowledgments** This research was partially supported by the NSFC grant (61201428), the Natural Science Foundation of China (61070130), the Key Project of Natural Science Foundation of Shandong Province (ZR2011FZ003), the Key Subject Research Foundation of Shandong Province, and the Shandong Provincial Key Laboratory of Network Based Intelligent Computing.

## References

1. Ito, T., Chiba, T., Ozawa, R., Yoshida, M., et al. (2001). A comprehensive two-hybrid analysis to explore the yeast protein interaction. *Proceedings of the National Academy of Sciences U S A*, 98(8), 4569–4574.

2. Ho, Y., Gruhler, A., Heibut, A., et al. (2002). Systematic identification of protein complexes in *Saccharomyces cerevisiae* by mass spectrometry. *Nature*, 415, 180–183.
3. Zhu, H., Bilgin, M., Bangham, R., et al. (2001). Global analysis of protein activities using proteome chips. *Science*, 293, 2101–2105.
4. Zhou, J. L., & Yi, M. C. (2008). Study of protein interactions in a review of calculation methods. *Journal of Computer Research and Development*, 45(12), 2129–2137.
5. Chang, C. C., & Lin, C. J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 2(3), 27.
6. Zhuo, M., Zhang, Y. Q., Li, M. L., et al. (2011). Based on the integrated learning approach for predicting protein-protein interactions. *Journal of Sichuan University (Engineering Science Edition)*, 43(3), 324–328.
7. Feng, T. N., Da, L., Jin, D. L., et al. (2011). Using lifting wavelet extracting protein-protein interaction characteristics. *Communications in Applied Mathematics and Computational Science*, 25(2), 12.
8. Guo, Y., Yu, L., Wen, Z., & Li, M. (2008). Using support vector machine combined with auto covariance to predict protein-protein interactions from protein sequences. *Nucleic Acids Research*, 36(9), 3025–3030.
9. Song, J. (2009). Prediction of protein-protein interaction using kernel nearest neighbor algorithm. *Application Research of Computers*, 26(11), 124–130.
10. Guo, Y., Li, M., Pu, X., et al. (2010). PRED-PPI: A server for predicting protein-protein interactions based on sequence data with probability assignment. *BMC Research Notes*, 3(1), 145–151.
11. Feng, T. N., Zhang, H., & Wang, Y. F. (2008). Prediction of protein SNR based on interaction. *Journal of Shanghai University (Natural Science)*, 14(6), 12.
12. Zhou, Z. R., Song, X. F., & Wang, M. B. (2010). Using a combination of classifiers for predicting protein-protein interactions. *Acta Electronica Sinica*, 38(6), 275–282.

# Chapter 7

## Application Analysis of Slot Allocation Algorithm for Link16

Hui Zeng, Qiang Chen, Xiaoqiang Li, and Jianguo Shen

**Abstract** In order to improve the efficiency of slot utilization of Link16, the traditional binary tree model of timeslot allocation is improved, and the performances of three kinds of timeslot assignment algorithm are compared and analyzed in the average transmission delay and message sending failure rate. When the user is determined and the message arrival rate is lower, the fixed timeslot allocation is simple and effective. When there are a lot of users or the user is uncertain, and the message arrival rate is very low, competitive slot allocation should be more efficient. When the number of users is small, the user message arrives sudden, and arrival rate is high, the performance of the dynamic slot allocation has a very distinct advantage. Finally, based on the characters of different timeslot assignment algorithms, the overall timeslot allocation scheme is presented for Link16 network.

### 7.1 Introduction

In the traction of modern information warfare demand, the data link has been significant developed in the field of military information, and Link16 is one of the most widely used tactical data link which uses the TDMA protocol. Because the platforms which need to access data link networks become more and more, timeslot allocation is very important for the efficiency to the Link16 network. In this chapter, the binary tree slot allocation model is analyzed. According to uniformity principle,

---

H. Zeng (✉) • J. Shen  
National Defense Information Academy, WuHan 430019, China  
e-mail: [zenghui\\_1974@163.com](mailto:zenghui_1974@163.com)

Q. Chen  
Institute of China Electronic System Engineering Corporation, Beijing 100039, China

X. Li  
Communications Training Base of the General Staff Headquarters, Xuanhua 075100, China

a method of slot allocation based on binary tree model [1] is proposed when the number of timeslots is not  $2^n$ .

Timeslot allocation algorithm of Link16 includes fixed allocation (FA) algorithm, competitive allocation (CA) algorithm, and dynamic allocation (DA) algorithm [2]. There are some researches in recent years of the performance of various algorithms, but their performance indicators are inconsistent, such as queuing delay of FA algorithm [3], collision probability of CA algorithm [4], and the throughput of DA algorithm [5], which cannot be an effective evaluation or meet the characteristic of Link16 application. In this chapter, the relations between the delay and the failure rate of message transmission and the load rate are analyzed for the three kinds of slot allocation algorithm. Each algorithm has advantages and disadvantages. According to the characteristics of the network participation group (NPG), different algorithms should be used for different NPG of Link16 network.

### 7.2 The Model of Timeslot Resource

One day is divided into 112.5 time units in Link16. Each unit is divided into 64 time frames, and the time frame is divided into 1,536 application timeslots. There is 7.8125 ms for every timeslot. Because the time unit is too long to describe the real-time demand of Link16, usually, the time frame is analyzed as a slot cycle in the process of timeslot allocation.

For the stability of the system work, the timeslot resource of every NPG should be uniformly distributed in the slot cycle. Binary tree model is commonly used in the slot block allocation of resources. 1,536 application timeslots of a time frame is divided into three groups (A, B, and C); each group consists of 512 timeslots. The timeslots of a group are staggered arranged as A-0, B-0, C-0, A-1, B-1, C-1, ..., A-511, B-511, and C-511. In order to simplify the identification and assignment, a slot block which contains  $2^n$  timeslots can be defined by the group of timeslots (A, B, or C), index (0~511), and recurrence rate number (RNN). As shown in Fig. 7.1, A-0-9 which contains all 512 slots in the A group is divided into A-0-8

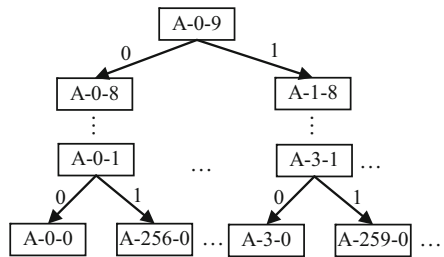


Fig. 7.1 The model of binary tree slot

and A-1-8. A-0-8 contains A-0, A-2, . . . , A-510, and A-1-8 contains A-1, A-3, . . . , A-511. Finally, the timeslots A-0-0, A-1-0, A-2-0, . . . , A-511-0, respectively, denote A-0, A-1, A-2, . . . , A-511.

As can be seen from the binary tree model diagram, if the timeslot block assigned to the NPG just by the power of 2 slots, it ensures that the timeslots belonging to a NPG are uniformly distributed in the slot ring, and the distance between the timeslots is  $3 \times 2^{9-RNN}$ . In the decomposition process of the model, if encoding the left represents 0 and the right represents 1, the tree node index number of the slots is the binary values of the coding which consists of a string from the leaf node to the root node of the direction of the binary code, and the number of timeslots of the slot block is  $2^{RNN}$ .

If the timeslot block assigned to the NPG is not a power of 2, then it should be decomposed into a group of timeslot subblocks which are the power of 2. Because the slots of timeslot block are not a power of 2, the slot cannot be uniformly distributed in the slot ring. In order to achieve the slot distribution as uniform as possible, the choice of timeslot subblock should comply with the principle that the common ancestor node is closest.

### 7.3 Performance Analysis of the TDMA Slot Allocation Algorithm

#### 7.3.1 Fixed Slot Allocation Algorithm for Link16

Based on the estimated amount of the user's business, fixed timeslot resources are assigned to each user by fixed timeslot allocation, and then the user can only occupy its own dedicated slot to send a message. Assuming that the messages arriving obey Poisson distribution with the parameter  $\lambda$ , the message transmission rate ( $\mu$ ) of the user is directly related to the amount of the dedicated timeslot; therefore, the message processing model is a M/D/1 queuing models. The message have to wait for the arrival of the transmit timeslot to send, so the arrival of the first message is the equivalent to the second message of the M/D/1 queuing models. Considering the change of the amount of messages in the send buffer  $\{X(t), t \geq 0\}$ , it is a typical birth and death process. State space defined as  $E = \{1, 2, 3, \dots\}$ , and the state transition matrix is

$$Q = \begin{bmatrix} -\lambda & \lambda & 0 & 0 & \dots \\ \mu & -\lambda - \mu & \lambda & 0 & \dots \\ 0 & \mu & -\lambda - \mu & \lambda & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \quad (7.1)$$

The stationary distribution is obtained as

$$\begin{cases} \pi_1 = \left[ \sum_{k=1}^{\infty} \left( \frac{\lambda}{\mu} \right)^k \right]^{-1} = 1 - \rho, & \left( \rho = \frac{\lambda}{\mu} < 1 \right) \\ \pi_k = \left( \frac{\lambda}{\mu} \right)^{k-1} \pi_1 = \rho^{k-1} (1 - \rho), & (k = 2, 3, \dots) \end{cases} \quad (7.2)$$

$\rho$  is called load rate. Therefore, the average number of the message to wait for transmission is obtained as

$$L = \sum_{k=2}^{\infty} (k-2) \pi_k = \frac{\rho^2}{1-\rho} \quad (7.3)$$

The average transmission delay is equal to

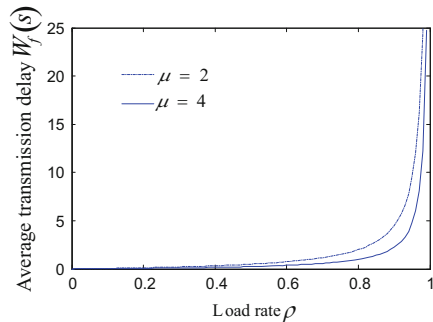
$$W_f = \frac{L}{\lambda} = \frac{\rho}{(1-\rho)\mu} \quad (7.4)$$

When the message in the transmission queue is more than  $N_g$ , it is assumed that the message will be discarded for the real-time requirements of sending a message, so message sending failure rate should be obtained as follows:

$$F_f = 1 - \sum_{k=1}^{N_g+1} \pi_k = \rho^{N_g+1} \quad (7.5)$$

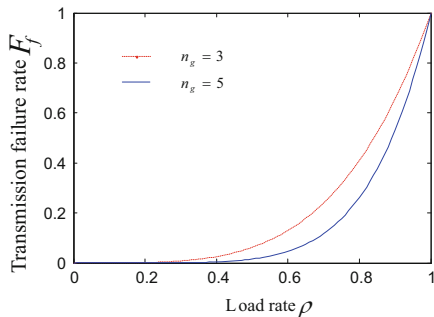
Figures 7.2 and 7.3 illustrate the relation between the system average waiting delay, message transmission failure rate, and user load rate, respectively.

As indicated in Fig. 7.2, the average waiting time is increased as the load rate increases. In the lower interval of  $\rho$ , the change of  $W_f$  is slow and gentle. However, when  $\rho$  increases to a certain extent, the value of  $W_f$  sharply becomes high.



**Fig. 7.2** Average transmission delay  $W_f$  vs load rate  $\rho$

**Fig. 7.3** Transmission failure rate  $F_f$  vs load rate  $\rho$



As shown in Fig. 7.3,  $F_f$  increase with the increasing of  $\rho$ . The major factor to reduce  $F_f$  is to decrease  $\rho$ , and the increasing of the transmit buffer can also be appropriate to reduce the  $F_f$ .

### 7.3.2 Competition Slot Allocation Algorithm for Link16

All users of the NPG randomly grab slot to send messages. Assumed all users have the same priority, which sent the message as soon as possible, so there may be a plurality of users competing for the same timeslot. Then the slot conflict will occur, which will result in a communication failure.

Assuming that the NPG number of users is  $N$ , the messages arriving obey Poisson distribution with the parameter  $\lambda$  and are independent of each other. The available slot resource of NPG is evenly distributed in the slot ring with the interval  $T$ . Then, the probability of  $n$  messages arriving in NPG is obtained as

$$P(n) = \frac{(N\lambda T)^n e^{-N\lambda T}}{n!} \quad (7.6)$$

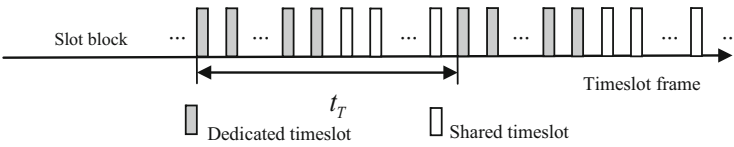
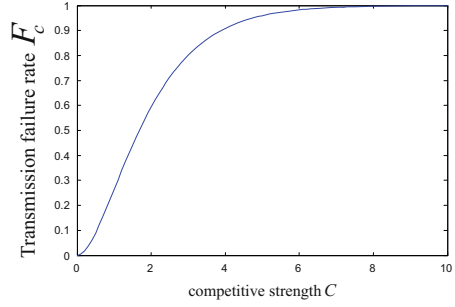
The probability of two or more messages arriving is much smaller than one for a user at time  $T$ , so we can ignore the probability that a user received two or more messages at time  $T$ , then the probability of the slot competition conflict which will result in communication failure can be obtained as

$$F_c = 1 - e^{-N\lambda t} - N\lambda t e^{-N\lambda t} = 1 - (1 + C)e^{-C}, \quad (C = N\lambda t) \quad (7.7)$$

$C = N\lambda t$  is defined as the competitive strength of user timeslot. The number of users, the user message arrival rate, and the slot interval have the equivalent effects on  $C$ . The relationship between  $F_c$  and  $C$  is shown in Fig. 7.4.

In this case, the message transmission delay also can be obtained by the formula (7.3). Because the load rate  $\rho$  is a user relative to the entire NPG and  $\mu$  is message transmission rate of the entire NPG, the average waiting delay is greatly reduced.

**Fig. 7.4** Transmission failure rate  $F_c$  vs competitive strength  $C$



**Fig. 7.5** Timeslot distribution of the dynamic slot allocation algorithm

### 7.3.3 Dynamic Slot Allocation Algorithm for Link16

The slot block is divided into a dedicated timeslot and shared timeslot. Dedicated timeslot is assigned to the specified user to send a message or make an appointment, and the shared timeslot is that all users can make an appointment to use. The distribution of timeslots in the timeslot frame is shown in Fig. 7.5.

Assuming message source can use ON/OFF model to describe, there are active period  $T_{on}$  and quiet period  $T_{off}$ , the message is generated. The messages arriving obey the Poisson distribution with parameter  $\lambda$  in active period, and no message arrives in the silent period. Active factor  $r$  is defined as

$$r = \frac{T_{on}}{T_{on} + T_{off}} \tag{7.8}$$

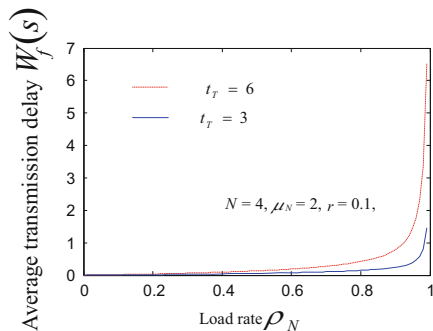
In this case, assuming that the length of  $T_{on}$  and  $T_{off}$  obeys negative exponential distribution, in a reservation period  $t_T$ , the probability that  $n$  active period is triggered by the all  $N$  independent users can be obtained as

$$P(n) = \frac{(Nrt_T)^n e^{-Nrt_T}}{n!} \tag{7.9}$$

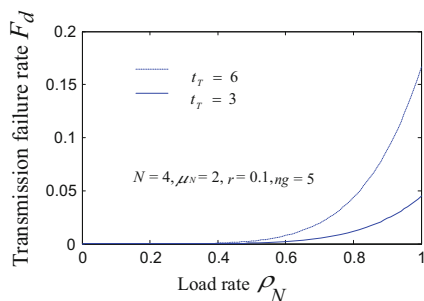
Shared slot averaging assigned to the user that enters the active phase, the average transmission delay of the message can be calculated by the formula (7.4) as follows:



**Fig. 7.6** Average transmission delay  $W_f$  vs load rate  $\rho_N$



**Fig. 7.7** Transmission failure rate  $F_d$  vs load rate  $\rho_N$



$$W_d = \sum_{n=1}^N P(n) \frac{\rho_n}{(1 - \rho_n)\mu_n} = \sum_{n=1}^N \frac{(Nr t_T)^n e^{-Nr t_T}}{n!} \frac{\frac{n\rho_N}{N}}{\left(1 - \frac{n\rho_N}{N}\right) \frac{N\mu_N}{n}} \quad (7.10)$$

$\rho_n$  and  $\mu_n$  are the load ratio and the message sending rate of  $n$  users that is triggered, respectively, wherein the condition of  $\mu_N > \lambda$  should be complied with.

The message sending failure rate can be calculated by the formula (7.5) as follows:

$$F_d = \sum_{n=1}^N P(n) \rho_n^{n_s+1} = \sum_{n=1}^N \frac{(Nr t_T)^n e^{-Nr t_T}}{n!} \left(\frac{n\rho_N}{N}\right)^{n_s+1} \quad (7.11)$$

The average transmission waiting delay and message transmission failure rate with the variation of the load ratio are shown in Figs. 7.6 and 7.7, respectively.

Since the dynamic slot assignment takes account of the burst of the message arrival and the sharing of resources, performance in both the average transmission waiting delay or message transmission failure rate is improved significantly. The analysis here, however, does not consider the impact of the dedicated timeslot occupying the timeslot resources. The amount of the timeslot resources occupied by

dedicated timeslot is determined by the number of users and the appointment cycle length. Dynamic slot allocation should give full consideration to the impact of the increase in the number of users and the shorter of the appointment cycle length. Specially, when the dedicated timeslots occupy the all timeslots of the slot block, it becomes fixed slot allocation algorithm. On the other hand, it becomes competition slot allocation algorithm if the all timeslots is defined as shared timeslots.

## 7.4 Link16 Slot Allocation Scheme

The network of Link16 is divided into different NPG. Firstly, the resources of the timeslot frame allocated to the NPG and then reallocated to the users of NPG.

The purpose of NPG division is to reduce the mutual influence of the communication of the different functional domains, so the resources of timeslot frame divide into timeslot blocks for NPG by fixed timeslot allocation algorithm.

There are about 20 major NPGs of Link16 [6]. Because the number and type of users and the characteristics of message are different from NPG, the timeslot block of NPG should be assigned to the user in different ways. From the above analysis, the slot allocation scheme for Link16 NPG is shown in Table 7.1.

**Table 7.1** The slot allocation scheme for the NPG of Link16

NPG-index	NPG-function	Timeslot allocation algorithm
1	Initial entry	FA CA
2	RTT-A	FA
3	RTT-B	CA
4	Network management	FA CA
5	PPLI-A	FA
6	PPLI-B	CA
7	Surveillance	FA DA
8	Mission management	FA CA
9	Air control	FA CA
10	Electronic warfare	FA
12	Voice A	CA
13	Voice B	CA
14	I-PPLI	FA DA
18	Weapons coordination	FA
19	Fighter-to-fighter net	FA CA
27	Joint PPLI	FA
28	Distributed network management	FA
29	Residual message	FA CA
30	IJMS P message	FA DA
31	IJMS T message	FA DA

## 7.5 Conclusion

When the user is determined and the message arrival rate is lower, the fixed timeslot allocation is a simple and effective allocation method. When there are a lot of users or the user is uncertain, and the message arrival rate is very low, competitive slot allocation should be more efficient. When the user message arrives sudden, the message arrival rate is high, and the number of users is small, the performance of the dynamic slot allocation has a very distinct advantage. Accordingly, the slot allocation method of the different NPG of Link16 should be adapted to its features, which help to reduce the delay of message transmission, improve the slot utilization, and reduce the probability of the message failed to be sent.

## References

1. Hou, X. (2008). *Data link dynamic slot allocation and its network planning technology research*. Chengdu: University of Electronic Science and Technology of China (In Chinese).
2. Commanding Officer Navy Centre for Tactical Systems Interoperability (NCTSI). (2001). *Understanding Link-16: A guidebook for new users*. San Diego: NCTSI.
3. He, Z.-x., Luo, X.-h. (2011). Towards fixed allocation timeslot number prediction method for TDMA tactical data link based on operation requirements. *International Journal of Advances in Computing Technology*, 3(6), 57–64 (In Chinese).
4. Xia, B., & Li, H. (2010). Collision analysis of time slot contention access in data-link. *Ship Electronic Engineering*, 30(6), 66–69 (In Chinese).
5. Zhang, L., Yin, H. (2008). *The Link16 throughput simulation based on DTDMA* (pp. 464–466). China-Ireland International Conference on Information and Communications Technologies, Beijing, China.
6. Mei, W.-h., & Cai, S.-f. (2004). *JTIDS/Link16 data link* (p. 103). Beijing: Higher Education Press (In Chinese).

# Chapter 8

## An Improved Cluster Head Algorithm for Wireless Sensor Network

Feng Yu, Wei Liu, and Gang Li

**Abstract** Routing algorithm is one of the key technologies of wireless sensor network. Based on the LEACH algorithm, the second choice of cluster head algorithm for WSN is proposed. It takes into account the cluster head as residual energy and distance to BS, and then it chooses a senior cluster head. The improved algorithm avoids direct communication between BS and the cluster head which has low energy and is far away from BS, which prolongs the lifetime of network and enhances the ability of data collection. The experiments show that this technique of event-driven can cut down the data transmission and further extend the lifetime of network.

### 8.1 Introduction

Wireless sensor network (WSN) is a convergence technology of computer network, communications, and sensor. WSN usually is widely applied in the military operations, environmental monitoring, medical rescue, traffic control, home automation, and other commercial applications [1, 2]. Study at home and abroad of WSN is not yet mature. Carrying out this cutting-edge technology research timely will bring great value and strategic significance to the development of the country. As the bright future in their application, a main goal in designing such WSNs is optimizing energy consumption to maximize the network lifetime. However, the influencing factors are various in impacting lifetime of network [3]. Actually, after deploying sensor nodes, effective routing control protocol is the key to improve the efficiency of energy. In the current study, the clustering algorithm is considered to be one of the most effective ways of efficient energy management. Typically, LEACH (Low Energy Adaptive Clustering Hierarchy) is the earliest cluster-organized

---

F. Yu (✉) • W. Liu • G. Li  
Network Center of Shenyang Jianzhu University, Shenyang 100168, China  
e-mail: [wind@sjzu.edu.cn](mailto:wind@sjzu.edu.cn)

routing protocol in clustering algorithm [4, 5]. Inspired from LEACH, many of clustering algorithms are created, such as LEACH-C [6, 7], PEGASIS [8], and TEEN [9]. In this paper, contraposing the deficiencies of LEACH, we present an improved routing protocol LEACH-SC, which put forward the cluster head of secondary selection algorithm and single-multi-hop routing algorithm to further reduce energy costs.

## 8.2 Related Works

### 8.2.1 LEACH Protocol

LEACH protocol is based on the hypothesis that all sensor nodes can communicate with sink node directly and run with many rounds. Each round begins with a setup phase when the clusters are organized, followed by a steady-state phase when data are transferred directly from the ordinary nodes to the cluster heads and then transmit to sink node or base station (BS). In the setup phase, each node chooses a random number between 0 and 1, if the number is lower than the threshold  $T(n)$ , then the node becomes a cluster head for current round. The threshold limit is defined by the following formula (8.1).

$$T(n) = \begin{cases} \frac{p}{1 - p * [r * \text{mod}(1/p)]}, & \text{if } n = G \\ 0, & \text{otherwise} \end{cases} \quad (8.1)$$

Where  $p$  denotes the percentage of cluster head in the total number of nodes,  $r$  is current rounds,  $G$  is a node set which have not been elected in the past  $1/p$  rounds. When one node is selected as cluster head, it will inform other nodes, or else it will choose a cluster to join in deciding as the distance to cluster heads.

### 8.2.2 Energy Model

LEACH employs the familiar radio model discussed in Lv's method [10], which is the first order radio model. If transmit an l-bit message over distance  $d$ , the energy consumption of the node is

$$E_{Tx}(k, d) = lE_{Tx-elec} + E_{Tx-amp}(l, d) = \begin{cases} kE_{elec} + k\epsilon_{fs}d^2, & d < d_0 \\ kE_{elec} + k\epsilon_{amp}d^4, & d \geq d_0 \end{cases} \quad (8.2)$$

$kE_{elec}$  is the energy consumption of radio dissipation, while the second item presents the energy consumption for amplifying radio. Depending on the transmission

distance, both the free space  $k\epsilon_{fs}$  and the multipath fading channel models  $k\epsilon_{amp}$  are used. When receiving this data, the radio expends:

$$E_{Rx}(l, d) = E_{Rx-elec}(l) = l \cdot E_{elec} \quad (8.3)$$

### 8.2.3 Problem Description

Cluster heads are selected randomly in LEACH. It is possible that the number of node joining the same cluster head could be too much or too little, which will eventually lead to cluster heads with the excessive number of nodes energy exhausted and die too fast [11]. But only a few nodes of the cluster heads energy will be surplus. Accordingly, the entire network load becomes imbalanced. In addition, because in LEACH protocol the cluster heads communicate with base stations by single-hop manner, its energy consuming and expandability is limited so it could not adapt to large network. In the following, we will discuss how to improve LEACH algorithm from two points: optimizing the method of cluster heads choosing and cluster heads routing.

## 8.3 The Second Choice of Cluster Head Algorithm

### 8.3.1 The Second Choice of Cluster Heads

Because the cluster heads of LEACH protocol are randomly generated by ordinary nodes in the network, from this way, there are some defects that cannot ensure the uniform distribution of cluster heads and guarantee the reasonable scale of the clusters. So the major scheme of LEACH-SC is to maintain the formula (8.1) without any changes and to propose the thought of second choice for cluster heads. After the clusters formed, the large scale of clusters will be broken, and then reselect cluster heads within the cluster; accordingly utilize multi-cluster heads to share the cluster together. Meanwhile, break up the small-scale clusters and cancel the cluster heads, and finally let the members of the cluster join other clusters. This can optimize the distribution of cluster heads, balance the scale of the clusters, and, to some certain extent, solve the problem of uneven distribution of cluster heads. However, we need to choose the number of cluster head and the scale of cluster reasonably, bringing them close or equal to the optimal number of cluster head. As is referred in Lv's paper [10], Eq. (8.4) was proposed.

$$k_{opt} = \sqrt{\frac{N}{2\pi}} \sqrt{\frac{\epsilon_{fs}}{\epsilon_{amp}}} \frac{M}{d_{bs}^2} \quad (8.4)$$

Where  $N$  is numbers of nodes in our experiment,  $\epsilon_{fs}$  and  $\epsilon_{amp}$  are amplifying coefficient.  $M$  is side length of square region, and  $d_{bs}^2$  denotes square of distance from cluster heads to base station. We choose  $\epsilon_{fs} = 10$  pJ/bit/m<sup>2</sup>,  $N = 100$ ,  $M = 100$ ,  $\epsilon_{amp} = 0.0013$  pJ/bit/m<sup>2</sup>,  $50 \leq d_{bs} \leq 71$ , next by the Eq. (8.4) we can get  $7 \leq kopt \leq 14$ . Accordingly, in the theory, the scale of each cluster is  $[N/14, N/7]$ , and that approximately means  $[7, 14]$ . So we can stipulate that when the cluster size is greater than 14 will be rebuilt and when its less than 7 will be broken.

### 8.3.2 Clusters Formation

After choice of the cluster heads have been completed, the message “joining cluster” will be broadcast within their communication range, to tell other nodes they are cluster heads. Typically, when receiving the broadcast message, the ordinary nodes will select a cluster joining in as the member of it.

In the LEACH algorithm, ordinary nodes choose which cluster to join in judging by their distance to the cluster heads. There being a shortcoming by this way. When the ordinary nodes distance to the sink node is far less than to the any cluster heads, they will do not join to the sink node, instead of the cluster heads, then the energy demand could be larger to communicate with the cluster heads than to communicate with the sink node directly. In the improved algorithm LEACH-SC, therefore, to join in a cluster for ordinary nodes, not only take into account the cluster heads, but think over the sink node, based on distance and energy consuming, and decide to join a cluster or as a stand-alone node communicating with the sink node directly.

### 8.3.3 Cluster Heads Routing

When without range of the sink nodes radio, the cluster heads cannot communicate with them directly, even if within the scope, resultingly far away from the sink nodes, it would consume greater energy to communicate between them. In the improved algorithm LEACH-SC, the cluster heads adopt the communication of single-multi-hop routing, which take the method of single-hop routing when close to the sink nodes, meanwhile, when the cluster heads farther from the sink nodes, they will use multi-hop manner to pass data to the sink nodes through other cluster heads.

### 8.4 Simulation and Analysis

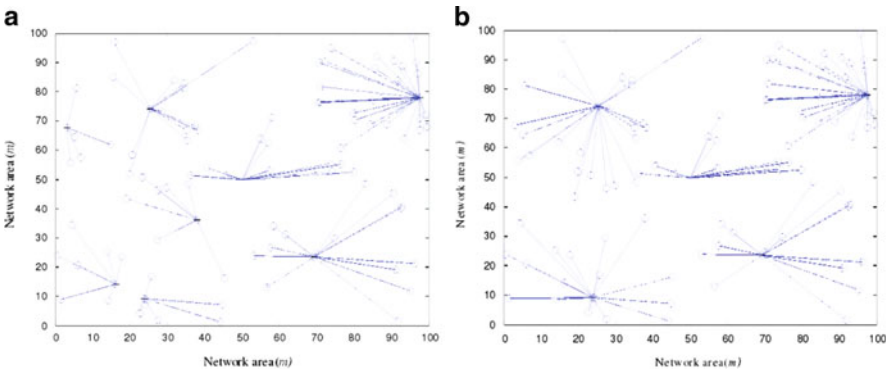
In order to evaluate the capability of the improved routing protocol, this paper makes a simulation for LEACH and LEACH-SC. For our experiments, in a zone with dimensions of  $\omega \times \omega$  and the distribution of  $N$  homogeneous sensor nodes are randomly simulated by MATLAB simulator. All nodes have the same initial energy  $E_0$  and energy consumption of radio  $E_{elec}$ . Table 8.1 shows the simulation parameters.

Figure 8.1a shows the building of clusters for the LEACH. From Fig. 8.1a, we can see the network appearing as many small clusters, which increase the number of cluster heads and cannot get them a reasonable distribution. If this kind of circumstance is appearing frequently, it will affect the overall energy distribution, so as to shorten the whole network of survival cycle. Figure 8.1b shows the building of clusters for the LEACH-SC. Compared to Fig. 8.1(a), Fig. 8.1(b) reveals that the number of cluster head is more reasonable and its cluster size is more balanced, which overcomes the disadvantage of LEACH clustering algorithm.

Figure 8.2 shows the total number of dead nodes over time, where the solid line represents LEACH-SC, and the forked line expresses LEACH. We assume the

**Table 8.1** Table of parameter value

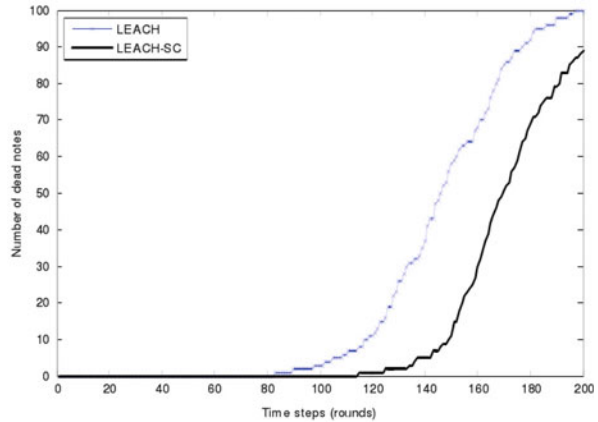
Description	Parameter	Value
Initial energy	$E_0$	0.5 J
Radio electronics energy	$E_{elec}$	50 nJ/bit
Radio amplifier energy for free space model	$\epsilon_{fs}$	10 pJ/bit/m <sup>2</sup>
Radio amplifier energy for multipath fading model	$\epsilon_{amp}$	0.0013 pJ/bit/m <sup>2</sup>
Desired probability of cluster heads	$P$	0.07
Data size	$l$	500 × 50 bit
Network size	$\omega \times \omega$	100 m × 100 m
Nodes	$N$	100



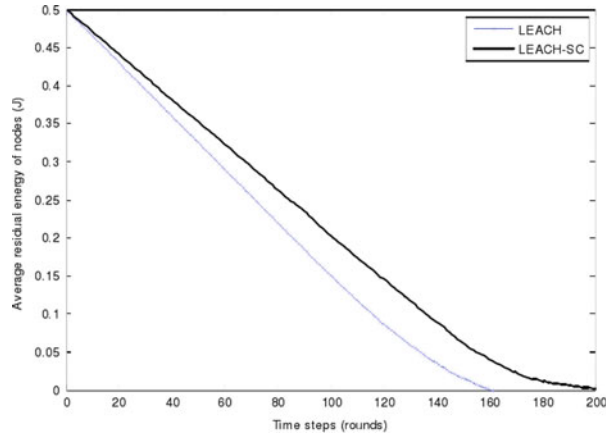
**Fig. 8.1** The cluster head choosing of LEACH (a) and LEACH-SC (b)



**Fig. 8.2** Number of dead nodes per rounds



**Fig. 8.3** Average residual energy of nodes per round



interval of First Node Dies (FND) as a steady period of network. Simulation results appear that  $FND = 85$  rounds for LEACH, whereas  $FND = 115$  for LEACH-SC. We can see that the improved algorithm increases 15 % of network stability on average comparing to LEACH. After running 200 rounds, the whole of network life increases 12 %.

Figure 8.3 is the result of average residual energy simulation. Judging from the figure, we can see the improved protocol have the more average residual energy of nodes than LEACH at any time, reflecting the greater energy-saving advantage. After running 160 rounds, the LEACH network of energy exhausted in advance, but the LEACH-SC still has a surplus energy to keep the network running and to extend the lifetime of the entire network up to 20 %. Thus, the improved protocol adopts a more reasonable means of cluster head election and intercluster communication, ultimately, effectively reducing the overall consumption of the network.

## 8.5 Conclusion

Cluster head is a chief undertaker of maximum energy consumption, and it is important to consider the energy consumption balance among cluster heads for WSNs. This paper introduced the improved routing algorithm LEACH-SC, which optimized the method of cluster heads choosing and cluster heads routing, realizing efficient energy distribution and prolonging the lifetime of network.

## References

1. Wu, C., & Hu, Y. (2009). Improvement of LEACH in wireless sensor networks. *Computer Technology and Development*, 19(3), 270–274.
2. Martirosyan, A., Boukerche, A., & Pazzi, R. W. N. (2008). Energy-aware and quality of service-based routing in wireless sensor networks and vehicular ad hoc networks. *Annals of Telecommunications*, 35(11), 669–681.
3. Wei, X., Wang, Y., & Yu, X. (2009). Cluster-heads multi-hop algorithm of LEACH protocol in wireless sensor network. *Journal of Guangxi Academy of Sciences*, 21(4), 89–92.
4. Wang, H.-Y., & Liu, S. (2010). A clustering algorithm based on fusion and transmission cost for WSN. *Dalian Ligong Daxue Xuebao in Chinese*, 50(4), 591–596.
5. Mao, S., & Zhao, C.-L. (2011). Unequal clustering algorithm for WSN based on fuzzy logic and improved ACO. *Journal of China Universities of Posts and Telecommunications in Chinese*, 18(6), 89–97.
6. Heinzelman, W. B. (2002). An application-specific architectures for wireless networks. *IEEE Transactions On Wireless Communications*, 1(4), 660–670.
7. Siv D. Murugana., Daniel C.F. Ma., Rolly I. Bhasin., Abraham O. Fapojuwo. (2005). A centralized energy-efficient routing protocol for wireless sensor networks. *IEEE Communications Magazine*, 43(3), 8–13.
8. Lindsey, S., & Raghavendra, C. S. (2002). PEGASIS: Power efficient gathering in sensor information systems. In *Proceedings of IEEE aerospace conference* (pp. 1125–1130). New York: The Aerospace Corporation.
9. Manjeshwar, A., & Agrawal, D. P. (2001). TEEN: A routing protocol for enhanced efficiency in wireless sensor networks. In *Proceedings of the 15th parallel and distributed processing symposium* (pp. 2009–2015). San Francisco: IEEE Computer Society.
10. Lv, T., Zhu, Q.-x., & Zhang, L.-q. (2011). An improved LEACH algorithm in wireless sensor network. *Acta Electronica Sinica Journal*, 39(6), 1405–1409.
11. Owais, A., Ahthsham, S., & Aamir, M. M. (2011). Comparison of routing protocols to assess network lifetime of WSN. *International Journal of Computer Science Issues*, 8(6), 220–224.

# Chapter 9

## An Ant Colony System for Dynamic Voltage Scaling Problem in Heterogeneous System

Yan Kang, Ying Lin, Yifan Zhang, and He Lu

**Abstract** Dynamic voltage scaling is an effective energy minimization technique by conjointly changing the supply voltage and the operational frequency during run-time. In this chapter, an improved ant colony system is presented for distributed systems consisting dynamic voltage scalable processing elements. The energy saving can be obtained by using the DVS algorithm on the schedule obtained by the presented scheduling algorithm. The pheromone information of the ants and the heuristic information inspired by the list heuristic rule and energy consumption are combined together to guide the ants search. The parameter value of heuristic is varied from higher value to lower value to lessen its impact on ants search, while the parameter value of pheromone information is increased during the run of ant algorithm. And the elitist solution is discarded if it cannot be improved from generation to generation. By cooperating several generations of artificial ants, the ants search for the path with a minimum energy consumption cost, and the quality of the solution can be improved for minimizing the energy consumption. Experiments are implemented to demonstrate the performance of the algorithm.

---

Y. Kang (✉) • Y. Zhang

Department of Software Engineering, School of Software, Yunnan University,  
Kunming 650091, Yunnan, China

e-mail: [kangyan@ynu.edu.cn](mailto:kangyan@ynu.edu.cn)

Y. Lin

Department of Information Security, School of Software, Yunnan University,  
Kunming 650091, Yunnan, China

H. Lu

School of Software, Yunnan University, Kunming 650091, Yunnan, China

## 9.1 Introduction

Energy efficiency has become an important issue of the distributed system especially the embedded system which is composed of processing elements (PES) normally powered by batteries. The power consumption depends on function executed, resulting in the various processing elements characteristics if gated clocks switch off unused circuit parts during idle periods. Energy consumption is at least a quadratic function of the supply voltage (hence CPU speed). Modern processors can reduce the power consumption by using dynamic voltage scaling (DVS) to utilize the idle intervals between the deadline and the real finishing time. It is shown that further energy consumption can be reduced by using the process element power characteristic during the voltage selection [1]. The voltage scales is investigated that DVS can decrease the power consumption by up to 10 times when executing real-life applications by taking full use of the intervals to process tasks as slowly as possible [2].

Yao presented one of the earliest theoretical models for DVS and presented an  $O(n^3)$  algorithm for computing a characterization of the minimum energy DVS schedule [3]. This optimal schedule did not make any special assumption on the power consumption function except convexity, and it has been referenced widely as it gave a main benchmark for evaluating other scheduling algorithms in both theoretical and simulation work.

There are static and dynamic techniques to using dynamic voltage scaling to decide appropriate operating voltage/speed. Schmitz researched static techniques by taking advantage of off-line parameters, such as periods and worst-case execution cycles [4]. Dynamic techniques (based on slack reclamation) exploit early completions of tasks to further decrease the speed and energy consumption. Here, we investigate static one. Scordino and Lipari used dynamic method to reduce the speed based on these predicted values and achieve more energy saving than static ones [5]. But the interest in static techniques is still high since static approaches can be enhanced to develop dynamic ones.

Pillai and Shin computed a single-optimal speed off-line and obtain the minimal speed to make a task set schedulable under earliest deadline first, and proposed a near-optimal method under rate monotonous [6]. Without fixing the processor speed, Saewong and Rajkumar assumed that the speed of the processor can be changed continuously in a given range and presented an algorithm to find the optimal speed value based on fixed priority assignments [7]. Based on the task parameters, Aydin proposed some methods to decide statically processor speed and assign different speed to every task before system execution [8, 9]. Liu and Mok gave a more general scheme to choose the speed switching instants more freely during the activation/deadline of some jobs [10].

Recently, there are several DVS-based algorithms proposed for slack allocation for a multiprocessor real-time system, while extensive researches on DVS scheduling algorithms have been proposed for independent tasks in a single processor real-time system.

In this chapter, we combine the schedule algorithm and dynamic scaling voltage strategy together to save the energy consumption effectively. Complicate task sets are concurrently processed on the distributed system with different degrees of parallelism. The presented scheduling algorithm wants to utilize the nonuniform workloads of processors and generates promising schedule pattern to reduce the energy consumption by using DVS strategy. Dynamic scaling voltage can reduce the energy consumption by exploiting the idle or slack time among the schedule pattern without obeying the precedence restricts and concurrence situation. By using ant colony optimization and heuristic strategy together to accumulate the pheromone information on path, the searches move towards the path with high trail, and find the solution with smaller energy consumption by iteratively ant generations. The algorithm includes three important co-synthesis steps: (a) assigning, determining the assignment of computational tasks to machines; (b) allocating, determining the execution order (sequencing) of tasks mapped to machines and communications cost; and (c) evaluation, determining the quality of the implementation candidate (energy consumption). The experiments are implemented to verify the performance of the algorithm.

This chapter is organized as follows: The next section briefly describe the problem model which includes task model and energy model. In Sect. 9.3, an ACO algorithm is presented while its performance is studied in Sect. 9.4. In Sect. 9.5, a conclusion is presented.

## 9.2 Task and Energy Mode

In the task scheduling problem, a deterministic set of tasks  $ST = \{st_1, st_2, \dots, st_n\}$  that are divided from an application is processed on a available set of processors  $SP = \{sp_1, sp_2, \dots, sp_m\}$  with fixed computing times  $SC = \{sc_{11}, sc_{12}, \dots, sc_{nm}\}$ . A directed acyclic graph (DAG) is given to demonstrate the set of tasks and the precedence restriction among them. The precedence constraints mean the successor task can be assigned after all its processor tasks have been completed. The communication cost  $cm_{i,j}$  represents the transfer time from precedence task  $t_i$  to its successor task  $t_j$  if they assigned on the different processors.

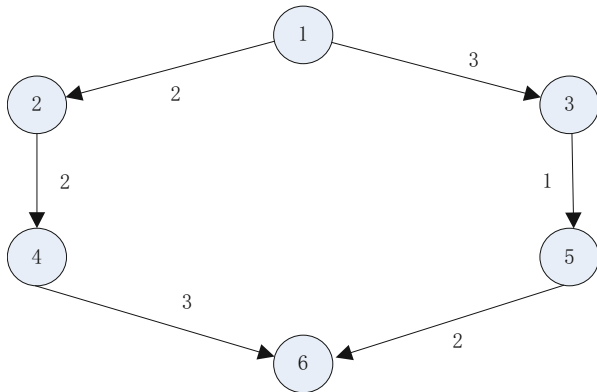
The task model is a fixed model, i.e., the execution time and the precedence constraint among the tasks are deterministic. Each task can be run on different processors with different processing time and cannot be interrupted before it finishes. All the processors can execute a task at a time, and they are continuously available and can process at most one task at a time. A feasible solution to the task scheduling problem satisfies the precedence constraint among the tasks.

It is shown that the power of a processing element is the sum of each CMOS circuit power which is composed of the static power and the dynamic one [4]. Most research groups have focused on reducing the dynamic power while the static power is negligible with respect to the dynamic power. The power consumption of a processor is calculated as

**Table 9.1** Computation time of Fig. 9.1

1	2	3	4	5	6
8 (32)	10 (30)	19 (47)	7 (19)	15 (42)	24 (61)
11 (41)	9 (27)	25 (52)	6 (43)	19 (36)	13 (41)

**Fig. 9.1** Example of a task graph with six tasks



$$E = P_{dv} \cdot t = C_{ef} \cdot V_{ds}^2 \cdot c \tag{9.1}$$

where  $P_{dv}$ ,  $C_{ef}$ ,  $V_{ds}$  are the dynamic power, the workload capacitance, and the supply voltage, respectively.

It is noted that various task scheduling patterns generate different initial energy consumptions and the slack and idle times. While a lot of researches focus on how to scale the tasks' execution speed to appropriate operating voltage/speed, we research the impact of task scheduling algorithm on the energy consumption.

First row in the table shows the tasks, second row shows the information about the first machine, and third row shows the information about the second machine. Table 9.1 gives the processing time (initial energy consumption) of the tasks in Fig. 9.1. Tables 9.2 and 9.3 show the normal execution time and scaling execution time of each task with the deadline 90, and the system consumption decreases from 2,482.56 to 945.19. It is demonstrated that energy saving obtained by DVS strategy is different based on the various schedule patterns.

The objective function of the scheduling problem by using DVS strategy is to find a feasible schedule with minimum energy consumption.

### 9.3 ACO–DVS Algorithm

Huang presented an ACO (ant colony optimization) algorithm combined with taboo search for the job shop scheduling problem [11]. In this chapter, a real number matrix ( $PT_{n \times m}$ ) is used to imitate the pheromone trail of the real ants and is updated during the run of the algorithm. When a good solution is found by assigning task  $st_i$

**Table 9.2** Energy consumption of Fig. 9.1 on one specific scheduling

1	2	3	4	5	6
0–8.45 (0–8.0)	8.45–18.72 (8.0–18.0)	18.72–41.41 (18.0–37.0)		41.41–58.7 (37.0–52.0)	58.7–90 (52.0–76.0)
			20.72–55.7 (20.0–26.0)		

**Table 9.3** Energy consumption of Fig. 9.1

1	2	3	4	5	6
0–11.96 (0–8.0)		11.96–43.11 (8.0–27.0)		43.11–67.02 (27.0–42.0)	
	13.96–44.83 (10.0–19.0)		44.83–69.02 (19.0–25.0)		69.02–90.0 (44.0–57.0)

on processor  $sp_j$ , pheromone is added to an element of the pheromone matrix  $T_{ij}$ . The following ants of the next generation directly use the value of  $T_{ij}$  to evaluate the promise of placing task on the processor  $sp_j$ .  $T_{ij}$  is actually simplified by omitting the iteration counter  $k$  from  $T_{ij}(k)$  which the amount of pheromone depends on previously found solution and will be changed after the following iteration. In our ACO, each task is considered as an ant, and the pheromone trail of ants is imitated by the solution obtained by the list heuristic strategy.

### 9.3.1 Initialization Phase

In ACO, the generated schedules by artificial ants may be so coarse that they should be improved by some complementary local search method. The reason that the earliest application of ACO to the problem generates unsatisfactory results may be due to the lack of an appropriate local search.

A set of artificial ants is initially created according to initial schedule and the pheromone information obtained by using the list scheduling heuristics. The energy consumption of each task and the total energy consumption of the schedule are used to generate the initial pheromones to all ants.

$$pst_i = \begin{cases} c + a/ec_{ij} + b/initec & \text{if } sch(i,j) = 1 \\ c & \text{else} \end{cases} \quad (9.2)$$

where  $sch(i,j) = 1$  means that task  $st_i$  is processing on processor  $sp_j$ , and  $ec_{ij}$  is the initial energy consumption of task  $st_i$ , and  $initec$  is the total energy consumption of the schedule obtained by the DVS algorithm.

### 9.3.2 Construction and DVS Phase

As the DVS slow down certain tasks by using available idle and slack time obtained following, the execution sequence of tasks is optimized by list scheduling strategy. The list scheduling heuristic is preferable than other heuristics in terms of the solution quality and time performance. The order of the tasks is obtained by using heuristic strategy, and then leads the ant to explore the new neighborhood based on the pheromone.

The execution sequence of the tasks is obtained by ordering the priority of tasks. The priority of tasks  $t_i$  is obtained as follows:

$$pst_i = se_{ij} + \max_{t_k \in tsu(t_i)} (ES_{ij}, se_{kj} + sc_{ij}) \quad (9.3)$$

where  $tsu(t_i)$  is the conjunctive and disjunctive successors of task  $st_i$ ,  $ES_{ij}$  is the earliest start time of task  $st_i$  if it is executed on processor  $sp_j$ .

The schedule is constructed according to the execution order obtained by the heuristic strategy, and then the processor is select based on the pheromone trail of the ants. The probability  $pr$ , a real number uniformly distributed in  $[0,1]$ , is generated randomly. The pheromone trails for task  $st_i$  is calculated as follows:

$$pt_{ij} = \begin{cases} \max \left\{ [pt_{ij}]^\alpha [ech_{ij}]^{ij} \beta \right\} & \text{if } pr < thre \\ [pt_{ij}]^\alpha [ech_{ij}]^{ij} \beta / \sum_{j=1}^m [pt_{ij}]^\alpha [ph_{ij}]^{ij} \beta & \text{else} \end{cases} \quad (9.4)$$

where the heuristic information  $ech_{ij}$  is the sum of earliest finish time and energy consumption for assigning task  $st_i$  on processor  $sp_j$ . The value of heuristic is decreased from generation to generation and is increased the pheromone values to further improve the solution quality it has found so far.

When extended by a time quantum which is a slice of the slack, the energy consumption of task  $t_i$  during the iterative  $j$  is computed as:

$$E_i^k(e_i + \Delta e_i) = E_i^k(e_i) \cdot V_{ids}^2 / div \cdot V_{iUp}^2 \quad (9.5)$$

where  $div$ , the number of the processors which run the direct successor of the task  $t_i$ , is used to evaluate the extend of energy saving.  $V_{ids}$  is the scaling voltage of task  $t_i$  and is given as:

$$V_{ids} = V_{it} + V_{iC} / rd_i + \sqrt{(V_{it} + V_{iC} / rd)^2 - V_{it}^2} \quad (9.6)$$

The strategy allocates the slack or idle time to the task with most perspective to optimize the power consumption of the tasks. And then it can reduce the processors power consumption more effectively by unevenly dividing the slack or idle interval



among the tasks. In the construction phase, each task is considered as an artificial ant, and the heuristic value is generated by some problem-dependent heuristic.

For lessening the influence of the old generations of ants, the corresponding pheromone information is updated by applying the local pheromone update rule as follows:

$$pst_{ij} = (1 - r) \cdot pst_{ij} + r \cdot C0 \tag{9.7}$$

Based on the best solution found so far and the evaporation of some portion of pheromone, the global pheromone information is updated as follows:

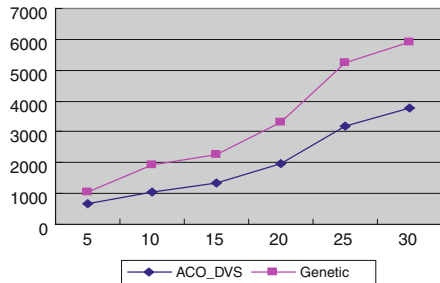
$$pst_{ij} = (1 - k)pst_{ij} + k/Btec \tag{9.8}$$

*Btec* is the best found energy consumption of the obtained schedule, and *r* and *k* are the pheromone evaporation rate.

### 9.4 Experiments

The algorithm was implemented in eclipse and run on Intel Xeon processors with 1 GHz speed with 1 GB of memory. The performance of ACO-DVS algorithm is compared with and genetic algorithm. For this purpose, we consider randomly generated tests and some specific sets by using various graph sizes and the energy consumptions. Figure 9.2 demonstrates the comparison of the energy consumption obtained by ACO-DVS and genetic algorithm. The horizontal line represents the number of the tasks, and the vertical line represents the average energy consumption. The average energy consumption obtained by ACO-DVS is smaller than that obtained by genetic algorithm for graph size from 5 to 20.

The results show that the ACO-DVS outperforms the general genetic algorithm. And the energy saving is larger when the number of tasks is increased.



**Fig. 9.2** Comparison of the energy consumption obtained by ACO-DVS and genetic algorithm

## 9.5 Conclusion

This chapter presented a combination of ACO algorithm and the DVS strategy for minimizing the energy consumption of the set of tasks executed on the heterogeneous distributed system. The heuristic information is obtained according to the energy consumption of the schedule system. The pheromone information and the list heuristic strategy are used to guide the search of the ants and generate the scheduling pattern. The factor of the heuristic and pheromone information is the variable based on the execution of the ant algorithm. Parameters of the algorithm are assigned by various values in the procedure of the ACO–DVS algorithm as parameter determines the convergence speed of the algorithm. At the beginning of the algorithm, parameter values are set to explore different regions of the search space and do not focus the search too early on a small region. At the end of the algorithm, parameter values are set to focus search near the best solution that has been found so far. The computational results on a set of random and specific instances testify the performance of the presented approach.

**Acknowledgments** This work has been supported by the Open Foundation of Key Laboratory in Software Engineering of Yunnan Province under Grant No. 2011SE03, Digital media technology and visualization of innovative communication platform under Grant No. 2012EI05, “CDIO-based software system modelling and design research and implementation” (Grant No. Rj14), and National Natural Science Foundation of China (Grant No. 60763008).

## References

1. Yun, H. S., & Kim, J. (2003). On energy-optimal voltage scheduling for fixed priority hard real-time systems. *ACM Transactions on Embedded Computing Systems*, 2(3), 393–430.
2. Burd, T. D., Pering, T. A., Stratakos, A. J., & Brodersen, R. W. (2000). A dynamic voltage scaled microprocessor system. *IEEE Solid-State Circuits*, 35(11), 1571–1580.
3. Li, M. N., & Yao, F. (2005). An efficient algorithm for computing optimal discrete voltage schedules. *SIAM Journal on Computing*, 35(3), 658–671.
4. Schmitz, M. T., Al-Hashimi, B. M., & Eles, P. (2004). Iterative schedule optimization for voltage scalable distributed embedded system. *ACM Transactions on Embedded Computing Systems*, 3(1), 182–217.
5. Scordino, C., & Lipari, G. (2006). A resource reservation algorithm for power-aware scheduling of periodic and aperiodic real-time tasks. *IEEE Transactions on Computers*, 12(55), 1509–1522.
6. Pillai, P., & Shin, K. G. (2001). Real-time dynamic voltage scaling for low-power embedded operating systems. In *Proceedings of the 18th ACM symposium on operating system principles* (pp. 89–201). New York, NY: ACM.
7. Saewong, S., & Rajkumar, R. (2003). Practical voltage-scaling for fixed-priority RT-systems. In *Proceedings of the 9th IEEE real-time and embedded technology and applications symposium* (pp. 106–115). Washington, DC: IEEE CS Press.
8. Aydin, H., Devadas, V., & Zhu, D. (2006). System-level energy management for periodic real time tasks. In *Proceedings of the 27th IEEE international real-time systems symposium* (pp. 313–322). Washington, DC: IEEE CS Press.

9. Aydin, H., Melhem, R., Mossé, D., & Mejia-Alvarez, P. (2004). Power-aware scheduling for periodic real-time tasks. *IEEE Transaction on Computers*, 53(5), 584–600.
10. Liu, Y., & Mok, A.K. (2003). An integrated approach for applying dynamic voltage scaling to hard real-time systems. In *Proceedings of the 9th IEEE real-time and embedded technology and applications symposium* (pp. 116–123). Washington, DC: IEEE CS Press.
11. Huang, K. L., & Liao, C. J. (2008). Ant colony optimization combined with taboo search for the job shop scheduling problem. *Computers and Operation Research*, 35(1), 1030–1046.

# Chapter 10

## An Improved Ant Colony System for Task Scheduling Problem in Heterogeneous Distributed System

Yan Kang, Yifan Zhang, Ying Lin, and He Lu

**Abstract** Task scheduling problem is a major issue of distributed system. An improved ant colony algorithm is presented to solve the scheduling problem in heterogeneous distributed system whose complexity is known to be NP-complete in general cases. To speed up the converging rate of the algorithm, elite initial solutions are generated by using an adaptable list heuristic algorithm which is a good tradeoff between the computation complexity and solution quality. A novel representation of pheromone can make effectively use of the task fitness value to accumulate pheromone in ACO (ant colony optimization) algorithm. To improve the self-adaptability of the algorithm, the ACO algorithm and the heuristic rule are combined together to adjust the searching space on the progress of the algorithm. Finally, local and global neighborhood searching are performed on the best solution obtained in iterations. Simulation results show that the performance of the improve ACO algorithm is better on finding optimal or near-optimal solutions than general genetic algorithm.

---

Y. Kang (✉) • Y. Zhang  
Department of Software Engineering, School of Software, Yunnan University,  
Kunming 650091, Yunnan, China  
e-mail: [kangyan@ynu.edu.cn](mailto:kangyan@ynu.edu.cn)

Y. Lin  
Department of Information Security, School of Software, Yunnan University,  
Kunming 650091, Yunnan, China

H. Lu  
School of Software, Yunnan University, Kunming 650091, Yunnan, China

## 10.1 Introduction

Scheduling for the tasks of an application represented as a directed acyclic graph (DAG) is very important in both fields of scientific computation and commercial application. The task scheduling problem is a branch of scheduling problem, which is complex combinatorial optimization problem and is very difficult to solve. It is well known that task scheduling problem is NP-complete, and its optimal solutions are hard to be achieved with traditional optimization approaches owing to the high computational complexity [1]. The classical task scheduling problem consists in scheduling a set of tasks on a set of processors with the objective to minimize the makespan of the schedule, subject to the precedence constraints among the tasks.

Various approaches have been proposed to solve task scheduling problem in the homogeneous system as this problem has been widely studied in the literature. Recently, many researches are focus on the application scheduling for achieving high performance in heterogeneous systems with nonuniform task processing time [2]. Since task scheduling problem in heterogeneous systems cannot be solved to guarantee optimality even under simplified assumption, many heuristics have been proposed for giving a suboptimal in polynomial time.

In this chapter, we improve the effectiveness and efficiency of the algorithm by proposing a hybrid algorithm for the task scheduling problem. The proposed algorithm combines a new list heuristic rule and ACO approach for the task scheduling problem. The list scheduling heuristic generates good initial solution by using characteristics of the heterogeneous multiprocessor scheduling problem, and ACO enhances it iteratively. The proposed hybrid algorithm has high performance in terms of both performance metrics (schedule length ratio, speedup, efficiency, and frequency of best results) and a cost metric (scheduling time).

This chapter is organized as follows: the next sections briefly describe the scheduling problem and a presented list scheduling algorithms, respectively. In Sect. 10.4, we propose an improved ant algorithm. In Sect. 10.5, experiments are given to investigate the performance of the algorithm which Sect. 10.6 concludes the chapter.

## 10.2 Task Scheduling Problem Model

In the task scheduling problem, a finite set of tasks  $T = \{t_1, t_2, \dots, t_n\}$  that are divided from an application is processed on a finite set of machines  $N = \{n_1, n_2, \dots, n_m\}$  with deterministic processing times  $E = \{e_{11}, e_{12}, \dots, e_{nm}\}$ .  $e_{i,j}$  is the execution time of task  $t_i$  on machine  $n_j$  that may be different on different processor depending on the processors computational capability. All tasks and their precedence constraints are represented as a DAG, and the precedence restriction means the successor task cannot be scheduled until all its processor tasks have been finished. The communication cost  $CM = \{cm_{11}, cm_{12}, \dots, cm_{nm}\}$ ,  $cm_{i,j}$  represents the cost of the data transfer from task  $t_i$  to its successor task  $t_j$ .

## 10.3 An Adaptable Scheduling Heuristic

### 10.3.1 Related Work

These heuristics are classified into a variety of categories such as list scheduling algorithms, clustering algorithms [3], genetic algorithms [4, 5], and task duplication-based algorithms [6, 7].

In list scheduling algorithms, the task in a list is constructed by assigning priority to it, and each task is assigned to the processor based on its priority. Several variant list scheduling algorithms have been proposed to deal with heterogeneous system, for example, Mapping Heuristic (MH), Levelized MinTime (LMT), Dynamic-Level Scheduling (DLS), Heterogeneous Earliest Finish Time (HEFT), and Critical Path On a processor (CPOP) [8, 9].

Ant colony optimization (ACO), one of the population-based metaheuristics dedicated to combinatorial optimization problems, has been successfully applied to a large number of discrete optimization problems, such as the traveling salesman problem and the vehicle routing problem, for which ACO was shown to be very competitive to other metaheuristics. Also, ACO has been applied successfully to task scheduling problems [10, 11].

With respect to list scheduling algorithms, it has been shown that minimizing the makespan of the tasks throughout the schedule is preferable to sequence the tasks in the descend order of their previously computed priorities, and then assign tasks on machines according to their sequence.

### 10.3.2 An Adaptable Scheduling Heuristic

If a task can be assigned to more than one predecessor and the computation cost of the task varies in the system, we think that the performance of the scheduling algorithm will be improved by considering the change ratio of the computation cost of the tasks. We present an elastic factor to imply the elastic of the task, i.e., the flexibility of the task performance. The priority of task  $t_i$  is given as

$$pt_i = \left( w_i + \max_{v_j \in succ(v_i)} (pt_j + cm_{ij}) \right)^* a + ef_i^* b, \quad (10.1)$$

where  $succ(t_i)$  is the set of immediate successors of task  $t_i$ ,  $w_i$ , the time-weight of task  $t_i$ , is the average computation cost of task  $t_i$ .  $a$ ,  $b$  are coefficients whose values are smaller than 1, and the sum of  $a, b$  is 1. And the value of the elastic factor  $ef_i$  is the computation time difference between the maximum computation time and the minimum computation time of task  $t_i$ .

The execution order (sequencing) of tasks is based on the priorities of all tasks. To assign priorities to all tasks, the upward rank of a task is computed as the critical path of that task.

Each task is scheduled onto the processor that gives the earliest completion time for the task. For the tasks to be scheduled, the value is computed recursively as shown in Eq. (10.2); all immediate predecessor tasks of  $t_i$  must have been scheduled.

$$EC_i = \min_{j \in n_i} \left( te_{ij} + \max_{t_j \in succ(v_i)} (AV_{ij}, e_{jk} + cm_{ij}) \right), \quad (10.2)$$

where  $AV_{ij}$  is the earliest time that processor  $n_k$  completed the execution of the last assigned task or the idle slot between the assigned tasks enough to schedule task  $t_i$  as early as possible. The inner max block in the  $EC$  equation returns the ready time, i.e., the time when all the data needed by  $t_i$  has arrived at processor  $n_j$ . The elastic factor of a task is given as a valid heuristic to minimize the schedule length. The solution is obtained by scheduling the tasks  $t_i$  onto processor  $n_j$  that gives the earliest finish time for the task.

## 10.4 Improved Ant Colony Optimization Algorithm

The general principle of ant algorithms is that the pheromone information reflects the outcomes of the decisions which have been made by former ants that found good solutions. In case of task scheduling problems, the decisions are which task to put on which place in the schedule. In our ACO, each task is considered as an ant, and the pheromone trail of ants is imitated by the solution obtained by the adaptable heuristic strategy. This pheromone trail, the executing situation of the processor, is updated during the run of the algorithm. The task is selected according to the order given by task prioritizing phase of the adaptable heuristic strategy. And ants select the processor based on its pheromone information and heuristic information. This is done to make explicit that the amount of pheromone depends on the current iteration and changes during the run of the algorithm. Yet, for simplicity, we omit the iteration counter.

### 10.4.1 Initialization Phase and Heuristic Rule

In ACO, the generated schedules by artificial ants may be so coarse that they should be improved by some complementary local search method. The reason that the earliest application of ACO to the problem generates unsatisfactory results may be due to the lack of an appropriate local search.

A set of artificial ants is initially created according to initial schedule and the pheromone information obtained by using the adaptable list scheduling heuristics given in Sect. 10.2. The execution situations of tasks give the pheromones to all ants, especially the execution situation of all tasks scheduled onto the same processor.

As in the task scheduling problem, the absolute position of a task is of importance; we use the following list scheduling strategy to generate the positions of tasks that represent the desire of setting task at the position in the sequence. And then ants build up new schedule by extending the already fixed prefix of the schedule and profit from existing list scheduling heuristics. The priority of task  $t_i$  is computed according to the former scheduling pattern as

$$pt_i = e_{ij} + \max_{t_k \in asu(v_i)} (AV_{ij}, e_{kj} + cm_{ij}), \quad (10.3)$$

where  $asu(t_i)$  is the set of immediate successors of task  $t_i$  and the immediate task running after task  $t_i$  on the same processor  $n_j$ ,  $cm_{ij}$  is the communication time between task  $t_i$  and its successor task  $t_j$ .

### 10.4.2 Construction Phase and Update Pheromone Rule

We construct a schedule by choosing the tasks based on the order given by the adaptable list heuristic strategy, and then a task can only be selected if all its predecessors are already in the partial schedule in each construction step. To assign each task onto a particular machine, an important issue is to define the pheromone trails for the machines to be assigned. When the random probability  $rp$ , a real number uniformly distributed in  $[0,1]$ , is less than the given probability  $gp$ , an ant chooses the machine with maximal pheromone for the task to be scheduled.

$$vt_i = \max \left\{ [phr(i,j)][heu(i,j)]^\beta \right\} \quad (10.4)$$

When the random probability  $rp$  is larger than the given probability  $gp$ , a processor is chosen according to the following probability distribution.

$$pt_i = \frac{[phr(i,j)][heu(i,j)]^\beta}{\sum_{j=1}^n [phr(i,j)][heu(i,j)]^\beta}, \quad (10.5)$$

where the pheromone information  $phr(i,j)$  is the last completion time for the processor  $n_j$ . And the heuristic information  $heu(i,j)$  is the earliest start time for the task  $t_i$  on the processor  $n_j$ , which is given as



$$heu(i, j) = e_{ij} + AV_{ij} + \max_{t_k \in succ(t_i)} (e_{kj} + cm_{ik}), \quad (10.6)$$

where  $AV_{ij}$  is the earliest start time that processor  $n_j$  can process task  $t_i$ .

The heuristic and pheromone information are indicators of how good it seems to put task  $t_i$  on the processor  $n_j$ . Parameters  $\beta$  determine the relative influence of the pheromone values and the heuristic values on the decision of the ant.

After an artificial ant has been assigned to a machine, the corresponding pheromone information is updated by applying the local pheromone update rule as follows:

$$phr(i, j) = (1 - \rho)phr(i, j) + \rho \cdot C, \quad (10.7)$$

where  $0 < \rho < 1$  is the pheromone evaporation rate,  $C$  is the initial pheromone level. The effect of the local pheromone update rule is to make the choice of putting  $t_i$  on the processor  $n_j$  less desirable for other ants to achieve diversification. Consequently, this mechanism favors the exploration of different schedules.

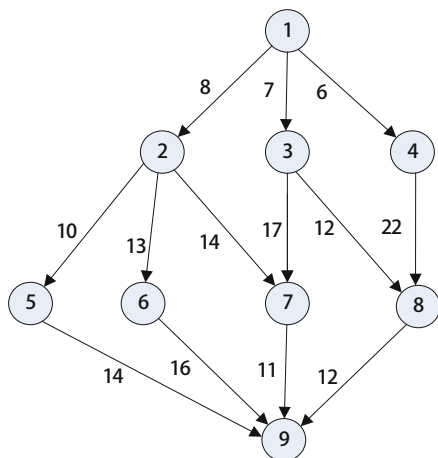
## 10.5 Experiment and Results

The algorithms described in Sect. 10.4 were implemented in eclipse-jee-indigo-SR1-win32 and run on Intel Xeon processors with 1 GHz speed with 1 GB of memory. Figure 10.1 demonstrates a DAG with 10 tasks and 15 edges on which the communication times of each task are labeled and assumed to be the same on the whole network.

Table 10.1 shows the processing times of each task on three available processors in the heterogeneous computing system. Table 10.2 shows the minimum results obtained by the IACO algorithm, i.e., begin time and completion time of all tasks. Table 10.3 shows the minimum results obtained by the genetic algorithm. Comparison between the genetic algorithm and IACO algorithm is shown in Tables 10.2 and 10.3 under the used number of ants is 52 and initial genetic individuals is 52 after 10 iterative. The schedule length obtained by IACO algorithm is 47 which is less than the schedule length obtained by genetic algorithm is 66. The improvement ratio between the difference between two results and the larger result is 28.8 %.

The performances and cost of the algorithms were compared with respect to set of experiments with various graph characteristics. The results show that the IACO algorithm is more effective than genetic algorithm with improvement is larger than 10 %.

**Fig. 10.1** Example of a task graph with 10 tasks



**Table 10.1** Communication and computation time of Fig. 10.1

V	1	2	3	4	5	6	7	8	9
P1	8	16	9	15	8	8	11	9	12
P2	7	17	11	12	9	14	17	13	14
P3	10	12	15	14	7	16	13	15	19

**Table 10.2** Schedule of Fig. 10.1 by IACO algorithm

V	1	2	3	4	5	6	7	8	9
P1			7-16		27-35	19-27			12
P2	0-7			7-19				19-32	14
P3		7-19					19-32		19

**Table 10.3** Schedule of Fig. 10.1 by genetic algorithm

V	1	2	3	4	5	6	7	8	9
P1	0-8			8-23			36-47		12
P2		8-25	25-36		36-45				14
P3						38-54		23-38	19

## 10.6 Conclusion

This chapter presented a combination of ACO algorithm and the list heuristic algorithm for the task scheduling algorithm in the heterogeneous multiprocessor system. We present a heuristic strategy to generate the initial solution for the ACO algorithm. And then new ant can generate by using pheromone and heuristic information based on the characteristic of the problem. The feasible solution is obtained by sequence the task based on the priority given by the heuristic rule. The computational results on a set of random and specific instances testify the

performance of the approach. And it is also demonstrated that the percentage of final schedule length is less than the initial one and the average improvement ratio are both sensitive to the graph structure and the initial one.

**Acknowledgments** This work has been supported by the Open Foundation of Key Laboratory in Software Engineering of Yunnan Province under Grant No. 2011SE03. “Digital media technology and visualization of innovative communication platform” under Grant No. 2012EI05, and National Natural Science Foundation of China (Grant No. 60763008).

## References

1. Cassavant, T., & Kuhl, J. A. (1988). A taxonomy of scheduling in general purpose distributed memory systems. *IEEE Transactions on Software Engineering*, 14(2), 141–154.
2. Ilavarasan, E., & Thambidurai, P. (2007). Low complexity performance effective task scheduling algorithm for heterogeneous computing environments. *Journal of Computer Sciences*, 3(2), 94–103.
3. Kafil, M., & Ahmed, I. (1998). Optimal task assignment in heterogeneous distributed computing systems. *IEEE Concurrency*, 6(3), 42–51.
4. Wu, A. S., & Jin, S. Y. (2004). An incremental genetic algorithm approach to multiprocessor scheduling. *IEEE Transactions on Parallel and Distributed Systems*, 15(9), 824–834.
5. Kaur, K., Chhabra, A., & Singh, G. (2010). Heuristics based genetic algorithm for scheduling static tasks in homogeneous parallel system. *International Journal of Computer Science and Security*, 4(2), 183–189.
6. Ahmed, I., & Kwok, Y. (1998). On exploiting task duplication in parallel program scheduling. *IEEE Transactions on Parallel and Distributed Systems*, 9(9), 872–892.
7. Bajaj, R., & Agrawal, D. P. (2004). Improving scheduling of tasks in a heterogeneous environments. *IEEE Transactions on Parallel and Distributed Systems*, 15(7), 107–118.
8. Iverson, M., Ozguner, F., & Follen, G. (1995). Parallelizing existing applications in a distributed heterogeneous environments. In *Proceedings of heterogeneous computing workshop* (pp. 93–100). Washington, DC: IEEE CS Press.
9. Topcuoglu, H., Hariri, S., & Wu, M. Y. (2002). Performance effective and low-complexity task scheduling for heterogeneous computing. *IEEE Transactions on Parallel and Distributed Systems*, 13(3), 260–274.
10. Chen, L., & Pan, Q. (2011). Improved ant colony algorithm to solve identical paraHel machine task scheduling problem. *Computer Engineering and Applications*, 47(6), 44–48.
11. Deng, R., Cheng, H. Z., Wang, B., Wang, X. M., & Li, C. (2010). Permutation ant colony system for heterogeneous DAG scheduling problem. *Journal of Computer Science*, 12(37), 193–196.

# Chapter 11

## Optimization of Green Agri-Food Supply Chain Network Using Particle Swarm Optimization Algorithm

Qian Tao, Zhexue Huang, Chunqin Gu, and Chenxin Zhang

**Abstract** The green agri-food supply chain network (GASCN) design is critical to reduce the total transportation cost for efficient and effective supply chain management. This paper proposes a new solution based on particle swarm optimization (PSO) to find optimal solution for GASCN problem. PSO adopts transforming operator to modify particles in the population. The novelty of the transforming operator is that it can avoid applying the penalty function so that the diversity of populations is decreased. To show the efficacy of the algorithm, PSO is also tested on three cases. Results show that the proposed algorithm is promising and outperforms GA by both optimization speed and solution quality, especially when the scale of problem is large.

---

Q. Tao (✉)

Department of Computer Science, Guangdong University of Education,  
Guangzhou 201305, China

Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences,  
Shenzhen 201305, China

e-mail: [taoalex66@gmail.com](mailto:taoalex66@gmail.com)

Z. Huang

Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences,  
Shenzhen 201305, China

C. Gu

Department of Computer Science, Zhongkai University of Agriculture  
and Engineering, Guangzhou 201305, China

C. Zhang

School of Software Engineering, Sun Yat-Sen University,  
Guangzhou 201304, China

## 11.1 Introduction

With the increasing demand for green agri-food, the green agri-food supply chain network (GASCN) design problem has been gaining focus of more and more researchers recently. Due to increasing competitiveness, the logistics firms are obliged to maintain high customer service levels at the same time they are forced to reduce total transportation cost for profit maximization.

As any other SCN, the GASCN is a network of organizations in order to bring agri-food to customers. The network involves planting bases (PB), distribution centers (DC), and points of sales (POS) with purpose of satisfying POS's demands, beginning with the PB of green agri-food, following with DC and ending with POS such as vegetable markets and supermarkets. Traditionally, manufacturing, distributing, and marketing along the supply chain operated independently. These organizations have their own objectives, and these objectives are often conflicting. But there is a need for a mechanism through which these different functions can be integrated together [1].

As for the SCN problem, the models and solutions for industrial products are relatively massive, but the solutions of the SCN about green agri-food are limited. In literature, there are many different studies on the design problem of supply networks, and these studies have been reviewed by Erenguc et al. [2] and Pontrandolfo et al. [3]. Amiri [4] has presented a Lagrangian relaxation approach to minimize the total cost of two-stage supply chain. Costa et al. [5] have worked on three stages of SCN optimization problem. Most of the researchers have concentrated on the improvement of the supply chain performance, and very few have considered the performance improvement of the algorithm concurrently [6].

Due to the NP hardness of SCN problems [7] and the large-sized problems in the real world, intelligent algorithms such as genetic algorithms (GA) [8] have been proposed to solve the SCN problems. Particle swarm optimization (PSO) [9], which is an important optimization tool as well as a continuous optimization method, was proposed by Kennedy and Eberhart in 1995. As a swarm intelligence algorithm, PSO simulates preying behaviors of bird flocking and fish schooling and searches for the optimal solution iteratively. Since PSO was put forward, it has attracted much attention from many scholars especially in the research fields of multi-objective problems [10]. The standard PSO has great advantage in searching for non-inferior solutions in a multidimensional complex space.

In this paper, an enhanced PSO is proposed in order to solve the GASCN problem for optimization. The proposed algorithm can be applied to SCN problems. The distinct feature of PSO is that it adopts transforming operator to modify chromosomes in the population and uses effective genetic operations. The transforming operator can assure that the solutions are always feasible. The performance of the proposed PSO has been compared with the state-of-the-art GA [8]. Results show that the proposed algorithm can achieve high-quality solutions with a much faster optimization speed.

The remainder of this paper is organized as follows. Section 11.2 presents the statement of the optimization problem. Section 11.3 describes the implementation of the proposed algorithm in detail, including the encoding method of particles, the design of transforming operator during the process of initialization of chromosomes, the design of fitness function, the identification of the local best position and global best position, and updating the particle's velocity and position. In Sect. 11.4, a series of experiments are conducted, and the results are analyzed to illustrate the performance of the proposed algorithm. Finally, in Sect. 11.5, this paper is concluded, and suggestions are given for future research.

## 11.2 Problem Statement

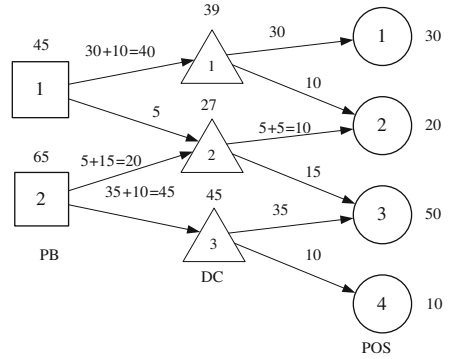
In order to optimize performance of GASCN, we should design the distribution network strategy that will satisfy demand requirement for the products imposed by points of sales. The problem is a SCN design problem for single green agri-food. The main objective to solve such an optimization problem of GASCN is to evaluate the selection of different planting bases or set of the planting bases and different distribution centers or set of the distribution centers, whereas the performance criteria are the minimization of the total transportation cost. Simultaneously, the demand for green agri-food from each point of sales must be satisfied. The assumptions used in this problem are as follows.

Planting bases, distribution centers, and points of sales are known. The requirement for green agri-food from each point of sales is known. Green agri-food and distribution centers are enough for distribution. The transportation costs of green agri-food on the path from each planting base  $pb_i$  to each distribution center  $dc_j$  and from each  $dc_j$  to each point of sales  $pos_k$  are known. Green agri-food distributed to  $pos_k$  are supplied by a  $dc_j$  or multi- $dc_j$ , and green agri-food distributed to  $dc_j$  are supplied by a  $pb_i$  or multi- $pb_i$ .

Suppose there are planting bases  $PB = \{pb_1, pb_2, \dots, pb_{|PB|}\}$ ; distribution centers  $DC = \{dc_1, dc_2, \dots, dc_{|DC|}\}$ , points of sales  $POS = \{pos_1, pos_2, \dots, pos_{|POS|}\}$ ; a green agri-food SCN  $GASCN \subseteq PB \times DC \cup DC \times POS$ , where “ $\times$ ” is Cartesian product; and demand of POS  $Req_{POS} = \{Req_{pos_1}, Req_{pos_2}, \dots, Req_{|POS|}\}$ , the objective of optimization of GASCN is to find  $PB_i = \{pb_{i1}, pb_{i2}, \dots, pb_{i|PB_i|}\} \subseteq PB$ ,  $DC_j = \{dc_{j1}, dc_{j2}, \dots, dc_{j|DC_j|}\} \subseteq DC$ ,  $GASCN = \{(pb_i, dc_j), (dc_j, pos_k) | pb_i \in PB_i, dc_j \in DC_j, pos_k \in POS\}$ , satisfying

$$\min f = TQ_i^j \times Cos t_{pb_i}^{dc_j} + TQ_j^k Cos t_{dc_j}^{pos_k}$$

**Fig. 11.1** A GASCN application used in our experiments



**Table 11.1** Notations

Symbol	Descriptions
$PB$	Set of planting bases
$pb_i$	Planting base $i$
$DC$	Set of distribution centers
$dc_j$	Distribution centers $j$
$POS$	Set of points of sales
$pos_k$	Points of sales $k$
$TQ_i^j$	Transportation quantity from node $i$ to node $j$
$X_i$	Particle $i$
$x_i^n$	The position of dimension $n$ particle $i$
$M$	Number of particles in the population
$N$	Total number of impossible paths
$n$	The index of paths, $n \in [1, N]$
$Req_{pos_k}$	The requirement of point of sales $k$
$f\_pb(n)$	The map from $n$ to a planting base
$g\_dc(n)$	The map from $n$ to a distribution center
$h\_pos(n)$	The map from $n$ to a point of sales
$Cost_i^j$	Unit transportation cost from node $i$ to node $j$

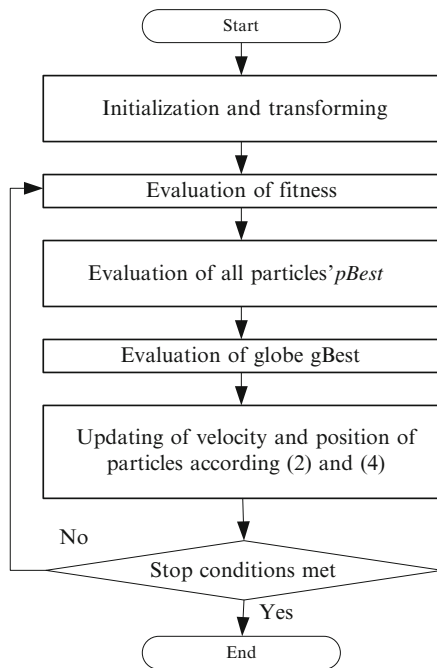
$$Req_{pos_k} = TQ_{dc_j}^{pos_k}$$

A GASCN is shown in Fig. 11.1. Table 11.1 lists the notations used in this paper.

### 11.3 Proposed PSO Algorithm

This paper proposes a PSO algorithm for optimization of green agri-food supply chain networks. In PSO, a novel transforming operator is applied. In addition, the method of representation is the critical to solve the optimization problem. In this

**Fig. 11.2** Flowchart of the proposed PSO



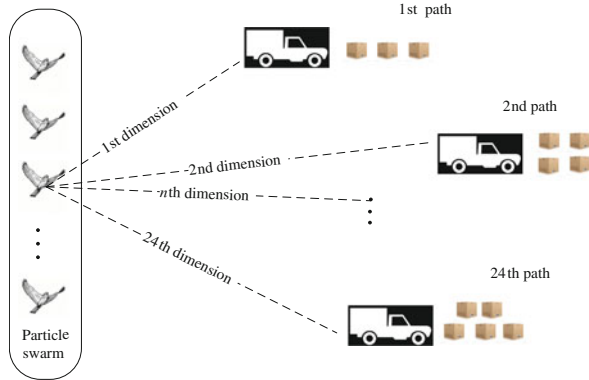
section, the particle’s representation is firstly described. Then the particle’s initialization, the evaluation of fitness, and the identification of the local and global best position, updating the particle’s velocity and position, are presented. A complete flowchart of the proposed PSO is shown in Fig. 11.2.

### 11.3.1 Encoding Method of a Particle

The encoding of particles is the premise of using particle swarm to search for an optimal solution in a GASCN problem. In the PSO, the number of particle dimensions is equal to the number of paths from *PB* to *POS* by *DC*, and each dimension corresponds to the transportation quantity of green agri-food on each path. In Fig. 11.2, the GASCN has 2 *PB*, 3 *DC*, and 4 *POS*, and the particles are encoded by 24 ( $2 \times 3 \times 4 = 24$ ) dimensions. The number of solution in the searching space of a particles is  $|PB| \times |DC| \times |POS|$ . Figure 11.3 gives an illustration of a searching of particle swarm in a 24-dimensional space.



**Fig. 11.3** An illustration of searching of particle swarm in a 24-dimensional space



### 11.3.2 Evaluation of Fitness

The fitness function of a particle  $X_i$  in the population is defined as

$$fitness_i = \sum_{n=1}^N x_i^n \times \left( Cost_{f-pb(n)}^{g-dc(n)} + Cost_{g-dc(n)}^{h-pos(n)} \right) \quad (11.1)$$

where  $f-pb(n) = \frac{n}{|DC| \times |POS|}$ ,  $g-dc(n) = \frac{n}{|POS|} \% |DC|$ ,  $h-pos(n) = n \% |POS|$ ,  $Cost_{f-pb(n)}^{g-dc(n)}$  is the transportation cost of green agri-food from  $PB_{f-pb(n)}$  to  $DC_{g-dc(n)}$ , and  $Cost_{g-dc(n)}^{h-pos(n)}$  is the transportation cost from  $DC_{g-dc(n)}$  to  $POS_{h-pos(n)}$ . Moreover, every particle is evaluated according to Eq. (11.1).

### 11.3.3 Updating the Particle's Velocity and Position

$$v_i^n(t+1) = \omega v_i^n(t) + c_1 \cdot r_1 \cdot (pBest_i^n(t) - x_i^n(t)) + c_2 \cdot r_2 (gBest_n(t) - x_i^n(t)) \quad (11.2)$$

$$\omega = 1.0 - gen^*0.6/499 \quad (11.3)$$

$$x_i^n(t+1) = x_i^n(t) + v_i^n(t+1) \quad (11.4)$$

The position and velocity of the particle are updated by Eqs. (11.2) and (11.4), respectively. In Eq. (11.2),  $\omega$  is the inertia weight and is calculated by Eq. (11.3),  $c_1, c_2$  are acceleration constants (also known as cognitive parameters and social parameters), and  $r_1, r_2$  are the independent random variables evenly distributed within [0, 1]. If the new particle does not satisfy the requirement of points of sales, the transforming operator is applied.

### 11.4 Experimental Study and Discussions

The proposed PSO for optimization of green agri-food supply chain network was performed on a machine with Pentium D 3.00GHz CPU and 1024MB of RAM. The operating system was MS Windows XP, and the compiler was VC++ 6.0. We tested the PSO on Case\_1, Case\_2, and Case\_3 from small scale to large scale (Table 11.2).

The following parameters were used in the PSO:  $M = 20$ ,  $c_1 + c_2 = 1.8 + 1.8 < 4$ . To validate the proposed PSO, we compared the PSO with GA [10] to tackle the various problems in GASCN. Experiments were independently simulated 30 times. Table 11.3 presents the comparison of GA and PSO. To validate the reliability and robustness of PSO, we tested two algorithms with a longer running time and 10,000 iterations. The mean fitness, best fitness, the average time in second used for obtaining the best result in each run of the two algorithms are compared with each other. The statistical results are shown in Table 11.3. From Table 11.3, GA outperforms PSO on Case\_1 in speed, but as the increasing of cases scale, PSO outperforms GA both in the solution quality and in the optimization speed.

**Table 11.2** Test cases

Test cases	$ PB $	$ LC $	$ ST $	Scale	$Req_{st_k}$	$Cost_t^i$
Case_1	2	3	4	24	[10, 50]	[1, 10]
Case_2	10	15	16	2,400	[10, 50]	[1, 10]
Case_3	15	20	25	7,500	[100, 500]	[1, 10]

**Table 11.3** Convergence speed and the solution accuracy comparisons

Test cases	GA			PSO		
	Mean fitness	Best fitness	Time (s)	Mean fitness	Best fitness	Time (s)
Case_1	758.978	750.229	17.656	590	590	24.757
Case_2	2,878.92	2,874.12	421.563	1,253.598	1,110	171.829
Case_3	78,239.3	76,546.6	2,867.44	39,694.79	36,667.6	689.686

<sup>a</sup>Convergence speed being measured on the mean CPU time needed to reach an acceptable solution; solution accuracy being measured on the best fitness and the mean fitness on 30 independent trials

## 11.5 Conclusion

In this paper, we have proposed a PSO for optimization of green agri-food supply chain network. PSO uses float-point representation to encode the SCN problems. In PSO, a transforming operator is used to keep solutions in feasible ones when total transportation number does not meet the demands of each POS. Experiments were conducted on three cases, and the experimental results demonstrate the superiority and effectiveness of the proposed algorithm for green agri-food supply chain network. In our future research, to the best of our knowledge, we will study more effective and efficient solution methodology based PSO algorithm.

**Acknowledgments** This work is supported by the Special Funds for Doctoral Research Project of Guangdong University of Education under Grant No. 2012ARF05, the Natural Science Foundation of Guangdong Province of China under Grant No. S2012020011067 and the National High Technology Research and Development Program of China (863) under Grant No. 2012AA101701.

## References

1. Altıparmak, F., & Gen, M. (2006). A genetic algorithm approach for multi-objective optimization of supply chain networks. *Computers and Industrial Engineering*, 51(1), 197–216.
2. Erenguc, S. S., Simpson, N. C., & Vakharia, A. J. (1999). Integrated production/distribution planning in supply chains: an invited review. *European Journal of Operational Research*, 115(2), 219–236.
3. Pontrandolfo, P., & Okogbaa, O. G. (1999). Global manufacturing: A review and a framework for planning in a global corporation. *International Journal of Production Economics*, 37(1), 1–19.
4. Amiri, A. (2006). Designing a distribution network in a supply chain system: Formulation and efficient solution procedure. *European Journal of Operational Research*, 171(2), 567–576.
5. Costa, A., Celano, G., Fichera, S., & Trovato, E. (2010). A new efficient encoding/decoding procedure for the design of a supply chain network with genetic algorithms. *Computers and Industrial Engineering*, 59(4), 986–999.
6. Prakash, A., Chan, F. T. S., Liao, H., & Deshmukh, S. G. (2012). Network optimization in supply chain: A KBGA approach. *Decision Support Systems*, 52(2), 528–538.
7. Altıparmak, F., Gen, M., Lin, L., et al. (2006). A genetic algorithm approach for multi-objective optimization of supply chain networks. *Computers and Industrial Engineering*, 51(1), 196–215.
8. Henry, C. W. L. (2009). Cost optimization of the supply chain network using Genetic Algorithms. *IEEE Transactions on Knowledge and Data Engineering*, 99(1), 1–36.
9. Kennedy, J., & Eberhart, R. C. (1995). Particle swarm optimization. In *Proceedings of IEEE international conference on neural networks* (pp. 1942–1948). Piscataway, NJ: IEEE Press.
10. Heo, J., Lee, K., & Garduno-Ramirez, R. (2006). Multiobjective control of power plants using particle swarm optimization techniques. *IEEE Transactions on Energy Conversion*, 21(2), 552–561.

# Chapter 12

## A New Model for Short-Term Power System Load Forecasting Using Wavelet Transform Fuzzy RBF Neural Network

Jingduan Dong, Changhao Xia, and Wei Zhang

**Abstract** Power load changes periodically. And the effects of climatic (precipitation, relative humidity, temperature, wind speed) on the load should be fuzzy. In order to solve the problem, this chapter presents a method combining wavelet transform, fuzzy set concept, and neural networks for short-term load forecasting. Through the wavelet transform, the load sequence decomposes into subsequences consisting of different wavelet coefficients. On the other side, by the fuzzy neural network, the samples of five meteorological factors influencing power load are transformed into fuzzy input with the subsequences, and then, the suitable RBF neural networks for the forecasting are selected. Finally, the load forecasting sequence is obtained by the reconstruction of the forecasted results from the subsequences. The simulation results demonstrate the proposed method possesses validity and practicability with a mean absolute error below 1.5 %.

### 12.1 Introduction

Power load forecasting is to study the change rule and variation trend of the electric load. Accurate power load forecasting can improve the security and stability as well as economy of the power systems. Because of the factors such as climate, economy, price, and policy many more efforts have been made towards the application and improvement of power load forecasting techniques in order to adapt to the impact of these factors. Short-term load forecasting is an essential part of power system operation mainly used in optimal unit commitment, generation schedule, hydro-thermal coordination, power exchange schedule, and so on [1]. Accurate prediction can forecast the load variation and lower power loss for power grid in advance.

---

J. Dong (✉) • C. Xia • W. Zhang  
College of Electrical Engineering and New Energy, China Three  
Gorges University, Yichang 443002, China  
e-mail: [76471786@qq.com](mailto:76471786@qq.com)

However, the power load is uncertain and changeable by external factors. The study shows that the impact of different climatic factors on the electricity load is also different from each other. Climate change is one of the important reasons of load fluctuations which make further research on improving the prediction methods [2].

There are various methods used in load forecasting over the years. The traditional short-term load forecasting methods are time series, autoregressive average, exponential smoothing, etc. The prediction models in these mathematical expressions cannot meet the requirements of power systems nowadays. The modern prediction algorithm artificial neural network (ANN) will be introduced in this chapter for the reason that the network can approximate any nonlinear function then model the performance and learn characteristics. ANN has been widely applied in short-term load forecasting in power system.

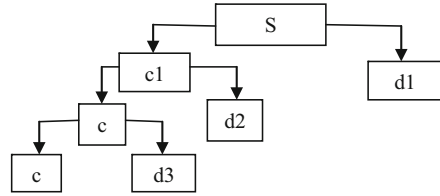
This chapter presents a wavelet transform fuzzy neural network model. It uses the subsequences of the load and the transformed climate fuzzy variable as the input of the neural network. The subsequences of the load are decomposed by wavelet transform, and the climate variables are transformed by membership functions. Then, the predicted results are obtained by the reconstruction through inverse wavelet changes to get the load sequence of the forecast day. The simulated results demonstrated that the accuracy meets the demand.

## 12.2 Wavelet Decomposition and Reconstruction

To process data scientifically and rationally is the most basic part to improve the load forecasting precision. For the short-term load forecasting, load data also need the process in an appropriate way. It is because of the power load changes periodically and will also fluctuate with the users' contingency change and the variety of climate conditions (precipitation, relative humidity, temperature, wind speed) at the same time that the various types of load signal generally performs as a continuous spectrum. This load type can be transformed into nonperiodic components, and these load components have different frequency characteristics including low and high frequencies [3]. By this way, the load transforms into different scales of wavelet coefficients. Then, the feature vector composed by the wavelet coefficients replaces the load data for neural network training. Finally, the load forecasting sequence was obtained by the reconstruction of the forecasted results from the output.

The historical power load data can be treated as time series signals. An original signal can be expressed by a multi-resolution representation. In order to give a brief introduction on multi-resolution decomposition, we need to take a three-tier multi-resolution decomposition, for example. The signal decomposition processes is shown in Fig. 12.1, where  $S$  means the original signal,  $c$  is low-frequency component, and  $d$  is high-frequency component.

**Fig. 12.1** The decomposition process



As we can see from the figure, multi-resolution analysis is just the decomposition of low-frequency part, and the high-frequency part will not be considered further. What is more, the multi-resolution analysis is essentially an iterative process of signal analysis between the two adjacent resolutions [4].

Mallat algorithm is employed in this chapter for wavelet decomposition and reconstruction. Using Mallat algorithm, the known signal can be decomposed into different scales. Mallat decomposition algorithm uses the extraction method that the subsequence is shortened to half of the original sequence. So it greatly reduces the amount of computation of the wavelet transform that improves computing speed which is conducive to processing large amounts of information. Briefly, in the space of discrete sequence, under the prerequisite of Shannon sampling theorem, the Mallat algorithms use a low-pass filter  $h(k)$  and a high-pass filter  $g(k)$  to the sampling signal binary band division. Through observing the signal in different frequency bands, the signal feature can be extracted. Besides, Daubechies function is selected as the mother wavelet.

The decomposition algorithm has the form of

$$c_{j+1,k} = \sum_{l \in \mathbb{Z}} c_{j,l} \overline{h_{l-2k}} \tag{12.1}$$

$$d_{j+1,k} = \sum_{l \in \mathbb{Z}} c_{j,l} \overline{g_{l-2k}} \tag{12.2}$$

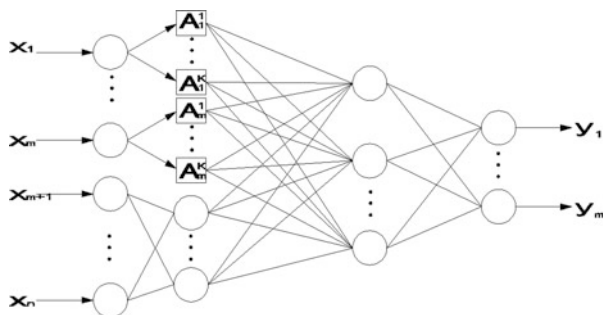
On the contrary, the algorithm also provides the simple way to reconstruct the original single:

$$c_{j,k} = \sum_{l \in \mathbb{Z}} h_{k-2l} c_{j+1,l} + \sum_{j \in \mathbb{Z}} g_{k-2l} d_{j+1,l} \tag{12.3}$$

### 12.3 Wavelet Transform Fuzzy Neural Network Forecasting Model

The neural network model based on the method of this chapter is shown in Fig. 12.2 where  $x_1 \sim x_m$  are original load inputs and  $x_{m+1} \sim x_n$  are original climatic factors inputs. The way to deal with load sequence via wavelet transform will not be given

**Fig. 12.2** Wavelet transform combined with fuzzy RBF neural network forecasting model



unnecessary details here. As for climate data, they are turned into fuzzy input via membership functions. The learning network chooses radial basis function (RBF) neural network which is similar to nonlinear continuous function with uniform approximation so as to speed learning process.

This new model considers the climatic factors affecting load, the analysis of the uncertainty factors, load data processing, and the basic forecast model. It combines these algorithms in an optimal way which ensures the accuracy of the prediction.

### 12.3.1 Fuzzification of Meteorological Factors

Fuzzy neural network combines fuzzy theory with neural network. The fuzzy logic is inserted into the fuzzy neural network instead of Boolean logic, so that the network can obtain the ability of fuzzy inference to improve the learning speed. Fuzzy neural network just blurs some of elements and retains the basic properties and structure.

The fuzzy set is the collection of elements that with some specific characters in some degree. The characteristic function which describes the elements' degree of membership in a fuzzy set is membership function. Furthermore, fuzzy set theory is the method of mapping from region to region while ANN mapping from point to point. Through combining the merits of the two, the network can be applied to different power systems.

Natural factors (typical micro factors such as temperature, precipitation, and wind speed) will lead to short-term relative volatility of the load [5]. Lots of load forecasting models just consider the influence of temperature on load. Obviously, temperature is an important indicator of the load fluctuation [6]. However, some other natural factors will affect human comfort so that cause fluctuations of power load. The fact is people tie between meteorological factors and load change. Specifically, meteorological factors affect the comfort of the human and thereby affect human behavior which will lead the load change [7].

There are five primary factors (i.e., daily maximum temperature, daily minimum temperature, precipitation, relative humidity, wind speed) which will be considered as the fuzzy input. The selection of the membership functions is based on the output and input functions linear or not. Due to the inputs above are mainly linear factors, we elect 2~3 membership functions for simplified precondition here. For these inputs, use fuzzy number of triangular or trapezoidal distribution to express the input spatial distribution. Therefore, we can simultaneously obtain multiple fuzzy membership degree of the input if a specific input value has been given.

### 12.3.2 Radial Basis Function Neural Network for Training

The RBF neural network involves three layers with only one hidden layer which applies a nonlinear transformation from input space to hidden space in its most basic form. Comparing to BP network, it has a higher learning efficiency and function approximation.

The input in this chapter is a feature vector constituted by the load coefficients of wavelet and fuzzy meteorological factors. And the input-output mapping of RBF network has the following form:

$$f(x) = \sum_{j=1}^m w_j \varphi(\|x - x_j\|) \quad (12.4)$$

where  $m$  is the number of hidden nodes,  $w_j$  is the adjustment weight,  $\varphi(\|x - x_j\|)$  is a set of  $m$  arbitrary functions, and  $\|\cdot\|$  refers to Euclidean norm.

For regularization RBF network itself, while it is in the process of learning, the hidden layer's activation functions evolve slowly in accordance with some nonlinear optimizations strategies, and output layer's weights adjust themselves rapidly through a linear optimization strategy. Here, take the method of fixed center selected at random for learning and employ an isotropic Gaussian function whose standard deviation is fixed according to the spread of the centers. In effect, the standard deviation of the Gaussian radial basis functions is fixed at

$$\delta = \frac{d_{\max}}{\sqrt{2P}} \quad (12.5)$$

where  $d_{\max}$  is the maximum distance between the chosen centers and  $P$  is the number of centers. This formula ensures that the individual RBF is not too peaked or too flat. Both of these two extreme conditions should be avoided.

Then, the weights of output layer adopt least mean square algorithm (LMS), and the weight adjustment takes the following form:



$$\Delta W_k = \eta(d_k - W_k^T \Phi) \Phi \quad (12.6)$$

In formula (12.6),  $W_k$  is the linear weight in the output layer of the network,  $d_k$  is the desired respond, and  $\Phi$  is an  $m$ -by- $m$  matrix with elements.

Only a small number of connection weights need to be adjusted by training. It is because of this characteristic that the RBF neural network has a relatively higher learning speed. And RBF network proves to be effective in power load forecasting [8].

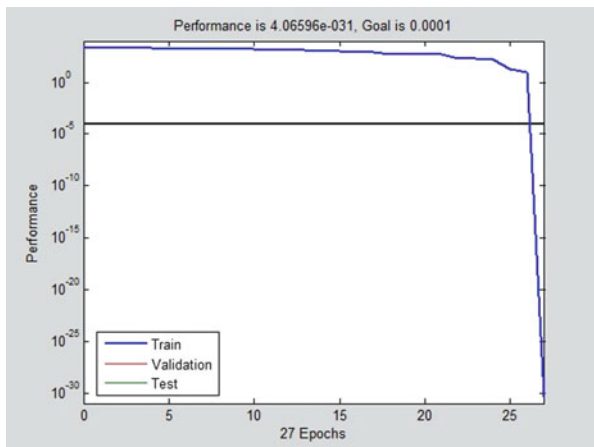
## 12.4 Practical Example

In this practical example, the method above is used to deal with 24-h-ahead load forecasting. In order to verify the reliability of the algorithm, we select the vagaries of climate on season transition. Hubei Yichang is just the city where the weather is changeable and rainy. Choose the historical hourly load data (24 points a day, a total of 720 points, load measured in MW) and corresponding climate data from October 1, 2005, to October 30, 2005, of Yichang, Hubei, China. The previous 29 days' data are taken as the training samples into the forecasting model for learning, and then, the data sequence of October 29 is used to forecast the power load data of October 30.

The load sequences adopt Mallat algorithm with layer 3 scale decomposition. Besides, take the method of periodic extension to get biorthogonal compactly supported wavelets db4 (Daubechies function as the mother wavelet) thus obtaining the wavelet coefficients. At the same time, select five climate factors (daily maximum temperature, daily minimum temperature, precipitation, relative humidity, and wind speed) related to the load fluctuation and choose respectively appropriate trapezoidal membership functions or triangular membership functions to obtain the fuzzy inputs. This step can be achieved easily by Matlab programming language. The neural network model is built via Matlab R2010a simulation software for simulation training and prediction.

The training error curve is shown in Fig. 12.3. As we can see from the figure, the network has a high convergence speed, and the number of iteration is just 27. It is a short time before the target error is met. And this is one of the reasons why we choose the RBF neural network. After the simulating, the mean absolute percentage error (MAPE) is used to appraise the performance of the network on the test. From the simulation result we can easily get that the network has a mean absolute error of 1.4775 % whose forecast value is almost identical with the actual value. This forecast accuracy meets the power system operation needs. The specific experiment results are shown in Table 12.1 and Fig. 12.4.

**Fig. 12.3** The error curve of the network



**Table 12.1** Comparison between the forecast load and actual load

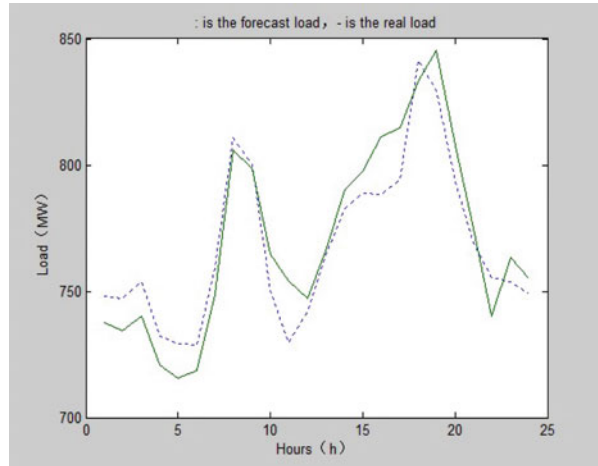
Hours	Actual load	Forecasting load	Error (%)	Hours	Actual load	Forecasting load	Error (%)
1	737	748.1	1.44	13	766.2	764.3	-0.25
2	734.5	747.3	1.74	14	790.3	782.8	-0.95
3	740.1	753.8	1.85	15	797.5	788.6	-1.12
4	720.8	732.4	1.61	16	811.4	788.3	-2.85
5	715.5	729.2	1.91	17	814.9	793.9	-2.58
6	718.5	728.7	1.42	18	833.0	841.6	1.03
7	749.1	759.6	1.40	19	845.6	829.4	-1.92
8	806.0	810.9	0.61	20	807.8	793.1	-1.82
9	798.3	800.3	0.25	21	774.7	769.1	0.72
10	764.5	750.4	-1.84	22	740.1	755.5	2.08
11	747.0	729.5	-3.27	23	763.3	753.8	1.24
12	766.2	741.7	-0.71	24	755.2	748.8	0.85

MAPE = 1.4775

## 12.5 Conclusion

The predictive model established in this chapter for short-term load forecasting combines wavelet transform with fuzzy neural network. The short-term power system load is periodic and nonlinear that the chapter puts forward a model considering different scales of load as well as the impact made by climate. This method highlights the relevance of the power systems and outside environment impact.

**Fig. 12.4** Curves of the forecast and actual load



It proves to be viable and applicable through the practical example. The forecasting precision meets the demand for power system operation. Although the impacts of climate have been considered in this model, the degrees of influences vary from each other. Future work is to deal with the diversity to improve the prediction accuracy of the forecasting.

## References

1. Kang, C., Xia, Q., & Zhang, B. (2004). Review of power system load forecasting and its development. *Automation of Electric Power Systems*, 28(17), 1–11 (in Chinese).
2. Khotanzad, A., Zhou, E., & Elragal, H. (2002). A neuro-fuzzy approach to short-term load forecasting in a price-sensitive environment. *IEEE Transactions on Power System*, 17(4), 1273–1282.
3. Tai, N., & Hou, Z. (2003). New short-term load principle with the wavelet transform fuzzy neural network for the power systems. *Proceedings of the CSEE*, 24(1), 24–29 (in Chinese).
4. Hu, C., Li, G., & Zhou, T. (2008). *Systems analysis and design based on MATLAB7.x* (pp. 25–27). Xi'an: University of Electronic Science and Technology Press (in Chinese).
5. Guiqin, F., & Li, Y. (2008). Influence analysis of meteorological conditions on electric loads. *Meteorological Science and Technology*, 36(6), 795–800 (in Chinese).
6. Soozanchi-K, Z. (2010). Modeling and forecasting short-term electricity load based on adaptive neural-fuzzy inference system by using temperature. *IEEE Signal processing systems*, 18–22.
7. Du, Y., & Lin, L. (2006). Analysis of the effect of composite meteorology index on power load. *Journal of Chongqing University*, 29(12), 56–60 (in Chinese).
8. Han, M., Xu, Z., & Yu, Y. (1994). Electric load-forecasting using RBF neural networks. *Journal of North China Electric Power University*, 21(4), 1–6 (in Chinese).

# Chapter 13

## Energy-Effective Frequency-Based Adaptive Sampling Algorithm for Clustered Wireless Sensor Network

Meiyan Zhang, Wenyu Cai, Liping Zhou, and Jilai Liu

**Abstract** The objective of wireless sensor networks is to extract the synoptic structures (spatiotemporal sequence) of the phenomena of ROI (region of interest) in order to make effective predictive and analytical characterizations. Energy limitation is one of the main obstacles to the universal application of wireless sensor networks. Recently, adaptive sampling strategy is regarded as a much promising method for improving energy efficiency. In this paper, we dedicate to investigating how to regulate sampling frequency of sensor nodes in different clusters dynamically following the change of signal frequency. The adaptive frequency-based sampling (FAS) algorithm proposed in this literature is to measure periodic signal frequency online in different clustered region, afterwards regulate signal sampling frequency following with minimal necessary frequency criterion; as a result, the previous desired level of accuracy is achieved, and the energy consumption is decreased. The simulation results are compared with that of fixed sampling rate approach with respect to energy conservation.

### 13.1 Introduction

Wireless sensor networks have received considerable academia research attention in present years [1]. Wireless sensor networks consist of a large number of tiny sensor nodes deployed over a geographical area; each node is a low-power device that integrates computing, communication, and sensing abilities. The key application of

---

M. Zhang • L. Zhou • J. Liu  
Electric Engineering Department, Zhejiang University of Water Resources  
and Electric Power, Hangzhou 310018, China  
e-mail: [Meiyan19831109@163.com](mailto:Meiyan19831109@163.com)

W. Cai (✉)  
School of Information Engineering, Hangzhou Dianzi University, Hangzhou 310018, China  
e-mail: [dreampp2000@163.com](mailto:dreampp2000@163.com)

wireless sensor networks is monitoring physical phenomena and acquiring environment information. Typically, each sensor node collects raw sensory data to the users through network interconnection for further analysis. The simplest way is to permit each sensor node to deliver its raw sensory data to the base station periodically, where the data can be assembled for subsequent analysis. However, this approach results in excessive communication, and the energy is wasteful. Energy limitation is one of main obstacles to the universal application of wireless sensor networks. In recent years, several energy management schemes have been proposed to reduce power consumption in the literatures, the novel one is adaptive sampling. It is emphasized that reducing the amount of acquired data by using adaptive sampling techniques also reduces the energy consumption for communication. A detailed survey shows that data acquisition and processing have energy consumption, which is significantly lower than that of communication [2].

Generally, data acquisition and processing consume energy that is significantly lower than that of communication. Therefore, traditional researches are concerned with conserving energy as much as possible by reducing transmission capability. Unfortunately, this assumption does not always hold in a number of practical applications. Therefore, this paper proposes a frequency-based adaptive sampling (FAS) algorithm to solve the problem of how to regulate sampling frequency of sensor nodes dynamically following the change of signal frequency. The key idea of this paper is to measure periodic signal frequency online, afterwards adjust signal sampling frequency to the real needs of the physical phenomena under observation. As a result, these two complementary requirements can be achieved: the previous desired level of accuracy is achieved through guaranteeing certain minimal sampling rate for reconstruction, e.g., Nyquist sampling frequency; moreover, the energy consumption is decreased through reducing sampling frequency rate of sensor nodes as much as possible. The FAS algorithm does not assume any hypothesis with regard to the observed signal; therefore, it is of more general applicability. The sensory signal evolution rates of different clustering regions are different for large-scale wireless sensor networks with many clusters; therefore, signal sampling rate ought to adjust subsequently following signal evolution tendency of different clusters.

## 13.2 Related Works

The problem of energy-efficient transmission has been investigated with certain technology such as mathematical optimization in the current papers [3]. However, most researches are concerned with energy-efficient transmission, not energy-efficient sampling. Unfortunately, data acquisition or sampling will consume much more energy than data transmission in a number of practical applications because acquisition times are typically longer than transmission or some sophisticated acquisition process such as multimedia sensor networks. Therefore, some researchers have been investigating energy-efficient sampling scheme. Temporal

correlation is used in an adaptive sampling algorithm for minimizing the energy consumption of a snow sensor [4]. A similar approach has been suggested that the sampling rate is adapted based on the outcome of a Kalman filter [5]. Adaptive sampling is also proposed that a flood alerting system is presented. The system includes a flood predictor that is used to adjust the reporting rate of individual node [6]. Other researchers are discussing how to perform energy-efficient sensory data sampling [7].

Generally, adaptive sampling schemes can be divided in two categories: one is adaptive spatial sampling scheme and the other is adaptive temporal sampling scheme. Adaptive spatial sampling schemes assure monitor accuracy using region location adjustment or wake-up state scheduling. Adaptive temporal sampling schemes assure monitoring accuracy using sampling frequency adjustment or online model estimation of signal tendency. Certainly, there are a few mutational adaptive sampling schemes such as multi-scale adaptive sampling, which provides multi-resolution sensory information as possible as required.

There are only a few researches concerned to temporal sampling schemes for wireless sensor networks. Lygouras presented a velocity-adaptive measurement system for closed-loop position control that relies on the adaptation of the sampling frequency to improve the response time [8]. Aplippi proposed an adaptive sampling algorithm (ASA) that adapted the sampling frequencies of the sensors to the evolving dynamics of the process. The ASA reduced the power consumption of the measurement phase by adapting the real needs of the physical phenomena under observation; however, it is a centralized approach but not adaptable for large-scale wireless sensor networks [9]. In this paper, we dedicate to studying adaptive temporal sampling by regulating sampling rate adaptively and proposing a novel adaptive frequency-based sampling algorithm. To the best of our knowledge, this is the first novel approach for clustered adaptive frequency-based technology. By adaptively sampling ROI of different clusters, the sampling rate decreases energy consumption; moreover, the necessary sensory accuracy is guaranteed.

### 13.3 FAS Algorithm

As we know, following the most famous Nyquist law in signal processing field, if the highest frequency  $F_{\max}$  of signal is given, the minimum sampling frequency  $F_N$  is more than at least twice the highest frequency so that signal reconstruction is guaranteed [10]. In particular, it is common to pick a sampling frequency 3–5 times higher than the signal maximum frequency, i.e.,  $F_N = cF_{\max}$ , where  $c = 3\text{--}5$  generally [11].

First of all, the principle of adaptive frequency sampling algorithm proposed in this paper is explained with Fig. 13.1, where a specific signal with amplitude shifting with time is figured. Fixed sampling rate by Nyquist law is illustrated in Fig. 13.1a, and over-sampling and under-sampling cases are illustrated in

**Fig. 13.1** Over-sampling and under-sampling

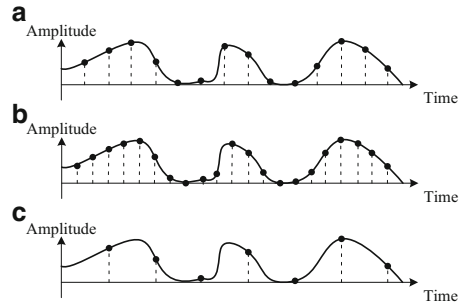


Fig. 13.1b, c. Obviously, over-sampling denotes that the sampling rate is much higher than necessary Nyquist sampling rate, and under-sampling denotes that the sampling rate is low than necessary Nyquist sampling rate. Consequently, design appropriate and necessary sampling rate to satisfy certain accuracy and minimal energy consumption is a promising but difficult mission.

Unfortunately, signal maximum frequency is not available a priori and changes over time in a nonstationary process. Consequently, the sampling frequency has to change adaptively following change tendency. For large-scale wireless sensor networks with many clusters, signal evolution rates of different clustering regions are different; therefore, signal sampling rate ought to adjust subsequently following signal evolution tendency of different clusters. Therefore, our proposed FAS algorithm must be operated on different clustering regions. We will illustrate the details of three procedures that form FAS approach.

### 13.3.1 Clusters Construction

In the first phase, sensor nodes establish different clusters autonomously and elect cluster head (CH) in a fully distributed fashion. Our design proposes a simple distributed clustering algorithm named maximum energy and minimum distance (MEMD) clustering algorithm to establish one-hop clusters, which is described below. First of all, the weight value of sensor node to elect CHs is defined as

$$\omega(i) > \omega(j) \Leftrightarrow E(i) > E(j) \text{ or } E(i) = E(j) \ \&\& \ id(i) > id(j) \quad (13.1)$$

where  $E(i)$  and  $E(j)$  denote the residual energy of sensor node  $i$  and  $j$ , respectively, and  $id(i)$  and  $id(j)$  denote ID number of sensor node  $i$  and  $j$ , respectively. It is obvious that the defined weight is positive correlation with node residual energy and the ID value of sensor node is the second factor. More specially, sensor nodes with more residual energy within all the neighbor nodes should be chosen to be cluster heads with higher probability, thus implementing maximum energy first principle. After cluster heads election, each cluster member sensor node will select

the nearest CH and join that cluster. It is required that the distance between each cluster member and its nearest CH must be smaller than maximal transmission power radius  $R_{\max}$ .

### 13.3.2 Intra-cluster Sampling Frequency Regulation

Within each cluster, sensor nodes are divided into two categories: one particular sensor node is elected as CH, which is responsible for data aggregation, and the other sensor nodes as cluster members (CMs) that are responsible for data acquisition. The most important assumption of this paper is the region of interest (ROI) within each cluster is data correlated and variation identical. Therefore, the sampling frequencies of sensor nodes in the same cluster ought to be same. The distributed nature of FAS algorithm is that different sampling frequency regulation processes are operated in different clusters at the same time; therefore, the FAS algorithm is appropriate even for large-scale sensor networks.

The primary and difficult problem is to estimate the maximal frequency of phenomena signal within certain cluster. Unfortunately, precise estimation of signal maximal frequency is hard, so we use a heuristic method for sampling frequency modification in this paper. In each iteration, CH broadcasts a certain sampling frequency to all CMs which is derived by current sensory accuracy. If the acquired sensory accuracy is lower than the threshold value  $\theta$ , then increases sampling rate  $f_c = f_c + \beta$  ( $\beta = (Th_{\max} - Th_{\min})/M$ ). On the contrary, if the acquired sensory accuracy is larger than the threshold value  $\theta$ , then decreases the sampling rate  $f_c = \alpha \times f_c$  ( $\alpha = 0.5 - 0.8$ ). The sampling frequency regulation scheme obeys multiplicative increase additive decrease (MIAD) rule. The details are illustrated in Fig. 13.2. The parameters used in this algorithm comprise of  $\alpha$ ,  $\beta$ ,  $Th_{\max}$ ,  $Th_{\min}$ ,  $M$ , and  $T$ . The sensory accuracy value  $\theta$  is a prefixed threshold and calculated by model-based correlation mechanism, which will be illustrated in our further paper.

In order to improve transfer reliability in wireless sensor networks, several retransmission manners such as ACK are used. Therefore, the impact of communication unreliability will influence the performance of FAS algorithm; nevertheless, the influence of the communication reliability has not been studied in this paper.

### 13.3.3 Inter-clusters Data Delivery

The final phase is inter-clusters data delivery from CHs to the base station. Although sampling frequencies of different clusters are different, CHs transmit aggregated data to the base station with the same frequency  $T$ . Actually, the delivery ratios of different CHs can be different with different regions, because the signal evolution over time in each region is different. The relay method between sensor nodes is the same with traditional routing schemes.



**Fig. 13.2** Pseudo-code of intra-cluster sampling frequency regulation

```

# Cluster i Sampling Frequency Regulation procedure
begin
Given initial sampling frequency  $f_i$ ;
CH broadcast  $f_i$ ;
while(each Time interval  $T$ )
{
CH receive data and aggregation;
CH calculate sensory accuracy;
if(sensory accuracy  $\geq \theta$ )
 $f_c = \alpha \times f_c$  ( $\alpha = 0.5 - 0.8$ )
else
 $f_c = f_c + \beta$  ( $\beta = \frac{Th_{max} - Th_{min}}{M}$ )
if( $f_c \geq Th_{max}$ )  $f_c = Th_{max}$ ; if( $f_c \leq Th_{min}$ )  $f_c = Th_{min}$ 
CH broadcast  $f_c$ ;
}
end

```

### 13.4 Performance Evaluation

In this section, we describe the evaluation settings and the metrics we have chosen for the evaluation. In order to verify the performance of FAS algorithm, some simulations are carried out with a series of fancied data with mutable and limited signal frequency. The evaluation metrics comprise of sampling frequency  $f_c$  and energy consumption rate comparing with fixed sampling rate. In our simulations, some important parameters of FAS algorithm are defined in Table 13.1.

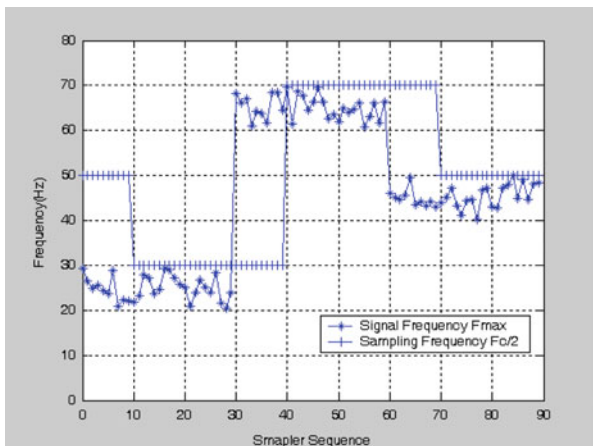
The half value of actual sampling frequency  $f_c/2$  and the signal frequency are described in Fig. 13.3. The original sampling frequency is defined as  $(Th_{max} + Th_{min})/2 = 50$  Hz. It is obvious that the change of sampling frequency drops behind that of signal maximal frequency with windows  $M = 10$  samples from Fig. 13.3. Moreover, the actual sampling frequency approaches the maximal frequency of signal. Therefore, the sampling frequency derived from FAS algorithm is better than fixed sampling frequency.

The energy consumption ratio is defined as the rate of actual sampling frequency with maximal frequency threshold, which is set as 100 Hz in our simulations. The results are illustrated in Fig. 13.4. As we know, if the actual signal frequency increases, the sampling rate ought to increase, thus consuming much more energy; therefore, energy consumption ratio will increase because of the fixed maximal frequency threshold  $Th_{max}$ .

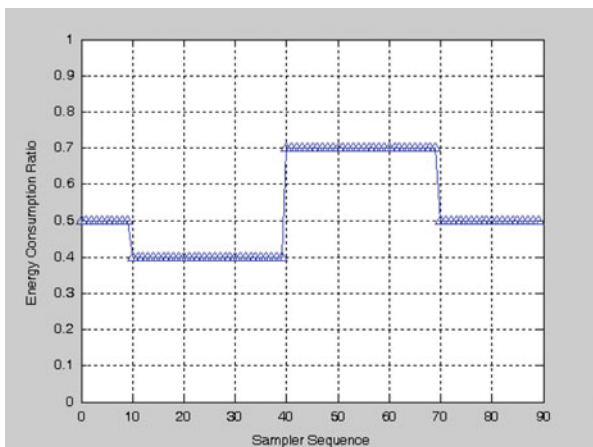
**Table 13.1** FAS parameters for simulations

Parameter	Value
$Th_{max}$	80 Hz
$Th_{min}$	20 Hz
$M$	10
$\alpha$	0.5–0.8
$T$	10 samples

**Fig. 13.3** Sampling frequency following signal frequency evaluation



**Fig. 13.4** Energy consumption ratio



### 13.5 Conclusion

In this paper, we proposed an energy-efficient adaptive sampling algorithm based on adaptive sampling frequency, which regulates signal sampling frequency of related sensor nodes so as to reduce energy consumption intelligently. However, most of the proposed solutions are limited to either temporal or spatial correlation.

Our future work is dedicated to finding more energy-efficient approach using the spatiotemporal correlation of wireless sensor networks. Moreover, in our next simulations, we will use real sensory data, which is derived from Intel Berkeley Research lab.

**Acknowledgements** This research was supported by the National Natural Science Foundation of China (No. 61102067) and the Natural Science Foundation of Zhejiang Province (No. LQ12F03006).

## References

1. Akyildiz, I. F., Weilian, S., Sankarasubramaniam, Y., & Cayirci, E. (2002). A survey on sensor networks. *IEEE Communication Magazine*, 40(8), 102–114.
2. Anastasi, G., Conti, M., Di Francesco, M., & Passarella, A. (2009). Energy conservation in wireless sensor networks. *Ad Hoc Networks*, 7(3), 537–568.
3. Ordonez, F., & Krishnamachari, B. (2004). Optimal information extraction in energy-limited wireless sensor networks. *IEEE Journal on Selected Areas in Communications*, 22(6), 1121–1129. Special Issue Fundamental Performance Limits Wireless Sensor Networks.
4. Alippi, C., Anastasi, G., Galperti, C., Mancini, F., & Roveri, M. (2007). Adaptive sampling for energy conservation in wireless sensor networks for snow monitoring applications. In *Proceedings of the MASS 2007* (pp. 1–6). Pisa, Italy: IEEE.
5. Gedik, B., Lu, L., & Yu, P. S. (2004). ASAP: An adaptive sampling approach to data collection in sensor networks. *IEEE Transactions on Parallel and Distributed Systems*, 18(12), 10–16.
6. Zhou, J., & De Roure, D. (2007). FloodNet: Coupling adaptive sampling with energy aware routing in a flood warning system. *Computer Science Technology*, 22(6), 121–130.
7. Gedik, B., Liu, L., & Yu, P. S. (2007). ASAP: An adaptive sampling approach to data collection in sensor networks. *IEEE Transactions on Parallel Distributed Systems*, 18(12), 1766–1783.
8. Lygouras, J. N., Lalakos, K. A., & Tsalides, P. G. (1998). High-performance position detection and velocity adaptive measurement for closed loop position control. *IEEE Transactions on Instrumentation and Measurement*, 47(4), 978–985.
9. Aplippi, C., Anastasi, G., Francesco, M. D., & Roveri, M. (2010). An adaptive sampling algorithm for effective energy management in wireless sensor networks with energy-hungry sensors networks. *IEEE Transactions on Instrumentation and Measurement*, 59(2), 335–344.
10. Jerri, A. J. (1977). The Shannon sampling theorem-its various extensions and applications: A tutorial review. *Proceedings of the IEEE*, 65(11), 1565–1595.
11. Rauth, D. A., & Randal, V. T. (2005). Analog-to-digital conversion part 5. *IEEE Instrumentation and Measurement Magazine*, 8(4), 44–54.

# Chapter 14

## An Indoor Three-Dimensional Positioning Algorithm Based on Difference Received Signal Strength in WiFi

Yibo Li and Xiting Liu

**Abstract** To further solve the problem that using the positioning algorithm directly based on received signal strength (RSS) in WiFi technology has lower positioning accuracy because the characteristics of the wireless channel affect the signal attenuation largely and randomly, an indoor three-dimensional positioning algorithm based on difference received signal strength (DRSS) is proposed through the analysis that two close propagation paths have similar interference. It can reduce the large positioning error caused by time-varying interference and directly use the fluctuant values of the interfered received signal. Meanwhile, the wireless signal attenuation model with a parameter of time-varying environment factor is used. A method of real-time estimating and modifying the parameters by least square estimation (LS) and the way of average is proposed to solve the problem that the model cannot describe the real-time changes of signal attenuation accurately. The environmental test results show that this method not only can obtain a more accurate model but also has higher positioning accuracy in three-dimensional multi-interference environment.

### 14.1 Introduction

Indoor positioning in WiFi technology has become a hot research with the growing popularity of WiFi access points. It can be used in many fields such as indoor airship positioning and personnel positioning [1]. The walls, pillars, personnel walking, and humidity can influence the signal attenuation [2]. The difficulty of indoor positioning is how to locate accurately based on large fluctuations of received signal strength values.

---

Y. Li (✉) • X. Liu

College of Automation, Shenyang Aerospace University, Shenyang 110136, China  
e-mail: [lyb20040612@aliyun.com](mailto:lyb20040612@aliyun.com)

The positioning based on received signal strength (RSS) in WiFi includes model-based method and model-free method. The model-free method mainly uses fingerprint database and ray-tracing technology [3, 4]. The model-based method is used widely and mostly improved by separating section or adding different compensation factors to describe the signal attenuation more accurately [1, 5]. The estimation of the parameters in the model is very important. Usually, the empirical values and the fixed values which are estimated before positioning are used [6]. In recent years, parameters are dynamically revised by mean value method [7, 8].

To further improve the positioning accuracy, the algorithm based on difference received signal strength (DRSS) is proposed through the analysis that two close propagation paths have similar interference. It can reduce the large positioning error caused by directly using the interfered received signal values. A method of real-timely modifying the parameters by least square estimation (LS) and the way of average is also proposed to make the model describe the real-time changes of signal accurately. The environmental test results show that this method has higher positioning accuracy in three-dimensional multi-interference environment.

## 14.2 Analysis of the Model

The signal attenuation model with an environmental factor is used, as below.

$$P_r(d) = P_r(d_0) - 10n \log_{10} \left( \frac{d}{d_0} \right) + h \quad (14.1)$$

Thereinto,  $P_r(d)$  and  $P_r(d_0)$  express the received signal strength at the distance of  $d$  and  $d_0$  from the transmitter, and they are measured in decibel.  $d_0$  expresses the distance of the fixed receiver to the transmitter, and the value is generally set to 1 m.  $h$  is environment factor which compensates the real-time changes of the environment. The exponent  $n$  is path loss exponent which means the degree of the signal strength attenuating with the propagation distance. The value is two in ideal free space, but it increases in practical environment. Meanwhile, there are extra losses because of the uncertain multi-interference factors such as the temperature, humidity, and personnel walking. So the values of  $n$  and  $h$  are time-varying and influence the positioning accuracy largely.

### 14.3 Proposed Approach

#### 14.3.1 Real-Time Estimation and Correction of Parameters

Multigroup signal strength values received by the fixed reference receivers from the fixed transmitters with known position are used to estimate the values of  $n$  and  $h$  by least square estimation (LS) and the way of average. Then, use the present parameter values in the model to calculate the current position. To describe conveniently,  $a$  access points and  $b$  fixed receivers are set with known positions, and a mobile receiver is used as the object to be located. The access points are divided into  $m$  ( $m \geq 4$ ) groups, and each group has two access points with little distance. Use  $AP_{ij}$  with the coordinate  $(x_{ij}, y_{ij}, z_{ij})$  to express the  $i$ th group and  $j$ th access point ( $i = 1, 2, \dots, m; j = 1, 2$ ).  $FR_\lambda$  with the coordinate  $(x_\lambda, y_\lambda, z_\lambda)$  expresses the  $\lambda$ th ( $\lambda = 1, 2, \dots, b, \lambda \geq m$ ) fixed receiver, and each fixed receiver has a group of access points at a distance of approximately 1 m at least. MR expresses the mobile receiver. The structure of the positioning is shown in Fig. 14.1.

The steps of estimating the model parameters are introduced as follows.

- (1) The nearest group of access points to the  $FR_\lambda$  is found out and set to  $AP_i$  which consists of  $AP_{i1}$  and  $AP_{i2}$ .
- (2) According to the above model, the signal strength value received from  $AP_{i1}$  by  $FR_\lambda$  is used as  $P_r(d_0)$ , and the distance between  $FR_\lambda$  and  $AP_{i1}$  is used as  $d_0$ . The values received from other  $2(m - 1)$  groups of access points by  $FR_\lambda$  are used as  $P_r(d)$ , and the corresponding distance is used as  $d$ . Then,  $C_{2(m-1)}^2$  values of  $n$  and  $h$  can be obtained through the model by least square estimation, and they can be expressed as  $\phi_{\lambda 1} = \{n_1 \ n_2 \ \dots \ n_{C_{2(m-1)}^2}\}$  and  $\psi_{\lambda 1} = \{h_1 \ h_2 \ \dots \ h_{C_{2(m-1)}^2}\}$ . In the same way, other  $C_{2(m-1)}^2$  values of

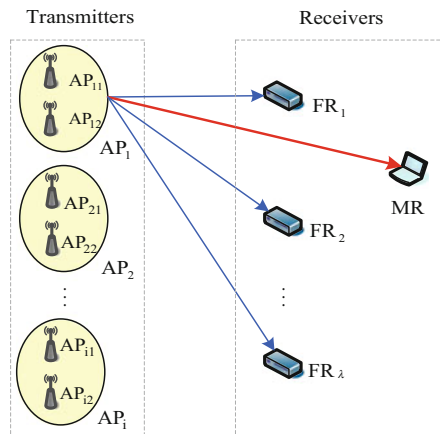


Fig. 14.1 The structure of positioning

$n$  and  $h$  can be obtained by using the values received from  $AP_{i2}$  by  $FR_\lambda$  as  $P_r(d_0)$ . They can be expressed as  $\phi_{\lambda 2} = \left\{ n_1 \quad n_2 \quad \dots \quad n_{C_{2(m-1)}^2} \right\}$  and  $\psi_{\lambda 2} = \left\{ h_1 \quad h_2 \quad \dots \quad h_{C_{2(m-1)}^2} \right\}$ . Then, the averages of  $n$  and  $h$  can be calculated based on only one fixed receiver  $FR_\lambda$  by Eq. (14.2).

$$n_\lambda = \frac{\sum \phi_{\lambda 1} + \sum \phi_{\lambda 2}}{2C_{2(m-1)}^2}, \quad h_\lambda = \frac{\sum \psi_{\lambda 1} + \sum \psi_{\lambda 2}}{2C_{2(m-1)}^2} \quad (14.2)$$

(3) Because  $b$  fixed receivers are used, the current values of the path loss exponent and the environment factor can be estimated by Eq. (14.3).

$$n = \frac{\sum_{\lambda=1}^b n_\lambda}{b}, \quad \zeta = \frac{\sum_{\lambda=1}^b \zeta_\lambda}{b} \quad (14.3)$$

Then, the model with time-varying environment factor and path loss exponent can be obtained by the substitution of the estimated values in real time.

### 14.3.2 The Positioning Algorithm Based on Difference Received Signal Strength

Directly using the signal strength value can cause large error because the value includes theoretical value and interference value which is difficult to estimate, as below.

$$P_r(d) = P_{rth}(d) + P_{rin}(d) \quad (14.4)$$

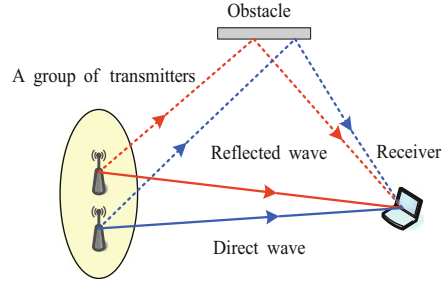
Thereinto,  $P_{rth}(d)$  expresses the theoretical value, and  $P_{rin}(d)$  expresses the interference value. In general indoor positioning, the wireless signal transmitted by a transmitter will arrive at a receiver through some paths. So the wireless signal transmitted by two very close transmitters will arrive at the same receiver through nearly the same paths, and they will have almost the same interferences and the interference values, as shown in Fig. 14.2.

Based on this idea, the difference received signal strength is presented to reduce the time-varying and random influences, as below.

$$P_r(d_{i1}) - P_r(d_{i2}) = [P_{rth}(d_{i1}) - P_{rth}(d_{i2})] + [(P_{rin}(d_{i1}) - P_{rin}(d_{i2}))] \quad (14.5)$$

Thereinto,  $P_r(d_{i1})$  and  $P_r(d_{i2})$  express the signal strength values received by mobile receiver from two transmitters  $AP_{i1}$  and  $AP_{i2}$ .  $d_{i1}$  and  $d_{i2}$  express the

**Fig. 14.2** The signal propagation paths of two close transmitters



distance between mobile receiver and  $AP_{i1}$ ,  $AP_{i2}$ . So, combining with the model, Eq. (14.6) can be obtained.

$$\frac{d_{i2}}{d_{i1}} = \frac{1}{\eta_i} \cdot 10^{\frac{RSSI_{e_{i12}}}{10n}} \tag{14.6}$$

Thereinto,  $\eta_i = \frac{d_{i01}}{d_{i02}}$ .  $d_{i01}$  and  $d_{i02}$  separately express the distance between the nearest fixed receiver and  $AP_{i1}$  and  $AP_{i2}$ .  $RSSI_{e_{i12}} = RSSI_{i12} - RSSI_{i012} - \zeta_{i12}$ ,  $RSSI_{i12} = P_r(d_{i1}) - P_r(d_{i2})$ ,  $RSSI_{i012} = P_r(d_{i01}) - P_r(d_{i02})$ ,  $\zeta_{i12} = \zeta_{i1} - \zeta_{i2}$ .  $P_r(d_{i01})$  and  $P_r(d_{i02})$  express the signal strength values received by a fixed receiver from  $AP_{i1}$  and  $AP_{i2}$ .  $\zeta_{i1}$  and  $\zeta_{i2}$  express the environment factors of the signal received from  $AP_{i1}$  and  $AP_{i2}$ .

According to the relation between distance and coordinates shown in Eq. (14.7),

$$d_{ij} = \sqrt{(x_{ij} - x)^2 + (y_{ij} - y)^2 + (z_{ij} - z)^2} \tag{14.7}$$

Equation (14.8) is obtained through a further derivation of Eq. (14.6).

$$\begin{aligned} R^2 + \frac{2(x_{i2} - \gamma_i^2 x_{i1})}{(\gamma_i^2 - 1)} x + \frac{2(y_{i2} - \gamma_i^2 y_{i1})}{(\gamma_i^2 - 1)} y + \frac{2(z_{i2} - \gamma_i^2 z_{i1})}{(\gamma_i^2 - 1)} z \\ = \frac{x_{i2}^2 + y_{i2}^2 + z_{i2}^2 - \gamma_i^2(x_{i1}^2 + y_{i1}^2 + z_{i1}^2)}{(\gamma_i^2 - 1)} \end{aligned} \tag{14.8}$$

Thereinto,  $(x,y,z)$  expresses the coordinate of mobile receiver.  $R^2 = x^2 + y^2 + z^2$ .  $\gamma_i = \frac{d_{i2}}{d_{i1}}$ .

Because this algorithm needs four groups of access points at least, for convenience, four groups are set and substituted in Eq. (14.8), as below.



$$\begin{cases} R^2 + \frac{2(x_{12} - \gamma_1^2 x_{11})}{(\gamma_1^2 - 1)}x + \frac{2(y_{12} - \gamma_1^2 y_{11})}{(\gamma_1^2 - 1)}y + \frac{2(z_{12} - \gamma_1^2 z_{11})}{(\gamma_1^2 - 1)}z = \frac{x_{12}^2 + y_{12}^2 + z_{12}^2 - \gamma_1^2(x_{11}^2 + y_{11}^2 + z_{11}^2)}{(\gamma_1^2 - 1)} \\ R^2 + \frac{2(x_{22} - \gamma_2^2 x_{21})}{(\gamma_2^2 - 1)}x + \frac{2(y_{22} - \gamma_2^2 y_{21})}{(\gamma_2^2 - 1)}y + \frac{2(z_{22} - \gamma_2^2 z_{21})}{(\gamma_2^2 - 1)}z = \frac{x_{22}^2 + y_{22}^2 + z_{22}^2 - \gamma_2^2(x_{21}^2 + y_{21}^2 + z_{21}^2)}{(\gamma_2^2 - 1)} \\ R^2 + \frac{2(x_{32} - \gamma_3^2 x_{31})}{(\gamma_3^2 - 1)}x + \frac{2(y_{32} - \gamma_3^2 y_{31})}{(\gamma_3^2 - 1)}y + \frac{2(z_{32} - \gamma_3^2 z_{31})}{(\gamma_3^2 - 1)}z = \frac{x_{32}^2 + y_{32}^2 + z_{32}^2 - \gamma_3^2(x_{31}^2 + y_{31}^2 + z_{31}^2)}{(\gamma_3^2 - 1)} \\ R^2 + \frac{2(x_{42} - \gamma_4^2 x_{41})}{(\gamma_4^2 - 1)}x + \frac{2(y_{42} - \gamma_4^2 y_{41})}{(\gamma_4^2 - 1)}y + \frac{2(z_{42} - \gamma_4^2 z_{41})}{(\gamma_4^2 - 1)}z = \frac{x_{42}^2 + y_{42}^2 + z_{42}^2 - \gamma_4^2(x_{41}^2 + y_{41}^2 + z_{41}^2)}{(\gamma_4^2 - 1)} \end{cases} \quad (14.9)$$

Using the first formula to subtract other ones separately can get another equation including three formulas, and it can be expressed in the form of matrix equation, as below.

$$A\theta = \beta \quad (14.10)$$

Thereinto,

$$A = \begin{bmatrix} \frac{(x_{12} - \gamma_1^2 x_{11})}{(\gamma_1^2 - 1)} - \frac{(x_{22} - \gamma_2^2 x_{21})}{(\gamma_2^2 - 1)} & \frac{(y_{12} - \gamma_1^2 y_{11})}{(\gamma_1^2 - 1)} - \frac{(y_{22} - \gamma_2^2 y_{21})}{(\gamma_2^2 - 1)} & \frac{(z_{12} - \gamma_1^2 z_{11})}{(\gamma_1^2 - 1)} - \frac{(z_{22} - \gamma_2^2 z_{21})}{(\gamma_2^2 - 1)} \\ \frac{(x_{12} - \gamma_1^2 x_{11})}{(\gamma_1^2 - 1)} - \frac{(x_{32} - \gamma_3^2 x_{31})}{(\gamma_3^2 - 1)} & \frac{(y_{12} - \gamma_1^2 y_{11})}{(\gamma_1^2 - 1)} - \frac{(y_{32} - \gamma_3^2 y_{31})}{(\gamma_3^2 - 1)} & \frac{(z_{12} - \gamma_1^2 z_{11})}{(\gamma_1^2 - 1)} - \frac{(z_{32} - \gamma_3^2 z_{31})}{(\gamma_3^2 - 1)} \\ \frac{(x_{12} - \gamma_1^2 x_{11})}{(\gamma_1^2 - 1)} - \frac{(x_{42} - \gamma_4^2 x_{41})}{(\gamma_4^2 - 1)} & \frac{(y_{12} - \gamma_1^2 y_{11})}{(\gamma_1^2 - 1)} - \frac{(y_{42} - \gamma_4^2 y_{41})}{(\gamma_4^2 - 1)} & \frac{(z_{12} - \gamma_1^2 z_{11})}{(\gamma_1^2 - 1)} - \frac{(z_{42} - \gamma_4^2 z_{41})}{(\gamma_4^2 - 1)} \end{bmatrix}$$

$$\beta = \begin{bmatrix} \frac{x_{12}^2 + y_{12}^2 + z_{12}^2 - \gamma_1^2(x_{11}^2 + y_{11}^2 + z_{11}^2)}{2(\gamma_1^2 - 1)} - \frac{x_{22}^2 + y_{22}^2 + z_{22}^2 - \gamma_2^2(x_{21}^2 + y_{21}^2 + z_{21}^2)}{2(\gamma_2^2 - 1)} \\ \frac{x_{12}^2 + y_{12}^2 + z_{12}^2 - \gamma_1^2(x_{11}^2 + y_{11}^2 + z_{11}^2)}{2(\gamma_1^2 - 1)} - \frac{x_{32}^2 + y_{32}^2 + z_{32}^2 - \gamma_3^2(x_{31}^2 + y_{31}^2 + z_{31}^2)}{2(\gamma_3^2 - 1)} \\ \frac{x_{12}^2 + y_{12}^2 + z_{12}^2 - \gamma_1^2(x_{11}^2 + y_{11}^2 + z_{11}^2)}{2(\gamma_1^2 - 1)} - \frac{x_{42}^2 + y_{42}^2 + z_{42}^2 - \gamma_4^2(x_{41}^2 + y_{41}^2 + z_{41}^2)}{2(\gamma_4^2 - 1)} \end{bmatrix},$$

$$\theta = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

Then, the value of the real-time coordinate of the mobile receiver can be estimated by least squares estimation, as below.

$$\widehat{\theta}_{LS} = (A^T A)^{-1} A^T \beta \quad (14.11)$$

### 14.4 Experimental Test and Validation

To evaluate the performance of the proposed method, the test environment which is about  $28\text{ m} \times 19\text{ m} \times 8\text{ m}$  is set up at first floor lobby of laboratory building, as shown in Fig. 14.3. Four groups of access points are placed near the middle of four walls. The distance between the two access points in one group is set to  $0.2\text{ m}$ . Computers are used as the mobile receiver and the fixed receivers. The fixed receivers have to send the real-time received signal to the mobile receiver to modify the parameters and evaluate the coordinates. Twenty coordinates are selected to analyze the results. Figure 14.4 shows

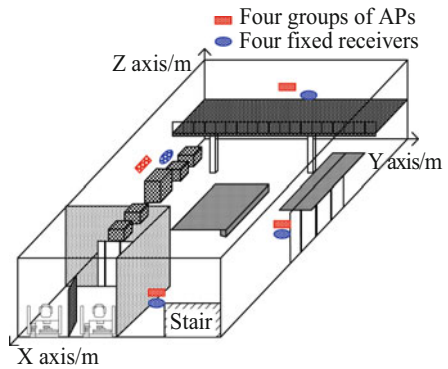


Fig. 14.3 The schematic plot of test environment

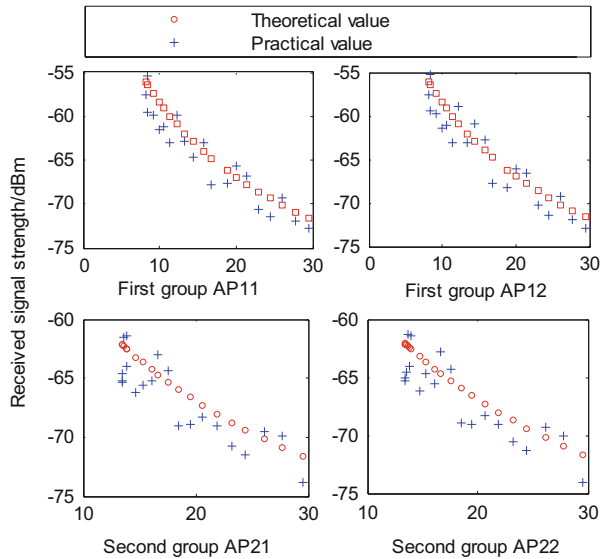
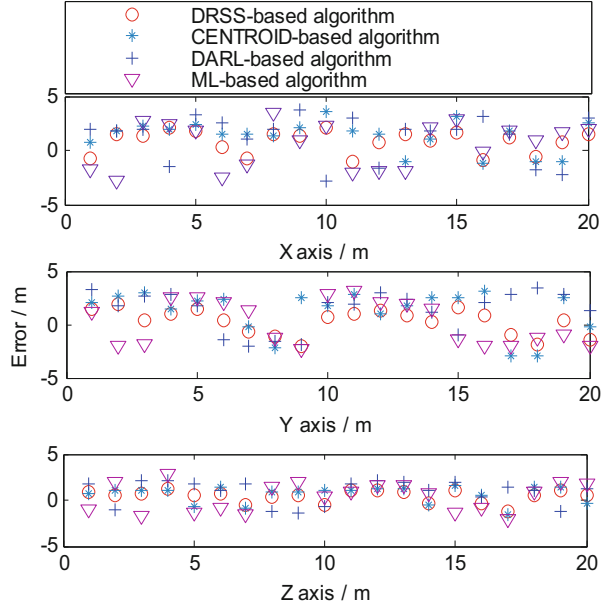


Fig. 14.4 The comparison of practical received signal strength values with theoretical values

**Fig. 14.5** The positioning error on X, Y, Z coordinate axis

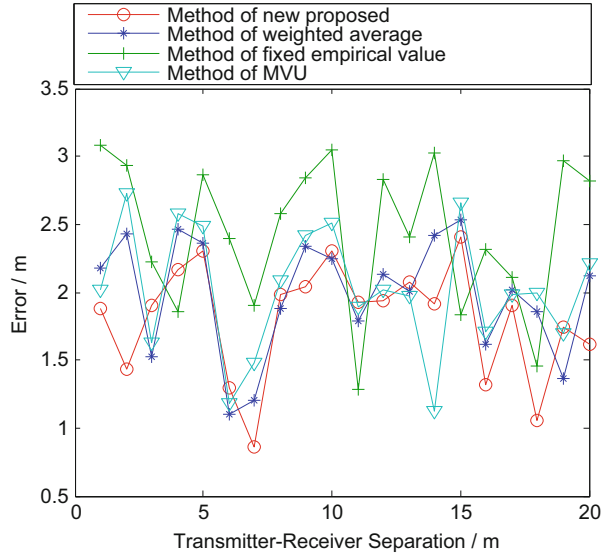


the comparison of the theoretical signal strength values with the practical ones which are received from two groups of access points. It can be seen that the practical values fluctuate largely and irregularly around theoretical ones and this will result in large positioning error. But the values received from one group of access points have almost the same fluctuations.

To know the positioning performance in the experiment, the proposed algorithm is compared with other three received signal strength-based algorithms including dynamic correction and centroid-based algorithm, DARL (dominant access point RSS location)-based algorithm, and ML (maximum likelihood)-based algorithm [6, 8, 9]. The positioning errors which are analyzed on X, Y, Z coordinate axis are shown in Fig. 14.5. We can see that most errors of proposed algorithm are less than 2.6 m and smaller than that of other algorithms. At present, the positioning error is generally about 3–4 m in similar research (based on WiFi technology), and the better ones can control the most errors within 3 m.

Another experiment is made to know the improvement degree of using new proposed method of estimating model parameters. Other three methods, weighted average, MVU (minimum variance unbiased), and fixed empirical value [1, 7, 8], are also used to compare with it, and the comparison results are shown in Fig. 14.6. The positioning accuracy is obviously improved when the real-time values are used and the new proposed method has better effect than other three ones.

**Fig. 14.6** The positioning error with different estimation methods of parameter



## 14.5 Conclusion

The algorithm of indoor three-dimensional positioning based on difference received signal strength (DRSS) is presented in this paper, and a new method of the real-time estimation and correction of parameters of the model with an environment factor is also used at the same time. As we have seen from the environmental test, the effect of the positioning based on this method is more sound. The multipath propagation, multi-interference factors, nonlinearity, and time variation of the wireless channel bring great influence on the positioning. So this method can use the difference received signal strength to weaken it to some degree, and the indoor propagation model of wireless signal can describe the real-time changes of the attenuation more accurately through the correction. Further research will be made to improve the performance of indoor positioning based on WiFi access points.

## References

1. Shi, W. R., Xiong, Z. G., & Xu, L. (2010). In-building RSSI-based user localization algorithm. *Computer Engineering and Applications*, 46(17), 232–235 (In Chinese).
2. Kaemarungsi, K., & Krishnamurthy, P. (2012). Analysis of WLAN's received signal strength indication for indoor location fingerprinting. *Pervasive and Mobile Computing*, 8(2), 292–316.
3. Zaruba, G. V., Huber, M., Kamangar, F. A., & Chlamtac, I. (2007). Indoor location tracking using RSSI readings from a single Wi-Fi access point. *Wireless Networks*, 13(2), 221–235.

4. Swangmuang, N., & Krishnamurthy, P. (2008). An effective location fingerprint model for wireless indoor localization. *Pervasive and Mobile Computing*, 4(6), 836–850.
5. Narzullaev, A., Park, Y., Yoo, K., & Yu, J. (2011). A fast and accurate calibration algorithm for real-time locating systems based on the received signal strength indication. *AEU-International Journal of Electronics and Communications*, 65(4), 305–311.
6. Park, D. W., & Park, J. G. (2011). An enhanced ranging scheme using WiFi RSSI measurements for ubiquitous location. *IEEE CNSI*, 296–301. doi:[10.1109/CNSI.2011.29](https://doi.org/10.1109/CNSI.2011.29)
7. Xu, F. Y., Shan, H. G., & Wang, Z. X. (2008). An indoor location algorithm based on received signal strength with parameter estimation. *Journal of Microwaves*, 24(2), 67–72 (In Chinese).
8. Xu, R. M., Zhuang, C. Y., & Yu, B. (2010). Dynamic correction algorithm for indoor wireless location based on RSSI. *Computer Knowledge and Technology*, 6(3), 686–688 (In Chinese).
9. Ni, W., & Wang, Z. X. (2006). Indoor location algorithm based on the measurement of the received signal strength. *Frontiers of Electrical and Electronic Engineering in China*, 1(1), 48–52.

# Chapter 15

## The Universal Approximation Capability of Double Flexible Approximate Identity Neural Networks

Saeed Panahian Fard and Zarita Zainuddin

**Abstract** This study investigates the universal approximation capability of three-layer feedforward double flexible approximate identity neural networks in the space of continuous functions with two variables. First, we propose double flexible approximate identity functions, which are a combination of double approximate identity functions and flexible approximate identity functions as investigated in our previous studies. Then, we prove that any continuous function  $f$  with two variables will converge to itself if it convolves with double flexible approximate identity. Finally, we prove a main theorem by using the obtained results.

### 15.1 Introduction

Universal approximation capability is the first question in the approximation theory of neural networks. That is, neural networks can approximate any continuous function in the topological space w.r.t. to a proper norm to any accuracy. Thus, the most fundamental theoretical studies for neural networks are related to the universal approximation capability of a new class of neural networks.

In this study, the main results of the universal approximation capability of artificial neural networks are surveyed in the space of continuous functions. It has been shown that any continuous function can be approximated on a compact set with uniform topology by sigmoidal neural networks [1, 2]. It has been proved that Gaussian radial basis function neural networks are universal approximators [3]. Moreover, it has been obtained that radial basis function neural networks can uniformly approximate any continuous function on a compact set provided that the activation function is continuous almost everywhere, locally

---

S.P. Fard (✉) • Z. Zainuddin  
School of Mathematical Sciences, Universiti Sains Malaysia,  
11800 USM Pulau Pinang, Malaysia  
e-mail: [saeedpanahian@yahoo.com](mailto:saeedpanahian@yahoo.com); [zarita@cs.usm.my](mailto:zarita@cs.usm.my); <http://math.usm.my/>

essentially bounded and not a polynomial [4]. Furthermore, it has been surveyed the universal approximation capability of feedforward neural networks [5].

On the other hand, we have introduced double approximate identity neural networks [6]. Double approximate identity neural networks are the generalization of Plane-Gaussian neural networks which were introduced in [7]. Then, we have proved that flexible approximate identity neural networks are universal approximators [8, 9]. Flexible approximate identity neural networks are the generalization of generalized Gaussian RBF neural networks which were presented in [10]. In addition, in our latest result, it has been shown that Mellin approximate identity neural networks are universal approximators [11].

The motivation of this study is to combine double approximate identity functions and flexible approximate identity functions in order to obtain an extension class of approximate identity. This extension is called double flexible approximate identity functions. Thus, a new class of neural networks is obtained by using double flexible approximate identity as activation functions.

The approach of this study is illustrated below. First, a theorem is proved by using the fundamental concept of the double flexible approximate identity functions. The proof of this result is based on the double convolution linear operator in the space of the continuous functions with two variables. Then, by using this result, a main theorem is obtained. The proof of the main theorem is in the framework of the theory of  $\varepsilon$ -net. This theorem shows the universal approximation capability of a three-layer feedforward double approximate identity artificial neural networks in the space of the continuous functions with two variables.

The remainder of this study is organized as follows: in Sect. 15.2, some basic definitions are studied which will be used in the following sections. In Sect. 15.3, a theoretical result is presented which will be constructed the fundamental structure of the next section. In Sect. 15.4, a main theoretical result is given. Finally, in Sect. 15.5, conclusions are derived.

## 15.2 Basic Definitions

In this section, a certain number of definitions are reviewed which will be used in the succeeding sections. As the fundamental concept, the definition of the double flexible approximate identity functions is presented.

**Definition 1** Let  $A = A(a_1, \dots, a_m)$ ,  $a_i \in \mathbb{R}$ ,  $i = 1, \dots, m$  be any parameters,  $\{\phi_n(x, y, A)\}_{n=1}^\infty$ ,  $\phi_n : \mathbb{R}^2 \rightarrow \mathbb{R}$ , be a sequence of two-dimensional flexible functions. The sequence is called a double flexible approximate identity if it satisfies the following conditions:

- 1)  $\int_{\mathbb{R}} \int_{\mathbb{R}} \phi_n(x, y, A) dx dy = 1$ ;
- 2) for any  $\varepsilon > 0$  and  $\delta > 0$ , there exists a number  $N$  such that if  $n \geq N$  it results:

- (i)  $\int_{|y| < \delta} \int_{|x| \geq \delta} |\phi_n(x, y, A)| dx dy \leq \epsilon,$
- (ii)  $\int_{|y| \geq \delta} \int_{|x| < \delta} |\phi_n(x, y, A)| dx dy \leq \epsilon,$
- (iii)  $\int_{|y| \geq \delta} \int_{|x| \geq \delta} |\phi_n(x, y, A)| dx dy \leq \epsilon.$

The above definition will be used in Theorems 2 and 4. In the following definition, double convolution is reviewed which will be used in Theorem 2.

**Definition 2 ([12])** A double convolution of two integrable functions  $f$  and  $g$  is defined as

$$f(x, y) *^x *^y g(x, y) = \int_{\mathbb{R}} \int_{\mathbb{R}} f(u, v) g(x - u, y - v) du dv.$$

Now, the definitions of  $\epsilon$ -net and finite  $\epsilon$ -net will be studied, respectively. These definitions will be used in Theorems 3 and 4.

**Definition 3 ([13])** Let  $\epsilon > 0$ . A set  $V_\epsilon \subset C(I, J)$  is called  $\epsilon$ -net of a set  $V$ , if  $\tilde{f} \in V_\epsilon$  can be found for  $\forall f \in V$  such that  $\|f - \tilde{f}\|_{C(I, J)} < \epsilon$ .

**Definition 4 ([13])** The  $\epsilon$ -net is said to be finite if it is a finite set of elements.

In the next section, a theoretical result in the space of continuous functions with two variables will be given.

### 15.3 Theoretical Result for Continuous Functions with Two Variables

In this section, the following theorem is studied which will be used in the proof of Theorem 2.

**Theorem 1 ([14])** Let  $I, J \subseteq \mathbb{R}$  be closed and bounded intervals and let  $f : I \times J \rightarrow \mathbb{R}$  be a continuous function with two variables. Then  $f$  is uniformly continuous.

Now, Theorem 2 is presented. This theorem shows a very important property of double approximate identity functions in the space of continuous functions with two variables.

**Theorem 2** Let  $A = A(a_1, \dots, a_m), a_i \in \mathbb{R}, i = 1, \dots, m$  be any parameters,  $\{\phi_n(x, y, A)\}_{n \in \mathbb{N}}, \phi_n(x, y, A) : \mathbb{R}^2 \rightarrow \mathbb{R}$  be a double approximate identity. Let  $f$  be a function on  $C(I, J)$ . Then  $\phi_n *^x *^y f$  uniformly converges to  $f$  on  $C(I, J)$ .

*Proof.* Let  $x \in I, y \in J$  and  $\epsilon > 0$ . Based on Theorem 1, there exists a  $\delta > 0$  such that  $|f(x, y) - f(u, v)| < \frac{\epsilon}{2\|\phi\|_{L^1(\mathbb{R}^2)}}$  for all  $u, v, |x - u| < \delta, |y - v| < \delta$ . Let us define  $\{\phi_n *^x *^y f\}_{n \in \mathbb{N}}$  by  $\phi_n(x, y, A) = n^2 \phi(n x, n y, A)$ . We consider



$$\begin{aligned}
& \phi_n^{*x *y} f(x, y) - f(x, y) \\
&= \int_{\mathbb{R}} \int_{\mathbb{R}} \phi_n(u, v, A) \{f(x - u, y - v) - f(x, y)\} dudv \\
&= \left( \int_{|y| < \delta} \int_{|x| < \delta} + \int_{|y| < \delta} \int_{|x| \geq \delta} + \int_{|y| \geq \delta} \int_{|x| < \delta} + \int_{|y| \geq \delta} \int_{|x| \geq \delta} \right) n^2 \cdot \\
&\quad \phi(nu, nv, A) \{f(x - u, y - v) - f(x, y)\} dudv \\
&= I_1 + I_2 + I_3 + I_4.
\end{aligned}$$

We calculate  $I_1$ ,  $I_2$ ,  $I_3$ , and  $I_4$  as follows:

$$\begin{aligned}
|I_1| &\leq \int_{|y| < \delta} \int_{|x| < \delta} n^2 |\phi(nu, nv, A) \{f(x - u, y - v) - f(x, y)\}| dudv \\
&< \frac{\epsilon}{4 \|\phi\|_{L^1(\mathbb{R}^2)}} \int_{|y| < \delta} \int_{|x| < \delta} n^2 |\phi(nu, nv, A)| dudv \\
&= \frac{\epsilon}{4 \|\phi\|_{L^1(\mathbb{R}^2)}} \int_{|t| < n\delta} \int_{|s| < n\delta} |\phi(s, t, A)| dsdt \\
&\leq \frac{\epsilon}{4 \|\phi\|_{L^1(\mathbb{R}^2)}} \int_{\mathbb{R}} \int_{\mathbb{R}} |\phi(s, t, A)| dsdt = \frac{\epsilon}{4}.
\end{aligned}$$

For  $I_2$ , we have

$$\begin{aligned}
|I_2| &\leq 2 \|f\|_{C(I, J)} \int_{|y| < \delta} \int_{|x| \geq \delta} n^2 |\phi(nu, nv, A)| dudv \\
&= 2 \|f\|_{C(I, J)} \int_{|t| < n\delta} \int_{|s| \geq n\delta} |\phi(s, t, A)| dsdt.
\end{aligned}$$

Since

$$\lim_{n \rightarrow \infty} \int_{|t| < n\delta} \int_{|s| \geq n\delta} |\phi(s, t, A)| dsdt = 0,$$

there exist  $n_0 \in \mathbb{N}$  such that for all  $n \geq n_0$ ,

$$\int_{|t| < n\delta} \int_{|s| \geq n\delta} |\phi(s, t, A)| dsdt < \frac{\epsilon}{8 \|f\|_{C(I, J)}}.$$

For  $I_3$ , we have

$$\begin{aligned}
|I_3| &\leq 2 \|f\|_{C(I, J)} \int_{|y| \geq \delta} \int_{|x| < \delta} n^2 |\phi(nu, nv, A)| dudv \\
&= 2 \|f\|_{C(I, J)} \int_{|t| \geq n\delta} \int_{|s| < n\delta} |\phi(s, t, A)| dsdt.
\end{aligned}$$

Since

$$\lim_{n \rightarrow \infty} \int_{|t| \geq n\delta} \int_{|s| \geq \delta} |\phi(s, t, A)| ds dt = 0,$$

there exist  $n_0 \in \mathbb{N}$  such that for all  $n \geq n_0$ ,

$$\int_{|t| \geq n\delta} \int_{|s| < n\delta} |\phi(s, t, A)| ds dt < \frac{\epsilon}{8\|f\|_{C(I, J)}}.$$

For  $I_4$ , we have

$$\begin{aligned} |I_4| &\leq 2\|f\|_{C(I, J)} \int_{|y| \geq \delta} \int_{|x| \geq \delta} n^2 |\phi(nu, nv, A)| dudv \\ &= 2\|f\|_{C(I, J)} \int_{|t| \geq n\delta} \int_{|s| \geq n\delta} |\phi(s, t, A)| ds dt. \end{aligned}$$

Since

$$\lim_{n \rightarrow \infty} \int_{|t| \geq n\delta} \int_{|s| \geq n\delta} |\phi(s, t, A)| ds dt = 0,$$

there exists an  $n_0 \in \mathbb{N}$  such that for all  $n \geq n_0$ ,

$$\int_{|t| \geq n\delta} \int_{|s| \geq n\delta} |\phi(s, t, A)| ds dt < \frac{\epsilon}{8\|f\|_{C(I, J)}}.$$

Combining  $I_1, I_2, I_3$ , and  $I_4$  for  $n \geq n_0$ , we have

$$\|\phi_n *^x *^y f(x, y) - f(x, y)\|_{C(I, J)} < \epsilon.$$

The above theorem will be used in the next section. The next section shows a main result for continuous functions with two variables

## 15.4 Main Result for Continuous Functions with Two Variables

In this section, Theorem 3 is reviewed which will be used in the proof of the Theorem 4.

**Theorem 3 ([13])** *A set  $V$  in  $C(I, J)$  is compact iff  $\forall \epsilon > 0$  in  $\mathbb{R}$  there exists a finite  $\epsilon$ -net.*

Now, Theorem 4 is presented. This main result shows the universal approximation capability of double flexible approximate identity neural networks in the space of continuous functions with two variables.

**Theorem 4** *Let  $I, J \subseteq \mathbb{R}$  be closed and bounded intervals. Let  $C(I, J)$  be the real linear space of all continuous functions on any compact subset of real two-dimensional space, and  $V \subset C(I, J)$  a compact set. Let  $A = A(a_1, \dots, a_m)$ ,  $a_i \in \mathbb{R}, i = 1, \dots, m$  be any parameters,  $\{\phi_n(x, y, A)\}_{n \in \mathbb{N}}, \phi_n : \mathbb{R}^2 \rightarrow \mathbb{R}$  be a double flexible approximate identity. Let the family of functions  $\{\sum_{j=1}^M \lambda_j \phi_j(x, y) | \lambda_j \in \mathbb{R}, x \in \mathbb{R}, y \in \mathbb{R}, M \in \mathbb{N}\}$ , be dense in  $C(I, J)$ , and given  $\varepsilon > 0$ . Then there exists  $N \in \mathbb{N}$  which depends on  $V$  and  $\varepsilon$  but not on  $f$ , such that for any  $f \in V$ , there exist weights  $c_k = c_k(f, V, \varepsilon)$  satisfying*

$$\left\| f(x, y) - \sum_{k=1}^N c_k \phi_k(x, y, A) \right\|_{C(I, J)} < \varepsilon.$$

Moreover, every  $c_k$  is a continuous function of  $f \in V$ .

*Proof* Since  $V$  is compact, for any  $\varepsilon > 0$ , there is a finite  $\frac{\varepsilon}{2}$ -net  $\{f^1, \dots, f^M\}$  for  $V$ . This implies that for any  $f \in V$ , there is an  $f^j$  such that  $\|f - f^j\|_{C(I, J)} < \frac{\varepsilon}{2}$ . For any  $f^j$ , by assumption of the theorem, there are  $\lambda_i^j \in \mathbb{R}, N_j \in \mathbb{N}$ , and  $\phi_i^j(x, y, A)$  such that

$$\left\| f^j(x) - \sum_{i=1}^{N_j} \lambda_i^j \phi_i^j(x, y, A) \right\|_{C(I, J)} < \frac{\varepsilon}{2}. \quad (15.1)$$

For any  $f \in V$ , we define

$$\begin{aligned} F_-(f) &= \left\{ j \mid \|f - f^j\|_{C(I, J)} < \frac{\varepsilon}{2} \right\}, \\ F_0(f) &= \left\{ j \mid \|f - f^j\|_{C(I, J)} = \frac{\varepsilon}{2} \right\}, \\ F_+(f) &= \left\{ j \mid \|f - f^j\|_{C(I, J)} > \frac{\varepsilon}{2} \right\}. \end{aligned}$$

Therefore,  $F_-(f)$  is not empty according to the definition of  $\frac{\varepsilon}{2}$ -net. If  $\tilde{f} \in V$  approaches  $f$  such that  $\|\tilde{f} - f\|_{C(I, J)}$  is small enough, then we have  $F_-(f) \subset F_-(\tilde{f})$  and  $F_+(f) \subset F_+(\tilde{f})$ . Thus  $F_-(\tilde{f}) \cap F_+(f) \subset F_-(\tilde{f}) \cap F_+(\tilde{f}) = \emptyset$ , which implies  $F_-(\tilde{f}) \subset F_-(f) \cup F_0(f)$ . We conclude the following:

$$F_-(f) \subset F_-(\tilde{f}) \subset F_-(f) \cup F_0(f). \quad (15.2)$$

Define

$$d(f) = \left[ \sum_{j \in F_-(f)} \left( \frac{\epsilon}{2} - \|f - f^j\|_{C(I,J)} \right) \right]^{-1}$$

and

$$f_h = \sum_{j \in F_-(f)} \sum_{i=1}^{N_j} d(f) \left( \frac{\epsilon}{2} - \|f - f^j\|_{C(I,J)} \right) \lambda_i^j \phi_i^j(x, y, A) \quad (15.3)$$

then  $f_h \in \left\{ \sum_{j=1}^M \lambda_j \phi_j(x, y, A) \right\}$  approximates  $f$  with accuracy  $\epsilon$  :

$$\begin{aligned} & \|f - f_h\|_{C(I,J)} \\ &= \left\| \sum_{j \in F_-(f)} d(f) \left( \frac{\epsilon}{2} - \|f - f^j\|_{C(I,J)} \right) \left( f - \sum_{i=1}^{N_j} \lambda_i^j \phi_i^j(x, y, A) \right) \right\|_{C(I,J)} \\ &= \left\| \sum_{j \in F_-(f)} d(f) \left( \frac{\epsilon}{2} - \|f - f^j\|_{C(I,J)} \right) \left( f - f^j + f^j - \sum_{i=1}^{N_j} \lambda_i^j \phi_i^j(x, y, A) \right) \right\|_{C(I,J)} \\ &\leq \sum_{j \in F_-(f)} d(f) \left( \frac{\epsilon}{2} - \|f - f^j\|_{C(I,J)} \right) \left( \|f - f^j\|_{C(I,J)} + \left\| f^j - \sum_{i=1}^{N_j} \lambda_i^j \phi_i^j(x, y, A) \right\|_{C(I,J)} \right) \\ &\leq \sum_{j \in F_-(f)} d(f) \left( \frac{\epsilon}{2} - \|f - f^j\|_{C(I,J)} \right) \left( \frac{\epsilon}{2} + \frac{\epsilon}{2} \right) = \epsilon. \end{aligned} \quad (15.4)$$

In the next step, we prove the continuity of  $c_k$ . For the proof, we use Eq. (15.2) to obtain

$$\begin{aligned} & \sum_{j \in F_-(\tilde{f})} \left( \frac{\epsilon}{2} - \|\tilde{f} - f^j\|_{C(I,J)} \right) \\ &\leq \sum_{j \in F_-(\tilde{f})} \left( \frac{\epsilon}{2} - \|\tilde{f} - f\|_{C(I,J)} \right) \\ &\leq \sum_{j \in F_-(\tilde{f})} \left( \frac{\epsilon}{2} - \|\tilde{f} - f^j\|_{C(I,J)} \right) + \sum_{j \in F_0(\tilde{f})} \left( \frac{\epsilon}{2} - \|\tilde{f} - f^j\|_{C(I,J)} \right). \end{aligned} \quad (15.5)$$

Let  $\tilde{f} \rightarrow f$  in Eq. (15.5), then we have

$$\sum_{j \in F_-(\tilde{f})} \left( \frac{\epsilon}{2} - \|\tilde{f} - f^j\|_{C(I,J)} \right) \rightarrow \sum_{j \in F_-(f)} \left( \frac{\epsilon}{2} - \|f - f^j\|_{C(I,J)} \right). \quad (15.6)$$

This obviously demonstrates that  $d(\tilde{f}) \rightarrow d(f)$ . Thus,  $\tilde{f} \rightarrow f$  results

$$d(\tilde{f}) \left( \frac{\epsilon}{2} - \|\tilde{f} - f^j\|_{C(I,J)} \right) \lambda_i^j \rightarrow d(f) \left( \frac{\epsilon}{2} - \|f - f^j\|_{C(I,J)} \right) \lambda_i^j. \quad (15.7)$$

Let  $N = \sum_{j \in F_-(f)} N_j$  and define  $c_k$  in terms of

$$\begin{aligned} f_h &= \sum_{j \in F_-(f)} \sum_{i=1}^{N_j} d(f) \left( \frac{\epsilon}{2} - \|f - f^j\|_{C(I,J)} \right) \lambda_i^j \phi_i^j(x, y, A) \\ &\equiv \sum_{k=1}^N c_k \phi_k(x, y, A) \end{aligned}$$

From Eq. (15.7),  $c_k$  is a continuous functional of  $f$ . Thus the approximation result follows.

## 15.5 Conclusions

In this study, the universal approximation capability of three-layer feedforward double approximate identity neural networks was investigated in the space of continuous functions with two variables. First, the definition of double flexible approximate identity functions was proposed. Then, Theorem 2 was obtained through this definition. Theorem 2 showed that any continuous function  $f$  with two variables convolved with double flexible approximate identity functions converges to itself. Through the obtained result, Theorem 4 was proved as the main result. The proof of Theorem 4 was in the framework of theory of  $\epsilon$ -net and similar to the proof of Theorem 1 [15]. Theorem 4 showed that three-layer feedforward neural networks were capable of uniformly approximating any continuous function  $f$  with two variables on a compact set to arbitrary accuracy.

## References

1. Cybenko, G. (1989). Approximation by superpositions of sigmoidal function. *Mathematics of Control, Signals and Systems*, 2(4), 303–314.
2. Funahashi, K. (1989). On the approximate realization of continuous mappings by neural network. *Neural Networks*, 2(3), 183–192.
3. Park, J., & Sandberg, I. W. (1991). Universal approximation using radial-basis-function networks. *Neural Computation*, 3(2), 246–257.
4. Liao, Y., Fang, S. C., & Nuttle, H. L. W. (2003). Relaxed conditions for radial-basis function networks to be universal approximators. *Neural Networks*, 16(7), 1019–1028.
5. Sanguineti, M. (2008). Universal approximation by ridge Computational models and neural networks: a survey. *The Open Applied Mathematics Journal*, 2(1), 31–58.
6. Zainuddin, Z., & Panahian Fard, S. (2012). Double approximate identity neural networks universal approximation in real Lebesgue spaces. *Neural information processing*. Lecture Notes in Computer Science, vol. 7663, (pp. 409–415). Berlin, Heidelberg: Springer.
7. Yang, X., Chen, S., & Chen, B. (2012). Plane-Gaussian artificial neural network. *Neural Computing and Applications*, 21(2), 305–317.
8. Panahian Fard, S., & Zainuddin, Z. (2013). On the universal approximation capability of flexible approximate identity neural networks. *Emerging technologies for information systems, computing, and management*. Lecture Notes in Electrical Engineering, vol. 236, (pp. 201–27). New York: Springer.
9. Panahian Fard, S., & Zainuddin, Z. (2013). Analyses for  $L^p$  [a, b]-norm approximation capability of flexible approximate identity neural networks. *Neural Computing and Applications*, DOI 10.1007/s00521-013-1493-9
10. Fernández, N. F., Hervás, M. C., Sanchez, M. J., & Gutiérrez, P. A. (2011). MELM-GRBF: A modified version of the extreme learning machine for generalized radial basis function neural networks. *Neurocomputing*, 74(16), 2502–2510.
11. Panahian Fard, S., & Zainuddin, Z. (2013). The universal approximation capabilities of Mellin approximate identity neural networks. *Advances in Neural Networks- ISNN 2013*. Lecture Notes in Computer Science, vol. 7951, (pp. 205–213). Berlin, Heidelberg: Springer.
12. Ditkin, V. A., & Prudnikov, A. P. (1962). *Operation calculus in two variable and its applications*. New York: Pergamon press.
13. Lebedev, V. (1997). *An introduction to functional analysis and computational mathematics*. Boston: Birkhäuser.
14. Jones, F. (1997). *Lebesgue integration on Euclidean space*. Boston: Jones and Bartlett.
15. Wu, W., Nan, D., Li, Z., & Long, J. (August 2007). Approximation to compact set of functions by feedforward neural networks. In 20th International Joint Conference on Neural Networks, Orlando, FL, USA, pp. 1222–1225.

# Chapter 16

## A Novel and Real-Time Hand Tracking Algorithm for Gesture Manipulation

Zhiqin Zhang

**Abstract** Direct use of the hand as an input device is an attractive method for providing natural human–computer interaction (HCI). Computer vision (CV) has the potential to provide more natural, noncontact solutions. As a result, there have been considerable research efforts to use the hand as an input device for HCI in recent years. Hand tracking is the most important procedure for HCI. This chapter presents a novel hand tracking algorithm which can track a hand stable and is real time, and we review on the latter hand tracking research direction, which is a very challenging problem in the context of HCI. Our algorithm is based on mean-shift and we improved it to fit for robust hand tracking by using integrated GIH and skin color mask, our improved algorithm can track hand reliably even in clutter environments. Finally, we demonstrate the benefits of our approach in contrast to existing methods.

### 16.1 Introduction

There has been a great emphasis lately in HCI research to create easier to use interfaces by directly employing natural communication and manipulation skills of humans. Adopting direct sensing in HCI will allow the deployment of a wide range of applications in more sophisticated computing environments such as Virtual Environments (VEs) or Augmented Reality (AR) systems. The development of these systems involves addressing challenging research problems including effective input/output techniques, interaction styles, and evaluation methods [1, 2]. In the input domain, the direct sensing approach requires capturing and interpreting the motion of head, eye gaze, face, hand, arms, or even the whole body [3].

---

Z. Zhang (✉)

School of Computer Science, Wuhan Donghu University, Wuhan 430212, China  
e-mail: [zzq9908@sohu.com](mailto:zzq9908@sohu.com)

Among different body parts, the hand is the most effective, general-purpose interaction tool due to its dexterous functionality in communication and manipulation. Object manipulation interfaces [4–6] utilize the hand for navigation, selection, and manipulation tasks in VEs. A reliable hand tracking method can make the manipulation tasks easy-to-use and stable.

Target tracking is a difficult problem which can be affected by environments, occlusion, movement speed, scaling, and rotation. A reliable hand tracking should solve several pivotal problems: deformation, background disturbance, and scaling. Our approach has been inspired by an established trend in hand tracking, mean-shift is a good tracking method which can track deformation targets, but the traditional mean-shift is not reliable in complicated background. In this study, we present a novel algorithm which is a significant improvement comparing with existing methods, experiments show that our algorithm can achieve similar or greater performance at a lower computational cost in comparison to existing methods, and the hand tracking system can be real time.

## 16.2 Background

There exist many reviews on target tracking [7–10], but the reviews involved in hand tracking is seldom, latest of which only covers studies up to 100.

A recent survey includes a very interesting overview of the use of color for face (and, therefore skin color) detection. A major decision towards providing a model of skin color is the selection of the color space to be employed. Several color spaces have been proposed including RGB [11], normalized RGB, HSV, YCrCb, YUV, etc. Color spaces efficiently separating the chrominance from the luminance components of color are typically considered preferable. This is due to the fact that by employing chrominance-dependent components of color only, some degree of robustness to illumination changes can be achieved. Terrillon et al. [12] review different skin chrominance models and evaluate their performance.

Another approach is to assume that the probabilities of skin colors follow a distribution that can be learned either offline or by employing an online iterative method [13].

The skin colors information is sensitive to the background and local luminance, the presented method aims to solve this disturbance by using both color information and shape information.

The remainder of this chapter is organized as follows: Sect. 16.3 describes the traditional mean-shift tracking algorithm, Sect. 16.4 describes our method details, Sect. 16.5 shows our experiments results, and the final section makes a conclusion.



### 16.3 Mean-Shift Tracker

The heart of the mean-shift algorithm is computation of an offset from location vector  $x$  to a new location according to the mean-shift vector

$$\Delta_x = \frac{\sum_a K(a-x)w(a)a}{\sum_a K(a-x)w(a)} - x = \frac{\sum_a K(a-x)w(a)(a-x)}{\sum_a K(a-x)w(a)} \quad (16.1)$$

where  $K$  is a suitable kernel function and the summations are performed over a local window of pixels  $a$  around the current location  $x$ . A “suitable” kernel  $K$  is one that can be written in terms of a profile function  $k$  such that

$$K(x) = k\left(\|x\|^2\right) \quad (16.2)$$

and profile  $k$  is nonnegative, nonincreasing, piecewise continuous, and  $\int_0^\infty k(r)dr < \infty$ .

An important theoretical property of the mean-shift algorithm is that the local mean-shift offset  $x$  computed at position  $x$  using kernel  $K$  points opposite to the gradient direction of the convolution surface.

$$C(x) = \sum_a H(a-x)w(a) \quad (16.3)$$

Kernels  $K$  and shadow kernels  $H$  must satisfy the relationship

$$h'(r) = -ck(r) \quad (16.4)$$

where  $h$  and  $k$  are the respective profiles of  $H$  and  $K$ ,  $r = \|a-x\|^2$ , and  $c > 0$  is some constant.

Note that the mean-shift vector computed by Eq. (16.1) is invariant to scaling of the sample weights  $w(a)$  by a positive constant  $c$ , that is, if each  $w(a)$  is replaced by  $cw(a)$  for  $c > 0$ . Note that the mean-shift vector is not invariant to a constant offset  $w(a) + c$ . However, all mean-shift kernels explored to date are symmetric about the origin, such that  $K(x) = K(-x)$ , in which case

$$\begin{aligned} \Delta_x &= \frac{\sum K(a-x)(w(a)+c)(a-x)}{\sum K(a-x)(w(a)+c)} = \frac{\sum K(a-x)w(a)(a-x) + c \sum K(a-x)(a-x)}{\sum K(a-x)(w(a)+c)} \\ &= \frac{\sum K(a-x)w(a)(a-x)}{\sum K(a-x)(w(a)+c)} \end{aligned} \quad (16.5)$$

showing that the direction of the mean-shift vector is invariant and just the step size changes [14].

The mean-shift target tracking procedure as follows:

1. Initialize the location of the target, in the current frame with  $\hat{\mathbf{y}}_0$ , compute the distribution  $\{\hat{\mathbf{p}}(\hat{\mathbf{y}}_0)\}_{u=1\dots m}$  and evaluate

$$\rho[\hat{\mathbf{p}}(\hat{\mathbf{y}}_0), \hat{\mathbf{q}}] = \sum_{u=1}^m \sqrt{\hat{\mathbf{p}}_u(\hat{\mathbf{y}}_0) \hat{\mathbf{q}}_u} \quad (16.6)$$

2. Derive the weights  $\{w_i\}_{i=1\dots n_h}$  according to Eq. (16.7)

$$w_i = \sum_{u=1}^m \delta[\mathbf{b}(x_i) - \mathbf{u}] \sqrt{\frac{\hat{\mathbf{q}}_u}{\hat{\mathbf{p}}_u(\hat{\mathbf{y}}_0)}} \quad (16.7)$$

3. Based on the mean-shift vector, derive the new location of the target

$$\hat{\mathbf{y}}_1 = \frac{\sum_{i=1}^{n_h} x_i w_i \mathbf{g}\left(\left\|\frac{\hat{\mathbf{y}}_0 - x_i}{h}\right\|^2\right)}{\sum_{i=1}^{n_h} w_i \mathbf{g}\left(\left\|\frac{\hat{\mathbf{y}}_0 - x_i}{h}\right\|^2\right)} \quad (16.8)$$

Update  $\{\hat{\mathbf{p}}_u(\hat{\mathbf{y}}_1)\}_{u=1\dots m}$  and evaluate  $\rho[\hat{\mathbf{p}}(\hat{\mathbf{y}}_1), \hat{\mathbf{q}}] = \sum_{u=1}^m \sqrt{\hat{\mathbf{p}}_u(\hat{\mathbf{y}}_1) \hat{\mathbf{q}}_u}$

4. While  $\rho[\hat{\mathbf{p}}(\hat{\mathbf{y}}_1), \hat{\mathbf{q}}] < \rho[\hat{\mathbf{p}}(\hat{\mathbf{y}}_0), \hat{\mathbf{q}}]$  Do  $\hat{\mathbf{y}}_1 \leftarrow \frac{1}{2}(\hat{\mathbf{y}}_0 + \hat{\mathbf{y}}_1)$
5. If  $\|(\hat{\mathbf{y}}_1 - \hat{\mathbf{y}}_0)\| < \epsilon$  Stop.

Otherwise Set  $\hat{\mathbf{y}}_0 \leftarrow \hat{\mathbf{y}}_1$  and go to step 1. The tracking consists in running for each frame of the optimization algorithm described above. Thus, given the target model, the new location of the target in the current frame minimizes the distance in the neighborhood of the previous location estimate [15].

## 16.4 Our Hand Tracking Method

### 16.4.1 Features Selecting

Traditional mean-shift algorithm is not reliable in complicated environments and is easy to produce track offset error because of falling into local optimum. Using the skin color detection to track hand is also easy to track wrong targets because of local luminance and background similar to skin color. The kernel reason of producing track error is that the selected features is not sufficient to identify the hand and the other fault objects, so we extract the skin color features and extended histogram features to represent the hand. HSV is a good color space which can distinguish efficiently hand color from background color [13], so we extract the integrated features in HSV.

### 16.4.2 Features Extracting

The three channel color features is computed in HSV, every channel responds to two thresholds which can be used to distinguish hand color and background color:

$$f_i(x) = k(x - \Phi_i) \{i = 1, 2, 3\} \quad (16.9)$$

$k(x)$  is a signal function and if  $(x - \Phi_i) > 0$ ,  $k(x) = 1$ , else  $k(x) = 0$ .  $\Phi_i$  represents for three color channel thresholds and the thresholds were calculated from experiment results. After the procedure we can get the mask image which represents for the most probability candidate hand regions, then the extended histograms can be calculated based on the three mask images. The single channel image histogram is not sufficient to represents for the hand shape information, which inspires us to use GIH (gradient-integrated histogram) features to extend the mean-shift tracking.

### 16.4.3 Algorithm Procedure

The algorithm can be described as follows:

1. Calculating the three channels mask images using Eq. (16.8).
2. Using the channel mask images and original HSV images to extract the hand candidate region images.
3. Weighting the three mask image using Eq. (16.10) and get the single integrated result mask image:

$$F(x) = w1 * h(x) + w2 * s(x) + w3 * v(x); \quad (16.10)$$

where  $w3 = 0.6$ ,  $w2 = 0.3$ ,  $w1 = 0.1$ , and  $G(x)$  is the result mask image,  $x = \{x,y\}$  is image coordinates, and  $h(x)$ ,  $s(x)$ ,  $v(x)$  correspond to the three channel mask images.

4. Normalized the result mask image  $F(x)$  using Eq. (16.11)

$$F_N = \max(0, F(x) / (\max(F(x) + 1))); \quad (16.11)$$

5. Get the gradient image  $G(x)$ , where  $G(x) = F_N'(x)$ . According to the hand shape distribution, we calculate the gradient features using only vertical and two 45 degree diagonal gradient features:  $G_v$ ,  $G_{d1}$ ,  $G_{d2}$ , then  $G(x)$  can be calculated using the following equation:

$$G(x) = w3 * G_v + w4 * G_{d1} + w5 * G_{d2} + w6 * F_N; \quad (16.12)$$

where  $w3 = 0.3$ ,  $w4 = 0.15$ ,  $w5 = 0.15$ ,  $w6 = 0.4$ .

6. Using the  $G(x)$  as the input image to calculate the mean-shift iterative as the above Sect. 16.3 mentioned.

## 16.5 Experimental Results

In this section, representative results from a prototype implementation of the proposed tracker are provided. The reported experiment consists of five long (total 12,815 frames) sequences that has been acquired and processed on-line and in real-time on a Pentium 4 laptop computer running MS Windows at 2.00 GHz. A camera with an usb2.0 interface has been used for this experiment.

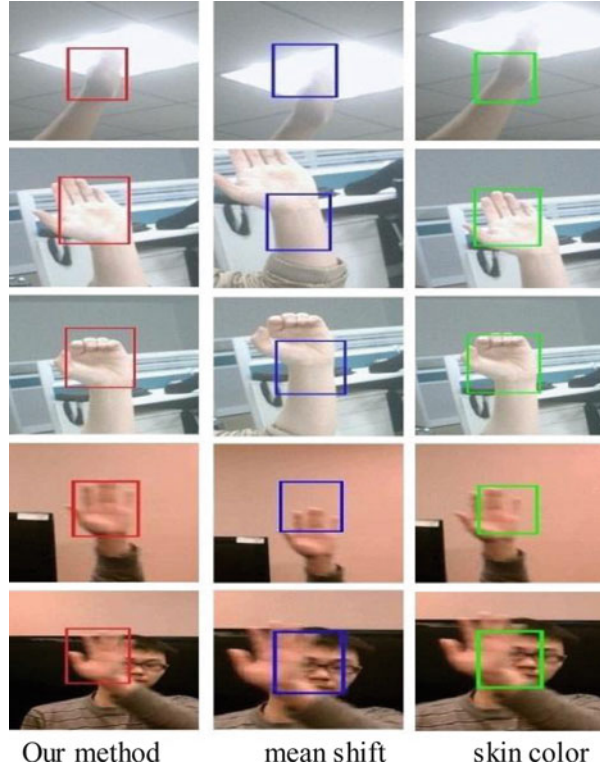
Figure 16.1 provides a few characteristic snapshots of the experiment. For visualization purposes, the bounding rectangle of tracked object is shown. The first column shows our method tracking results, the second column shows the traditional mean-shift tracking results, and the third column shows the skin color method tracking results.

The tracker is initialized using a hand detector such as haar detector or some others color detectors, then the tracker can work well on the whole tracking procedure even encountering the clutter background and drastic hand deformation.

As it can be verified from the snapshots, the labeling of the object hypotheses is consistent throughout the whole sequence, which indicates that they are correctly tracked. Thus, the proposed tracker performs very well in all the above cases, some of which are challenging.

Besides the reported example, the proposed tracker has also been extensively tested with different cameras and in different settings involving different scenes and humans.

**Fig. 16.1** Comparing of several tracking results



The second column and the third column in Fig. 16.1 shows the snapshots of the traditional mean-shift tracking and skin color tracker, respectively, results using the same cameral recording video, the two trackers are lost under clutter environments. From the results we can make a conclusion that our method can achieve good performance improvement comparing to the existing mean-shift tracker and skin color tracker even under clutter background.

Table 16.1 shows the cost of our algorithm and existing methods, the cost is almost the same and real time. Moreover, the cost can be reduced greatly by using parallel algorithm design [16].

## 16.6 Conclusion

In this chapter a new method for tracking hand has been presented. The proposed method can cope successfully with hands moving in complex patterns and local illumination environments using a still camera. The experiments verified that the proposed method is fast and reliable comparing to the existing methods. Since the tracker is not based on explicit background modeling and subtraction, it may

**Table 16.1** The cost comparison table of several methods

Video no.	Mean-shift (traditional)	Proposed method	Skin color
1.	8 ms	10 ms	20 ms
2.	10 ms	13 ms	28 ms

operate even with images acquired by a moving camera. Ongoing research efforts are currently focused on improving the reliability of tracking with a moving camera.

## References

- Oikonomidis, I., Kyriazis, N., & Argyros, A. (2011). Efficient model-based 3d tracking of hand articulations using Kinect. In *Proceedings of BMVC'2011* (pp. 101.1–101.11). BMVA Press.
- Oikonomidis, I., Kyriazis, N., & Argyros, A. A. (2012). Tracking the articulated motion of two strongly interacting hands. In *Proceedings of CVPR'2012* (pp. 1862–1869). NJ: IEEE Conference Publications.
- Erol, A., & Bebis, G. (2007). Vision-based hand pose estimation: A review. *Computer Vision and Image Understanding*, 108(1), 52–73.
- Bowman, D. (2002). Principles for the design of performance-oriented interaction techniques. In K. M. Stanney (Ed.), *Handbook of virtual environments: Design, implementation, and applications* (pp. 201–207). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gabbard, J. (1997). *A taxonomy of usability characteristics in virtual environments*. Master Dissertation of University of Western Australia, Australia.
- Buchmann, V., Violich, S., Billinghamurst, M., & Cockburn, A. (2004). FingARtips: Gesture based direct manipulation in augmented reality. In *Proceedings of GRAPHITE '04: 2nd International Conference on Computer Graphics and Interactive Techniques in Australasia and South East Asia* (pp. 212–221). New York, NY: ACM Press.
- Comaniciu, D., & Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5), 603–619.
- Levinshtein, A., Stere, A., Kutulakos, K., Fleet, D., Dickinson, S., & Siddiqi, K. (2009). Turbopixels: Fast superpixels using geometric flows. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 31(12), 2290–2297.
- Paragios, N., & Deriche, R. (2000). Geodesic active contours and level sets for the detection and tracking of moving objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(3), 266–280.
- Kwon, J., & Lee, K. M. (2010). Visual tracking decomposition. In *Proceedings of CVPR'2010* (pp. 1269–1276). NJ: IEEE Conference Publications.
- Jebara, T. S., & Pentland, A. (1997). Parameterized structure from motion for 3d adaptive feedback tracking of faces. In *Proceedings of CVPR'97* (pp. 144–150). NJ: IEEE Conference Publications.
- Terrillon, J. C., Shirazi, M. N., Fukamachi, H., & Akamatsu, S. (2000). Comparative performance of different skin chrominance models and chrominance spaces for the automatic detection of human faces in color images. In *Proceedings of Automatic Face and Gesture Recognition'2000* (pp. 54–61). NJ: IEEE Conference Publications.
- Saxe, D., & Foulds, R. (1996). Toward robust skin identification in video images. In *Proceedings of Automatic Face and Gesture Recognition'1996* (pp. 379–384). NJ: IEEE Conference Publications.

14. Collins, R. T. (2003). Mean-shift blob tracking through scale space. In *Proceedings of CVPR'2003* (pp. 234–240). NJ: IEEE Conference Publications.
15. Comaniciu, D., Ramesh, V., & Meer, P. (2000). Real-time tracking of non-rigid objects using mean shift. In *Proceedings of CVPR'2000* (pp. 142–149). NJ: IEEE Conference Publications.
16. Vijayanarasimhan, S., & Kapoor, A. (2010). Visual recognition and detection under bounded computational resources. In *Proceedings of CVPR'2010* (pp. 1006–1013). NJ: IEEE Conference Publications.

# Chapter 17

## A Transforming Quantum-Inspired Genetic Algorithm for Optimization of Green Agricultural Products Supply Chain Network

Chunqin Gu and Qian Tao

**Abstract** The green agricultural products supply chain network (GAP-SCN) design provides an optimal platform for efficient and effective supply chain management. This chapter proposes a new solution based on transforming quantum-inspired genetic algorithm (TQGA) to find optimal solution for the GAP-SCN problem. TQGA adopts transforming representation to convert the Q-bit representation to float-point number, and the float-point number to Q-bit representation, using transforming operator to modify chromosomes. The novelty of the transforming operator, that it can avoid the diversity of populations, is decreased. To show the efficacy of the algorithm, TQGA is tested on two cases. Results show that the proposed algorithm is promising and outperforms the classic GA by both optimization speed and solution quality.

### 17.1 Introduction

In recent years, the demand for green agricultural products is increasing. The green agricultural products supply chain network (GAP-SCN) design problem has been gaining importance due to increasing competitiveness. The logistics firms are obliged to maintain high customer service levels while at the same time they are forced to reduce total transportation cost.

The GAP-SCN, as any other supply chain, is a network of organizations. The network involves production bases (PB), logistics centers (LC), and sale terminals (ST) with purpose of satisfying sale terminals' demands, beginning with the PB of

---

C. Gu (✉)

Department of Computer Science, Zhongkai University of Agriculture and Engineering, Guangzhou 201305, China  
e-mail: [guchunqin@gmail.com](mailto:guchunqin@gmail.com)

Q. Tao

Department of Computer Science,  
Guangdong University of Education, Guangzhou 201305, China



green agricultural products, following with LC, and ending with ST such as supermarkets and vegetable markets. Traditionally, manufacturing, distribution, and marketing along the supply chain operated independently. These organizations have their own objectives, and these objectives are often conflicting. But there is a need for a mechanism through which these different functions can be integrated together [1].

As for the SCN problem, the models and solutions for industrial products are relatively massive, but the solutions of the SCN about green agricultural products are limited. In literature, there are many different studies on the design problem of supply networks, and these studies have been reviewed by Erenguc et al. [2] and Pontrandolfo et al. [3]. Amiri [4] has presented a Lagrangian relaxation approach to minimize the total cost of two-stage supply chain. Costa et al. [5] have worked on three stages of SCN optimization problem. Moreover, there are also some intelligent algorithms for SCN problem, such as genetic algorithms (GA) [6, 7]. Most of the researchers have concentrated on the improvement of the supply chain performance, and very few have considered the performance improvement of the algorithm concurrently. Recently, Han and Kim [8] proposed several quantum-inspired GAs (QGAs). Obviously, it is a challenge to apply the algorithm for other different problems. Since the solution representation of the GAP-SCN problem is completely different from that of the knapsack problem, the original QGAs cannot be directly applied to the GAP-SCN problem. The main contribution of the study is to propose a novel transforming QGA in order to solve GAP-SCN problem for performance optimization.

The remainder of this chapter is organized as follows. Section 17.2 presents the statement of the optimization problem. Section 17.3 describes the preliminary knowledge of QGA. Section 17.4 gives the implementation of the proposed algorithm in detail. In Sect. 17.5, a series of experiments are conducted, and the results are analyzed to illustrate the performance of the proposed algorithm. Finally, in Sect. 17.6, this chapter is concluded.

## 17.2 Problem Statement

In order to optimize the performance of GAP-SCN, we should design the distribution network strategy that will satisfy demand requirement for the products imposed by sale terminals. The main objective to solve such a performance optimization problem of GAP-SCN is to evaluate the selection of different production bases or set of the production bases and different logistics centers or set of the logistics centers, whereas the performance criteria are the minimization of the total transportation cost. Simultaneously, the demand for GAP from each sale terminal must be satisfied. The assumptions used in this problem are as follows:

1. Production bases, logistics centers, and sale terminals are known.
2. The requirement for GAP from each sale terminal is known.

**Table 17.1** Notations

Symbol	Descriptions	Symbol	Descriptions
$PB$	Set of production bases	$p'_l$	The $l$ th individual in the $g$ th generation with the Q-bit representation
$pb_i$	Production base $i$	$M$	Number of chromosomes in the population
$LC$	Set of logistics centers	$m$	Length of Q-bit gene
$lc_j$	Logistics center $j$	$N$	Total number of impossible paths
$ST$	Set of sale terminals	$n$	The index of paths, $n \in [1, N]$
$st_k$	Sale terminal $k$	$Req_{st_k}$	The requirement of sale terminal $k$
$TQ_i^j$	Transportation quantity from node $i$ to node $j$	$f_{-pb}(n)$	The map from $n$ to a production base
$C_t$	Chromosome $t$	$g_{-lc}(n)$	The map from $n$ to a logistics center
$QC_t$	Quantum chromosome $t$	$h_{-st}(n)$	The map from $n$ to a sale terminal
$x_t^n$	$n$ th gene of chromosome $i$	$Cost_i^j$	Unit transportation cost from node $i$ to node $j$

- GAP and logistics centers are enough for distribution.
- The transportation costs of green agricultural products on the path from each production base  $pb_i$  to each logistics center  $lc_j$  and from each  $lc_j$  to each sale terminal  $st_k$  are known.
- GAP distributed to  $st_k$  are supplied by a  $lc_j$  or multi- $lc_j$ , and GAP distributed to  $lc_j$  are supplied by a  $pb_i$  or multi- $pb_i$ .

Suppose there is a set  $PB = \{pb_1, pb_2, \dots, pb_{|PB|}\}$  of production bases, a set  $LC = \{lc_1, lc_2, \dots, lc_{|LC|}\}$  of logistics centers, a set  $ST = \{st_1, st_2, \dots, st_{|ST|}\}$  of sale terminals, a  $GAP - SCN \subseteq PB \times LC \cup LC \times ST$ , where “ $\times$ ” is a Cartesian product, and a set  $Req_{ST} = \{Req_{st_1}, Req_{st_2}, \dots, Req_{|ST|}\}$  of requirement of sale terminal  $ST$ ; the objective of performance optimization of GAP-SCN is to find  $PB_i = \{pb_{i1}, pb_{i2}, \dots, pb_{i|PB_i|}\} \subseteq PB$ ,  $LC_j = \{lc_{j1}, lc_{j2}, \dots, lc_{j|LC_j|}\} \subseteq LC$ ,  $ST$ , and  $SCN = \{(pb_i, lc_j), (lc_j, st_k) | pb_i \in PB_i, lc_j \in LC_j, st_k \in ST\}$ , satisfying  $\min f = TQ_i^j \times Cos t_{pb_i}^{lc_j} + TQ_j^k Cos t_{lc_j}^{st_k}$ ,  $Req_{st_k} = TQ_{lc_j}^{st_k}$  (Table 17.1).

### 17.3 Preliminary Knowledge of QGA

The smallest unit in a two-state quantum computer is called a Q-bit. The Q-bit may be in the “1” state or in the “0” state, or in any superposition of the two. The state of a Q-bit can be represented as

$$|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad (17.1)$$

where  $\alpha$  and  $\beta$  are complex numbers that specify the probability amplitudes of the corresponding states.  $|\alpha|^2$  and  $|\beta|^2$  denote the probability that the Q-bit will be found

in the “0” state and “1” state, respectively. Normalization of the state to the unity guarantees  $|\alpha|^2 + |\beta|^2 = 1$ .

A Q-bit individual as a string of  $m$  Q-bit is defined as

$$\begin{bmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_m \\ \beta_1 & \beta_2 & \dots & \beta_m \end{bmatrix}, \quad (17.2)$$

where  $|\alpha_i|^2 + |\beta_i|^2 = 1, i = 1, 2, \dots, m$ .

Q-bit representation has the advantage that it can represent a linear superposition of states. Because the Q-bit representation and represent linear superposition of the state’s probability, evolutionary computing with Q-bit representation has a better characteristic of population diversity than other representation. The state of a Q-bit can be updated by the quantum gate operation  $U$ .

In QGA, rotation quantum gate is selected. The quantum gate is adjusted as follows:

$$\begin{bmatrix} \alpha'_i \\ \beta'_i \end{bmatrix} = U(\theta_i) \begin{bmatrix} \alpha_i \\ \beta_i \end{bmatrix} = \begin{bmatrix} \cos(\theta_i) & -\sin(\theta_i) \\ \sin(\theta_i) & \cos(\theta_i) \end{bmatrix} \begin{bmatrix} \alpha_i \\ \beta_i \end{bmatrix}, \quad (17.3)$$

where  $(\alpha_i, \beta_i)$  is the  $i$ th Q-bit and  $\theta_i$  is the rotation angle of each Q-bit toward either 0 or 1 state depending on its sign.

Inspired by the concept of quantum computing, QGA applies the probabilistic amplitude representation of Q-bit to the coding of chromosome, based on the representation of the quantum vector, so that one chromosome can represent any linear superposition.

## 17.4 Proposed Transforming Quantum-Inspired Genetic Algorithm

This chapter proposes a TQGA for performance optimization of green agricultural products supply chain networks. In TQGA, the transforming representation and transforming operator are applied. In addition, the chromosome representation is critical to solve the optimization problem. In this section, the chromosome representation is firstly described. Then, the chromosome initialization, the evaluation of fitness, and updating operator by rotary quantum gate are presented.

### 17.4.1 Representation of Chromosomes

Each gene in the chromosome is mapped to a path from PB to ST by LC. The gene value indicates transportation quantity of green agricultural products on a path.

Based on this idea, we use float-point representation to encode the SCN problems. Each chromosome  $C_t$  in the population is represented as

$$C_t = (x_t^1, x_t^2, \dots, x_t^n, x_t^{n+1}, \dots, x_t^N), \quad (17.4)$$

where  $x_t^n \in [0, Req_{h\_st(n)}]$  represents the transportation quantity of green agricultural products on path  $n$ ,  $n = 1, 2, \dots, N$ ,  $N = |PB| \times |LC| \times |ST|$  is the total maximum number of impossible paths,  $t = 1, 2, \dots, M$ .  $Req_{h\_st(n)}$  is the requirement of sale terminal  $h\_st(n)$  and is applied to clamp the maximum gene value.

Each quantum chromosome  $QC_t$  in the population is represented as

$$QC_t = (p_t^{11}, p_t^{12}, \dots, p_t^{1m}, p_t^{21}, p_t^{22}, \dots, p_t^{2m}, \dots, p_t^{N1}, p_t^{N2}, \dots, p_t^{Nm})$$

### 17.4.2 Transforming of the Encoding

After  $QC_t$  is created, we should transform it into float-number representation for evaluating the fitness.  $x_t^n$  is calculated by Eq. (17.5).

$$x_t^n = U_{\min} + \left( \sum_{i=1}^m b_i^l \times 2^{i-1} \right) \times \frac{U_{\max} - U_{\min}}{2^m - 1}, \quad (17.5)$$

where  $U_{\min} = 0$ ,  $U_{\max} = Req_{h\_st(n)}$ .

Before applying rotation quantum gate to updating quantum chromosome, the float-number representation chromosome also should be transformed into quantum chromosome.  $\sum_{i=1}^m b_i^l \times 2^{i-1}$  is calculated by Eq. (17.6); then,  $b_i^l$  is calculated by method of successive division

$$\sum_{i=1}^m b_i^l \times 2^{i-1} = (x_t^n - U_{\min}) \times \frac{2^m - 1}{U_{\max} - U_{\min}} \quad (17.6)$$

### 17.4.3 Initialization of Chromosomes

Initially, a population with  $M$  Q-chromosomes is created.  $\alpha_t^l$  and  $\beta_t^l$ ,  $l = 1, 2, \dots, m$ , are initialized with  $1/\sqrt{2}$ . If the gene value of some chromosome is more or less than the maximum requirement, an additional operator, called transforming, is applied on chromosomes. After transforming operator, the initial population  $C_1, C_2, \dots, C_M$  is generated.

### 17.4.4 Evaluation of Fitness

The fitness function of a chromosome  $C_i$  in the population is defined as

$$fitness_i = \sum_{n=1}^N x_i^n \times \left( Cost_{f-pb(n)}^{g-lc(n)} + Cost_{g-lc(n)}^{h-st(n)} \right), \quad (17.7)$$

where  $f-pb(n) = \frac{n}{|LC| \times |ST|}$ ,  $g-lc(n) = \frac{n}{|ST|} \% |LC|$ ,  $h-st(n) = n \% |ST|$ ,  $Cost_{f-pb(n)}^{g-lc(n)}$  is the transportation cost of green agricultural products from  $PB_{f-pb(n)}$  to  $LC_{g-lc(n)}$ , and  $Cost_{g-lc(n)}^{h-st(n)}$  is the transportation cost from  $LC_{g-lc(n)}$  to  $ST_{h-st(n)}$ .

### 17.4.5 Rotation Operation

The rotation operation is only used in QGA to adjust the probability amplitudes of each quantum gene. According to the rotation operation in Eq. (17.3), a quantum gate  $U(\theta_i)$  is a function of  $\theta_i = s(\alpha_i, \beta_i) \cdot \Delta\theta_i$ , where  $s(\alpha_i, \beta_i)$  is the sign of  $\theta_i$  that determines the direction and  $\Delta\theta_i$  is the magnitude of rotation angle. The lookup table of  $\Delta\theta_i$  is proposed by [8].

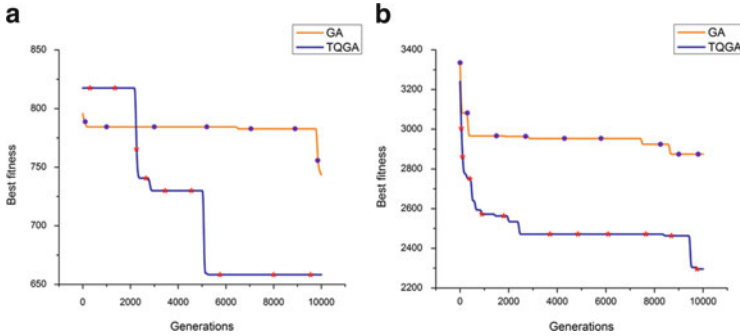
## 17.5 Experimental Study and Discussions

The proposed TQGA for performance optimization of green agricultural products supply chain logistics network was performed on a machine with Pentium D 3.00GHz CPU and 1024MB of RAM. The operating system was MS Windows XP, and the compiler was VC++ 6.0. We tested the TQGA on Case\_1 and Case\_2 as given in Table 17.2.

The following parameters were used in the GA:  $M = 50$ ; in the TQGA,  $M = 20$ . To validate the proposed TQGA, we compared the TQGA with GA to tackle the various problems in GAP-SCN. Experiments were independently simulated 30 times. Figure 17.1 presents the comparison of GA and TQGA. To validate the reliability and robustness of TQGA, we tested two GAs with a longer running time and 10,000 iterations. It can be seen from Fig. 17.1 that TQGA has higher solution accuracy and quicker convergence speed than GA. The advantage of optimization speed of TQGA is outstanding. TQGA outperforms GA both in the solution quality and in the optimization speed, especially in the optimization speed.

**Table 17.2** Test cases

Test cases	$ PB $	$ LC $	$ ST $	Scale	$Req_{st_k}$	$Cost_i^j$
Case_1	2	3	3	18	[10,50]	[1,10]
Case_2	10	15	16	2400	[10,50]	[1,10]



**Fig. 17.1** Comparison of GA and TQGA. (a) Case\_1, (b) Case\_2

## 17.6 Conclusion

In this chapter, we have proposed a TQGA for performance optimization of GAP-SCN. In TQGA, a transforming operator is used to keep solutions in feasible ones when total transportation number does not meet the demands of each ST. A transforming of encoding between float-number representation and quantum representation is described. Experimental results demonstrate the superiority and effectiveness of the proposed algorithm for green agricultural products supply chain network.

**Acknowledgments** This work is supported by the Natural Science Foundation of Guangdong Province of China under Grant No. S2012020011067, the National High Technology Research and Development Program of China (863) under Grant No. 2012AA101701, and the Natural Science Foundation of Zhongkai University of Agriculture and Engineering under Grant No. G3100013.

## References

1. Altiparmak, F., & Gen, M. (2006). A genetic algorithm approach for multi-objective optimization of supply chain networks. *Computers and Industrial Engineering*, 51(1), 197–216.
2. Erenguc, S. S., Simpson, N. C., & Vakharia, A. J. (1999). Integrated production/distribution planning in supply chains: An invited review. *European Journal of Operational Research*, 115 (2), 219–236.
3. Pontrandolfo, P., & Okogbaa, O. G. (1999). Global manufacturing: A review and a framework for planning in a global corporation. *International Journal of Production Economics*, 37(1), 1–19.

4. Amiri, A. (2006). Designing a distribution network in a supply chain system: Formulation and efficient solution procedure. *European Journal of Operational Research*, 171(2), 567–576.
5. Costa, A., Celano, G., Fichera, S., & Trovato, E. (2010). A new efficient encoding/decoding procedure for the design of a supply chain network with genetic algorithms. *Computers and Industrial Engineering*, 59(4), 986–999.
6. Altıparmak, F., Gen, M., Lin, L., & Paksoy, T. (2006). A genetic algorithm approach for multi-objective optimization of supply chain networks. *Computers and Industrial Engineering*, 51(1), 196–215.
7. Xu, J., Liu, Q., & Wang, R. (2008). A class of multi-objective supply chain networks optimal model under random fuzzy environment and its application to the industry of Chinese liquor. *Information Sciences*, 178(8), 2022–2043.
8. Han, K. H., & Kim, J. H. (2002). Quantum-inspired evolutionary algorithm for a class of combinatorial optimization. *IEEE Transactions on Evolutionary Computation*, 6(6), 580–593.

# Chapter 18

## A Shortest Path Algorithm Suitable for Navigation Software

Peng Luo, Qizhi Qiu, Wenyan Zhou, and Pei Fang

**Abstract** In allusion to the shortage of hardware configuration in the mobile devices and high time-complexity of Dijkstra algorithm, the chapter comes up with a shortest path algorithm based on cut-corner for restricted searching area. This algorithm aims at shrinking the smallest searching area quickly and considers the advantages of Ellipse algorithm and Rectangle algorithm. When tested in the simulator, we find that the time-complexity of Cut-corner algorithm is reduced by 5–20 % compared with that of other conventional algorithms. Thus, it has better effect when used in navigation software of low-end mobile device.

### 18.1 Introduction

The shortest path algorithm plays an important role in the research of intelligent transportation systems [1], and with the high development of mobile devices, it can be used more widely in navigation software.

The Dijkstra algorithm [2] is a conventional single-source shortest path algorithm. Dijkstra algorithm uses blind search from the source  $S$  nearly with concentric circles. It does not stop until the radius reaches the destination  $D$  [3]. It is no use to search the nodes, which are deviated far from the destination orientation in the algorithm. Navigation software has high requirements of time. Hence Dijkstra algorithm is not suitable.

A large amount of studies focused on searching strategy, data structure, and restricted searching area to improve the efficiency of searching. But the improvement of data structure is very complicated. By contrast, the improved algorithm for restricted searching area is more mature.

---

P. Luo • Q. Qiu • W. Zhou (✉) • P. Fang  
Department of Computer Science and Technology, Wuhan University of Technology,  
Wuhan 430000, China  
e-mail: [nancyzhouwhut@gmail.com](mailto:nancyzhouwhut@gmail.com)



Nordbeck [4] first came up with the idea of shortest path algorithm based on ellipse for restricted searching area (hereafter referred as Ellipse algorithm). Then many scholars studied on the Ellipse algorithm and made some progress. Lu [5] proposed shortest path algorithm based on rectangle for restricted searching area (hereafter referred as Rectangle algorithm). In succession, shortest path algorithm based on orientation rectangle for restricted searching area [6] (hereafter referred as Orientation Rectangle algorithm) was proposed. Wang [7] delivered Ellipse-based shortest path algorithm for typical urban road networks to concentrate on the relationship between the Euclidean distance and the restricted searching area.

The main idea of restricted searching area is to shrink the area to reduce the total time-complexity.

In this chapter, we come up with a shortest path algorithm based on cut-corner for restricted searching area (hereafter referred as Cut-corner algorithm). The experiment results show that the time-complexity of Cut-corner algorithm is obviously less than other algorithms. So, it is more suitable for navigation software whose capacity of hardware configuration is limited.

## 18.2 Relevant Work

### 18.2.1 *Hardware Configuration of Mobile Device*

Mobile device is used more widely in our daily life because of its portability, and it also promotes the spread of navigation software. Nowadays, navigation software is widely applied in Android mobile device. We have investigated 57 kinds of Android mobile phones. Among them, 32 phones are with the memory of 512 MB and 31 phones are the clock speed of 1 GHz.

### 18.2.2 *Conventional Shortest Path Algorithm for Restricted Searching Area*

In order to explicate the algorithms, symbolic interpretation is listed.

For the given source  $S$  and destination  $D$ :

$|SD|$ : the Euclidean distance between  $S$  and  $D$ .  $d$ : the shortest path length between  $S$  and  $D$ .  $r$ : the ratio of shortest path length to Euclidean distance,  $r=d/|SD|$ .

The main idea of Ellipse algorithm is as follows: each node in the shortest path ought to fall in the ellipse (or on the boundary) whose focal points are  $S$ ,  $D$  and long axis is  $r \times |SD|$ .

Thus for the given source and destination at random, the key of Ellipse algorithm is to determine the value of  $r$ . Current idea is to calculate  $r$  by statistics and the

method is as follows. For the given road network  $G$ , through testing a certain number of samples at random, calculate the ratio  $r_i$  of shortest path length to Euclidean distance of each sample and then determine a value of  $r$  whose confidence interval is 95 % when  $r$  is greater to  $r_i$ , that is, the  $r$  corresponding to  $G$ .

Rectangle algorithm aims at avoiding the high time-complexity of nonlinear operations in Ellipse algorithm. Ellipse is a nonlinear conic section; thus, it needs a lot of involution and evolution operations [8] when judging whether a new added node is inside the ellipse or not. Hence, Lu [5] came up with Rectangle algorithm which constructed minimum circumscribed rectangle contained ellipse. Rectangle algorithm expands the searching area leading to more time of searching the shortest path compared to Ellipse algorithm while its linear operation reduces the time of constructing searching area.

Orientation Rectangle algorithm was proposed to ulteriorly shrink the searching area of Rectangle algorithm. The orientation from the source to the destination roughly represents the orientation of the shortest path. Rectangle algorithm pays no attention to the orientation of the source and the destination, thus Han [6] put forward the Orientation Rectangle algorithm. It also only needs simple linear operations just as Rectangle algorithm does. In one hand, it spends more time to rotate the coordinate system. In the other hand, it reduces the time of searching the shortest path because of fewer nodes inside. In all, the time-complexity of Orientation Rectangle algorithm may be lower than that of Rectangle algorithm.

### ***18.2.3 Time-Complexity of Conventional Algorithms by Comparison***

In order to evaluate the performance of the mentioned algorithms, we construct a virtual road network area of 4,115 nodes and 7,743 roads, which are even distribution. We test Dijkstra algorithm, Ellipse algorithm, Rectangle algorithm, and Orientation Rectangle algorithm in this area. The experiment environment is Eclipse 3.7.2 and Android simulator. The memory of simulator sets as 512 MB. Adjacency list is adopted to store road network.

Table 18.1 shows the experiment results of the shortest path from the source ID 1023 to the destination ID 1193, and Table 18.2 reveals that from 1830 to 2102.  $T_1$  (ms) is the time of constructing searching area,  $T_2$  (ms) is the time of searching shortest path,  $L$  is the length of shortest path, and  $N$  is the number of nodes in the searching area.

Tables 18.1 and 18.2 reveal the following facts:

- Each algorithm can work efficiently. All these three restricted searching area algorithms can find the same accurate shortest path.
- Each algorithm can improve the performance greatly when compared with Dijkstra algorithm. The method to shrink searching area can obviously reduce the time-complexity of algorithms. The searching area of Ellipse algorithm is the

**Table 18.1** Experiment result of the source ID 1023 and destination ID 1193

	Dijkstra algorithm	Ellipse algorithm	Rectangle algorithm	Orientation rectangle algorithm
T <sub>1</sub>	0	51.5	18.6	28.7
T <sub>2</sub>	3,223.4	43.6	57.6	54.8
L	370.2	370.2	370.2	370.2
N	–	81	107	99

**Table 18.2** Experiment result of the source ID 1830 and destination ID 2102

	Dijkstra algorithm	Ellipse algorithm	Rectangle algorithm	Orientation rectangle algorithm
T <sub>1</sub>	0	53.5	22.0	30.1
T <sub>2</sub>	8,099.5	171.9	313.5	270.7
L	541.2	541.2	541.2	541.2
N	–	175	257	228

smallest and is markedly superior to that of other two algorithms. The number of the searching nodes can embody the fact. The searching node number of Ellipse algorithm is reduced by 24.3 % and 18.2 %, respectively, when compared with Rectangle algorithm and Orientation Rectangle algorithm in Table 18.1. In Table 18.2, the rate is 31.9 % and 23.2 %.

- Linear operations can greatly reduce the time of constructing searching area. In Table 18.1, Rectangle algorithm reduces the time by 63.8 % compared with Ellipse algorithm; meanwhile, Orientation Rectangle algorithm reduces 44.3 %. The reduction is 58.9 % and 43.7 % in Table 18.2.

## 18.3 A Shortest Path Algorithm Based on Cut-Corner for Restricted Searching Area

### 18.3.1 The Idea of Algorithm

The conclusions in Sect. 18.2.3 throw light on the key to the algorithm for navigation software. It is more important to use lower-complicated operations. The time-complexity of Dijkstra algorithm is proportional to the square of the number of the nodes. Both Rectangle algorithm and Orientation Rectangle algorithm expand the searching area leading to more searching time than Ellipse algorithm. With the increase of the length between  $S$  and  $D$  leading to the larger constructed area, the time reduced by simple linear operations is less than that increased by path searching, shown in Table 18.2. Therefore, the chapter focuses on the following two criteria to judge whether a restricted searching area algorithm is suitable when applied in navigation software.

- The time-complexity of the algorithm should be much lower. Tables 18.1 and 18.2 show that Rectangle algorithm and Orientation Rectangle algorithm spend

less time than Ellipse algorithm, because they only execute the linear operations to compare the node coordinate with the boundary of rectangle when judging whether a new added node is inside the restricted area.

- The searching area should be constructed as small as possible in the supposed even-distribution area. Through the results conveyed by Tables 18.1 and 18.2, we can find that the searching area of Ellipse algorithm is the smallest. The number of searching nodes also accounts for this.

According to the above two points, we propose Cut-corner algorithm so as to construct polygon whose area is approaching ellipse as restricted searching area to the utmost extent. Its main idea is to cut-corner based on Orientation Rectangle to make the restricted area close to that of ellipse. Cut-corner algorithm uses the advantages of both the minimum restricted searching area and simple linear operations.

### 18.3.2 Algorithm Description

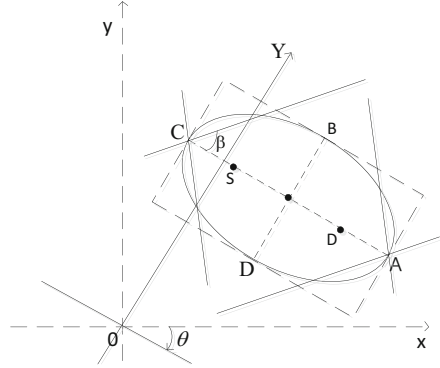
The procedure of Cut-corner algorithm is as follows:

- To rotate the origin coordinate system until the new X-axis parallels with the concatenated line between  $S$  and  $D$ . The new rotation coordinate is called XOY.
- To define two containers  $listN$  and  $listR$  in order to respectively store the nodes and roads in restricted searching area.
- To calculate the Euclidean distance  $|SD|$  between the source  $S$  and destination  $D$ , half major axis of ellipse  $a$ , half minor axis of ellipse  $b$ , and the value of cut-angle  $\beta$  to construct the searching area.
- To judge whether all the nodes in road network is traversed, if it is, go to step 6.
- To get next coordinate of node, change it into the corresponding coordinate in XOY; judge whether this node is inside the cut-corner area. If it is, add it into  $listN$  or go to step 4.
- To judge whether all the roads in road network is traversed, if it is, go to step 8.
- To get the next road, judge whether it is in the cut-corner area. If it is, put the road into  $listR$  or go to step 6.
- To use Dijkstra algorithm to search the shortest path and print it out.

### 18.3.3 Algorithm Implementation

Figure 18.1 illustrates the Cut-corner algorithm. In Fig. 18.1,  $S$  is the source and  $D$  is the destination,  $x_0y_0$  is the original coordinate system, and the angle between the line  $SD$  and X-axis is  $\theta$ . Then rotate the  $x_0y_0$  coordinate system to XOY by  $\theta$ , and the concatenated line of  $S$ ,  $D$  parallels with the X-axis in the rotated coordinate system.  $P$  is the midpoint of  $S$  and  $D$ , and  $A$ ,  $B$ ,  $C$ ,  $E$  are the four tangent

**Fig. 18.1** The sketch of constructing searching area based on Cut-corner algorithm



intersections of orientation rectangle and ellipse.  $\beta$  is the value of cut-angle in Cut-corner algorithm.

In the  $xOy$  coordinate system, assume that the coordinate of the source  $S$  is  $(x_1, y_1)$  and that of the destination  $D$  is  $(x_2, y_2)$  and then come to the conclusion:

The Euclidean distance of  $S$  and  $D$  is

$$|SD| = \sqrt{(x_1 - x_2) \times (x_1 - x_2) + (y_1 - y_2) \times (y_1 - y_2)} \quad (18.1)$$

The rotated angle of coordinate system  $\theta$  is

$$\theta = \arctan((y_2 - y_1)/(x_2 - x_1)), \theta \in (-\pi/2, \pi/2) \quad (18.2)$$

According to the rotation equation of coordinate system (18.3),

$$\begin{aligned} X &= x \cos(-\theta) - y \sin(-\theta) = x \cos \theta + y \sin \theta \\ Y &= x \sin(-\theta) + y \cos(-\theta) = y \cos \theta - x \sin \theta \end{aligned} \quad (18.3)$$

We can gain the coordinates of the source  $S$  and destination  $D$  in the  $XOY$  coordinate system:

$$S = (x_1 \cos \theta + y_1 \sin \theta, y_1 \cos \theta - x_1 \sin \theta) \quad (18.4)$$

$$D = (x_2 \cos \theta + y_2 \sin \theta, y_2 \cos \theta - x_2 \sin \theta) \quad (18.5)$$

Then, the coordinate of midpoint  $P$  of  $S$  and  $D$ :

$$\begin{aligned} P = (m, n) &= ((x_1 \cos \theta + y_1 \sin \theta + x_2 \cos \theta + y_2 \sin \theta)/2 \\ &\quad (y_1 \cos \theta - x_1 \sin \theta + y_2 \cos \theta - x_2 \sin \theta)/2) \end{aligned} \quad (18.6)$$

Assume that  $a$  is the half major axis of ellipse,  $b$  is the half minor axis of ellipse, and  $c$  is the half focal length of ellipse.

Define the proportional coefficient  $r = a/c$ , and  $r$  can be gained by statistics [7]. Thus,

$$a = r \times c = r \times |SD|/2 \quad (18.7)$$

According to equation of ellipse  $a^2 = b^2 + c^2$ , combined with Eq. (18.7), we can get

$$b = \sqrt{a^2 - c^2} = \sqrt{r^2 - 1} \times c = \sqrt{r^2 - 1} \times |SD|/2 \quad (18.8)$$

The searching area of Cut-corner algorithm is supposed to equal to that of Ellipse algorithm. Thus we can get the relationship of cut-angle  $\beta$  and proportional coefficient  $r$ . The derivation refers to Zhou et al. ("An improved shortest path algorithm based on orientation rectangle for restricted searching area," Unpublished) which concentrates on the theoretical research. But this chapter focuses more on practical application.

$$\beta = \arctan\left(\left(2 - \sqrt{r^2 - 1}\right)/((4 - \pi) \times r)\right) \quad (18.9)$$

The horizontal coordinate of  $A$  is  $m+a$ . The vertical coordinate of  $B$  is  $n+b$ . The horizontal coordinate of  $C$  is  $m-a$ . The vertical coordinate of  $E$  is  $n-b$ . Thus the searching area of cut-corner is

$$Z = \{(x, y) | m - a \leq x \leq m + a, n - b \leq y \leq n + b, \arctan(|(y - n)/(x - (m - a))|) \leq \beta, \arctan(|(y - n)/(x - (m + a))|) \leq \beta\} \quad (18.10)$$

According to Eqs. (18.1), (18.2), (18.6), (18.7), (18.8), (18.9), and (18.10), the searching area of Cut-corner algorithm can be solved.

## 18.4 The Analysis of Experiment Results

In the same experiment environment as Sect. 18.2.3, we test Cut-corner algorithm. Table 18.3 shows the experiment result; the meaning of each parameter is the same as Sect. 18.2.3. The second column represents the experiment result from the source ID 1023 to the destination ID 1193, and the third column represents that from ID 1830 to ID 2102.

We can learn from Tables 18.1, 18.2, and 18.3.

- The feasibility of algorithms: as mentioned in Sect. 18.2.3, all the algorithms can find the accurate shortest path, so does the Cut-corner algorithm.

**Table 18.3** The experiment results of Cut-corner algorithm

	ID 1023 to ID 1193	ID 1830 to ID 2102
T <sub>1</sub>	32.6	39.5
T <sub>2</sub>	40.1	174.1
L	370.22	541.2
N	82	176

- The time of constructing searching area: the tables show the time–cost relationship of all the discussed algorithms, which is simplified as:

Rectangle algorithm  $\leq$  Orientation Rectangle algorithm  $\leq$  Cut-corner algorithm  $\leq$  Ellipse algorithm.

For Rectangle algorithm, all the operations are linear and simple, so it spends the least time. While Orientation Rectangle algorithm needs more time to rotate the coordinate system, based on orientation rectangle, Cut-corner algorithm requires additional comparison. Therefore, the time reveals the incremental relationship. However, the three algorithms are all linear operations, which are simpler than the nonlinear operations of Ellipse algorithm. Thus the constructing time is lower than that of Ellipse. For example, in the case of the source ID 1023 and destination ID 1193, Cut-corner algorithm reduces the time by 36.7 % compared with Ellipse algorithm. On the contrary, it increases the time by 75.3 %, 13.6 % compared with Rectangle algorithm and Orientation Rectangle algorithm.

- The time of path searching: the time of Cut-corner algorithm and that of Ellipse algorithm are nearly the same. However, they are both lower than that of Rectangle algorithm or Orientation Rectangle algorithm. So, it costs less time in Cut-corner algorithm. For example, in the case of the source ID 1023 and the destination ID 1193, Cut-corner algorithm reduces the time of searching shortest path by 8.0 %, 30.4 %, and 26.8 %, respectively, compared with Ellipse algorithm, Rectangle algorithm, and Orientation algorithm.
- Among all the algorithms, the total time of Cut-corner algorithm is the least. In the case of the source ID 1023 and destination ID 1193, it reduces the total time by 23.6 %, 4.6 %, and 12.9 %, respectively, compared with Ellipse algorithm, Rectangle algorithm, and Orientation algorithm.

From the above results, when the Euclidean distance between the source and the destination is relatively short, Rectangle algorithm and Orientation Rectangle algorithm are better than Ellipse algorithm. But when the Euclidean distance becomes longer, those two algorithms cannot reflect superiority compared with Ellipse algorithm. Cut-corner algorithm is a combination of advantages of Ellipse algorithm, which has a relatively small searching area, and Rectangle algorithm and Orientation Rectangle algorithm that have the simple linear operations. The time in each situation is the least, which coincides with the main idea, proposed in Cut-corner algorithm in Sect. 18.3.1. Navigation software has high requirement of time. So it values the time-complexity of algorithm more than the shortest path.

In fact, the confidence interval for Ellipse algorithm to search the shortest path is 95 %, which completely satisfy the veracity requirements of navigation software. So Cut-corner algorithm is more suitable for navigation software.

## 18.5 Conclusion

Searching the shortest path is an extremely important function in the navigation software. The fast development of mobile device makes the navigation software more popular. However, because of lacking in hardware resource, the navigation software in mobile device needs more effective algorithm. The chapter aims at solving the shortage of conventional restricted searching area algorithm and proposes the Cut-corner algorithm. It constructs a simulated area and experiments several shortest path algorithm for restricted searching area in the Android simulator. The experiment results show that the Cut-corner algorithm combines the advantage of small searching area and simple linear operations. Its time-complexity is lower than that of Ellipse, Rectangle, and Orientation Rectangle algorithms. Thus it is more suitable for navigation software in mobile device.

## References

1. Wang, Y. B. (2007). Efficient implementation to Dijkstra arithmetic in ITS. *Computer Engineering*, 33(6), 256–258 (In Chinese).
2. Dijkstra, E. W. (1959). A note on two problems in connection with graphs. *Numerische Mathematik*, 1, 269–271.
3. Zhou, Y., Cao, H., & Li, J. X. (2007). A shortest route-planning algorithm within a restricted area. *Microelectronics and Computer*, 24(8), 110–112 (In Chinese).
4. Nordbeck, S., & Rystedt, B. (1969). *Computer cartography shortest route programs* (pp. 28–32). Sweden: The Royal University of Lund.
5. Lu, F., Lu, D. M., & Cui, W. H. (1999). Time shortest path algorithm for restricted searching area in transportation networks. *Journal of Image and Graphics*, 4A(10), 849–853 (In Chinese).
6. Han, G., & Jiang, J. (2003). *The optimal path algorithm with orientation rectangle restriction in vehicle navigation system* (pp. 66–70). In The 7th Proceeding of China GPS Association, Shenzhen (In Chinese).
7. Wang, S. M., Xing, J. P., Zhang, Y. T., & Bai, B. H. (2011). Ellipse-based shortest path algorithm for typical urban road networks. *Systems Engineering – Theory and Practice*, 31(6), 1158–1164 (In Chinese).
8. Fu, M. Y., Li, J., & Deng, Z. H. (2004). A practical route planning algorithm for vehicle navigation system. In *Proceedings of the 5th world congress on intelligent control and automation* (pp. 5326–5329). New York: Institute of Electrical and Electronics Engineers Inc.



# Chapter 19

## An Energy-Balanced Clustering Routing Algorithm for Wireless Sensor Networks

Mingqiang Chen and Xianhai Tan

**Abstract** Aimed at the problem of nodes energy imbalance, which is caused by the heavy burden of cluster heads in clustering wireless sensor networks, an uneven clustering routing algorithm based on multihop communication has been proposed for wireless sensor networks. An election algorithm is used for reasonable selection of cluster heads based on candidate threshold and time driven, the independent nodes are introduced to reduce burden of the cluster heads, and the multihop routing based on angle is applied to optimize the intercluster routing algorithm. Simulation results show that the algorithm can save the network energy effectively and balance the energy consumption.

### 19.1 Introduction

Wireless sensor network (WSN) is a self-organizing and distributed network which is comprised of a large number of microsensor nodes and some base stations (BS) owning wireless communication and computing capacity. WSN has been widely used in many fields, such as environmental monitoring, traffic management, health care, and national defense [1]. Sensor nodes generally come with a disposable battery powered, and their energy is very limited; therefore, the design of routing protocol is the key. In recent years, many researchers have proposed a variety of WSN routing algorithms. Among them, the clustering protocols are concerned by the researchers. LEACH is a clustering routing algorithm [2], in which each node has a certain probability of becoming a cluster head per round, and data deliver directly to BS by single-hop. Previous research has proved that

---

M. Chen (✉) • X. Tan  
School of Information Science and Technology, Southwest Jiaotong University,  
Chengdu 610000, China  
e-mail: [newseem@qq.com](mailto:newseem@qq.com)

multihop routing communication can save more energy [3], but it easily leads to dissipation of energy and the hot spots problem [4]. An energy-efficient unequal clustering (EEUC) algorithm was proposed [5], but it still has the following problems (1) the candidate cluster heads are selected by a preset probability, sensors with low residual energy can still become candidate cluster heads, and their energy is consumed in the election of final cluster heads. The energy of nodes is more imbalanced in the network. (2) The ordinary nodes join the cluster whose signal is the strongest, but it does not consider that some nodes near BS which send directly to BS will consume smaller energy than indirectly. (3) The method of multihop routing is unreasonable, resulting in waste of energy.

In view of the above problems, this chapter presents an energy-balanced uneven clustering routing algorithm for wireless sensor networks (EBUC). It owns the following advantages (1) differentiated from randomly producing candidate cluster heads, the EBUC selects candidate cluster head whose residual energy is not less than the cluster average energy, and the residual energy of nodes is quantified as the time for driving the cluster heads election. The selected cluster heads own high energy, and the message complexity of the cluster formation algorithm is reduced. (2) By defining the independent nodes, the loads of some cluster heads are reduced nearby BS. (3) The routing algorithm based on angle is proposed, reducing the intercluster energy consumption.

## 19.2 Network Model

The network consists of  $N$  sensor nodes, which are randomly deployed over a vast field. Each node has a unique ID in the network, and we assume it has the following characteristics.

1. All sensor nodes are energy heterogeneous, whose positions are fixed, and have the capability of sensing signal emission angle.
2. BS has endless energy, a strong computing and storage capacity.
3. If transmission power is known, the distance between the sender and the receiver can be calculated [6].
4. Sensors can use power control to vary the amount of transmission power which depends on the receiver.

We use a simplified model in which the radio hardware energy dissipation consists of the sending circuit loss and the power amplifier loss [7]. As shown in Eq. (19.1), both the free space ( $d^2$  power loss) and the multipath fading ( $d^4$  power loss) channel models are used in the model, depending on the distance between the transmitter and receiver. The energy spent for transmission of an  $l$ -bit packet over distance  $d$  is

$$E_{Tx}(l, d) = \begin{cases} lE_{elec} + l\epsilon_{fs}d^2, & d < d_o \\ lE_{elec} + l\epsilon_{mp}d^4, & d \geq d_o \end{cases} \quad (19.1)$$

Equation (19.2) shows energy consumption, when the node receives this packet.

$$E_{Rx}(l) = lE_{elec} \quad (19.2)$$

A cluster head consumes  $E_{DA}$  amount of energy for data aggregation. We also assume that the sensed information is highly correlated, so the cluster heads can always aggregate the data received from its members into a single length-fixed packet.

### 19.3 The EBUC Mechanism

After the deployment of sensor nodes, BS broadcasts BS\_MSG to the entire network with the determined power. Each node calculates the distance to BS by the receiving signal strength. We need to control the range of competition radius in the network and set  $R_c$  of  $s_i$  as Eq. (19.3).

$$s_i.R_c = \left(1 - c \cdot \frac{d_{\max} - d_{s_i-BS}}{d_{\max} - d_{\min}}\right) \cdot R_c^0 \quad (19.3)$$

$c$  is a constant coefficient between 0 and 1.  $d_{\max}$  and  $d_{\min}$  represent the maximum and minimum distance between sensor nodes to BS.  $d_{s_i-BS}$  is the distance from  $s_i$  to the BS.  $R_c^0$  is the maximum competition radius which is predefined. According to Eq. (19.3),  $R_c$  varies from  $(1 - c)R_c^0$  to  $R_c^0$ . As an example, if  $c$  is set to  $1/3$ ,  $R_c$  varies from  $2/3 R_c^0$  to  $R_c^0$  according to its distance to BS.

#### 19.3.1 Election of Cluster Heads

In many WSNs applications, the network node density is large, so it is not necessary for all nodes to become the candidate cluster heads.  $E_{ave}$  is the energy threshold, which is the cluster average remaining energy in the last round. If the residual energy of some node is not less than  $E_{ave}$ , it will become a candidate cluster head, and its competition time  $t$  is calculated in Eq. (19.4).

$$t = \alpha \cdot T_{CH} \cdot \frac{E_{\max} - E_i}{E_{\max} - E_{ave}}, E_i \geq E_{ave} \quad (19.4)$$

$T_{CH}$  is a period of time for the cluster head election.  $E_i$  represents the residual energy of  $s_i$ .  $E_{\max}$  means the maximum residual energy of the intercluster nodes in the last round.  $\alpha \in (0.9, 1)$  is a random number which is used to avoid the collisions of HEAD\_MSGs which are caused by the nodes owning same residual energy.

In the election, the candidate cluster heads broadcast HEAD\_MSGs (containing the node ID, the competition radius  $R_c$ , and competition time  $t$ ) by radius  $R_c^0$  at their competition time  $t$ , in order to declare that they win. The adjacent candidate cluster head set of  $s_i$  is defined as  $s_i.N_c = \{s_j | d_{s_i-s_j} < \max(s_i.R_c, s_j.R_c)\}$ . When the sensors receive HEAD\_MSGs from  $N_c$ , they will quit their own election process. After  $T_{CH}$ , all cluster heads are selected, and then they broadcast CH\_ADV\_MSGs.

### 19.3.2 The Formation of Clusters

After all cluster heads are selected, the nodes are divided into the ordinary nodes and the cluster heads.  $T_r$  is a threshold for distinguishing the ordinary nodes. If the distance between an ordinary node and BS is less than  $T_r$ , the ordinary node is divided into the first class nodes. Otherwise, it is divided into the second class nodes. We assume that  $s_i$  is one of the first class nodes, CH is the nearest cluster head to  $s_i$ , and the distance between the two is recorded as  $d_{CH-BS}$ . According to Eqs. (19.1) and (19.2), when  $k$ -bit packet is transmitted to BS by single-hop, the energy consumption is shown in Eq. (19.5).

$$s_i.E_{direct} = k \cdot E_{elec} + k \cdot \epsilon_{fs} \cdot d_{s_i-BS}^2 \quad (19.5)$$

If  $s_i$  sends  $k$ -bit packet to BS via CH, the energy consumption is shown in Eq. (19.6).

$$s_i.E_{CH} = k \cdot (E_{elec} + \epsilon_{fs} \cdot d_{s_i-CH}^2) + k \cdot E_{elec} + k \cdot E_{DA} \quad (19.6)$$

Comparing  $E_{direct}$  with  $E_{CH}$ , we know that by which energy consumption is lower.

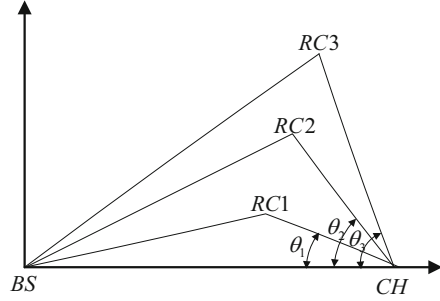
$$s_i.E_{CH} - s_i.E_{direct} = k \cdot (E_{elec} + E_{DA} + \epsilon_{fs} \cdot d_{s_i-CH}^2 - \epsilon_{fs} \cdot d_{s_i-BS}^2) \quad (19.7)$$

Order  $H = (E_{elec} + E_{DA})/\epsilon_{fs}$ , then:

1. If  $d_{s_i-BS}^2 - d_{s_i-CH}^2 \leq H$ , that is  $s_i.E_{CH} - s_i.E_{direct} \geq 0$ . Energy consumption that  $s_i$  sends data directly to BS is smaller than indirectly.  $s_i$  is defined as an independent node and sends a JOIN\_BS\_MSG to BS.
2. If  $d_{s_i-BS}^2 - d_{s_i-CH}^2 > H$ , that is  $s_i.E_{CH} - s_i.E_{direct} < 0$ . Energy consumption that  $s_i$  sends packet indirectly to BS is lower than directly, and  $s_i$  joins the nearest cluster.

For far from BS, the energy consumption is extremely high, when the second class nodes send packet directly to BS, so they send JOIN\_CH\_MSGs to join the nearest cluster. When all clusters are established, a Voronoi diagram [5] is formed in the network.

**Fig. 19.1** The relay node model



### 19.3.3 Data Transmission

From the adjacent cluster head set of CH, the cluster heads which satisfy the condition of  $0 \leq \theta < \frac{\pi}{2}$  within the range of  $n \times R_c^0$  are selected to constitute  $N_{RC}$  which is the candidate relay set of CH.  $n$  is the smallest positive integral, which makes that  $N_{RC}$  non-null.  $\theta$  is an angle and measured by CH. As shown in Fig. 19.1, we assume  $N_{RC}$  of CH contains RC1, RC2, and RC3 and  $\theta_1 < \theta_2 < \theta_3$ .

If CH chooses RC1 as the relay node, we assume a free space propagation channel model, and RC1 communicates with BS directly. To deliver an  $l$ -length packet to BS, the energy consumed by RC1 and CH is

$$\begin{aligned} E_{2hop} &= E_{Tx}(l, d_{CH-RC1}) + E_{Rx} + E_{Tx}(l, d_{RC1-BS}) \\ &= 3lE_{elec} + l\epsilon_{fs}(d_{CH-RC1}^2 + d_{RC1-BS}^2) \\ &= 3lE_{elec} + l\epsilon_{fs}(d_{CH-BS}^2 + 2d_{CH-RC1}^2 - 2d_{CH-BS}d_{CH-RC1}\cos\theta_1) \end{aligned} \quad (19.8)$$

According to Eqs. (19.1) and (19.2), thus we define

$$\cos t1 = d_{CH-RC1}^2 - d_{CH-BS}d_{CH-RC1}\cos\theta_1 \quad (19.9)$$

Among them,  $d_{CH-RC1}$  means the distance between CH and RC1. After CH is placed in position,  $d_{CH-BS}$  is a constant, which means the distance between CH and BS. Similarly, we can calculate  $\cos t2$  of RC2 and  $\cos t3$  of RC3. Because of  $\cos t1 < \cos t2 < \cos t3$ , RC1 is selected as the relay node of CH.

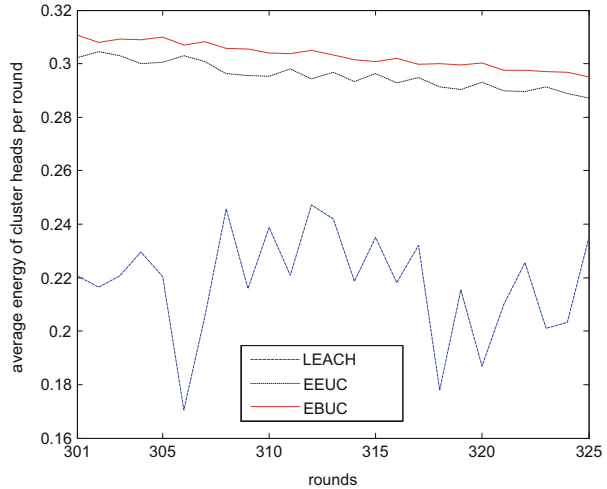
## 19.4 Algorithm Analysis and Simulation

### 19.4.1 Message Complexity

**Lemma 1** The message complexity of the cluster formation algorithm is  $O(N)$  in the network.

**Table 19.1** Simulation parameters

Parameter	Value	Parameter	Value
Network coverage	(0,0)~(200,200) m	$E_{elec}$	50 nJ/bit
BS location	(0,250)	$\epsilon_{fs}$	10 pJ/(bit·m <sup>2</sup> )
N	400	$\epsilon_{mp}$	0.0013 pJ/(bit·m <sup>4</sup> )
Initial energy	0.50 ± 0.01 J	$d_o$	87 m
Data packet size	4,000 bits	$E_{DA}$	5 nJ/(bit·signal)

**Fig. 19.2** Cluster heads average energy

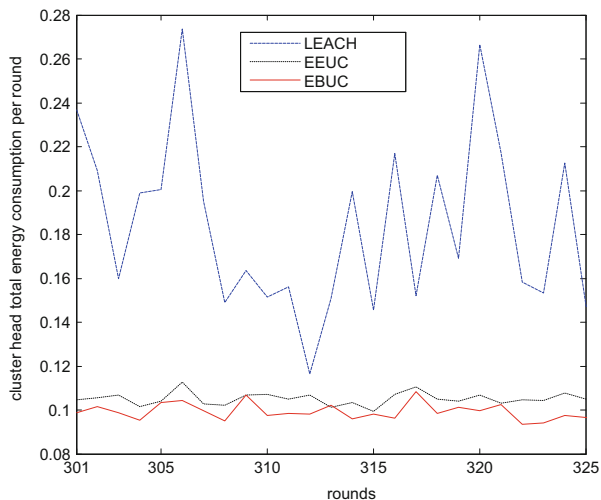
*Proof* At the beginning of the cluster head selection phase,  $N \times T$  candidate cluster heads are produced, and each of them broadcasts a HEAD\_MSG. If it has  $x$  cluster heads and  $y$  independent nodes, it will send out  $x$  CH\_ADV\_MSGs and  $y$  JOIN\_BS\_MSGs, and then the ordinary nodes will send  $N - x - y$  JOIN\_CH\_MSGs. So the messages add up to  $N \times T + x + y + N - x - y = N(T + 1)$  per round. The message overhead of EBUC is  $N \times T$  smaller than EEUC.

### 19.4.2 Simulation Analysis

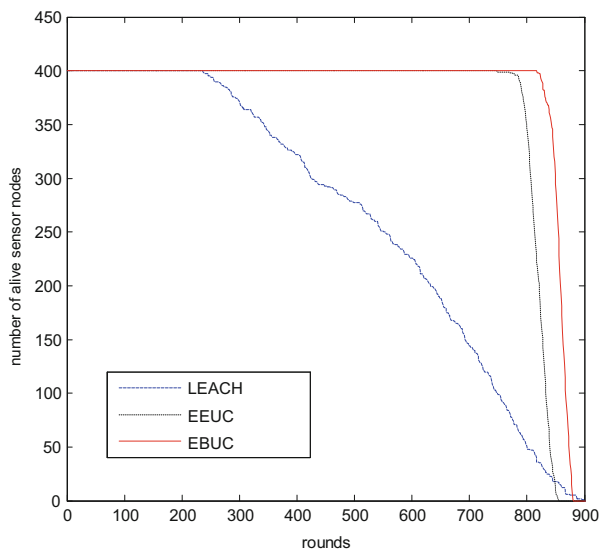
To illustrate the energy efficiency of EBUC, we use MATLAB to simulate LEACH, EEUC, and EBUC under an ideal condition, and packet loss and delays are ignored. The energy consumption of data aggregation and transmission is measured in the experiment. Firstly, survival time of the network is measured, and the energy efficiency of EBUC is analyzed. Secondly, the residual energy and energy consumption characteristic of cluster heads are analyzed. The simulation parameters are shown in Table 19.1.

Figure 19.2 is the cluster heads average energy curve of three protocols in 25 randomly selected rounds, and it shows that the cluster heads average energy

**Fig. 19.3** Cluster heads energy consumption



**Fig. 19.4** Number of survival nodes curve



of EBUC is the highest of the three algorithms. Figure 19.3 is the comparison of cluster heads total energy consumption of three protocols in 25 randomly selected rounds. It shows that the total energy consumption of EBUC is the lowest of the three algorithms.

Finally, we examine the network lifetime of three algorithms. Figure 19.4 shows that EBUC clearly improves the network lifetime over LEACH and EEUC. In EEUC, candidate cluster heads are randomly selected. Therefore, sensors with low residual energy can still become candidate cluster heads, and their energy is consumed in the election of final cluster heads. Such is avoided in EBUC. Multihop

routing based on angle is used in EBUC, so its energy consumption of EBUC is lower than EEUC as illustrated in Fig. 19.3. Comparing to EEUC, stable period of EBUC is increased by 10.6 %. Simulation results show (1) EBUC is able to effectively balance the energy consumption of nodes in the network and extend the network survival time. (2) The cluster heads of EBUC own the highest average energy and capacity of data transmission of the three algorithms. (3) The energy consumption of EBUC cluster heads is the smallest of the three algorithms.

## 19.5 Conclusion

The article proposes an energy-balanced uneven clustering WSN routing algorithm EBUC. The main idea is that the candidate cluster heads are produced by the candidate threshold and their own competition time, so it avoids that the candidate cluster heads are produced randomly and at the same time the message complexity is reduced. Then, it introduces the concept of the independent nodes to reduce some cluster head loads, and the intercluster route is optimized by the factor of multihop energy consumption for saving energy of cluster heads. Simulation results show that the cluster heads energy of EBUC is the highest of the three algorithms, and its energy consumption is the lowest, and it can efficiently prolong the network lifetime.

## References

1. Su, I. J., Tsai, C. C., & Sung, W. T. (2012). Area temperature system monitoring and computing based on adaptive fuzzy logic in wireless sensor networks. *Applied Soft Computing*, 12(5), 1532–1541.
2. Heinzelman, W., Chandrakasan, A., & Balakrishnan, H. (2000). Energy-efficient communication protocol for wireless microsensor networks. *IEEE System Sciences*, 3005–3014. doi:10.1109/HICSS.2000.926982
3. Mhatre, V., & Rosenberg, C. (2004). Design guidelines for wireless sensor networks: Communication, clustering and aggregation. *Ad Hoc Networks*, 2(1), 45–63.
4. Perillo, M., Cheng, Z., & Heinzelman, W. (2004). *On the problem of unbalanced load distribution in wireless sensor networks* (pp. 74–79). Global Telecommunications Conference Workshops 2004, IEEE, Texas.
5. Li, C., Mao, Y., Chen, G., & Wu, J. (2005). *An energy-efficient unequal clustering mechanism for wireless sensor networks* (Vol. 8, p. 604). Mobile Ad Hoc and Sensor Systems Conference, IEEE, Washington, DC.
6. Pei, Z. M., Deng, Z. D., Xu, S., & Xu, X. (2010). A new localization method for wireless sensor network nodes based on N-best rank sequence. *Acta Automatica Sinica*, 36(2), 119–207.
7. Heinzelman, W., Chandrakasan, A., & Balakrishnan, H. (2002). An application specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communications*, 1(4), 660–670.



# Chapter 20

## Simulation and Analysis of Binary Frequency Shift Keying Noise Cancel Adaptive Filter Based on Least Mean Square Error Algorithm

Zhongping Chen and Jinding Gao

**Abstract** Pseudorandom binary frequency shift keying (2FSK) sine wave signals with the frequency of 1,200 bit/s and 2,200 bit/s are produced by using the rand function in MATLAB. A noise cancel adaptive filter based on least mean square error LMS algorithm is designed and simulated on MATLAB. The relationships between the adaptive parameters (filter taps  $N$  and iterations step length  $\mu$ ) and the convergence speed and precision are analyzed. The optimized adaptive parameters which are not only sufficed for filtering performance but also suited for FPGA hardware implementation are found out, that is, filter taps  $N = 22$  and iteration step length  $\mu = 0.006$ . It may provide a good foundation for FPGA hardware implementation.

### 20.1 Introduction

Filtering is one of the most basic and important techniques in the field of signal and information processing. In the process of signal transmission, the signal usually will be disturbed such that signal transmission due to multipath delay caused by the inter-code interference and additive noise affects the communication quality and is the main cause of bit errors. The digital filter can change the relative proportion of the frequency component contained in the signal by a numerical calculation, to the purposes of suppressing or eliminating interference [1].

The adaptive filter belongs to the category of modern filter; when the statistical characteristics of the input signal and the noise are unknown or the statistical characteristics of the input process change, the adaptive filter can automatically

---

Z. Chen  
Hunan Engineering Polytechnic, Changsha 410151, China  
e-mail: [czpmcu@126.com](mailto:czpmcu@126.com)

J. Gao (✉)  
Hunan International Economics University, Changsha 410205, China  
e-mail: [jdgao@qq.com](mailto:jdgao@qq.com)

adjust its parameters to meet the requirements of certain criteria, its research mainly including adaptive algorithm and hardware implementation. Adaptive noise cancellation is one of the four classic applications [2, 3].

Binary frequency shift keying (2FSK) is one of the most widely used digital communication modulations; the International Telegraph and Telephone Consultative Committee (CCITT) recommends a data rate of lower than 1,200 bit/s with 2FSK modulation way in the voice band. In order to find out the optimal adaptive parameters that not only satisfy the filtering performance but also are suitable for FPGA hardware implementation, in this chapter, take the binary frequency shift keying signal (2FSK), for example, design and simulation of the 2FSK signal denoising adaptive filter based on LMS algorithm on MATLAB, analyze the performance of adaptive filter, obtain the optimum adaptive parameters that both can satisfy the requirement of filtering performance and can suit hardware implementation, and provide reference basis for the FPGA hardware implementation [4, 5].

## 20.2 Summary of the LMS Algorithm

There are two main adaptive algorithm: least mean square error (LMS) algorithm and the recursive least squares (RLS) algorithm. The least mean square error (LMS) algorithm proposed by Widrow and Hoff is a fast search algorithm using gradient estimated value instead of the gradient vector, because it has a small amount of calculation and easy-to-realize benefits that are widely used in practice. The basic idea is to adjust the filter weighting parameter to figure out the mean square error between the filter output signal and the least desired signal. As shown in Fig. 20.1, using the direct form FIR structure,  $x(k)$  is the input signal;  $y(k)$  is the output of the filter;  $d(k)$  is the desired signal, i.e., the reference input; and  $e(k)$  is the error signal  $W(k)$  for the variable filter coefficient.

The least mean square error algorithm iterative formula based on the steepest descent algorithm is as follows:

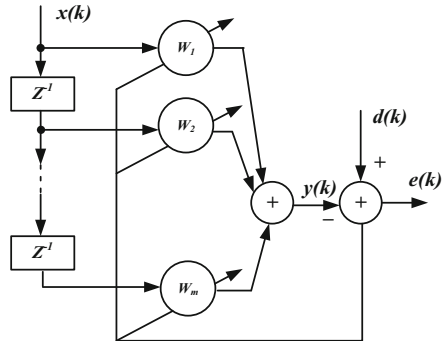
$$y(k) = \sum_{i=0}^N w_i(k)x(k-i) \quad (20.1)$$

$$e(k) = d(k) - y(k) \quad (20.2)$$

$$w(k+1) = w(k) + 2\mu e(k)x(k) \quad (20.3)$$

where  $N$  is the filter order and  $\mu$  is the iterative step. The LMS algorithm convergence speed and accuracy mainly depend on these two parameters. In order to ensure the convergence of the algorithm, the iterative step  $\mu$  requirements in  $0 < \mu < 1/\lambda_{\max}$ ,  $\lambda_{\max}$  are the maximum eigenvalue of the input signal autocorrelation matrix and the step size  $\mu$  that influences the convergence speed of the filter. The higher the order of the filter, the greater the accuracy; however, the higher

**Fig. 20.1** Structure of LMS adaptive filter



the order number, the more hardware implementation cost of the resources. In this chapter, by MATLAB simulation, find out the optimum adaptive parameters  $\mu$  and  $N$  that both can satisfy the requirement of filtering performance and can suit hardware implementation, to provide a reference for the FPGA hardware implementation.

## 20.3 Simulations and Analysis

### 20.3.1 2FSK Signal Generation

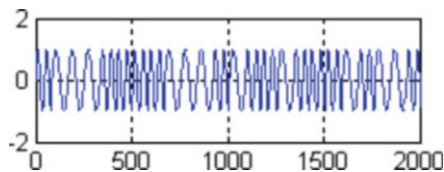
Rand (N) is the MATLAB built-in pseudorandom number generating function, the function value within (0.0 1.0), and obeys uniform distribution.

According to Rand (N), if the value is greater than or less than the mean value 0.5, then take 1 cycle of frequency 1,200 Hz for the sine signal or take one cycle of frequency 2,200 Hz for the sine wave; random 2FSK sequence generated from 2FSK pseudorandom signal is shown in Fig. 20.2.

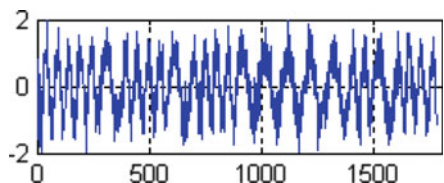
### 20.3.2 LMS Adaptive Filter Simulation and Analysis

Taking the 2FSK pseudorandom sequence as the desired signal  $d(k)$ , the input signal  $x(k)$  of filter is the 2FSK pseudorandom sequence superimposed with white Gaussian noise.  $e(k)$  is the error signal and  $y(k)$  is the filter output signal. According to the LMS iterative algorithm, written the m file in MATLAB, to reduce the FPGA implementation cost of hardware resources, the main thing is to meet the minimum requirements of the filtering performance of the adaptive filter order  $N$ . In this chapter, by changing the iteration step size  $\mu$  and order  $N$ , a series of simulation was conducted. Shown in Figure 20.3 is a filter input signal  $x(k)$ . Figures 20.4 and 20.5 compare how to select a few key parameters of the filter error simulation output;

**Fig. 20.2** Random 2FSK sequence



**Fig. 20.3** Filter input signal



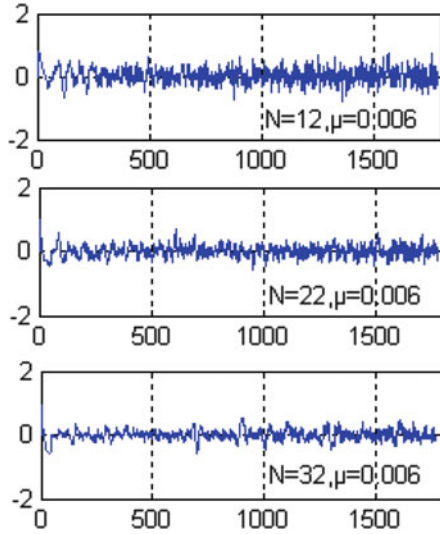
the abscissa is the number of sampling points, and the ordinates are the relative amplitudes of the size.

Shown in Figure 20.4 is the step size  $\mu = 0.006$ , the error output image of the various orders adaptive filter. It can be seen from the figure that when the filter step size  $\mu$  is fixed, with the increase of the filter order  $N$ , error output of the filter is smaller, filtering performance is better, and the accuracy of the adaptive filter is higher;  $N = 12$ , the larger error output of the filtering performance cannot meet the requirements;  $N = 22$  and  $N = 32$ , the relatively small error output can meet the performance requirements; and  $N = 32$ , error output smaller than  $N = 22$ ; however, taking into account the more orders of FPGA hardware implementation and the greater cost of hardware resources, so as long as they can meet the filter performance requirements,  $N = 22$  is more ideal.

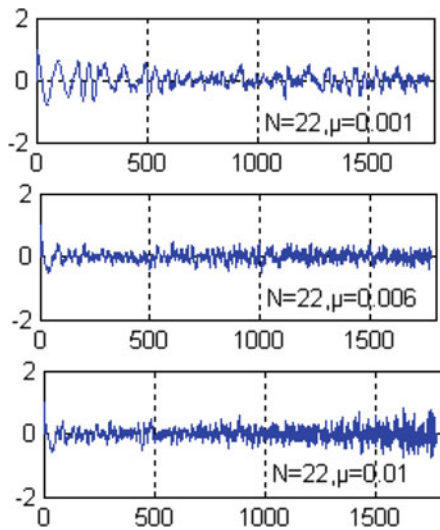
Shown in Figure 20.5 is a filter order  $N = 22$ , the error output image of the adaptive filter of the various iteration step. It can be seen in Fig. 20.5 that the iteration step size  $\mu$  has great influence of the convergence speed of the filter.  $\mu$  value is small; when  $\mu = 0.001$ , the error convergence speed is very slow and the filter cannot meet the filtering requirements; the greater the value of  $\mu$ , the faster the convergence speed. When  $\mu = 0.006$ , it can meet the filtering requirements, but as mentioned above, in order to ensure the convergence of the algorithm, the step size  $\mu$  requires a  $0 < \mu < 1/\lambda_{\max}$  range. When the iteration step length  $\mu = 0.01$  h, the filter output is unstable and even divergence.

After a series of simulation and analysis, the best adaptive parameters are identified, which meet not only the filtering requirements but also those for FPGA hardware implementation: filter order  $N = 22$  and the step size  $\mu = 0.006$ .

**Fig. 20.4** The step size  $\mu = 0.006$ , various orders of the adaptive filter error output



**Fig. 20.5** Order of  $N = 22$  different iterations of step adaptive filter error output



## 20.4 Conclusion

FPGA hardware implementation adaptive filter is an important way to meet the high-speed real-time signal processing requirements. LMS adaptive de-noising filter about pseudorandom binary frequency shift keying (2FSK) signal is designed and simulated on MATLAB. The relationships between the adaptive parameters (filter taps  $N$  and iterations step length  $\mu$ ) and the convergence speed and precision are analyzed in this paper. The convergence of the algorithm to identify the filtering

performance that meets the requirements for hardware is achieved. The best adaptive parameters are filter order  $N = 22$  and the step size  $\mu = 0.006$ ; it provides a reference for the FPGA hardware implementation.

**Acknowledgements** This work is partially supported by the Research Fund of Excellent Youth Project of Hunan Provincial Education Department (09B058), Innovative Special Projects of Research Conditions of Hunan Province (2013TT2037), and the Research Fund of Key Project of Hunan Provincial Education Department (10A068). The authors also gratefully acknowledge the helpful comments and suggestions of the reviewers, which have improved the presentation.

## References

1. Pan, S., & Wang, G. (2000). *VHDL practical tutorial* (pp. 52–56). Chengdu: University of Electronic Science and Technology of China Press.
2. Diniz, P. S. R. (2004). *Adaptive filtering: Algorithms and practical implementation* (2nd ed., pp. 12–159). Beijing: Publishing House of Electronics Industry.
3. Haykin, S. (2003). *Adaptive filter theory* (4th ed., pp. 1–169). Beijing: Publishing House of Electronics Industry.
4. Windrow, B., & Hoff, M. E. (1960). Adaptive switching circuits. *Proceedings of Wescon Convention Record, Part 4, 12(1)*, 96–140.
5. Gao, J. (2013). An improved delayed structure of least mean square adaptive filter and its field programmable gate array implementation. *ICIC Express Letters, Part B: Applications, 4(1)*, 69–73.

# Chapter 21

## Density-Sensitive Semi-supervised Affinity Propagation Clustering

Kunlun Li, Qi Meng, Shangzong Luo, Hexin Li, and Qian Wang

**Abstract** A density-sensitive semi-supervised affinity propagation clustering algorithm (DS-SAP) is proposed in this chapter. The DS-SAP uses supervised information of the pairwise constraints for adjusting data points distance matrix. Then we introduce a novelty similarity metric based on the characteristics of global and local data distribution. This metric can effectively reflect the reality of data distribution. The DS-SAP clustering algorithm is based on the frame of the traditional AP algorithm and has extended data processing capacity compared to the traditional AP algorithm. Experimental results show that the new algorithm is outperforming traditional AP clustering algorithm.

### 21.1 Introduction

Traditional clustering algorithms are sensitive to the choice of initial clustering center. Frey and others put forward a new clustering algorithm based on factor graph theory, the affinity propagation algorithm, which can avoid the above problems [1]. Different from many classic clustering algorithms, AP makes all the data points as the potential exemplars, which could avoid the initial clustering center influence of clustering results. It can offer a better performance with a higher efficiency of data work with no requirement of being symmetry made by the similarity matrices of data points [2]. Since 2007 AP algorithms have been proposed, many methods have been introduced to improve the clustering accuracy of AP. AP is applied in a wider range with the improved methods introduced.

Examples are the following: DS-AP algorithm by Prof. Lu makes dense data sparse and parallel based on AP clustering algorithm [3]. KMNC-AP by Prof. Xing increases the efficiency and accuracy of clustering by using the idea of K-mutual

---

K. Li (✉) • Q. Meng (✉) • S. Luo • H. Li • Q. Wang  
College of Electronic and Information Engineering, Hebei University, Baoding 071002, China  
e-mail: [likunlun@hbu.edu.cn](mailto:likunlun@hbu.edu.cn); [meng88219@163.com](mailto:meng88219@163.com)

nearest neighbor consistency to adjust the similarity between data points [4]. Dong proposed a new algorithm, which had the ability of describing the characters of data clustering effectively, and this method has extended data processing capacity compared with traditional AP [5]. Ahmad Akl applied the AP algorithm to the gesture recognition system, to achieve a high identification accuracy [6].

We propose a new similarity metric to improve AP in this chapter. With the introduction of semi-supervised learning strategy to improve similarity matrices, we take the pairwise constraints and a kind of density-sensitive metric into AP, to introduce the density-sensitive semi-supervised affinity propagation clustering algorithm.

## 21.2 Affinity Propagation

AP algorithm [1] is a new algorithm by B. Frey from Toronto University. The main characteristic of AP is that it makes all the data points as the initial exemplars, so that it can avoid the situation that clustering result is influenced by the choice of the initial selection of exemplars.

AP sets up a collection of real-valued similarities for all the points on how to attract others, which is the similarity metric between any of the two points  $x_i$  and  $x_k$ ; each similarity metric is set to a negative squared error (Euclidean distance):  $s(i,k) = -\|x_i - x_k\|^2$ . AP at first assumes that the possibility of all data points choosing to be the exemplar is the same, which means to set the same P for all  $s(i,i)$ . Meanwhile, Preference eventually decides number of clusters.

AP algorithm propagates two kinds of information between each two data points: the “responsibility”  $r(i,k)$  sent from data point  $x_i$  to data point  $x_k$ , which reflects how well  $x_k$  serves as the exemplar of  $x_i$  considering other potential exemplars for  $x_i$ , and the “availability”  $a(i,k)$  sent from data point  $x_k$  to data point  $x_i$ , which reflects how appropriate  $x_i$  chooses  $x_k$  as its exemplar considering other potential points that may choose  $x_k$  as their exemplars. The information is updated in an iterative way as

$$r(i,k) \leftarrow s(i,k) - \max_{k_s.t.k' \neq k} \left\{ a(i,k') + s(i,k') \right\} \quad (21.1)$$

$$\text{if } i \neq k, a(i,k) \leftarrow \min \left\{ 0, r(k,k) + \sum_{i'.s.t.i' \notin \{i,k\}} \max \left\{ 0, r(i',k) \right\} \right\} \quad (21.2)$$

$$a(k,k) \leftarrow \sum_{i'.s.t.i' \neq k} \max \left\{ 0, r(i',k) \right\} \quad (21.3)$$

The parameter  $\lambda$ , known as the damping factor, is also introduced in the information updating of AP,  $\lambda \in [0,1)$ . The default value of  $\lambda$  is 0.5. Damping factor could help AP to avoid shocking. Shocking can be avoided by enlarging the damping factor.



### 21.3 Density-Sensitive Measurement

In the processing of data propagation, two types of consistency can be found [7], which are also in accordance with the semi-supervised learning consistency of the data in a priori assumptions:

1. Global consistency, the high similarity between the data points on the same manifold
2. Local consistency, the high similarity between the neighboring data points

According to the analysis of data's space distribution, the data from the same clustering tend to be in a dense manifold area; however, the data from different clustering tend to be in the sparse area [8]. Generally, we can design a new similarity measure based on local data density characteristics. From two aspects, we introduced a kind of density-sensitive measure to improve similarity measure of AP algorithm.

According to the global consistency, data points in the center or around are more concentrated and more likely to be grouped into one type when clustering, so shortening their distance is needed. On the contrary, enlarging the distance should be applied to the data points which are far from the center. So we accordingly devise new measures, defined as follows:

$$DS(x_i, x_j) = \rho^{[D(x_i, x_j)/\theta_{ij}]} - 1 \quad (21.4)$$

Among the types above,  $D(x_i, x_j) = \frac{d(x_i, x_j)}{\max(d(x_i', x_j'))}$  is the normalized Euclidean distance. This distance makes evaluation results of different scales unified to the same scale,  $d(x_i, x_j)$  is the Euclidean distance of data points,  $\max(d(x_i', x_j'))$  is the MAX distance of data points, and we can set the  $\theta_{ij}$  and adjust the  $D(x_i, x_j)$  amplitude to more effectively reflect the reality of data distribution,  $\rho > 1$ , known as the damping factor.

According to local manifold properties, through the appropriate transforming of local comparability measurement between data points, we can convert the points on the same manifold into hyperellipsoidal or hypersphere convex data, so that the distribution characteristics of the data can be more precisely expressed. As a result, the clustering accuracy of AP is improved. Thus we can define a local manifold distance metric as follows:

$$DS(x_i, x_j) = \rho \left[ \frac{\epsilon}{\theta_{ij} \cdot \max(d(x_i, x_j))} \right] - 1 \quad (21.5)$$

The data points  $x_i$  and  $x_j$  are about  $\epsilon$  density that are linked together and on the same manifold in the data set  $X$  and  $x_i \neq x_j$ . When the value of  $\epsilon$  is small, which

leads the class number a few more, then we can adjust  $\rho$  to achieve better clustering results. So we can transform data points on the same manifold through the transformation to the approximation for the center with  $x_i$  hypersphere; of course, it is also easier to AP clustering algorithm identification to improve the algorithm accuracy.

According to the space distribution of data, data belonging to the same propagation group stay in a high dense area; however, data belonging to different propagation groups stay in sparse areas. Measurement of density sensitiveness can enlarge the distance of data points in different areas and reduce the distance of data points in the same area.

## 21.4 Density-Sensitive Semi-supervised AP

In the reality, it is easier to get pairwise constraints compared with labeling the unlabelled patterns. We can take some labeled data for semi-supervised clustering. Semi-supervised AP adjusts the similarity matrix using the priori known labeled data or pairwise constraints. The principle of pairwise constraints is that when the constraint points  $\{(x_i, x_j)\} \in M$ , we think the two points have high similarity, so we adjust the  $s(i, j) = 0$  and they belong to a same cluster; when the constraint points  $\{(x_i, x_j)\} \in C$ , we think two points have low similarity, then we adjust the  $s(i, j) = -\infty$  and divided them into different cluster. The details are as follows:

$$(x_i, x_j) \in M \Rightarrow s(i, j) = 0 \ \& \ s(j, i) = 0 \quad (21.6)$$

$$(x_i, x_k) \notin M \ \& \ (x_i, x_j) \in M \ \& \ (x_j, x_k) \in M \Rightarrow \\ s(i, k) = 0 \ \& \ s(k, i) = 0 \ \& \ M = (x_i, x_k) \cup M \quad (21.7)$$

$$(x_i, x_j) \in C \Rightarrow s(i, j) = -\infty \ \& \ s(j, i) = -\infty \quad (21.8)$$

$$(x_i, x_j) \notin \{M \cup C\} \ \& \ (x_i, x_k) \in C \ \& \ (x_k, x_j) \in M \Rightarrow \\ s(i, j) = -\infty \ \& \ s(j, i) = -\infty \quad (21.9)$$

The capacity for pairwise constraints helping the unsupervised clustering algorithms is limited; we should also utilize the priori knowledge belonging to data set to assist clustering. In this chapter, we use semi-supervised pairwise constraints to adjust data points to optimize the similarity matrix. Furthermore, we introduce density-sensitive measure to improve the data metric matrix between the points; therefore, the more closely the distribution of the various points within the same category, the more decentralized for various points between the different classes; so as to achieve better clustering results, an improved algorithm is proposed as follows:

*Algorithm:* Density-sensitive semi-supervised affinity propagation clustering algorithm (DS-SAP)

1.  $\forall x_i, x_j \in X$ , calculate the Euclidean distance between two points:

$$D_{ij} = \left( \|x_i - x_j\|^2 \right)^{\frac{1}{2}} \quad (21.10)$$

2. Add pairs limit information based on a priori pairwise constraints:

$$\begin{cases} D_{ij}, D_{ij} = 0, \text{ if } (x_i, x_j) \in ML \\ D_{ij}, D_{ij} = -\infty, \text{ if } (x_i, x_j) \in CL \end{cases} \quad (21.11)$$

3.  $\forall x_i, x_j \in X$ , calculated the density-sensitive measure between two points:

$$DS_{ij} = \begin{cases} \rho^{[D(x_i, x_j)/\theta_{ij}] - 1} & x_i, x_j \text{ cannot link to } \varepsilon \\ \rho \left[ \frac{\varepsilon}{\theta_{ij} \bullet \max(d(x_i, x_j))} \right] - 1 & x_i, x_j \text{ link to } \varepsilon \end{cases} \quad (21.12)$$

4. The similarity matrix is configured according to the similarity measure of the density-sensitive  $S \in R^{n \times n}$ :

$$S_{ij} = \frac{1}{DS_{ij} + 1}, S_{ii} = 0 \quad (21.13)$$

5. Based on the similarity matrix S of the above configuration, completed clustering by AP algorithm.

DS-SAP algorithm directly modifies the distance measure based on semi-supervised pairwise constraint a priori information, so the similar relationship will change with the constraint condition. The spatial consistency a priori information adjusts the distance between data points automatically; it plays an indirect role to modify the similarity matrix. Making full use of the two types of a priori information can maximize the guiding role of clustering search.

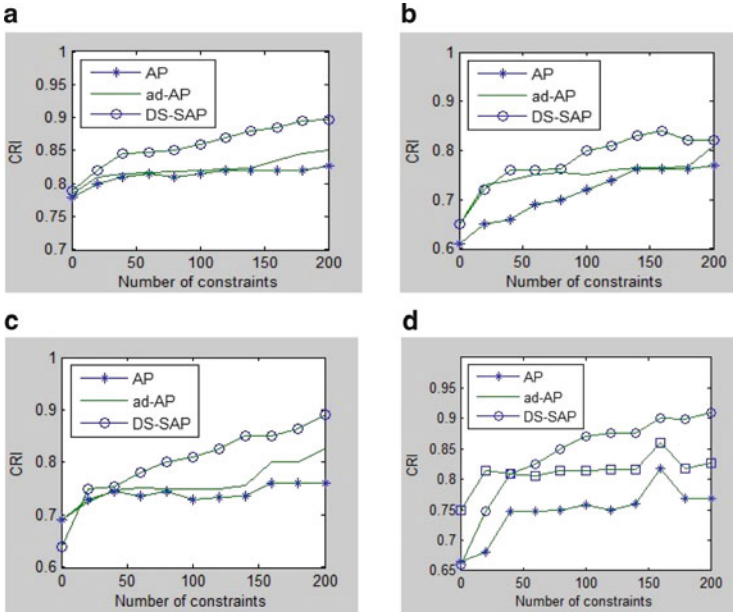
## 21.5 Experiments

The data sets used in this experiment are from UCI, including Iris, Ionosphere, Glass, and Sonar. We compare DS-SAP with ad-AP and AP clustering effects on the above data sets. Meanwhile, the running time of DS-SAP and AP is also included with Circles, Spirals, and Face Image.

Experimental environment: Processor Core2 2.2GHz, memory 2 GB, hard-drive 500G, Windows 7, parameters in the experiment are as: Maxits = 3,000; Conviits = 100; Lam = 0.8; Table 21.1 shows the properties of the dataset:

**Table 21.1** The experiment dataset

	Iris	Ionosphere	Glass	Sonar	Circles	Spirals	Face image
Instance	150	351	214	208	600	1,000	900
Dimension	4	34	9	60	2	2	50 × 50
Class	3	2	6	2	3	2	100



**Fig. 21.1** The CRI index of the three algorithms in the UCI data sets. (a) Iris, (b) Ionosphere, (c) Glass, and (d) Sonar

**Table 21.2** Run-time list unit/s

Algorithm	Circles	Spirals	Face image
DS-SAP	3.06	12.05	16.57
AP	3.25	11.38	17.69

During the experiment, the clustering result is assessed according to CRI:

$$CRI = \frac{\text{correct free decisions}}{\text{total free decisions}}$$

Among the type above,  $\text{total free decisions} = (n \times (n - 1))/2 - C_n$ , where  $n$  is the number of data points,  $C_n$  is the number of the pairwise constraints, and correct free decisions are divided in the correct data to minus the number of constraints to the division for the number of correct data. The experimental results are shown (Fig. 21.1; Table 21.2) as follows:

The experimental results show the DS-SAP clustering accuracy is improved. DS-SAP adjust the similarity matrix using pairwise constraints of

semi-supervised and the density-sensitive distance measure; it can express the nature of the internal structure of the data, so that it spontaneously improves performance of the algorithm. Meanwhile, the time complexity of the proposed DS-SAP algorithm is mainly depended on the similarity matrix and AP algorithm clustering iterative times. Although the building time of the similarity matrix is relatively increased, the similarity matrix is closer to the ideal after optimization. It can greatly reduce the number of iterations of the clustering, so the total time of our algorithm is not increased and the experiments have verified it. In summary, due to the optimization of the similarity matrix, DS-SAP has higher accuracy than AP and ad-AP, the running time of the algorithm generally flat or even less than the AP.

## 21.6 Conclusion

This chapter put forward an improved framework of affinity propagation based on semi-supervised (DS-SAP). In this algorithm we used semi-supervised pairwise constraints and introduced density-sensitive measure to improve the data metric matrix between the points, so as to achieve better clustering results. The experimental results on the four data sets of UCI and other three data sets showed that the algorithm clustering performance was improved.

**Acknowledgements** The project is supported by the National Natural Science Foundation of China (No. 61073121), National Science and Technology Support Plan Project (No. 2013BAK07B04), Natural Science Foundation of Hebei Province, China (No. F2013201170) and Medical Engineering Alternate Research Center Open Foundation of Hebei University (No. BM201102).

## References

1. Frey, B. J., & Dueck, D. (2007). Clustering by passing messages between data points. *Science*, 315(5814), 972–976.
2. Givoni, I. E., & Frey, B. J. (2009). A binary variable model for affinity propagation. *Neural Computation*, 21(6), 1589–1600.
3. Weiming, L., Chenyang, D., Baogang, W., Chunhui, S., & Zhenchao, Y. (2012). Distributed affinity propagation clustering based on map reduce. *Journal of Computer Research and Development*, 49(8), 1762–1772.
4. Yan, X., & Yong, Z. (2012). Affinity propagation based on K-mutual nearest neighbor consistency. *Application Research of Computers*, 29(7), 2524–2526.
5. Jun, D., Suoping, W., & Fanlun, X. (2010). Affinity propagation clustering based on variable-similarity measure. *Journal of Electronics and Information Technology*, 32(3), 509–514.
6. Akl, A., & Valae, S. (2010). *Accelerometer-based gesture recognition via dynamic-time warping, affinity propagation, and compressive sensing* (pp. 2270–2273). IEEE International Conference on Acoustics Speech and Signal Processing, Dallas, TX.

7. Zhai, D., Chang, H., Shan, S., Chen, X., & Gao, W. (2012). Multiview metric learning with global consistency and local smoothness. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 3(3), 53.1–53.22.
8. Gui, J., Zhao, Z., Hu, R., & Jia, W. (2013). Semi-supervised learning with local and global consistency by geodesic distance and sparse representation. In *Intelligent science and intelligent data engineering* (Vol. 7751, pp. 125–132). Berlin: Springer.

# Chapter 22

## The Implementation of a Hybrid Particle Swarm Optimization Algorithm Based on Three-Level Parallel Model

Yi Xiao and Yu Liu

**Abstract** In order to improve the efficiency of hybrid particle swarm optimization (PSO) algorithm, a PSO merging simulated annealing and hill climbing (SAHCPSO) is implemented based on a three-level parallel model to increase its convergence speed and to decrease the operation time. SAHCPSO can enhance the diversity of the population and avoid population premature convergence. By analyzing and optimizing the SAHCPSO, we complete the task mapping on the model and make full use of CPU/GPU heterogeneous cluster resources. Optimization for parallel accessing further improves the efficiency of the algorithm. The parallel SAHCPSO implements the coarse-grained parallelism between computation nodes and fine-grained parallelism within each node, greatly reducing the operation time. The experimental results show that with the increase of particle scale, higher speedup can be obtained. The high efficiency of the parallel strategy of the model makes the parallel SAHCPSO more easily to solve large-scale problems.

### 22.1 Introduction

The particle swarm optimization algorithm (PSO) is an evolutionary intelligent algorithm that was first proposed by Kennedy and Eberhart [1]. The idea comes from birds' predation. The advantages of PSO are less parameters and easy to implement. Many significant problems in the field of scientific research and engineering applications can be solved by using PSO. But PSO is easy to premature

---

Y. Xiao (✉)  
College of Information Science and Engineering, Guilin University of Technology,  
Guilin 541004, China  
e-mail: [louisxcode@yahoo.com](mailto:louisxcode@yahoo.com)

Y. Liu  
College of Mechanical and Control Engineering, Guilin University of Technology,  
Guilin 541004, China

convergence, and it lacks local search capability. A particle swarm algorithm based on dynamic inertia weight vector and dimension mutation was proposed to balance the local and global searching of particle swarm and increase the speed of convergence [2], but it still takes a long time to compute when solving large-scale optimization problems. In the field of scientific and engineering computing, parallel computing has become an important method to solve large-scale complex problems. With the development of science and technology, CPU/GPU heterogeneous architecture is gradually becoming an important parallel computing platform [3]. We apply a three-level parallel computing model to a hybrid particle swarm optimization algorithm merging simulated annealing and hill climbing (SAHCPSO), which can enhance the diversity of the population, avoid population premature convergence, and improve the efficiency of operations [4].

## 22.2 Three-Level Parallel Computing Model

### 22.2.1 *MPI, OpenMP, and CUDA*

MPI (Message Passing Interface) is a standard of message-passing system. It is a library and a language-independent communication protocol used to write parallel programs, and in fact, it has become a current mainstream programming model in distributed storage architecture. OpenMP (Open Multi-Processing) defines a collection which contains a set of compile-guide statements, runtime library functions, and environment variables. OpenMP completely describes the parallelism architecture and fully uses the advantage of shared memory architecture. It avoids the overhead of message passing and provides runtime scheduling, fine-grained operation mechanism, and coarse-grained operation mechanism. CUDA (Compute Unified Device Architecture) is a general purpose parallel computing architecture—with a new parallel programming model and instruction set architecture which leverages the parallel compute engine in both CPU and GPU to solve many complex computational problems in a more efficient way [5].

### 22.2.2 *Mixed Parallel Computing*

The three-level parallel computing model is based on TC (Thread Communication) mixed mode of MPI + OpenMP parallel model. In this model, CPU creates threads to control GPU running kernel function [6, 7]. The parallel model is shown in Fig. 22.1.

According to the parallel model, parallel algorithm use GPU as the central computing component. On the cluster, program processes are divided into subtasks by main thread. Main thread using message-passing function sends subtasks to all computing nodes. After each of the computing nodes receiving the task computing threads for logic operation, communication threads and control threads are created



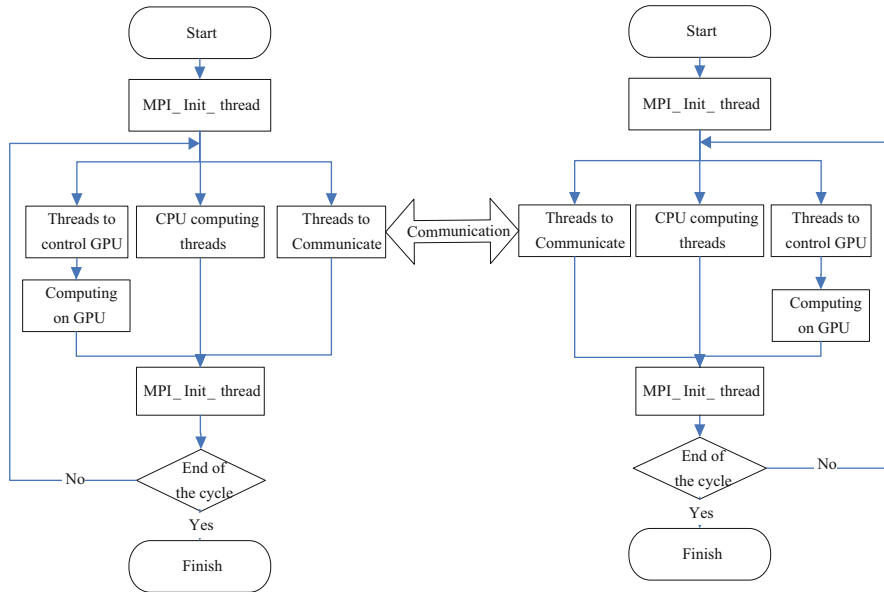


Fig. 22.1 Three-level parallel computing model

by OpenMP. Communication threads are used to communicate with other nodes through MPI function. Control threads are created to run CUDA kernel function [8].

### 22.3 Algorithm Parallelization Solving

#### 22.3.1 Analysis of SAHCPSO

SAHCPSO program will repeatedly execute the algorithm and calculate the correct result which is equal to the optimal value. The number of correct results determines the ability of convergence. SAHCPSO is divided into four stages: initialization of particles information, mutation of particles information, updating of particles information, and searching for the group best fitness.

Stage 1: SAHCPSO initializes the position ( $pos$ ) and velocity ( $v$ ) of all particles. For each particle, fitness can be calculated by using particle position information ( $x_i$ ) as follow:

$$fitness = \sum_{i=1}^n [x_i^2 - 10 \cos(2\pi x_i) + 10] \tag{22.1}$$

Each particle calculates their fitness value and saves as personal best value ( $p-Best$ ). Then SAHCPSO compares all particles' p-Best to find the optimal value which saves as group best value ( $g-Best$ ). When SAHCPSO update the

**Table 22.1** The proportion of each part of the serial algorithm running time

Stage	Initialization	Mutation	Updating	Searching
Proportion	1.3 %	39.39 %	43.56 %	4.2 %

value of  $p$ -Best and  $g$ -Best, corresponding particle position information will be saved. In order to avoid aimless search, position and velocity information will be limited to a definite range.

Stage 2: In this stage SAHCPSO create mutational position information based on the current particle position information according to Eq. (22.2). Where  $pos'$  is mutational position information, *Gaussian* is a Gaussian random number.

$$pos' = pos + (1 + Gaussian) \quad (22.2)$$

The mutation probability of each dimension of position information is 10 %. For each particle, fitness can be calculated by using mutational position information according to Eq. (22.1). If the fitness is better than the corresponding particle current fitness value, update the particle fitness and  $p$ -Best. Otherwise, the particle current fitness will be forced to updating the fitness and  $g$ -Best with the probability of 1 and 5 %.

Stage 3: SAHCPSO starts updating the particles information. The position and velocity information of all particles are updated according to Eqs. (22.3) and (22.4). Where  $w, r1, r2, c1, c2$  is algorithm parameters,  $k$  is the current number of iterations,  $p$ -BestPos and  $g$ -BestPos is the position information of current  $p$ -Best and  $g$ -Best.

$$v_i^{k+1} = wv_i^k + c_1r_1(p - BestPos_i^k - x_i^k) + c_2r_2(g - BestPos^k - x_i^k) \quad (22.3)$$

$$x_i^{k+1} = x_i^k + v_i^{k+1} \quad (22.4)$$

Stage 4: When  $p$ -Best has been updated, SAHCPSO compares all particles'  $p$ -Best and searches the optimal value saving as group best value ( $g$ -Best).

According to Amdahl's law, the impact of the overall system performance depends on proportion and acceleration efficiency of the program which can be optimized [9]. Table 22.1 shows the test of the original serial SAHCPSO for different stages of the average of the proportion of time in different iterations. Therefore, the stages we need to parallelize are particles information mutation and updating.

### 22.3.2 Parallel Implementation of SAHCPSO

For the characteristics of the procedure, there are two parallel strategies in the parallel computing of SAHCPSO on three-level parallel computing model:

1. Parallelization between the repeated algorithm processes of program
2. Parallelization between the calculations of each particle

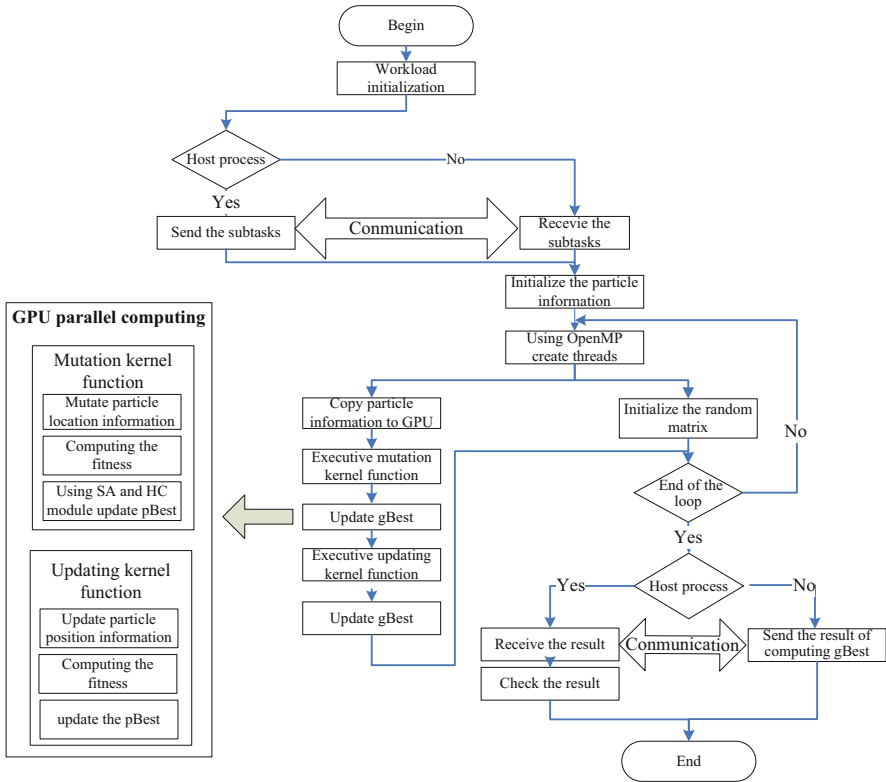


Fig. 22.2 The flow chart of the parallel SAHCPSO

According to the three-level parallel computing model, the repeated algorithm processes of program are mapped to cluster nodes which are coarse-grained parallelism; the processes of particle information calculation are mapped to a single node which uses the CPU and GPU to achieve fine-grained parallel computing. GPU threads are responsible for calculating mutation of particles information and updating of particle information. The random number which GPU required is calculated by CPU threads. For the lack of global synchronization mechanism, g-Best cannot be updated on GPU and the stages of mutation and updating have to be divided into two kernel functions.

The parallel SAHCPSO flow chart is shown in Fig. 22.2. In order to improve the speed of GPU parallel accessing, data structure of the parallel algorithm consists of a series of particles information matrixes and random matrixes which can provide data accessing patterns for memory coalescing [10]. Take particle position information matrix, for example, the number of particles is  $p\text{-Count}$  and the dimension is  $dim$ . Then we can get the particle position information matrix:  $PSO [p\text{-Count}] [dim]$ .

The contents of random number matrixes are prejudice on CPU according to the high speed of calculation for integer data on GPU. When a single-precision float

random number is generated, CPU compares the random number with the corresponding probability parameters and stores a specific integer number to random number matrix.

### 22.4 Experimental Results

The experiments are performed on cluster which has four nodes with one Intel Core i3 3.3GHz processor connecting to windows XP operating system (2 GB of memory). Each computing node is equipped with a NVIDIA GT405 card. The bandwidth of cluster network is 100 Mbps.

We execute the original serial SAHCPSO and improve parallel algorithm under different particle scale ( $p$ -Count) and iterations to test the impact of different conditions on execution time. Figure 22.3 shows the performance of serial algorithm and parallel algorithm computing on 2000 and 5000  $p$ -Count. With the expansion of the particles scale, the growth of the time also gradually expands. It is obvious that the parallel algorithm can greatly reduce the growth of time.

Speedup refers to the ratio of the time required by the serial algorithm running on the single node and the parallel algorithm running on the cluster. Figure 22.4 shows the test of speedup under different  $p$ -Count in the case of 2,000 iterations. With the expansion of particles scale, the growth in the time of the serial algorithm is greater than the parallel algorithm. Therefore, the speedup increases gradually.

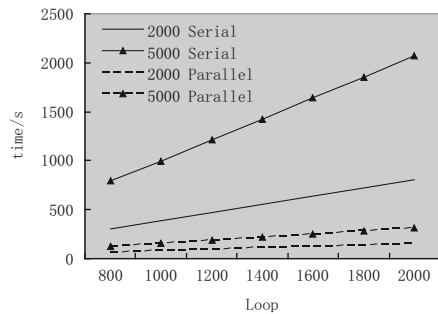


Fig. 22.3 The effect of different particle scale on execution time

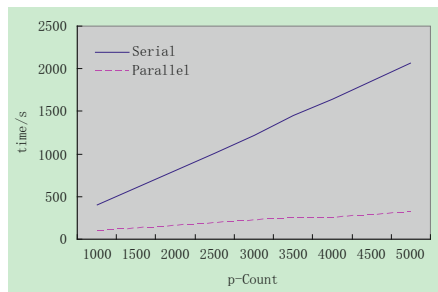


Fig. 22.4 Speedup of the different particle scale

## 22.5 Conclusion

The paper implemented the SAHCPSO parallel computation on a three-level parallel computing model. Practice proved that this hybrid model can make full use of CPU/GPU heterogeneous cluster resources. By analyzing the characteristics of the serial algorithm, parallelism embedded in SAHCPSO can be excavated. The calculation of particles' fitness can be mapped to GPU and CPU threads. The repeat of calculation result can be mapped to computing node processes. Therefore, most of SAHCPSO calculations can be executed in parallel. Optimization for parallel accessing further improves the efficiency of the algorithm. The results showed that the parallel algorithm is correct and efficient. The speedup we have achieved is up to 6.6. In addition, because the model has a good scalability, larger particles scale can be computed.

**Acknowledgements** This work has been supported by The National Natural Science Foundation of China under research project 41264005 and also supported by The Guangxi department of education under the research project 201102ZD018. We also thank master degree candidate Jiaxing You for his generous help to our work especially in SAHCPSO programming.

## References

1. Kennedy, J., & Eberhart R. (1995). *Particle swarm optimization* (pp. 1942–1948). IEEE International Conference on Neural Networks, IEEE, Perth.
2. Liang, X. M., Dong, S. H., Long, W., & Xiao X. F. (2011). PSO algorithm with dynamical inertial weight vector and dimension mutation. *Computer Engineering and Applications*, 47(5), 29–31 (In Chinese).
3. Tian, G., & Lu. F. S. (2011). Measuring MPI/OpenMP+CUDA high-performance computing environment configuration and application. *Silicon Valley*, 17(9), 118–119 (In Chinese).
4. Li, J. M., Wan, D. L., Chi, Z. X., & Hu X. P. (2006). A parallel particle swarm optimization algorithm based on fine-grained model with GPU-accelerating. *Journal of Harbin Institute of Technology*, 12(38), 2162–2836 (In Chinese).
5. Xu, Q. Y., & Chen Q. K. (2010). Research and implementation of MPI+CUDA model based on SMP clusters. *Computer Engineering and Design*, 15(31), 3408–3412 (In Chinese).
6. Liu, Q. K., Ma, M. W., & Yan W. C. (2011). Parallel matrix multiplication based on MPI + CUDA asynchronous model. *Journal of Computer Applications*, 12(31), 3327–3330 (In Chinese).
7. Teng, R. D., & Liu Q. K. (2010). Mixed CUDA, MPI and OpenMP in three mode parallel programming. *Microcomputer Applications*, 9(31), 63–69 (In Chinese).
8. Sanders, J., & Kandrot E. (2011). *CUDA by example: An introduction to general-purpose GPU programming* (pp. 162–165). NJ: Addison-Wesley.
9. Pacheco P. S. (2011). *An introduction to parallel programming* (pp. 60–62). Burlington: Morgan Kaufmann.
10. Kirk, D. B., & Hwu W. W. (2012). *Programming massively parallel processors: A hands-on approach* (2nd ed., pp. 135–137). Burlington: Morgan Kaufmann.

# Chapter 23

## Optimization of Inverse Planning Based on an Improved Non-dominated Neighbor-Based Selection in Intensity Modulated Radiation Therapy

Xiao Zhang, Guoli Li, and Zhizhong Li

**Abstract** Intensity modulated radiation therapy (IMRT) is a principal cancer treatment at present, and the optimization of inverse planning is the core to realize the IMRT treatment planning, so it is important to study how to optimize the inverse planning as much as possible. The optimization of inverse planning in IMRT refers to a number of parameters and requires rapid calculation speed clinically, so an improved NNIA for multi-objective is adopted in this paper. At first, according to the IMRT dose constraints of multiple targets, an average dose-based function is used, and then the optimization results compared with the SAGA algorithm under the water phantom show the feasibility and efficiency of the improved NNIA algorithm.

### 23.1 Introduction

Cancer has become the main disease in the world [1]. Intensity modulated radiation therapy (IMRT) as a precise radiotherapy technology is an effective method to treat cancer. The key step for IMRT is the inverse planning, and the purpose of IMRT inverse planning is to determine the ray type and energy, beam orientation and number, and so on [2]. Because of the complexity of the implementation for IMRT inverse planning, it usually needs computer-aided inverse planning to solve the problem. The optimization of inverse planning is the core to realize the IMRT treatment planning, so it has important meaning to research on this problem.

The optimization of inverse planning in IMRT refers to a number of parameters; therefore, it has multi-objective characteristic. But the heart of multi-objective optimization is to solve the contradiction between convergence speed and population diversity. According to the study of the optimization algorithms in IMRT

---

X. Zhang (✉) • G. Li • Z. Li  
College of Information Engineering, Zhejiang University of Technology,  
Hangzhou 310023, China  
e-mail: [tmny@zjut.edu.cn](mailto:tmny@zjut.edu.cn)

inverse planning, the optimization algorithms can be classified as deterministic algorithms and multi-objective evolutionary algorithms. The deterministic algorithms dominated by gradient method are with the advantage of rapid convergence, but may fall into the local minimum for non-convex problems. The multi-objective evolutionary algorithms can escape local minimum theoretically, but their calculating speed usually is relatively slower than deterministic algorithms.

In our previous study, several algorithms are used for the optimization of IMRT inverse planning, such as hybrid multi-objective gradient algorithm [3], multi-objective hybrid genetic approach based on simulated annealing (SAGA) [4], and improved fast non-dominated sorting genetic algorithm (NSGA-II) [5].

In this paper, in order to solve the problem of convergence speed for better, an improved multi-objective immune algorithm with non-dominated neighbor-based selection (NNIA) is applied to the optimization of IMRT inverse planning; the improved NNIA adopts the nonuniform mutation operator making the search step length adjusted adaptively during the different stages of evolution. Finally, an irregular field of segment at deep of 1.5 cm under the condition of 30 cm × 30 cm × 30 cm water phantom is used as the simulation case, through the comparison with SAGA algorithm; the feasibility and efficiency of the improved NNIA algorithm for IMRT inverse panning is to be proved.

## 23.2 Optimization Algorithm

Immune algorithm (IA) is a multi-objective evolutionary algorithm based on artificial immune system (AIS), in which the antigen and antibody are respectively defined as the optimization problem and the solution of problem; through immune operations, the antibodies will evolve; after evaluating the affinity between antigens and antibodies, the Pareto solutions are obtained.

In this paper, an improved NNIA has been adopted, which has rapid convergence speed and can effectively maintain the diversity of population.

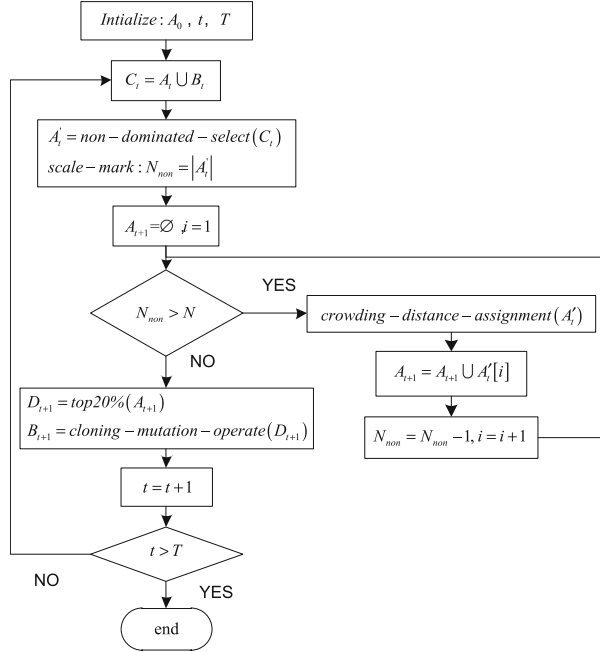
This improved NNIA is implemented by using the non-dominated neighbor-based selection, proportional cloning operator [6], nonuniform mutation operator, and elitism.

For the non-dominated neighbor-based selection, the calculation of crowding distance is a key problem. Through calculating the crowding distance of an individual, its probability of being selected will be determined. The crowding distance of an individual  $d$  belonged to the population  $D$  can be expressed as

$$I(d, D) = \sum_{i=1}^m \frac{I_i(d, D)}{f_i^{\max} - f_i^{\min}} \quad (23.1)$$

where  $f_i^{\max}$  and  $f_i^{\min}$  are the maximum and minimum for the  $i$ th subgoal.

**Fig. 23.1** The flow chart of improved NNIA



For the nonuniform mutation operator, the results of mutation operator are associated with its iteration  $t$ . The range of variation is relatively large in an early phase, and with the advance of the evolution, the range of variation becomes smaller and smaller, mainly for local search. The variable  $x_k \in [a_k, b_k]$  after mutation operation can be determined by following expressions:

$$x'_k = \begin{cases} x_k + \Delta(t, b_k - x_k), & \text{rnd}(0, 1) = 0 \\ x_k + \Delta(t, x_k - a_k), & \text{rnd}(0, 1) = 1 \end{cases} \quad (23.2)$$

where  $\Delta(t, y) = y(1 - r^{(1-t/T)^\lambda})$ ;  $\lambda$  is a unified parameter, it determines the dependence of random disturbance on the iteration  $t$ ;  $r \in [0, y]$ . The flow chart of improved NNIA algorithm is shown in Fig. 23.1.

### 23.3 The Objective Function

There are two main categories of objective functions: the model based on dose constraint and the model based on dose-volume constraint. For the objective function based on dose constraint, it has a simple concept of optimization and is easier to be used; for the objective function based on dose-volume constraint, it has



definite physical significances and its application is more flexible. But the form of the objective function can have a significant impact on the speed and effect of optimization.

In our study, an average dose-based function for the planning target volume (PTV) and the surrounding normal tissue (NT) are used in this paper; the mathematical expression is as follows:

$$\begin{aligned} y_1 = f_{PTV} &= \frac{1}{N_{PTV}} \sum_{i=1}^{N_{PTV}} \left( d_i^{PTV} - D_{ref}^{PTV} \right)^2 \\ y_2 = f_{NT} &= \frac{1}{N_{NT}} \sum_{i=1}^{N_{NT}} \left( d_i^{NT} - D_{ref}^{NT} \right)^2 \end{aligned} \quad (23.3)$$

where  $N^{PTV}$  and  $N^{NT}$  are the corresponding numbers of sampling points for PTV and NT, respectively.  $d_i$  is the calculated dose value at the  $i$ th sampling point.  $D_{ref}^{PTV}$  and  $D_{ref}^{NT}$  are the average prescription doses for the PTV and NT, respectively.

The fitness functions are expressed as follows:

$$fitness_i = \frac{1}{1 + y_i}, i = 1, 2 \quad (23.4)$$

The dose calculation method among inverse planning is a finite-size pencil beam (fsPB) model based on Monte Carlo (MC) [7], which comes from the previous research of our group.

## 23.4 Optimization Result

The simulation case is an irregular field of segment at deep of 1.5 cm under the condition of 30 cm × 30 cm × 30 cm water phantom. The slice of the irregular field is shown in Fig. 23.2. The part of shadow is PTV, and the other is NT, in which  $N_{PTV} = 13$  and  $N_{NT} = 87$ .

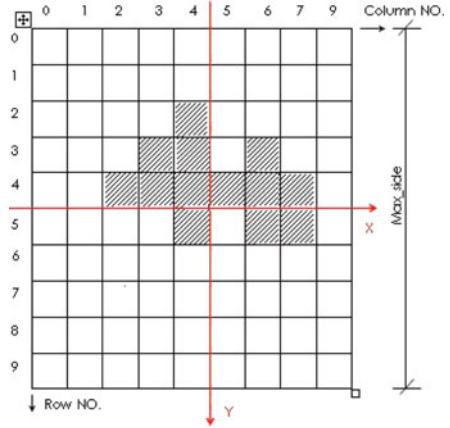
The optimization parameters are set as Table 23.1.

We also adopt SAGA algorithm which has studied in early stage for comparison; it uses the similar input condition, and the annealing conditions are set as initial temperature is 7 and random seed is 0.7.

The results of these two algorithms are shown as Table 23.2.

The intensity distribution maps of improved NNIA and SAGA are shown as Figs. 23.3 and 23.4.

**Fig. 23.2** The slice of the irregular field



**Table 23.1** Optimization parameters

Parameter name	Parameter value
$D_{ref}^{PTV}$	0.922
$D_{ref}^{NT}$	0.013
$fitness_i (i = 1, 2)$	$> 0.982$
Cloning scale	100
Population	100
Generations	500

**Table 23.2** Optimization results

Algorithm	$fitness_1$	$fitness_2$	Time of calculation
Improved NNIA	0.9972	0.9893	64s857 ms
SAGA	0.9899	0.9656	87s253 ms

**Fig. 23.3** Improved NNIA

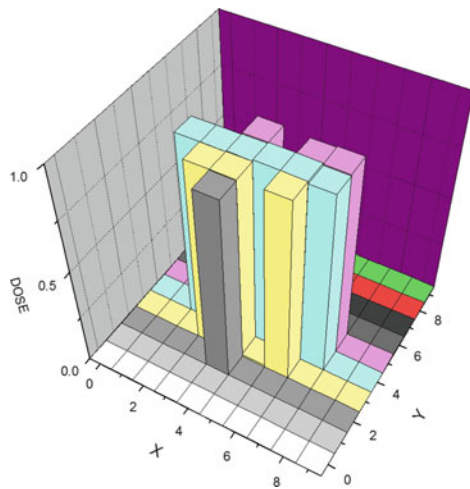
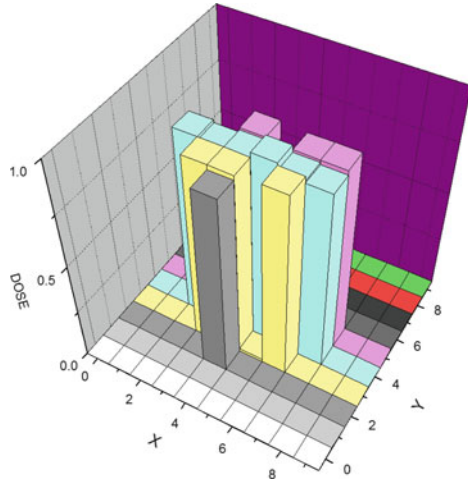


Fig. 23.4 SAGA



## 23.5 Conclusion

Through comparing the intensity distribution maps of Fig. 23.3 with that of Fig. 23.4 and judging the optimization results of two algorithms, it can be found that the improved NNIA can obtain solutions in a relatively short time, and the conformity of intensity distribution map is better. Therefore, a conclusion can be drawn: the improved NNIA is an efficient optimization algorithm for IMRT inverse planning. But when the algorithm is used in clinic, many practical problems need to be considered, such as the choice of treatment angle. Therefore the focus of our future work is how to apply this algorithm to clinic.

## References

1. Zheng, J., Chen, D., Yu, X., & Tian, X. (2013). Investigation of the cancer-related fatigue and quality of life in cancer patients with chemotherapy. *Shanxi Medical Journal*, 42(2), 141–143.
2. Hu, Y. (2005). Advances of intensity modulation therapy. *China Medical Device Information*, 11(2), 1–5.
3. Wang, Z. (2008). *Application of optimization with scaled conjugate gradient algorithm in IMRT* (pp. 751–754). In Proceedings of the 2nd International Conference on Bioinformatics and Biomedical Engineering, Shanghai.
4. Li, G., Li, Z., & Lin, L. (2009). *Multi-objective optimization for irregular field in IMRT based on hybrid SA algorithm* (pp. 1–4). In Proceedings of the International Conference on Information Engineering and Computer Science, Wuhan.
5. Lin, L. (2009). *An improved NSGA-II algorithm for the optimization of IMRT inverse planning* (pp. 1–3). In Proceedings of the 2nd International Conference on Biomedical Engineering and Informatics, Hangzhou.

6. Gong, M., & Jiao, L. (2008). Multi-objective immune algorithm with non-dominated neighbor-based selection. *Evolutionary Computation*, 16(2), 225–255.
7. Zhou, J. (2009). A fast size pencil beam algorithm based on Monte Carlo simulation. *Journal of University Science and Technology of China*, 39(5), 515–519.

# Chapter 24

## A Recommendation System for Paper Submission Based on Vertical Search Engine

Zhen Xu, Yi Yang, Fei Wang, Jiao Xu, Zhong Li, Fuqiang Mu, and Lian Li

**Abstract** In this work, the proposed orchestrating and sharing system for online paper aims at managing papers from information collecting, paper editing, paper type-setting, and paper submitting to paper sharing. In the five aspects above, there are many available tools which help science researchers write papers, but these tools work separately not cooperatively. Orchestrating and sharing system for online paper integrates functions of these tools, which offers one-stop service. As an important part of this system, the recommendation for paper submission is to provide valuable information about the latest international conferences and journal for paper publication. When papers are written, our system, a context-aware solution for paper, automatically obtains the keywords from context. Given that the recommendation for paper submission is subject-oriented search, we design a recommendation system for paper submission based on vertical search engine, which enhances the search accuracy by the improved URL-based filtering algorithm and the improved content-based filtering algorithm.

### 24.1 Introduction

PapersCloud [1] is an online paper orchestrating and sharing system that supports the whole life cycle of science papers (briefly called paper in this paper below). The basic meaning of the life cycle can be popularly understood as “the whole process from the cradle to the grave” (Cradle-to-Grave). In accordance with the definition of the life cycle, we propose a paper life cycle approach. Paper life cycle

---

Z. Xu • Y. Yang (✉) • F. Wang • J. Xu • L. Li  
School of Information Science and Engineering, Lanzhou University,  
Lanzhou 730000, China  
e-mail: [xuzhen1109@163.com](mailto:xuzhen1109@163.com); [yy@lzu.edu.cn](mailto:yy@lzu.edu.cn)

Z. Li • F. Mu  
Gansu Wanwei Company, Lanzhou 730000, China

is the process that from collection information, paper editing, paper type-setting, paper submission to paper sharing. In our system, paper life cycle includes references recommendation, paper editing, paper type-setting, paper submission recommendation, and paper sharing management. References recommendation is to offer some related papers for the interesting topics. Paper editing and type-setting get the input of the paper and type-setting according to a certain format. Paper submission recommendation gives a list of institutes for submitting the paper for publication. The last is to manage the available papers remotely. This chapter mainly aims at papers submission recommendation.

Paper submission recommendation gives information about paper publication for the authors after their accomplishment of their papers. How to always get better result is the main part of this paper. We can get the research field and the keywords from the paper, and we search the most related information for the author from our recommendation library. According to the preference of every author, we give out the result in different ways: for new authors we give out the most related and for old users we just give out the recently information such as the main dates of their preferences.

## 24.2 System Overview

This system is mainly designed to provide information of conferences and periodicals when users want to submit their papers. And the system continues to use the classical framework, it consists of web spider module, dumper module, index module, and query module. The main process of the system is shown as following Fig. 24.1. For instance, when user finishes a paper about cloud computing, our system will automatically obtains the keywords from context. When system obtains the keywords, web spiders crawl the web pages form Internet, after that we filter some useless pages and becoming the web page database. Then web page go through information extraction, calculation of the PageRank value, etc., and after these steps, data can be indexing. The last step is similarity calculation, and the query result will be shown to the user. In next, we will introduce the four modules in detail.

The first module of search engine is web spider module. It is the foundation of the search engine. It is in charge of downloading useful web pages from the World Wide Web. All search data are derived from the work of web spider module.

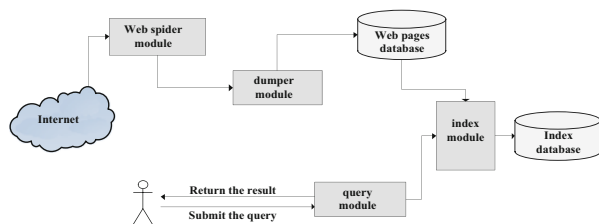
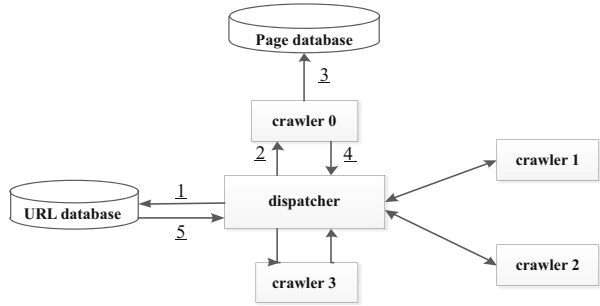
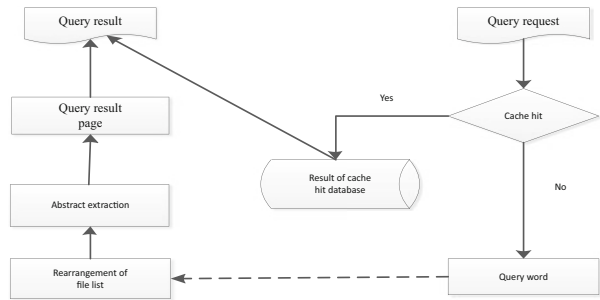


Fig. 24.1 The main process of the system

**Fig. 24.2** The crawlers working process



**Fig. 24.3** The process of query module



This is a huge project because there are tens of thousands of web pages [2] on the Internet. World Wide Web has bow-tie structure. Because of this, we select directory-type Website as the crawler start page. Figure 24.2 shows how crawlers work.

In Fig. 24.3, we can see crawler get URL from the URL database and then download the web page, and after the filter, useless pages form the page database.

In this module, our main consideration is the selection of the URL library. The choice of text search engine is portals and directory-type site. As for this, it get much more information about subsites or related sites of the portal. Thus, there is much more unconcerned information it get. Because the selection of URL library influences the search result to a large extent and we pay more attention on the search for international conferences and journals, we use both method of URL setting and portals for the crawler. To set the certain sites, we can get more accurate information than we get from the portals; on the contrary, we can get entire information from the method of portals.

The second module is dumper module. For dumper module, it basic and primary work is categorized extract [3] valuable information from the semi-structured page. And this information can represent the attributes of the page, such as anchor text, title, and content.

Because the information we get through the crawler are complicated and uncertain, we need to eliminate the unconcerned information and put the rest to our recommendation library. As to dumper module, it does the work to search results according to the conditions of the authors such as deadline of paper

submission, journals or conferences, SCI or EI, and the rest from our recommendation library.

The third one is index module. The index module is the search engine's data warehouse; it stores and indexes millions of thousands of web pages. The purpose of indexing the web page is convenient for querying in the next stage. We need to index the page crawl from the web that can accelerate the speed of query.

In this module, we will first progress the Chinese word segmentation. It is mainly to segment the sentence into a collection with suitable word. Then, we will calculate the PageRank value. The result of offline calculation will be returned a list of PageRank, including a PageRank value of every page and will be easily retrieved in the query module. At last, we will index the pages.

The last module is query module. Query module directly faces the users. It receives the query request by online users and gives the user result in accordance with the calculation by retrieving, sorting, and abstracting extract, and so on.

Figure 24.3 shows the process of query module. Firstly, the system receives the query request from the user and then compares the request with the cache hit. If the request is in it, we directly show the result to the user, and if not we query the word from index module. When the index module returns the result, rearrange the file and extract the abstract, form the result page to user. The whole query requirements not only are faster but also are able to provide users with available results.

## 24.3 Detailed Design

Our system provides the function that can offer some information about recommended conferences or periodicals for users, which requires the system to consider how to control the search results that will not be offset and filter the useless query information. These issues will be exhaustively described in next context.

### 24.3.1 *URL Filter Analysis and Implementation*

Search engine crawlers work mechanism's priority is to grab the web pages which are the highly relevant to the subject. Web pages were sorted according to the page ranking, and only keep themes that are above the URL threshold.

Currently widely used URL filtering algorithm consists of two classes, PageRank algorithm and HITS algorithm [4]. The basic idea of PageRank algorithm [5, 6] is that, web pages from a number of high-quality web links must have the high quality web pages. HITS algorithm bases on the idea that the really value of the page is highly relevant to the theme of the user's search content. HITS algorithm easily occurs that pages deviate from the core theme and irrelevant results returned.



Compared with the two algorithms, PageRank algorithm is in dominant position. So, we selected the PageRank algorithm and improved it. PageRank algorithm is simple, but it ignores the user’s understanding of web page. It gives different web pages with the same weight; this will search high value but have little relationship about the subject.

In order to improve the results, we improved PageRank algorithm by using the user’s visiting navigation path diagram to modify the traditional PageRank value.

Different web page has different probability to be visited by the users, so you can through the web page’s visiting probability express the initial PageRank value. The simple way is Eq. (24.1).  $A_p$  is the number that page p was visited;  $\sum A_p$  is the number that all pages were visited.

$$PR_1(P) = \frac{A_p}{\sum A_p} \tag{24.1}$$

We should consider the actual situation that different users access to different pages and unbalanced to set the current page’s ability to recommend the outlink web page. And combined with actual visiting condition, the PageRank value expresses as below.

$$PR_{n+1}(P) = (1 - d) * 1 + d * \sum \frac{PR_{n+1}(T_i) * W_p(T_i)}{C(T_i)} \tag{24.2}$$

In formula (24.2),  $W_p(T_i)$  is the weight that is obtained from page  $T_i$  visiting page A, and this weight is proportional to the number page  $T_i$  outlink page A, inverse proportion to the page  $T_i$  outlink all the pages. So  $W_p(T_i)$  is

$$W_p(T_i) = \frac{A_p^{T_i}}{A_{B(T_i)}} \tag{24.3}$$

In formula (24.3),  $A_p^{T_i}$  is the number that user through page  $T_i$  visited page P;  $A_{B(T_i)}$  is the number that user through page  $T_i$  all outlinks,  $B(T_i)$  is the page  $T_i$ ’s all outlinks.

According to the formula (24.1), the PageRank value expresses is

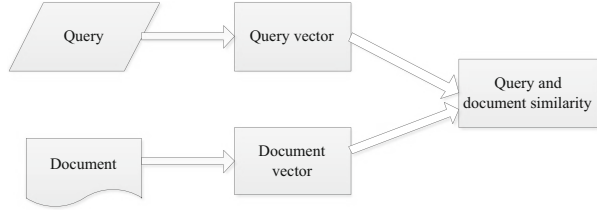
$$PR_{n+1}(P) = (1 - d) * W_p + d * \sum \frac{PR_{n+1}(T_i) * W_p(T_i)}{C(T_i)} \tag{24.4}$$

In formula (24.4),  $W_p$  is the weight that user random visit some page; the weight is proportional to the number that user not through other link visit page A and inversely proportional to the number that user visit page A; the expression is

$$W_p = \frac{A'_p}{A_p} \tag{24.5}$$

In formula (24.5),  $A'_p$  is the number that user not through other link visit page A, and  $A_p$  is the number that user visit page A.

**Fig. 24.4** The calculation method of vector space retrieval model



At last, we got the modified PageRank value:

$$PR_{n+1}(P) = (1 - d) * \frac{A'_P}{A_P} + d * \sum \frac{PR_{n+1}(T_i) * W_p(T_i)}{C(T_i)} \quad (24.6)$$

By improving the PageRank algorithm, the URL's filtering accuracy is improved. And through that we can get more useful page that related to the user's requirements.

### 24.3.2 Content Filter Analysis and Implementation

Content filtering needs to make sure the content filtering algorithm with contextual and real-time result; therefore, filtration precision and filtration velocity become a key content filtering criterion. Current algorithms for matching model include Boolean model, vector space model, and analysis semantic model. Boolean model is a strict matching model; its fast speed is very convenient to realize and suitable for structured information. Vector space model [7] is expressed as a vector to page document, and it will be submitted to customers with the search content. Semantic analysis model based on keyword matching will search for the link between the search item and the actual content to build for a semantic model.

In content-based filtering algorithm, we need to evaluate the similar degree between page and the search subject, that is to say the keywords [8] and the web page will sort by correlation. We would improve the algorithm to increase the search relevance of content. Specific ideas of the improve algorithm:

- Combined the keywords in the text with the frequency and location to determine the relative weights
- Web crawler is to analyze and collect the network data that have higher relative weights and filter irrelevant information
- Using vector space model to collected text for the N dimensional vector and calculate the similarity

Therefore, we simplified the user's query keyword and network resources to a vector which means weight. The method of vector space model works as Fig. 24.4.

Vector similarity algorithm is as follows: Suppose we have two text data which have relation to the author's paper and are expressed as D1, D2.  $w_{1k}$ ,  $w_{2k}$  indicate

text weight. Similarity  $sim(D_1, D_2)$  can be expressed with the distance between the vector:

$$sim(D_1, D_2) = \frac{\sum_{k=1}^n w_{1k} * w_{2k}}{\sqrt{\sum_{k=1}^n w_{1k}^2 \sum_{k=1}^n w_{2k}^2}} \tag{24.7}$$

Through the improved algorithm we can further filter out content which is not related to web pages or the related degree is not high, increase the precision of the system.

### 24.4 Experiment Studying

In this section, we describe an experiment to test our vertical search engine. We crawl about 26,756 web pages on the Internet by open source search engine Nutch. After that we transported the crawling web pages to database. The keywords, “cloud computing,” “data base,” and “mobile computing,” respectively, are chosen for the experiments. And we compute the accuracy. The comparison of the two results is as follows in Fig. 24.5. Obviously the accuracy of the improved one is higher to the original one.

### 24.5 Conclusion

In this chapter, one vertical search engine for paper submitted was designed, which can give users a help to search available conference or periodicals. We improved the accuracy through the selection of URL library, data filtering, better PageRank, and better similarity algorithm. The key problem, in detailed design, filtering

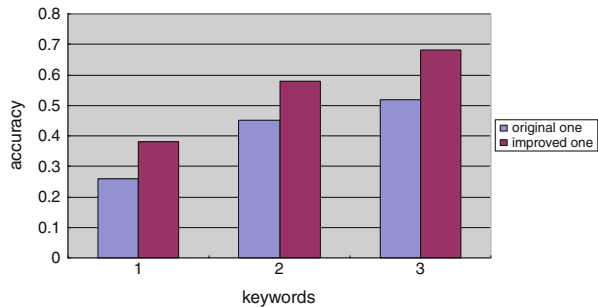


Fig. 24.5 The contrast result

algorithm was discussed in detail. With the improved algorithm, the search engine's query result is comprehensive and its precision is high and it gives users more available results.

There are also some further problems to solve, such as better user interface to display the result, more reasonable Chinese word segmentation and site revisit time, and so on. All above-mentioned issues deserve further research.

**Acknowledgments** This work was supported by the Natural Science Foundation of P. R. of China (90912003, 90812001 and 61073193), the Key Science and Technology Foundation of Gansu Province (1102FKDA010), Natural Science Foundation of Gansu Province (1107RJZA188), and the Fundamental Research Funds for the Central Universities (lzujbky-2012-47, lzujbky-2012-48).

## References

1. Zhang, L., Yang, L., & Xu, Z. (2012). *Paperscloud: A composing-free, collaborative editing platform for scientific papers* (pp. 286–291). The Proceeding of the 2012 IET International Conference on Frontier Computing-Theory, Technologies and Applications (IETFCA 2012), ISBN 978-1-849-19604-8.
2. Yang, J. C., & Ling, P. L. (2009). Improvement of PageRank algorithm for search engine. *Computer Engineering*, 35, 35–37.
3. Chau, M., & Chen, H. C. (2008). A machine learning approach to web page filtering using content and structure analysis. *Elsevier Science Direct*, 44(2), 482–494.
4. Kleinberg, J. M. (1999). Authoritative sources in a hyperlinked environment. *Journal of the ACM (JACM)*, 46(5), 604–632.
5. Hjjhjf, T. (2002). *Topic-sensitive PageRank* (pp. 517–526). Proceedings of the 11th International Conference on World Wide Web (WWW02), Honolulu, Hawaii.
6. L. Page, S. Brin, R. Motwani. (1998). *The page rank citation ranking: Bringing order to the web*. Stanford Digital Library Technologies Project 1998.
7. Zhang, Y. M., & Zhou, J. F. (2000). A trainable method for extracting Chinese entity names and their relations. *Proceedings of the Second Chinese Language Processing Workshop*, 12, 66–72.
8. Amit, C., & Jaewoo, K. (2007). *Selective approach to handling topic oriented tasks on the World Wide Web* (pp. 343–348). Proceeding of the 2007 I.E. Symposium on Computational Intelligence and Data Mining (CIDM2007), ISBN 1-4244-0705-2/07.

# Chapter 25

## Analysis and Improvement of SPRINT Algorithm Based on Hadoop

Shanshan Fei, Qiaoyan Wen, and Zhengping Jin

**Abstract** With the rapid development of computers and networks, the growth of data causes the data mining increasingly difficult. To solve this problem, this paper proposes an improved SPRINT algorithm based on the Hadoop platform. By analyzing the traditional SPRINT algorithm, we improve it in three aspects: eliminate unnecessary and repetitive calculations in the processing of discrete attributes; none presort of continuous attributes and split by line directly when splitting; and add the node field for attributes list in the data structure. For illustration, a performance test of acceleration and accuracy is executed to prove the effectiveness of the improved SPRINT algorithm. Compared to the original SPRINT algorithm, experimental result shows that the improved SPRINT algorithm guarantees the accuracy and reduces the computing time for the best split point thus accelerates the speed of decision-tree construction.

### 25.1 Introduction

The rapid popularity of the Internet has brought a wealth of information for people, but followed by a massive data explosion struck, that is, to say “information explosion.” Nowadays, the storage and process of large data sets has become the new challenge faced by many enterprises. It becomes a new target for data mining that how to mining valuable and understandable information from massive data in a fast, efficient, and cost-effective way [1].

The emergence and development of cloud computing has brought new opportunities and challenges for data mining. Cloud computing achieves the process for large data sets by evenly distributing storage and calculation to a multiple of storage

---

S. Fei (✉) • Q. Wen • Z. Jin  
Network Security Laboratory, Beijing University of Posts and Telecommunications,  
Beijing 100876, China  
e-mail: [shanshanfei@163.com](mailto:shanshanfei@163.com)

computing nodes in a cluster [2, 3]. It greatly reduces the cost owing to the low-cost computer clusters instead of the expensive servers. Data mining is going into the era of cloud based with the help of superior computing power of cloud computing.

Decision tree is the most basic and common classification algorithm in data mining [4]. The basic idea of decision-tree induction is the greedy algorithm, which takes the top-down way and crashes one by one. The construction of the decision tree usually consists of two stages: construction stage and pruning stage. In the construction stage, a fully grown decision tree is built by recursively calling the algorithm. This paper will focus on tree construction stage.

This paper will focus on the SPRINT algorithm that is a kind of decision-tree algorithms [5]. We propose an improved SPRINT algorithm based on the Hadoop platform and make a performance test. The performance test includes the writing of program code, the building of Hadoop cluster, the submitting of program, and the comparative analysis of results.

## 25.2 Hadoop and MapReduce

Our study is based on the Hadoop platform, while the writing of program code uses the MapReduce programming model [6]. Hadoop is an Apache open source project used to build the cloud platform. It allows users more easily to write and run applications for processing huge amounts of data. Because the cluster has high access speed and a good backup, we set up a Hadoop cluster to verify the effectiveness of the improved SPRINT algorithm.

MapReduce is Google's core computing model. It is a simple and linearly scalable model for data processing. MapReduce highly abstracts the complex parallel computing process running on large-scale clusters to two functions: Map and Reduce [7, 8]. Each of them defines a mapping from one set of key-value pairs to another. Map is responsible for breaking down tasks, while Reduce is responsible for merging the decomposed tasks. Programmers only need to specify the Map function and the Reduce function to write distributed parallel programs.

## 25.3 Basic Idea of SPRINT Algorithm

The selection of test attributes and division of the sample set are the key links in constructing a decision tree. Gini index can effectively search for the best split point [9]. The point that has minimum Gini index is chosen as the best split point with the maximum information gain. It is very beneficial to generate a good decision tree.

The Gini index is described as:

1. If the set  $S$  contains  $M$  records belonging to  $N$  categories, the Gini index is

$$Gini(S) = 1 - \sum_{i=1}^n P_i^2$$

where  $P_i$  is the frequency of class  $i$ .

2. If the set  $S$  is divided into  $S1$  and  $S2$ , respectively, corresponding to  $M1$  records and  $M2$  records, the Gini index is

$$Gini_{split}(S) = \frac{m_1}{m} Gini(S1) + \frac{m_2}{m} Gini(S2)$$

SPRINT algorithm uses two data structures: the attributes list and histograms [10]. When the data set is too large to fit entirely in memory, SPRINT algorithm can save the rest of the attributes list to the hard disk and only put the current attributes list into memory. The attributes list is split with the division of the node. It includes three fields: attributes value, attributes class label, and row index. Histogram is shown in two forms based on the attribute is continuous attribute or discrete attribute. When the attributes list is generated for the first time, it will be sorted firstly.

For continuous attributes, firstly presorted, assume that the results are  $v_1, v_2, \dots, v_i, \dots, v_n$ . Because the split point is between two adjacent points, so there are  $n-1$  possibilities. The split forms are two parts:  $V \leq v_i$  and  $V > v_i$ ; it usually takes the midpoint  $\frac{v_i+v_{i+1}}{2}$  as the candidate split point. Scanning the attributes list from top to bottom and calculating the Gini index of all candidate points, the best splitting point is the point that has the minimum Gini index. For discrete attributes, there are  $m$  mutually different values, the segmented nature is divided into two sets, and there are  $2^m$  possible divisions. We need to calculate the Gini index for each division and look for the best candidate split point. Ultimately, the best splitting point is the point with minimum Gini index and the corresponding attribute is the test attribute.

For example, in the “car insurance” case, the attributes list and the Gini index are shown in Fig. 25.1.

The process of SPRINT algorithm is shown below:

1. Generate initial attributes list: A. Create root node N. Presort A with continuous attribute. Attach A to N.
2. Create the decision tree for node N as following BuildTree algorithm.

The process of BuildTree is shown below:

1. If all attributes in A are of the same class, return; otherwise, go to 2.
2. Scan A and update histograms.
3. Compute the minimum Gini index for candidate split points to get the best split point.

<i>rid</i>	<i>age</i>	<i>carType</i>	<i>risk</i>
0	23	familyCar	high
1	17	racingCar	high
2	43	racingCar	high
3	68	familyCar	low
4	32	truck	low
5	20	familyCar	high

→

<i>rid</i>	<i>age</i>	<i>risk</i>
0	23	high
1	17	high
2	43	high
3	68	low
4	32	low
5	20	high

+

<i>rid</i>	<i>carType</i>	<i>risk</i>
0	familyCar	high
1	racingCar	high
2	racingCar	high
3	familyCar	low
4	truck	low
5	familyCar	high

<i>S1</i>	<i>S2</i>	<i>gini</i>
age ≤ 18.5	age > 18.5	0.4
age ≤ 21.5	age > 21.5	0.333
age ≤ 27.5	age > 27.5	0.222
age ≤ 37.5	age > 37.5	0.417
age ≤ 55.5	age > 5.5	0.267
age ≤ 55.5	age > 5.5	0.267

<i>S1</i>	<i>S2</i>	<i>gini</i>
{ }	{ familyCar , racingCar, truck }	0.444
{ familyCar }	{ racingCar, truck }	0.444
{ racingCar }	{ familyCar, truck }	0.333
{ truck }	{ familyCar , racingCar }	0.267
{ familyCar , racingCar }	{ truck }	0.267
{ familyCar, truck }	{ racingCar }	0.333
{ racingCar, truck }	{ familyCar }	0.444
{ familyCar , racingCar, truck }	{ }	0.444

Fig. 25.1 Attributes list generate and Gini index

4. Use the best split point to split A into A1 and A2 and associate them with N1 and N2.
5. Execute BuildTree recursively to create tree for N1 and N2.

## 25.4 Analysis and Improvement of SPRINT Algorithm

### 25.4.1 The Analysis of SPRINT Algorithm

SPRINT algorithm is completely free from memory limit, easy to parallel, and better scalability, acceleration, and expansion, and the resulting decision tree is more compact and accurate.

Through in-depth study of SPRINT algorithm, its shortcomings are as follows:

1. There are repetitive and unnecessary calculations when calculating the Gini index of discrete attributes.

For discrete attributes, assuming it has  $m$  values different from each other, the nature of the split is that  $m$  values are divided into two sets and there are  $2^m$  possible divisions. This means we must calculate  $2^m$  Gini values. From Fig. 25.2,



```

Algorithm SPRINT
Input: training sample S
Output: decision-tree
generate attributes list A according S
create node N and attach it to A
BuildTree(A,N)
{
    if(each attribute in A belong to the same class) then
        return
    for( $A_i$  in A)
        if  $A_i$  is continuous then
            take the midpoint of adjacent points as candidate split point
            obtain the minimum Gini index
        else if  $A_i$  is discrete then
            compute each Gini index of half the possible divisions
            obtain the minimum Gini index
        obtain the best split point of A and split A into A_Left and A_Right
        create nodes N_Left and N_Right
        attach N_Left and N_Right with A_Left and A_Right
        BuildTree(A_Left, N_Left)
        BuildTree(A_Right, N_Right)
}

```

Fig. 25.2 Process of the improved SPRINT algorithm

we can find that some divisions have the same Gini index. For example, the divisions  $S1 = \{\text{family car}\}$ ,  $S2 = \{\text{racing car, truck}\}$  and  $S1 = \{\text{racing car, truck}\}$ ,  $S2 = \{\text{family car}\}$  have the same value 0.444. In fact, when the set S is divided into S1 and S2, exchanging data in S1 and S2, the Gini index is unchanged according to the definition of Gini index. Therefore, the Gini index of nearly half of the candidate split points with that of the other half are same. Such calculations are repeated.

In addition, from the basic idea of SPRINT algorithm, we can see that when a subset is belonging to the same class, the division is terminated. If a set S needs to split, it means that the set S does not belong to the same class. So, the sets which contain 0 value and M values have no sense and the Gini index of them must have the maximum value. Therefore, such calculations are unnecessary.

2. The split of nondividing attributes becomes difficult after the continuous attributes perform presorted.

SPRINT algorithm will perform presorted for continuous attributes, and the discrete attributes do not need to be sorted when the attributes list is generated for the first time. If it utilizes presorted, when split by a continuous attribute, you need to loop scan to determine whether the row is in the left sub tree or in the right sub tree for the nondividing attributes. It increases the scan of attributes list and causes the large amount of computation.

3. The traditional data structure cannot meet the mapping function on the MapReduce framework.

The function of Map is mapping. By defining different petitioner functions, we can map different <key, value> pairs to different Reducers for processing. The attributes list on each node uses Map to be mapped to different Reducers for parallel processing. The original attributes list includes attributes value, class label, and the row index, which obviously does not meet the requirement. This requires a new data structure to solve this problem.

#### 25.4.2 *The Improvement of SPRINT Algorithm*

The original SPRINT algorithm is improved in the following three aspects:

1. Eliminate repetitive and unnecessary calculations of discrete attributes.

The repetitive and unnecessary calculations should be eliminated in the processing of discrete attributes. In general, you need to calculate  $N$  Gini index with the unimproved algorithm, but you only need to calculate  $\frac{N-2}{2}$  Gini index with the improved algorithm. Obviously, the improved algorithm reduces nearly half of the amount of computation for the best split point.

2. None presort of continuous attributes.

Continuous attributes do not perform presorted. We just mark the split index so the nondividing attributes can be split by line directly that saves split time spent on the nondividing especially when the data set is large. The cost of this improvement is that the time spent on the calculation of Gini index for continuous attributes is more than that using presorting technology. Overall, the effect is still relatively good. Maybe the input sample is not typical enough which requires further verification.

3. Add node field in attributes list.

The node field is added in the attributes list to record the current processing node that is used to perform the mapping. However, the improvements have not yet reached the purpose of reducing disk scans. The improvement of the data structure requires further study.

In addition, the current split attribute is deleted immediately after each split because the candidate split attributes are the attributes that never used to split. This saves space and time. The process of the improved SPRINT algorithm is shown in Fig. 25.2.

The detail procedure of improved SPRINT algorithm is shown as follows:

1. Generate initial attributes list: A. Create root node N. Attach A to N.
2. Create the decision tree for node N as following BuildTree algorithm.

The process of BuildTree is described as follows:

1. If all attributes in A belong to the same class, return.
2. For ( $A_i$  in A).
  - If  $A_i$  is continuous, the midpoint of the adjacent points is taken as the candidate split point and the Gini index is calculated. If  $A_i$  is discrete, the Gini index of half the possible divisions are calculated.
  - Compute the minimum Gini index for candidate split points to get the best split point. Use it to split A into A\_Left and A\_Right and associate them with N\_Left and N\_Right.
3. Recursively execute BuildTree(A\_Left, N\_Left) and BuildTree(A\_Right, N\_Right) until a decision tree is generated.

### 25.4.3 Performance Test

In this experiment, we implement the original and the improved SPRINT algorithm in MapReduce framework based on the Hadoop. In this part, we focus on analyzing the performance such as accuracy and acceleration of the improved SPRINT algorithm. The experiment uses the virtual machine to build a three-node Hadoop cluster [11].

The cluster includes three nodes: one Master and two Slaves. The Master machine is responsible for the implementation of the data distribution and task decomposition as NameNode and JobTracker. These two Slave machines are responsible for distributed data storage and the execution of tasks as DataNode and TaskTracker. The IP addresses of each node are shown in Table 25.1.

The comparison of results is shown in Table 25.2.

From Figs. 25.3 and 25.4, we can see the time spent on dealing with the same number of records decreased for the improved SPRINT algorithm and still achieves almost the same accuracy that proves the effectiveness of the improvement.

## 25.5 Conclusion

This paper proposes an improved SPRINT algorithm that improves in three aspects: eliminate unnecessary and repetitive calculations in the processing of discrete attributes; none presort of continuous attributes and split by line directly when splitting; and add the node field for attributes list in the data structure. Through the performance test, we prove that compared to the original SPRINT algorithm, the improved SPRINT algorithm reduces the computing time for the best split point

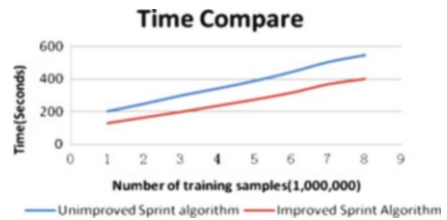
**Table 25.1** Hosts in cluster

Host name	IP address	Experimental environment
Master	192.168.1.133	Operating system: CentOS 6.0; Hadoop: Hadoop 1.0.4; IDE platform: Eclipse 3.3.0; Java runtime environment: JDK1.6.0.19
Slave1	192.168.1.134	
Slave2	192.168.1.135	

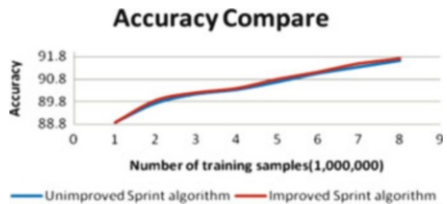
**Table 25.2** Comparison between unimproved and improved algorithm

Recorder/1,000,000	Unimproved		Improved		Save (s)
	Accuracy (%)	Time (s)	Accuracy (%)	Time (s)	
1	88.59	202	88.59	130	72
2	89.77	248	89.77	165	83
3	90.12	298	90.12	199	99
4	90.33	341	90.34	236	105
5	90.67	387	90.68	274	113
6	91.04	441	91.06	315	126
7	91.32	502	91.35	367	135
8	91.61	544	91.65	399	145

**Fig. 25.3** Time compares



**Fig. 25.4** Accuracy compares



thus accelerates the speed of decision-tree construction and still guarantees the accuracy. Of course, there are still some defects of the improved SPRINT algorithm, such as the attributes list is scanned many times and a better data structure is needed in order to save more time and space. This is also the next further study for us.

**Acknowledgements** National Natural Science Foundation of China (Grant Nos. 61170270, 61100203, 60903152, 61003286, 61121061) and Fundamental Research Funds for the Central Universities (Grant Nos. support this work BUPT2011YB01, BUPT2011RC0505, 2011PTB-00-29, 2011RCZJ15).

## References

1. Fayyad, U., Piatetsky-Shapiro, G., & Smyth, P. (1996). From data mining to knowledge discovery in databases. *AI Magazine*, 17(3), 37.
2. Armbrust, M., Fox, A., Griffith, R., Joseph, A.D., Katz, R., Konwinski, A., & Zaharia, M. (2010). A view of cloud computing. *Communications of the ACM*, 53(4), 50–58.
3. Chen, Q., & Deng, Q.-n. (2009). Cloud computing and its key techniques. *Journal of Computer Applications*, 29(9), 2565 (In Chinese).
4. Durkin, J., Jingfeng, C., & Zixing, C. (2005). Decision tree technique and its current research. *Control Engineering*, 12(1), 15–21 (In Chinese).
5. Liu, H., Chen, J., & Chen, G. (2002). Review of classification algorithms for data mining. *Journal of Tsinghua University: Science and Technology*, 42(6), 727–730 (In Chinese).
6. White, T. (2012). *Hadoop: The definitive guide* (pp. 15–38). Sebastopol, CA: O'Reilly Media.
7. Dean, J., & Ghemawat, S. (2008). MapReduce: Simplified data processing on large clusters. *Communications of the ACM*, 51(1), 107–113.
8. Dean, J. (2006). Experiences with MapReduce, an abstraction for large-scale computation. In *PACT: Proceedings of the 15th international conference on parallel architectures and compilation techniques* (Vol. 16, Issue 20, p. 1). Washington, DC: IEEE Computer Society.
9. Raileanu, L. E., & Kilian, S. (2004). Theoretical comparison between the gini index and information gain criteria. *Annals of Mathematics and Artificial Intelligence*, 41(1), 77–93.
10. Ganti, V., Gehrke, J., & Ramakrishnan, R. (1999). Mining very large database. *Computer*, 32(8), 38–45.
11. Lu, Q., & Cheng, X. (2012). The research of decision tree mining based on Hadoop. In *9th international conference on fuzzy systems and knowledge discovery (FSKD) 2012* (pp. 798–801). Piscataway, NJ: IEEE.

# Chapter 26

## Prediction Model for Trend of Web Sentiment Using Extension Neural Network and Nonparametric Auto-regression Method

Haitao Zhang, Binjun Wang, and Guangxuan Chen

**Abstract** In order to solve the problem of prediction for long-term web sentiment, a prediction model is built using the proposed method in this paper. First, a novel clustering method based on the extension neural network (ENN) is introduced to recognize the types of subclass of web sentiment. For each class of social events, the class model library of the development trend of web sentiment is established by cycle analysis and ENN clustering combined with nonparametric auto-regression analysis (NAR) method. Then the adaptive transformation is applied to the already known development trend of a new social event, and the min-sum of mean square error (MSE) from the library is selected to predict the future development trend of web sentiment. Empirical findings indicated that compared with the traditional methods, such as the GM (1,1) and least squares estimation (LS) method, the approach presented in this paper yields a higher correlation value in predicting the long-term development trend of web sentiment and can predict the turning points of the development trend more effectively. The ENN- and NAR-based prediction model can effectively solve the problem of prediction for long-term web sentiment.

### 26.1 Introduction

The World Wide Web and other textual databases provide a convenient platform for exchanging opinions. Reviews and blogs written with the purpose of conveying a particular opinion or sentiment are more and more popular in China. Therefore, prediction of web sentiment development is one of the most important subjects for plan and operation in the field of mass media and sociology. An accurate sentiment prediction is the basis of making decision on the trend of public security and public

---

H. Zhang (✉) • B. Wang • G. Chen  
People's Public Security University of China, Beijing 100038, China  
e-mail: [okhaitao@126.com](mailto:okhaitao@126.com)

opinions flow for the sociology researchers. The preliminary research in this paper has been carried out by literature survey and observation. Various models for web sentiment prediction have been reported in the literature. Detailed methods include statistical methods [1–3], the regression functions [4], and grey theory [5], neural network method [6]. Artificial neural networks are currently established as a promising approach to sentiment prediction since they are able to learn the functional relationship between system inputs and outputs through a training process. However, a limitation of the neural network (NN) approach is difficult to understand; the content of network memory and the traditional neural network approaches need large amounts of training data and thus require heavy computational efforts to create satisfactory models. Grey prediction models do not need large amounts of data [7]. They have been successfully applied in many fields, but the accuracy of grey prediction models still needs further improvement. Auto-regression method is difficult to solve nonlinear problems such as population growth forecasting due to the form of linear parametric. This paper proposed a novel prediction method; we use a new neural network topology, called the extension neural network (ENN) proposed in Wang’s paper [8], to recognize the trend types of web sentiment. Then, according to the actual data of every type, we use the nonparametric auto-regression analysis method (NAR) to build the prediction models of every type.

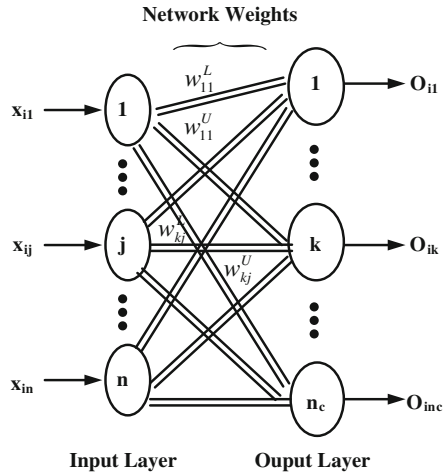
## 26.2 Extension Neural Network

Extension neural network is a combination of the neural network and the extension theory. The extension theory proves a novel distance measurement for classification processes, and the neural network can embed the salient features of parallel computation power and learning capability. The ENN permits classification of problems with range features, continuous input, and discrete output [9].

The schematic structure of the ENN is depicted in Fig. 26.1. It comprises of both the input layer and the output layer. The nodes in the input layer receive an input feature pattern and use a set of weighted parameters to generate an image of the input pattern. In this network, there are two connection values (weights) between input nodes and output nodes; one connection represents the lower bound for this classical domain of the features and the other connection represents the upper bound. The connections between the  $j$ th input node and the  $k$ th output node are  $w_{kj}^L$  and  $w_{kj}^U$ . This image is further enhanced in the process characterized by the output layer. The output layer is a competitive layer. There is one node in the output layer for each prototype pattern and only one output node with nonzero output to indicate the prototype pattern that is closest to the input vector.

The learning of the ENN can be seen as supervised learning; before the learning, several variables have to be defined. Let training set be  $X = \{X_1, X_2, \dots, X_{Np}\}$ ,

**Fig. 26.1** The structure of extension neural network (ENN)



where the total number of training patterns is  $N_p$ . The  $i$ th input vector is  $X_i^p = \{x_{i1}^p, x_{i2}^p, \dots, x_{in}^p\}$ , where  $n$  is the total number of the features. To evaluate the clustering performance, the total error number is set as  $N_m$ . The total number of the input pattern is set as  $N_p$  and the total error rate  $E_T$  is defined below:

$$E_T = \frac{N_m}{N_p} \tag{26.1}$$

The detailed supervised learning algorithm can be described as follows:

Step 1: Set the connection weights between input nodes and output nodes according to the range of classical domains. The range of classical domains can be directly obtained from previous experience or be determined from training data as follows:

$$w_{kj}^L = \min_{T_i \in k} \{x_{ij}\}; w_{kj}^U = \max_{T_i \in k} \{x_{ij}\} \tag{26.2}$$

Step 2: Read  $i$ th training pattern and its cluster number  $p$ :

$$X_i = \{x_{i1}, x_{i2}, \dots, x_{in}\} \tag{26.3}$$

Step 3: Use the extension distance (ED) to calculate the distance between the input pattern  $X_i$  and the  $k$ th cluster as follows:

$$ED_{ik} = \sum_{j=1}^n \left( \frac{|x_{ij} - (w_{kj}^U + w_{kj}^L)/2| - (w_{kj}^U - w_{kj}^L)/2}{(w_{kj}^U - w_{kj}^L)/2} + 1 \right) \tag{26.4}$$

The proposed distance in this paper is a modification of ED [10]. It can describe the distance between the  $x$  and a range  $\langle W^L, W^U \rangle$ . Different ranges of classical



domains can arrive at different distances due to different sensitivities. This is a significant advantage in classification applications. Usually, if the feature covers a large range, the requirement of data is fuzzy or low in sensitivity to distance. On the other hand, if the feature covers a small range, the requirement of data is precision or high sensitivity to distance.

Step 4: Find the  $m$ , such that  $ED_{im} = \text{Min}\{ED_{ik}\}$ . If  $m = p$ , then go to step 6; otherwise step 5.

Step 5: Update the weights of the  $p$ th and the  $m$ th clusters as follows:

$$\begin{cases} w_{pj}^{L(new)} = w_{pj}^{L(old)} + \eta \left( x_{ij} - \frac{w_{pj}^{L(old)} + w_{pj}^{U(old)}}{2} \right) \\ w_{pj}^{U(new)} = w_{pj}^{U(old)} + \eta \left( x_{ij} - \frac{w_{pj}^{L(old)} + w_{pj}^{U(old)}}{2} \right) \end{cases} \quad (26.5)$$

$$\begin{cases} w_{mj}^{L(new)} = w_{mj}^{L(old)} - \eta \left( x_{ij} - \frac{w_{mj}^{L(old)} + w_{mj}^{U(old)}}{2} \right) \\ w_{mj}^{U(new)} = w_{mj}^{U(old)} - \eta \left( x_{ij} - \frac{w_{mj}^{L(old)} + w_{mj}^{U(old)}}{2} \right) \end{cases} \quad (26.6)$$

where  $\eta$  is a learning rate. In this step, we can clearly see that the learning process is only to adjust the weights of the  $p$ th and the  $m$ th clusters.

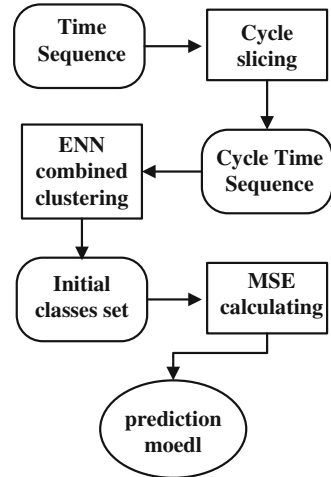
Step 6: Repeat step 2 to step 5; if all patterns have been classified, then a learning epoch is finished.

Step 7: Stop if the clustering process has converged or the total error has arrived at a preset value; otherwise, return to step 3.

## 26.3 The Proposed Method

Sentiment prediction is different from traditional text categorization because sentiments are ordinal variables in contrast to the categorical nature of topics; lots of contradicting opinions might coexist on the same event, which interact with each other to produce the global sentiment. Indeed, sentiment prediction is a much harder task than topic of text classification tasks. For many historic events or other great issues that happened in society, studies found that not only the same type of event development trend has a high similarity but also the development of the same event experiences the same cycle of history. In order to improve the fitting accuracy of the model and prediction ability, a series of procedures must be done.

**Fig. 26.2** Flow chart of prediction model for web sentiment



As showed in Fig. 26.2, firstly, this paper will make a classification of events and cycle slicing. Then establish the cycle class model recognized by ENN and NAR method for each type of event and create class model libraries. After that, pick up the class model, which has the minimum mean square error (MSE) for long-term prediction. Due to the nonlinear relation between input factors and output values, the prediction of sentiment is not an easy work. Past value of the time sequence will affect the future value.

In short, the main ideas of the proposed method are divided into two parts. First, we use a novel clustering method based on the ENN to recognize the type of every subclass of time sequences. When the dates are given, the changed ranges will be obtained from the proposed models. Second, using the nonparametric auto-regression analysis method to build the prediction model of web sentiment development of long term, when the dates are given, the prediction values can be calculated according to the calculation of MSE. Use the ENN to learn the clustering models of every cycle time sequence derived from the statistical data of Google trends. The detail learning method can refer to Sect. 26.2; the prediction model can be calculated by the nonparametric auto-regression analysis method as in Sect. 26.3; the typical prediction model in our problem can be written as a NAR model will be found on the basis of sequence  $\{\Delta Y_t\}$ ,  $\Delta Y_t = m(\Delta Y_{t-1}, \Delta Y_{t-2}, \dots, \Delta Y_{t-p}) + \varepsilon_t$ , and random error sequence  $\{\varepsilon_t\}$  is independent and identically distributed.  $E(\varepsilon_t) = 0, E(\varepsilon_t^2) = \sigma^2$ , and  $\varepsilon_t$  are independent with  $\Delta Y_{t-1}, \Delta Y_{t-2}, \dots, \Delta Y_{t-p}$ . The parameter  $p$  is defined according to cross-validation (CV) method. The value of CV ( $p$ ) is obtained though calculating using MATLAB programming. An orthogonal sequence estimation method will be operated on the model [11]. After abovementioned procedures, the prediction data should be input, and the growth type of the prediction will be determined by the ENN-based clustering method. Use the prediction model of this type as NAR method to calculate the detailed values of the prediction data.

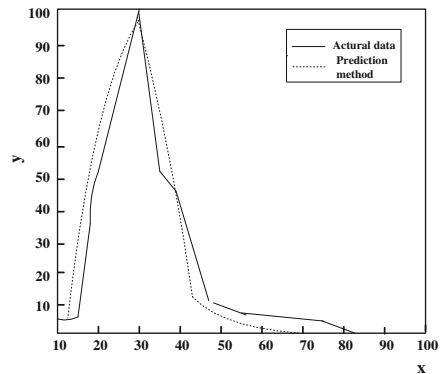
## 26.4 Case Studies and Discussions

It is clear that the relations between the input factors and output values are highly nonlinear curves. Therefore, if the prediction models are directly to use the NAR, the prediction error will be large, which is the main reason in the proposed method to delimit some growth types (or some sections) in the related curves. When the subtypes are delimited, every section can be seen as approximately linear, thus the sectioned prediction models will provide higher accuracy. This paper uses the historical data from Google trends and forums to build the prediction model to predict the bird flu event. In this case, we could make conclusion that when  $k = 1$ , the value of  $CV(k)$  is minimum, so, the optimal selection of lag is 1, in every recognized type the corresponding NAR model is  $\Delta Y_t = m(\Delta Y_{t-1}) + \varepsilon_t$ . Comparisons of prediction results using the grey GM (1,1) model and LS method are also conducted. In this paper, the accuracy of the prediction model is calculated with the value of correlation degree analysis. The compared prediction results with the grey GM (1,1) model and least square (LS) method show that the proposed method has better accuracy. Selecting the first 10 days of data for six different events, the compared results of modeling accuracy using different methods are shown in Table 26.1. The comparison between actual and prediction curve for the event of bird flu is shown in Fig. 26.3. Obviously, the compared results indicate that the proposed prediction method has the higher accuracy than both grey GM (1,1) and LS method.

**Table 26.1** Comparison between different methods on correlation degree

Event	Proposed method	LS	GM (1,1)
Bird flu	0.7910	0.5496	0.5189
Li Gang	0.8103	0.4834	0.3987
Needle shot	0.7637	0.5313	0.4897
Maiden suicide	0.8123	0.4720	0.5108
Dust lungs	0.8527	0.5435	0.4412
Drag racing	0.7879	0.5104	0.3925

**Fig. 26.3** Comparison between the actual data and long-term trend prediction curve



## 26.5 Conclusion

Web sentiment prediction has always been a great issue for the society. The sentiment of public reflects the attitude towards society and government at that time. In this paper, we address the ENN concept in details. As existing models are inadequate for a variety of reasons, we introduce the ENN- and NAR-based model that is suited to predict the trend of web sentiment. The proposed method not only can be used to predict the values of web sentiment, it also can be used with multi-input factors in nonlinear prediction problems. The main advantage of the proposed method is that it can give the range and values at the same time. Moreover, because of the simple structure of ENN, the computing time of the proposed method is also short. We also demonstrate the usefulness of the web sentiment representation for long-term prediction. Comparisons with the grey GM (1,1) model and LS method show that the proposed prediction method has better accuracy and provides more information that is scientific to sociology researchers.

## References

1. Hardle, W., & Chen, R. (1997). Nonparametric time analysis, a selective review with examples. *International Statistical Review*, 65(1), 49–72.
2. Smith, B. L., & Williams, B. M. (2002). Comparison of parametric and nonparametric models for traffic flow forecasting. *Transportation Research*, 10(4), 15–19.
3. Wang, L. (2007). *Public opinion study—theory, method and reality hotspot* (pp. 27–41). Tianjin: Tianjin Social Sciences Press.
4. Masry, E., & Fan, J. (1997). Local polynomial estimation of regression functions for mixing processes. *Scandinavian Journal of Statistics*, 24(2), 165–179.
5. Kayacan, E., & Ulutas, B. (2010). Grey system theory based models in time series prediction. *Expert System with Applications*, 37(2), 1784–1789.
6. Kaur, H., & Raghava, G. P. (2004). A neural network method for prediction of  $\beta$  turn types in proteins using evolutionary information. *Bioinformatics*, 20(16), 2751–2758.
7. Xie, N. M., & Liu, S. F. (2005). Discrete GM(1,1) and mechanism of grey forecasting model. *Systems and Engineering Theory and Practices*, 25(1), 95–97.
8. Wang, M. H., & Hung, C. P. (2005). Extension neural network-type2 and its applications. *IEEE Transactions on Neural Networks*, 16(6), 1352–1361.
9. Wang, M. H., Tseng, Y. F., Chen, H. C., & Chao, K. H. (2009). A novel clustering algorithm based on the extension theory and genetic algorithm. *Expert Systems with Applications*, 36(4), 8269–8276.
10. Cai, W. (1983). The extension set and incompatibility problem. *Journal of Scientific Exploration*, 1, 81–93.
11. Xia, Y., & Li, W. K. (2002). Asymptotic behavior of band width selected by the cross-validation method for local polynomial fitting. *Journal of Multivariate Analysis*, 83(2), 265–287.

# Chapter 27

## K-Optimal Chaos Ant Colony Algorithm and Its Application on Dynamic Route Guidance System

Hai Yang

**Abstract** Dynamic route guidance system is an important part of the intelligent transportation system; the core part of which is optimal path algorithm. This paper has analyzed the main influencing factors on the choice of optimal path, then provided an improved K-optimal chaos ant colony algorithm (K-CACA). The road impedance factor in K-CACA is based on the length, crowdedness, condition, and traffic load of the road sections. The optimizing procedure of the algorithm is speeded up by introducing the included angle threshold of direction. The chaos perturbation effectively refrains the algorithm from trapping into local optima. The results of simulation experiment show that K-CACA is effective and has much higher capacity of global optimization than Dijkstra algorithm and basic ant colony algorithm for optimal route choice.

### 27.1 Introduction

Dynamic route guidance system (DRGS) is an important part of the intelligent transportation system [1]. It offers the useful optimal route guidance information to the driver according to starting and destination point. The aim of DRGS is to achieve the goals of improving traffic system, voiding traffic jam, reducing the vehicles' travel time, and realizing the rational distribution of traffic flow on each road section by guiding the driver's travel decision [2]. So the core content of DRGS is the detection of optimal path in the traffic network.

With the constant enlargement of urban road network, the graph theoretic algorithms such as Dijkstra algorithm and mathematical programming methods are difficult to satisfy, the real-time requirements of DRGS. Initially proposed by

---

H. Yang (✉)

College of Information Science and Electricity Engineering, Shandong Jiaotong University,  
Jinan 250023, Shandong, China  
e-mail: [yanghai\\_sdjtu@163.com](mailto:yanghai_sdjtu@163.com)

Marco Dorigo in 1991, the ant colony algorithm (ACO) was aiming to search for an optimal path in a graph based on the behavior of ants seeking a path between ant nest and a source of food. ACO has been successfully applied to solve different combinatorial optimization problems such as TSP [3] and also is suitable for DRGS because of its robustness, positive feedback, and distributed computing.

## 27.2 The Influencing Factors of the Optimal Path Selection

### 27.2.1 The Length of Road Section

The length of road section is the most common weight during the procedure of finding the optimal path because the length is easy to access and presents the travel cost directly [4]. But the optimal path does not only mean the shortest path in dynamic route guidance system. The travel time, crowdedness, conditions, and traffic loadings of the roads also affect the driver's decision of optimal path.

### 27.2.2 The Travel Time and Crowdedness

The travel time is more important than the length of road section in the real urban traffic network when drivers decide which path they should go. The travel time factor is dynamic according to the vehicle number per unit time. So we can compute the crowdedness of road by the travel time factor [5].

The crowdedness of road refers to the situation that the vehicles on the road cannot run at the normal speed so that both the travel time and parking delay become longer. The crowdedness factor can be expressed as  $[T(i,j) - T_0(i,j)]/T_0(i,j)$ , where  $T(i,j)$  is the real travel time of road section  $(i,j)$  and  $T_0(i,j)$  is the normal travel time of road section  $(i,j)$  without jam. In general, the smaller  $T(i,j)$  is, the lower the crowdedness factor is, and vice versa.

### 27.2.3 Traffic Load

The traffic load refers to the number of vehicle staying on some road section for a moment. Particularly, the traffic load per kilometer presents the traffic density [6]. So the traffic load factor presents the change of traffic situation over time and its equation of state is as follows:

$$\frac{dx_{(i,j)}(t)}{dt} = I_{(i,j)}(t) - O_{(i,j)}(t) \quad (27.1)$$

where  $x_{(i,j)}(t)$  is the vehicle number on road section  $(i,j)$  at time  $t$ .  $I_{(i,j)}(t)$  is the number of vehicles coming into road section  $(i,j)$  per unit time.  $O_{(i,j)}(t)$  is the number of vehicles leaving from road section  $(i,j)$  per unit time.

### 27.3 K-Optimal Chaos Ant Colony Algorithm

The urban traffic network is usually abstracted as a weighted, directed graph  $G = (V,A,R)$ , with node set  $V$ , road section set  $A$ , and  $R = \{r_{ij}|(i,j) \in A\}$ ;  $r_{ij}$  denotes the road impedance factor, that is, heuristic information of road sections [7].

#### 27.3.1 The Selection of Subsequent Node Based on the Direction

Because the dynamic route guidance system provides optimal path information according to the specific starting point and destination point, the procedure of finding optimal path is obviously directional [8]. When the ant selects next node, the probability of succeeding in finding optimal path is larger, if the ant’s moving direction is close to the direction of destination point. Otherwise, the probability of succeeding in finding optimal path is lower, if the ant’s moving direction is far away from the direction of destination point.

Let  $v_s$  denote the starting node with the coordinate  $(x_s, y_s)$ ;  $v_d$  denotes the destination node with the coordinate  $(x_d, y_d)$ ;  $v_i$  denotes the current node where the ant stays with the coordinate  $(x_i, y_i)$ , and  $v_j$  denotes some node adjacent to  $v_i$  with the coordinate  $(x_j, y_j)$ . Let  $\overrightarrow{v_i v_d}$  denote the direction from current node  $v_i$  to destination node  $v_d$ , called the approximate destination direction. Thus, the included angle  $\theta_{ij}$  of road section  $(v_i, v_j)$  and  $\overrightarrow{v_i v_d}$  is as follows:

$$\theta_{ij} = \arctan \frac{k_d - k_{next}}{1 + k_d \cdot k_{next}} \tag{27.2}$$

where  $k_d = \frac{y_d - y_i}{x_d - x_i}$  denotes the slope of the approximate destination direction  $\overrightarrow{v_i v_d}$ .

$k_{next} = \frac{y_j - y_i}{x_j - x_i}$  denotes the slope of road section  $(v_i, v_j)$ .

In order to enlarge the probability of succeeding in finding optimal path, we can set an included angle threshold  $\phi$ . The subsequent nodes whose included angle is greater than the threshold  $\phi$  will be omitted. The convergence speed of K-optimal chaos ant colony algorithm (K-CACA) will be accelerated by decreasing  $\phi$ , and the solutions diversity will be enlarged by increasing  $\phi$ .

### 27.3.2 The Heuristic Information of Road Section

In the new algorithm, the road impedance factor, which is gained by using weighted summation of the length, crowdedness, condition, and traffic load of the road sections [9], presents the heuristic information of road sections.

Let  $r_{ij}$  denote the road impedance factor of road section  $(v_i, v_j)$ :

$$r_{ij} = \frac{\omega_1 \cdot \frac{1}{L_{ij}} + \omega_2 \cdot \frac{1}{C_{ij}} + \omega_3 \cdot Q_{ij} + \omega_4 \cdot S_{ij}}{\sum_{k=1}^4 \omega_k} \quad (27.3)$$

where  $L_{ij}$  denotes the length of road section  $(v_i, v_j)$ ;  $C_{ij}$  denotes the crowdedness of  $(v_i, v_j)$ ;  $Q_{ij}$  denotes the condition of  $(v_i, v_j)$ ;  $S_{ij}$  denotes the traffic load of  $(v_i, v_j)$ , and  $\omega_k (k = 1, 2, 3, 4)$  is weight factor.

### 27.3.3 Pheromone Updating Strategy

The pheromone updating strategy of the new algorithm is as follow:

$$\tau_{ij}(t+1) = (1 - \rho) \cdot \tau_{ij}(t) + \sum_{l=1}^L \Delta \tau_{ij}^l \quad i = 1, \dots, N \quad j = 1, \dots, K \quad (27.4)$$

where  $\rho$  is the evaporation rate of pheromone.

### 27.3.4 Chaos Selection Strategy

When ant colony algorithm initializes, the strength of pheromones on every road section is equal so that the possibility of every road section is equal. It is difficult for ant colony to find an optimization path; besides, the convergence speed of algorithm is slow. So using chaos operator is necessary because it can increase the searching efficiency by the random and ergodic of chaos [10].

Take the logistics mapping, for example, iterative formula is as follows:

$$z_{i+1} = \mu \cdot z_i \cdot (1 - z_i), \quad i = 0, 1, 2, \dots, \quad \mu \in (2, 4] \quad (27.5)$$

where  $\mu$  is the control parameter, with the domain between  $(2, 4]$ . When  $\mu = 4$ ,  $0 \leq z_0 \leq 1$ , logistic function is the full mapping in  $(0, 1)$ , which is at a totally chaos status. A chaos sequence will be created by the iteration and then be converted to a chaos ergodic parameter when solving optimization problems in space. After introducing chaos perturbation, the new status transition strategy becomes



$$p_{ij}^k(t) = \begin{cases} \frac{[\tau_{ij}^s(t)]^\alpha \cdot (r_{ij}^s)^\beta \cdot (1 + Z_{ij})^\gamma}{\sum_{j \in allowed_k} [\tau_{ij}^s(t)]^\alpha \cdot (r_{ij}^s)^\beta \cdot (1 + Z_{ij})^\gamma}, & j \in allowed_w \text{ and } \theta_{ij} \leq \phi \\ 0, & otherwise \end{cases} \tag{27.6}$$

In the formula above,  $\theta_{ij}$  is the included angle of road section  $(v_i, v_j)$  and the approximate destination direction  $\overrightarrow{v_i v_d}$ .  $\phi$  is the included angle threshold.

### 27.3.5 K-Optimal Paths Guidance

The K-optimal chaos ant colony algorithm proposed in this paper provides drivers with K-optimal paths, which are almost exactly the same. Drivers decide which path they want to adopt by themselves, according to the real conditions of roads and the travel preference. The value of parameter K is referred to the scale of the traffic network, generally,  $K = 3$ . The newly improved algorithm can relax the conflict of traffic jam due to that many drivers choose the same optimal path.

### 27.3.6 The Realization Steps of K-CACA

The realization steps of K-CACA are proposed as follows:

- Step 1: Set the iteration number  $n = 0$ , initialize the pheromone of all road section, and let  $K = 0$ .
- Step 2: Put the starting node into  $tabu_k(s)$ , compute the road impedance factors of every road, and then choose one subsequent node  $j$  to move according to  $p_{ij}^k$  and the included angle threshold, finally put node  $j$  into  $tabu_k(s)$ .
- Step 3: Update the pheromone of roads the ants passed by. Take record the current optimal path and let  $K = K + 1$ .
- Step 4: If  $K$  equals to the specific value, then stop and go to step 6.
- Step 5: Adjust the initial parameters according to the latest traffic situations, run this algorithm again in order to obtain new optimal paths. Go to step 2.
- Step 6: Output  $K$  optimal paths.

### 27.4 Experimental Results

We successfully achieve this arithmetic and carry out simulation experiments using the road network of Jinan City shown in Fig. 27.1 from Google Earth. The topology of the road network shown in Fig. 27.1 is demonstrated in Fig. 27.2. The road impedance factors demonstrated in Table 27.1 are obtained by calculating the weighted summation of the length, crowdedness, condition, and traffic load of the road sections at 8:30 in the morning. Let  $m$  denote the number of ants and set  $m = 20$ ,  $\alpha = 1$ ,  $\beta = 3$ , and  $\rho = 0.3$ , and use logistic mapping. The aim of



Fig. 27.1 The road network of Jinan City

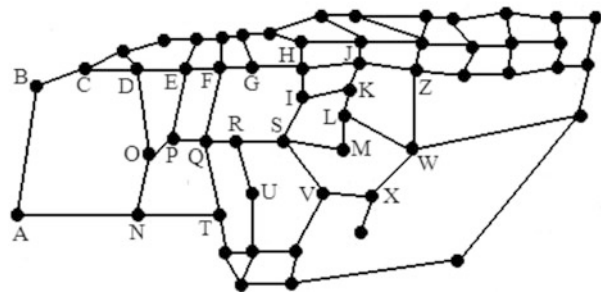


Fig. 27.2 The topology of road network of Jinan City

Table 27.1 The road impedance factor  $r_{ij}$  of road sections

Road section	$r_{ij}$	Road section	$r_{ij}$	Road section	$r_{ij}$	Road section	$r_{ij}$
(A, B)	4.96	(B, C)	2.16	(C, D)	2.02	(D, E)	3.13
(E, F)	2.87	(F, G)	3.01	(G, H)	3.08	(H, J)	3.97
(J, Z)	3.99	(Z, W)	4.01	(A, N)	4.86	(N, O)	3.22
(O, P)	2.13	(D, O)	3.99	(P, E)	3.91	(P, O)	2.71
(P, Q)	2.21	(E, P)	3.63	(Q, F)	3.70	(Q, R)	1.86
(N, T)	3.72	(T, Q)	3.91	(X, W)	3.77	(R, S)	3.64
(S, V)	3.71	(S, M)	3.57	(M, L)	2.28	(H, I)	1.13
(K, L)	1.97	(F, Q)	4.11	(L, W)	3.11	(V, X)	2.99

**Table 27.2** The results of three algorithms

Algorithm	Optimal path	Distance (km)	Average arrival time
K-CACA	A-N-T-Q-R-S-M-L-W	13.67	22 min and 7 s
Dijkstra	A-N-O-P-Q-R-S-M-L-W	14.75	26 min and 3 s
Basic ACO	A-B-C-D-E-F-G-H-J-Z-W	14.14	23 min and 21 s

simulation experiment is to find the optimal path from node A to node W using K-CACA, Dijkstra, and basic ant colony algorithm (ACO) separately. The results are shown in Table 27.2.

The results of simulation experiment show that K-CACA is effective and has much higher capacity of global optimization than Dijkstra algorithm and basic ant colony algorithm for optimal path choice.

## 27.5 Conclusion

In this article, K-optimal chaos ant colony algorithm is proposed for finding optimal path in dynamic route guidance system. The road impedance factor is introduced into K-CACA based on the length, crowdedness, condition, and traffic load of the road sections. The included angle threshold both accelerates the convergence speed and enlarges the solutions diversity. The chaos perturbation effectively refrains the algorithm from trapping into local optima. The experiment results show K-CACA is much more suitable for DRGS.

**Acknowledgements** The research is supported by Chinese Natural Science Foundation (61103022) and Scientific Research Fund Project of Shandong Jiaotong University (Z201213).

## References

1. Boyce, D. E., Ran, B., & Leblanc, L. J. (1995). Solving an instantaneous dynamic user optimal route choice model. *Transportation Science*, 29(2), 128–142.
2. Kuwahara, M., & Akamatsu, T. (1997). Dynamic user optimal assignment with physical queues for a many-to-many origin–destination pattern. *Transportation Research B*, 31(1), 1–10.
3. Dorigo, M. (1997). Ant colonies for the traveling salesman problem. *Biosystems*, 43(2), 73–81.
4. Ben-Akiva, M., Palama, A., & Kaysi, I. (1991). Dynamic network models and driver information systems. *Transportation Research A*, 25(5), 251–266.
5. Deek, H. L., & Kanafani, A. (1993). Modeling the benefits of advanced traveler information systems in corridors with incidents. *Transportation Research C*, 1(4), 303–324.
6. Daganzo, C. F. (1998). Queue spillovers in transportation networks with a route choice. *Transportation Science*, 32(1), 3–11.
7. Kim, J., Mitchell, J. S. B., & Polishchuk, V. (2012). Routing multi-class traffic flows in the plane. *Computational Geometry-Theory and Applications*, 45(3), 99–114.

8. Cordeau, J. F., & Maischberger, M. (2012). A parallel iterated tabu search heuristic for vehicle routing problems. *Computers and Operations Research*, 39(9), 2033–2050.
9. Blum, C., & Dorigo, M. (2004). The hyper-cube framework for ant colony optimization. *IEEE Transactions on Systems, Man and Cybernetics B*, 34(2), 1161–1172.
10. Yunwu, W. (2009). Application of chaos ant colony algorithm in web service composition based on QoS. *International Forum on Information Technology and Applications*, 172(2), 225–227.

# Chapter 28

## A Certainty-Based Active Learning Framework of Meeting Speech Summarization

Jian Zhang and Huaqiang Yuan

**Abstract** This paper proposes using a certainty-based active learning framework for extractive meeting speech summarization in order to reduce human effort in generating reference summaries. Active learning chooses a selective set of samples to be labeled by annotators. A combination of informativeness and representativeness criteria for sample selection is proposed. The results of summarizing parliamentary meeting speech show that the amount of labeled data needed for a given summarization accuracy can be reduced by more than 40 % compared to random sampling. The certainty-based active learning framework can effectively reduce the need of labeling samples for training. Furthermore, compared with lecture speech summarization task, the experiments show that the proposed active learning method of meeting speech summarization is obviously more affected by choice of different kinds of classifiers.

### 28.1 Introduction

The need for the summarization of spontaneous speech, such as classroom lectures, conference speeches, and parliamentary speeches, is ever increasing with the advent of remote learning, distributed collaboration, and electronic archiving. Short abstracts cannot sufficiently meet these user needs. State-of-the-art summarization systems are built by the extractive summarization method in a passive supervised learning framework, which compiles a summary from sentences or segments chosen from the transcribed document using some saliency criteria [1–3]. However, these supervised learning extractive summarization systems require a large amount of training data of reference summaries [4, 5]. There is no

---

J. Zhang (✉) • H. Yuan  
Engineering and Technology Institute, Dongguan University of Technology,  
Dongguan 523808, China  
e-mail: [zjian03@gmail.com](mailto:zjian03@gmail.com)

clear guideline for compiling stable and reproducible reference summaries. To minimize the human annotation efforts, yet still producing the same level of performance as a supervised learning approach, we study how to apply active learning approach for training extractive speech summarizer. We then propose three criteria and strategies in our certainty-based active learning framework of meeting speech summarization.

Being the first piece of work on active learning for lecture speech summarization task, we produced stable and reproducible reference summary and minimized the need for human annotation efforts, yet still producing the same level of performance as a supervised learning approach [6]. In this article, we further describe our certainty-based active learning framework in detail and verify its effectiveness on different genres of spontaneous speech.

The rest of this article is organized as follows: Sect. 28.2 describes our certainty-based active learning framework, the criteria, and the strategies for selecting samples. Section 28.3 first describes the parliamentary meeting speech corpus, for our experiments, and then outlines the acoustic/prosodic and lexical feature sets for representing each sentence of the transcriptions and then briefly depicts the probabilistic SVM classifier and naive Bayesian classifier as our extractive summarizers. Our experimental setup and the evaluation results are described in Sect. 28.4. Our conclusion follows in the end of this article.

## 28.2 Active Learning Framework and Sample Selection Strategies

### 28.2.1 Active Learning Algorithm

Human annotators label a small set of training data with summary labels. The training data is used to learn the initial model of the classifier. Then the classifier is used to predict labels for the sentences belonging to all transcribed documents from an unlabeled pool. Because human annotators usually annotate sentences of each document with summary labels by taking the context of the document into account, at each iteration they choose several unlabeled transcribed documents from the pool according to some sample selection strategies. Next the chosen transcribed documents are labeled by human annotators. These annotated transcribed documents are then added for retraining the classifier. This approach is described as follows:

Initialization

For an unlabeled data set:  $U_{\text{all}}, i = 0$

- (1) Randomly choose a small set of data  $X\{i\}$  from  $U_{\text{all}}$ ;  $U\{i\} = U_{\text{all}} - X\{i\}$ .
- (2) Label each sentence in  $X\{i\}$  as summary or non-summary using the RDTW-based semiautomatic annotation procedure proposed by Zhang et al. [5], and save these sentences and their labels in  $L\{i\}$ .

### Active Learning Process

- (3)  $X\{i\} = \text{null}$ .
- (4) Train the classifier  $M\{i\}$  using  $L\{i\}$ .
- (5) Test  $U\{i\}$  by  $M\{i\}$  and select the most useful documents from  $U_{\text{fig}}$  based on informativeness/representativeness strategies described in Sect. 28.2.3.
- (6) Save selected documents into  $X\{i\}$ .
- (7) Label each sentence in  $X\{i\}$  as summary or non-summary using the RDTW-based semiautomatic annotation procedure.
- (8)  $L\{i + 1\} = L\{i\} + X\{i\}$ ,  $U\{i + 1\} = U\{i\} - X\{i\}$ .
- (9) Evaluate  $M\{i\}$  on the testing set  $E$ .
- (10)  $i = i + 1$ , and repeat from (3) until  $U\{i\}$  is empty or  $M\{i\}$  obtains expected performance.
- (11)  $M\{i\}$  is produced and the process ends.

We propose the following two criteria for choosing the most useful samples  $X$  from the unlabeled data  $U$  for annotation.

## 28.2.2 Sampling Criteria

### 28.2.2.1 Criterion 1: Informativeness

All active learning scenarios involve evaluating the informativeness of unlabeled samples. The simplest and most commonly used query framework is uncertainty sampling [7]. The basic idea of informativeness criterion is that samples that the current model is most uncertain about are selected for annotation. This criterion is often straightforward for probabilistic learning models. When we use a probabilistic model for binary classification as the summarizer, an informativeness-based uncertainty sampling simply selects the samples whose posterior probability of being summary sentences (positive examples) are close to the classification hyperplane. This intuition is justified by D. Lewis and J. Catlett [8] and G. Schohn and D. Cohn [9] based on a version space analysis. They claim that labeling a sample that lies on or close to the hyperplane is guaranteed to have an effect on the model construction.

We measure the informativeness score of unlabeled transcribed document  $D = \{s_1, s_2, \dots, s_n, \dots, s_N\}$  by  $\text{Score}_{\text{inf}}(D) = (1/N)\sum \text{informative}(s_n)$ , which indicates the ratio of the number of informative sentences to that of all sentences in the document  $D$ .

If the sentence  $s_n$  satisfies the informativeness criterion,

$$P\left(c\left(\vec{s}_n\right) = 1 \mid D\right) \in [(1 - \beta)^*T, (1 + \beta)^*T]$$

where  $c\left(\vec{s}_n\right) = 1$  means the sentence  $s_n$  is a summary sentence, **then**  $\text{informative}(s_n)$  is equal to 1; **if not**,  $\text{informative}(s_n)$  is equal to 0.

We denote  $T = P\left(c\left(\vec{s}_n\right) = 1\right)$  as the prior probability of the sentence  $s_n$  which is a summary sentence without considering any information on document  $D$ . We have tried different values for  $T$  when we evaluate our method on the development set. We found  $T$  as 0.12 is suitable empirically for meeting compression tasks.  $\beta$  is tuned by evaluating the development set sentences to optimize the active learning algorithm. Sentences that satisfy the informativeness criterion are close to the classification hyperplane. Misclassification implies that the model is most uncertain about the samples.

### 28.2.2.2 Criterion 2: Representativeness

Another general active learning criterion [10] is to query the sample that would impart the greatest change to the current model if we knew its label. This criterion has also been applied to probabilistic sequence models like conditional random fields (CRFs) [10]. The intuition behind this criterion is that it prefers samples that are likely to most influence the model regardless of the resulting query label. This criterion has been shown to work well in empirical studies, though it will be computationally expensive if both the feature space and set of labels are very large [10].

We consider that the model will become more robust if it can be retrained by the samples selected by this representativeness criterion. The basic idea of representativeness criterion is that samples that are most generalized and are likely to be margin points are selected.

We measure the representativeness score of unlabeled transcribed document  $D = \{s_1, s_2, \dots, s_n, \dots, s_N\}$  by  $\text{Score}_{\text{repre}}(D) = (1/N)\Sigma \text{representative}(s_n)$ , which indicates the ratio of the number of representative sentences to that of all sentences in the document  $D$ .

If the sentence  $s_n$  satisfies the representativeness criterion,

$$P\left(c\left(\vec{s}_n\right) = 1|D\right) \in [0, \beta * T] \cup [(1 - \beta * T), 1],$$

**then**  $\text{representative}(s_n)$  is equal to 1; **if not**,  $\text{representative}(s_n)$  is equal to 0.

Sentences with classification scores in the range  $[0, \beta * T]$  are believed to be most likely non-summary sentences. Sentences with scores in the range  $[(1 - \beta * T), 1]$  are believed to belong most likely to the summary-sentence class. It means that if  $P\left(c\left(\vec{s}_n\right) = 1|D\right)$  is close to 1, the sentence  $s_n$  is most likely summary sentence. If the samples from the above range are misclassified, those samples would impart the greatest change to the current model. Adding those samples to the training data set for labeling will improve the robustness of the model.



### 28.2.3 Sample Selection Strategies

We apply the following sample selection strategies based on the above criteria for selecting the most useful unlabeled examples from the unlabeled pool at each active learning iteration.

*Random Selection Baseline:* A baseline sample selection strategy is to randomly choose  $K$  documents from  $U\{i\}$  to  $X\{i\}$ .

*Length Selection Baseline:* A baseline sample selection strategy is to choose  $K$  documents from  $U\{i\}$  to  $X\{i\}$  according to the average length of a sentence. The documents with larger average length of a sentence are chosen first.

*Sample Selection Strategy 1 (Informativeness):* We choose  $K$  documents with the highest value of  $\text{Score}_{\text{info}}(D)$  from  $U\{i\}$  to  $X\{i\}$  for annotation.

*Sample Selection Strategy 2 (Robustness):* We choose  $K$  documents with the highest value of  $\text{Score}_{\text{repre}}(D)$  from  $U\{i\}$  to  $X\{i\}$  for annotation.

*Sample Selection Strategy 3 (Hybrid):* We build a hybrid strategy by considering criterion 1 and criterion 2 together for striking a proper balance between the informativeness and representativeness criteria to reach the maximum effectiveness on speech summarization. We choose  $K$  documents with the highest value of  $\text{Score}(D)$ , defined as  $\text{Score}(D) = \text{Score}_{\text{info}}(D) * [\text{Score}_{\text{repre}}(D)]^\lambda$  from  $U\{i\}$  to  $X\{i\}$  for annotation, where  $\lambda$  ( $0 \leq \lambda \leq 1$ ) is tuned by evaluating the development set sentences to optimize the active learning algorithm.

## 28.3 Meeting Corpus, Features, and Summarizers

Our parliamentary meeting speech corpus contains meeting audio files, the Hansard transcriptions, and the meeting minutes from the Hong Kong Legislative Council. For our experiments, we use all 70 ordinary session meeting data from the year 2008 and the year 2009, including audio files, Hansards, and minutes.

We represent each sentence by a feature vector which consists of acoustic features and linguistic features as follows.

Similar to text summarization, the linguistic information can help us predict the summary sentences. We extract eight linguistic features from transcribed documents: len I (the number of words in the sentence), len II/len III (the previous/next sentence's len I value), TFIDF (the summation of the  $\text{tf} * \text{idf}$  value of each word in the sentence), and cosine (cosine similarity measure between two sentence vectors).

We extract all linguistic features from the manual and ASR-transcribed documents respectively. We segment Chinese words of these transcribed documents for calculating length features. We use an off-the-shelf Chinese lexical analysis system, the open source HIT IR Lab Chinese Word Segmenter [5] to process our corpora.

We had a detailed study of linguistic features based on the word segmentation result. We find that the errors produced by the word segmentation process have little effect on the summarization process. We also find the number of words as a feature is better than the number of characters because word can convey more unambiguous information.

Acoustic/prosodic features in speech summarization system are usually extracted from audio data. Researchers commonly use acoustic/prosodic variation—changes in pitch, intensity, speaking rate—and duration of pause for tagging the important contents of their speeches. We also investigate these features for their efficiency in predicting summary sentences of presentation speech and meeting speech.

Our acoustic feature set contains 12 features: Duration I (time duration of the sentence), SpeakingRate (average syllable duration), F0 (I–V) (F0 min, max, mean, slope, range), and E (I–V) (energy min, max, mean, slope, range).

We calculate Duration I from the annotated manual transcriptions that align the audio documents. We then obtain SpeakingRate by phonetic forced alignment by HTK [6]. Next, we extract F0 features and energy features from audio data by using the Praat tool [6].

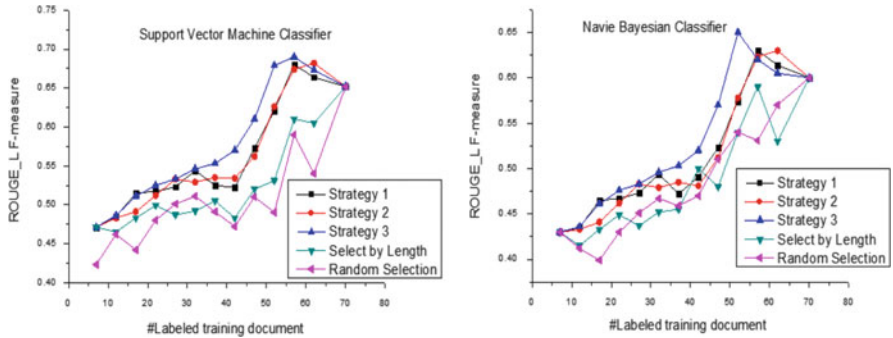
We then build a discriminative model—support vector machine (SVM) classifier as our summarizer based on these sentence feature vectors.

We apply a generative model—Naive Bayesian classifier implemented by ourselves for the summarizer to investigate whether our active learning method is affected by choice of different machine learning methods.

## 28.4 Experimental Results and Evaluation

We perform tenfold cross-validation experiments on manual transcriptions, one group for lecture speeches and another one for meeting speeches. First, we divide the 70 meeting speeches into ten subsets. Each subset has seven ones. We use nine subsets as the unlabeled data pool and use the remaining one for test.

We start our experiments by randomly choosing seven speeches from the unlabeled data pool as seeds for manual labeling. We then train the initial classifier using the seeds. We gradually increase the training data pool by choosing five more speeches from the unlabeled data pool each time for annotation. We separately carry out ten sets of experiments for each genre of speech for comparison according to the above three strategies. For each sample selection step, we also choose the unlabeled documents by random selection baseline strategy and length selection baseline strategy. In each group of experiment, we repeat the process of randomly choosing seven speeches as seeds, six times. We evaluate the summarizer by ROUGE-L (summary-level longest common subsequence) F-measure [5]. We calculate the mean value of ROUGE-L F-measure of each active learning step as the final performance of this step. For meeting speeches, all the evaluation results of the three active learning strategies and the two baseline strategies using the SVM



**Fig. 28.1** Average performance of the SVM classifier and the Naive Bayesian classifier trained by using the combination of acoustic features and linguistic features on parliamentary meeting speech corpus

classifier and the Naive Bayesian classifier are respectively shown in Fig. 28.1. We train the classifiers by using a combination of linguistic features and acoustic features.

We consider the extractive summarization problem as a binary classification problem. From Fig. 28.1, we find that for classification performance, random sample selection is always consistently and significantly outperformed by the active learning strategy using our proposed sample selection strategies. By using only 37 documents for training, the performance of the SVM classifier achieved by strategy 3 which combines the informativeness criterion and representativeness criterion for selecting samples is better than that of the SVM classifier trained by random sample selection using all 70 presentations (ROUGE-L F-measure of 0.512 vs. that of 0.502). This shows that our active learning approach requires 40 % less training data. We also find that using strategy 3, our active learning approach requires less training data to obtain the same level of summarization performance compared to that using strategy 1. In other words, strategy 3 is better than strategy 1. When we apply Naive Bayesian classifier as our summarizer, we obtain the same finding as that shown in Fig. 28.1. Besides, we find that the length sample selection method produces worse performance than our active learning sample selection strategies. This indicates that the performance gains of the active learning methods are due to their ability to choose the unlabeled transcribed documents which contain more informative and representative samples which help find more accurate classification hyperplane and improve the summarization performance. However, the random sample selection strategy and the length sample selection strategy cannot guarantee selecting those documents with more informative and representative samples.

We compare a discriminative model—SVM classifier with a generative model—Naive Bayesian classifier as shown in Fig. 28.1 to investigate the effectiveness of active learning using different kinds of machine learning methods. We find that our active learning method of meeting speech summarization is obviously affected

by choice of different classifiers. For meeting speech corpus, the best performance of SVM classifier is absolute a 4 % higher than that of the Naive Bayesian classifier (0.69 vs. 0.65).

## 28.5 Conclusion

This paper described an approach of active learning to reduce the need for human annotation for summarizing parliamentary meeting speech. We chose the unlabeled examples according to a combination of informativeness criterion and representativeness criterion. The summarization results showed an increasing learning curve, consistently higher than that by using randomly chosen training samples. Furthermore, the experiments showed that the proposed active learning method of meeting speech summarization was obviously affected by choice of different classifiers.

**Acknowledgements** Supported by the Natural Science Foundation of Guangdong Province of China (Grant No. S2012040007560), the Foundation of Guangdong Educational Committee (Grant No. 2012KJCX0099), and the National Natural Science Foundation of China (Grant No. 61300197).

## References

1. Fujii, Y., Yamamoto, K., Kitaoka, N., & Nakagawa, S. (2008). *Class lecture summarization taking into account consecutiveness of important sentences* (pp. 2438–2441). In Proceedings of Interspeech, IEEE.
2. Hori, C., & Furui, S. (2001). *Advances in automatic speech summarization* (pp. 1771–1774). In Proceedings of Eurospeech 2001.
3. Mrozinski, J., Whittaker, E., Chatain, P., & Furui, S. (2005). *Automatic sentence segmentation of speech for automatic summarization* (Vol. 1, Issue 5, p. 12). In Proceedings of ICASSP.
4. Kawahara, T., Nanjo, H., & Furui, S. (2001). *Automatic transcription of spontaneous lecture speech* (pp. 186–189). In Proceedings of the IEEE Workshop on Automatic Speech Recognition and Understanding, IEEE.
5. Zhang, J., Huang, S., & Fung, P. (2008). *RSHMM++ for extractive lecture speech summarization* (pp. 161–164). In Proceedings of 2008 I.E. Workshop on Spoken Language Technology, IEEE.
6. Zhang, J., & Fung, P. (2009). *Active learning of extractive reference summaries for lecture speech summarization* (pp. 23–26). In Proceedings of the 2nd Workshop on Building and Using Comparable Corpora (BUCC), Association for Computational Linguistics.
7. Schohn, G., & Cohn, D. (2000). *Less is more: Active learning with support vector machines* (pp. 839–846). In Machine Learning-International Workshop THEN Conference.
8. Tong, S., & Koller, D. (2002). Support vector machine active learning with applications to text classification. *The Journal of Machine Learning Research*, 2(1), 45–66.
9. Lewis, D., & Catlett, J. (1994). Heterogeneous uncertainty sampling for supervised learning. In *Proceedings of the eleventh international conference on machine learning* (pp. 148–156). Morgan Kaufmann.
10. Settles, B. (2009). *Active learning literature survey* (pp. 1648–1715). University of Wisconsin-Madison. Computer Sciences Technical Report.

# Chapter 29

## Application of Improved BP Neural Network in the Frequency Identification of Piano Tone

Xu Chen and Jun Tang

**Abstract** For the problems existing in the identification process of piano tone, this paper puts forward an MFCC-based (Mel Frequency Cepstrum Coefficient) feature extraction algorithm and a new piano tone identification method with BP neural network as the matching model. Using the MFCC feature extraction algorithm to extract parameters is a good alternative, which could improve the identification rate. Regarding the improved BP neural network as the matching model of tone identification consumes moderate training time and owns high recognition rate. Simulation results show that the piano tone identification combining the BP neural network with the MFCC algorithm is simple, fast, and highly accurate.

### 29.1 Introduction

With the rapid development of multimedia technology, using computer and electronic technology to carry out feature extraction, tone recognition and auxiliary music creation have been given more and more attention [1]. If the results of computer and electronic technologies could be used in music appreciation, learning, and creative fields [2], using computer to automatically identify the music and complete automatic evaluation not only can reduce the work intensity of music educators but also is conducive to improve the scientificity, impartiality, and objectivity of music evaluation.

As a branch of speech recognition [3], tone recognition technology has been paid much attention to in recent years. From the current research status, to construct the tone signal recognition system, analog circuits or digital circuits are frequently used to conduct filtering, adding, Fourier transform, spectral analysis and autocorrelation analysis, and so on for signals. The identification rate has been the greatest

---

X. Chen (✉) • J. Tang  
Hunan Urban Construction College, Xiangtan 411101, China  
e-mail: [xtchenxu@163.com](mailto:xtchenxu@163.com); [xttangjun@163.com](mailto:xttangjun@163.com)

impediment to limit its development. All the piano tones contain the information of 88 keys, while the speech recognition requires a very large amount of data for model training and match to ensure the recognition rate [4]. From every point of view, the theoretical property and practical operability of the tone recognition are better than speech recognition. Therefore, it is entirely possible to use speech recognition theory, combined with the characteristics of tone recognition, to study and process the tone recognition [5].

In this paper, MFCC (Mel Frequency Cepstrum Coefficient) [6] and BP (Back Propagation) neural network algorithm are regarded as the research basis to complete a set of feasible musical note recognition methods [7]. In MATLAB, the aforementioned algorithm is used to test the single key tone, and the simulation result is good; subsequently, validate the system by continuous tones. The results show that the algorithm can be used for a wide range of tone identification fields, which will promote the development of tone recognition.

## 29.2 Piano Tone Identification Methods

### 29.2.1 Physical Characteristics of Piano Tone

Set the tension of the piano string as  $T_0$ , the density as  $\rho$ , the cross-sectional area as  $A$ , and the length as  $l$ . The string is divided into  $n + 1$  segments, and the number of the division points is  $n$ . Then we can obtain the free response function of the string, as shown in Eq. (29.1).

$$y(x, t) = \sum_{n=1}^{\infty} \left\{ \frac{2}{l} \left[ \int_0^l f_1 \sin \frac{n\pi}{l} x dx \right] \cos P_n t + \frac{2}{n c \pi} \left[ \int_0^l f_2(x) \sin \frac{n\pi}{l} x dx \right] \sin P_n t \right\} \sin \frac{P_n}{c} x \quad (29.1)$$

The natural frequency  $P_n$  of the string is as shown in Eq. (29.2).

$$P_n = \frac{n\pi}{l} \sqrt{\frac{T_0}{\rho A}} \quad (n = 1, 2, 3, \dots) \quad (29.2)$$

When  $n = 1$ ,  $P_1 = \frac{n\pi}{l} \sqrt{\frac{T_0}{\rho A}}$  is called the baseband of string which determines the pitch of the string, the main frequency to be identified in this paper; when  $n = 2, 3, 4, \dots$ , the frequency  $P_n$  is called overtone and its value is an integer multiple of the frequency.

The frequency domain constitution of single notes (such as every piano key) is basically the same in the frequency segment, and the difference lies in the linear



**Fig. 29.1** Diagram of 88-key piano keyboard

combination between the harmonics. During the whole process from signal to signal generation to the signal ending, the physical form of harmonics and dominant frequency is smoothly weakened. That is, the composition is constant, and it is the direct amplitude of waveforms that takes changes, weakened from the maximum to zero [8].

### ***29.2.2 Music Theory Analysis of Piano Tone***

A piano consists of 88 keys, with 7 full octaves. The piano keyboard is shown in Fig. 29.1.

The piano musical alphabet is obtained in accordance with the law of averages of the 12-tone system [9]. Each octave is divided into a frequency doubling, and musical alphabet between each adjacent octave is a multiple relationship in which the frequency ratio of musical alphabet with the difference of one frequency doubling is 2. Each octave is divided into 12 half-degree syllables. The main frequencies of 88-key piano are shown in Table 29.1.

### ***29.2.3 Identification Method of Piano Music***

The recognition process of piano tone includes the acquisition and digitization, preprocessing, and feature extraction of musical tone signals and the study of tone recognition model.

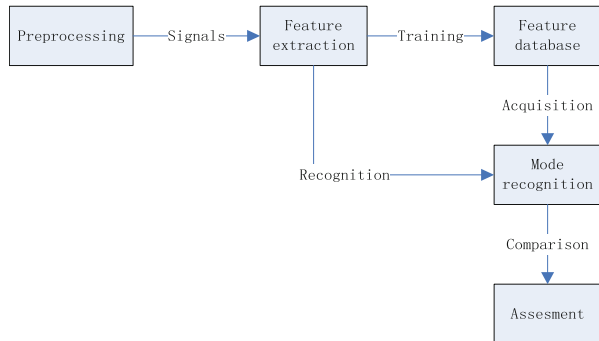
A special microphone and sound capture card are used to acquire a collection of 88 keys of the piano five times, with 440 tone data saved in WAV format. Sound capture card with 16-bit accuracy and sampling frequency of 22,050 Hz is used to digitize the signal for the operating system [10].

Use the feature extraction algorithm of MFCC and methods based on BP neural network recognition model to analyze and design a system for tone recognition starting from the analysis of the unique performance of tones in combination with a variety of recognition models of speech recognition system. Departing from the physical and musical characteristics of the musical tone, the main frequency and the musical alphabet of the tone signal are identified to complete the entire recognition system design. The whole tone recognition process is shown in Fig. 29.2.

**Table 29.1** Main frequency of 88-key piano

Sequence number	Main frequency	Frequency (Hz)
1	$c^1$	261.63
2	$d^1$	293.7
3	$e^1$	329.63
4	$f^1$	349.23
5	$g^1$	392.0
6	$a^1$	440
7	$b^1$	493.88

**Fig. 29.2** The whole tone recognition process



### 29.3 Mel Frequency Cepstrum Coefficient

The MFCC is used to conduct transformation for the spectrum of the time domain by the nonlinear spectrum and carry out studies after it is ultimately transformed in cepstrum spectrum. The specific calculation processes are shown as follows:

1. Conduct windowing frame selection for the tone signal to change the tone signal to the short-term signal; then adopt the Fast Fourier Transform to transform the time domain signal  $x(n)$  into a frequency domain signal  $X(m)$ , and calculate the short-time energy spectrum on this basis  $P(f)$ .
2. By the formula (29.3), the frequency spectrum of short-term energy spectrum  $P(f)$  in the frequency spectrum axis is transformed into the  $P(M)$  in the Mel axis.

$$F_{mel} = 2595 \lg(1 + f_{HZ}/700) \tag{29.3}$$

3. Set M band-pass filters  $H_m(k)$  in the spectral range of the tone, and calculate the input of band-pass filter with the energy spectrum  $P(M)$  as input, as shown in Eq. (29.4).

$$\theta(M_k) = \ln \left[ \sum_{k=1}^N |X(k)|^2 H_m(k) \right] \quad m = 1, 2, \dots, M \quad N = 5,000 \tag{29.4}$$



In this equation,  $N$  is the width of the windowed pour election. Set the coverage area of each filter is 0–5,000 Hz and they have the same MEL characteristics, namely, the span of filter group in the Mel coordinates is equal. Set the center frequency as  $f(m)$ . Based on the corresponding relationship between Mel coordinates and the time field coordinates, the span of band-pass filters in the time domain coordinates increases with the rise of  $m$ , i.e., the spacing between adjacent  $f(m)$  is continuously increasing. The transfer function of band-pass filter system is as shown in Eq. (29.5).

$$H_m(k) = \begin{cases} 0 & k < f(m - 1) \\ \frac{2[k - f(m - 1)]}{[f(m) - f(m - 1)][f(m + 1) - f(m - 1)]} & f(m - 1) \leq k \leq f(m) \\ \frac{2[f(m + 1) - k]}{[f(m + 1) - f(m)][f(m + 1) - f(m - 1)]} & f(m) \leq k \leq f(m + 1) \\ 0 & k > f(m + 1) \end{cases} \quad (29.5)$$

The definition of center frequency  $f(m)$  is as shown in Eq. (29.6):

$$f(m) = \frac{N}{F_s} F_{mel}^{-1} \left[ F_{mel} \left( f_l + m \frac{F_{mel}(f_h) - F_{mel}(f_l)}{M + 1} \right) \right] \quad (29.6)$$

wherein,  $f_l, f_h$ , respectively, represents the lowest frequency and highest frequency in the filter frequency range,  $F_s$  is the sampling frequency, and  $F_{mel}^{-1}$  is the inverse function of  $F_{mel}$ ,  $F_{mel}^{-1}(b) = 700(e^{b/2595} - 1)$ .

4. Output energy of the  $m$ th filter, noted as  $\theta(M_k)$ . Use the modified discrete cosine transform (IDCT) to obtain the Mel frequency cepstral  $c(n)$  on the left side of the Mel, as shown in Eq. (29.7).

$$c(n) = \sum_{k=1}^M \theta(M_k) \cos \left( n(k - 0.5) \frac{\pi}{M} \right) \quad n = 1, 2, \dots, P \quad (29.7)$$

In this equation,  $P$  is the order of MFCC parameters, and in this system, it is set as 12 orders.

### 29.4 BP Neural Network Model

12 neurons are adopted in input layer to correspond to the 12-dimensional characteristic parameters of the tone signal; 1 neuron is adopted in output layer, corresponding to the main frequency of piano tone; the number of hidden layer neurons is determined through experimental comparison.

MLP mode predictor in BP network is selected as the system neuron. Its working principle is to use the neuron output  $f(t - \tau), \dots, f(t - 1)$  before time  $t$  as the input of time  $t$  to give the output of feature parameter  $\hat{f}(t)$  at this time.

The relationship between the feature input and prediction output is as shown in Eqs. (29.8) and (29.9).

$$h(t) = F\left(\sum_{s=1}^r \omega_s^{(1)} f(t-s)\right) \quad (29.8)$$

$$\hat{f}(t) = \omega_0^{(2)} h(t) \quad (29.9)$$

In the equation,  $\omega_s^{(1)}$  and  $\omega_0^{(2)}$  are weight matrixes between neurons, and  $h(t)$  is used as the output of the hidden layer, while  $f(t)$  refers to the sigmoid function.

Each neuron network node is replaced by an MLP forecast period. Assuming that there are  $N_n$  predictive models, and for musical tone recognition, we can split it into  $N_n$  sides and use the corresponding forecasting models to predict each piece of musical tone signals. Finally, the optimization of musical tone segmentation is achieved through the calculation of a minimum cumulative prediction residual, as shown in Eq. (29.10).

$$D(n) = \min_{\{i(t)\}} \sum_{t=1}^T \left\| \hat{f}(n, i(t)) - f(t) \right\|^2 \quad (29.10)$$

In the formula,  $\hat{f}(n, i(t))$  represents the feature component predicted by  $s$  the  $i$ th predictor. The recurrence of algorithm is as shown in Eqs. (29.11), (29.12), and (29.13).

$$g(t, i) = d(t, i) + \min\{g(t-1, i), g(t-1, i-1)\} \quad (29.11)$$

$$d(t, i) = \left\| \hat{f}(t, i) - f(t) \right\|^2 \quad (29.12)$$

$$D(n) = g(T, Nn) \quad (29.13)$$

$D(n)$  is the total distance between the input musical tone signal and the template, and the tone recognition will be completed if the minimum  $D(n)$  is obtained.

The mean of accumulated prediction residual can be used to judge, as shown in Eq. (29.14).

$$\bar{D}(n) = \frac{1}{M} \sum_{m=1}^M D(n, m) \quad (29.14)$$

In the equation,  $M$  represents the number of samples in the training sample database.

## 29.5 Experimental Results

In light of the fact that the piano keys contained only 88 keys, the data is relatively less than normal, so we extracted the musical tone of each key in different environments five times to get a total of 440 sets of data constructing the input unit the neural network, while the output unit of training is the main frequency value corresponding to each note. After that, the input data is normalized, which enables the input of certain exceptions to fall within the expected range, enhancing the adaptability of the system to the musical tone data.

Through the setting of relevant decomposition filter bank in MATLAB, Daubechies4 order wavelet decomposition is conducted on musical signals. MFCC is used to extract the feature parameters of music, including the use of a 12-order extraction algorithm, 256 audio signal FFT transform length, 20,500 Hz sampling frequency, and the frame length of 256. First, read the preprocessed musical tone file mentioned above; and then conduct the normalization of Mel filter bank coefficients by the programmed melbankm function, followed by the solution to DCT parameters, the filtering of the preemphasis filter, as well as the reframing of the musical tone signal; and finally calculate the MFCC parameters of musical tone signals. The 8-order MFCC parameters of “A11.wav” are shown in Table 29.2.

After obtaining the trained BP neural network model, we randomly selected eight musical tone signals of keys to have them verified and corrected again.

As can be seen from Table 29.3, musical tone signals whose BP neural network matches are conducted using MFCC parameters as the feature parameters have the actual value of the signals consistent with the expected value, which shows a relatively high recognition rate.

**Table 29.2** The 12-order MFCC parameters of “A11.wav”

Sequence number	Feature parameters					
1	50.214	8.991	21.365	1.5712	16.298	4.709
	20.287	4.4121	5.7611	2.6902	5.021	0.1187
2	51.098	9.6791	24.461	2.8612	16.675	3.1981
	17.2016	1.0015	9.1417	0.4098	4.1005	0.01941
3	49.4572	7.7453	26.7091	4.5546	16.4416	2.8717
	14.4908	4.0198	13.1121	1.7789	2.9103	0.7659
4	54.5476	13.1154	22.3224	2.7621	18.5532	1.9012
	11.3910	2.5846	10.2217	0.50012	4.4199	0.1876

**Table 29.3** Recognition results of partial musical tones of 88-key piano

Actual musical alphabet	Expected value	Actual value	Error (%)	Expected musical alphabet
$e^4$	2,594	2,561	98.73	$e^4$
# $A_2$	28.91	27.83	96.26	# $A_2$
$A_1$	55	53.92	98.04	$A_1$
# $A$	116	114.87	99.03	# $A$
$e^1$	337	336.13	99.74	$e^1$
$d$	155	155.72	99.54	$d$
$b^2$	988	987.06	99.90	$b^2$
$c^2$	523	521.37	99.69	$c^2$

## 29.6 Conclusion

This paper proposed piano tone recognition methods including signal de-noising of high-frequency tone, endpoint detection, single note segmentation, feature parameters extraction, and recognition of model matching. The MFCC method with good recognition performance and noise resistance has been adopted to extract feature parameters of piano musical tone, improving the efficiency and accuracy of feature extraction. With an improved BP neural network model to identify the piano musical tone, a highly efficient neural network was built. The experimental analysis proves the correctness and feasibility of musical tone recognition system which improves the recognition rate.

**Acknowledgement** This paper is supported by Scientific Research Fund of Hunan Provincial Education Department (11C0231, 12C0995).

## References

1. Abeyratne, U. R., Kinouchi, Y., Oki, H., Okada, J., Shichijo, F., & Matsumoto, K. (1991). Artificial neural networks for source localization in the human brain. *Brain Topography*, 4(1), 3–21.
2. Zhang, Q., Nagashino, H., & Kinouchi, Y. (2003). Accuracy of single dipole source localization by BP neural networks from 18-channel EEGs. *IEICE Transactions on Information and Systems*, 86(4), 1447–1455.
3. Anguita, D., Parodi, G., & Zunino, R. (1994). An efficient implementation of BP on RISC-based workstations. *Neurocomputing*, 6(2), 57–65.
4. Ye, H., & Young, S. (2006). Quality-enhanced voice morphing using maximum likelihood transformations. *IEEE Transactions on Audio, Speech, and Language Processing*, 14(4), 1301–1312.
5. Curry, B., & Morgan, P. (1997). Neural networks: A need for caution. *Omega, International Journal of Management Sciences*, 25(8), 123–133.
6. Chen, D., & Mohler, R. R. (2003). Neural-network-based load modeling and its use in voltage stability analysis. *IEEE Transactions on Control Systems Technology*, 11(4), 460–470.

7. Baldi, P. (1996). Gradient descent learning algorithm overview: A general dynamical systems perspective. *IEEE Transactions on Neural Networks*, 6(1), 182–195.
8. Moller, M. (1993). A scaled conjugate gradient algorithm for fast supervised learning. *Neural Networks*, 6(10), 525–533.
9. Hagan, M., & Menhaj, M. (1994). Training feed-forward networks with the Marquadt algorithm. *IEEE Transactions on Neural Networks*, 5(6), 152–157.
10. Storn, R., & Price, K. (1997). Differential evolution — a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11(6), 341–359.

# **Part II**

## **Data Processing**

# Chapter 30

## Implicit Factoring with Shared Middle Discrete Bits

Meng Shi, Xianghui Liu, and Wenbao Han

**Abstract** We study the problem of implicit factoring presented by May and Ritzenhofen in 2009 and apply it to more general settings, where prime factors of both integers are only known by implicit information of middle discrete bits. Consider two integers  $N_1 = p_1q_1$  and  $N_2 = p_2q_2$  where  $p_1, p_2, q_1,$  and  $q_2$  are primes and  $q_1, q_2 \approx N^\alpha$ . In the case of  $t \log_2 N$  bits shared in one consecutive middle block, we describe a novel lattice-based method that leads to the factorization of two integers in polynomial time as soon as  $t > 4\alpha$ . Moreover, we use much lower lattice dimensions and obtain a great speedup. Subsequently, we heuristically generalize the method to an arbitrary number  $n$  of shared blocks. The experimental results show that the constructed lattices work well in practical attacks.

### 30.1 Introduction

Efficient factoring of large integers is one of the most fundamental problems in algorithmic number theory and plays an essential role in RSA's security. Since RSA system was proposed in 1977, many researchers have been trying to solve large integer factoring problem. For classical computing model, the quadratic sieve, the elliptic curve method, and the number field sieve have been presented to optimize the factorization complexity. It is not known if there exist integer factorization algorithms whose running time is polynomial in the bit size of  $N$ .

In 2009, May and Ritzenhofen first presented the large integer factorization problem with implicit information [1]. Consider two integers  $N_1 = p_1q_1$  and  $N_2 = p_2q_2$ , where  $p_1, p_2, q_1, q_2$  are primes and  $p_1, p_2$  share  $t \log_2 N$  least significant bits (LSBs). They proved that when  $q_1, q_2 \approx N^\alpha$ ,  $N_1, N_2$  can be factored in polynomial time if  $t \geq 2\alpha$ . These results can further be extended to

---

M. Shi (✉) • X. Liu • W. Han

Department of Applied Mathematics, Zhengzhou Information Science and Technology Institute, Zhengzhou 450002, China  
e-mail: [bagoubj@yahoo.cn](mailto:bagoubj@yahoo.cn)

the situation when  $N_1 = p_1q_1$ ,  $N_2 = p_2q_2, \dots, N_k = p_kq_k$ , and all the  $p_i$ 's share  $t \log 2N$  LSBs. Note that the additional information here is only implicit: the attacker only knows that  $p_1$  and  $p_2$  share them. Subsequently, Sarkar and Maitra did further research applying Coppersmith and Gröbner-basis techniques [2]. Contrary to May's result, their method works when  $p_1, p_2$  share either  $t \log 2N$  LSBs or MSBs or bits in the middle. In 2010, Faugere, Marinier, and Renault presented a method different from previous one and improved their conclusions [3].

However, this attack works on a condition that  $p_i$  is sharing the MSBs, LSBs, or bits in the middle. Unfortunately, the unknown part is in general not located in one consecutive bit block but widely spread over the whole bit string [4]. Hence, whether we can extend our method to the general scenario is a work worth researching.

*Our contribution:* Instead of the strategy of finding roots of multivariate polynomials applied in Sarkar's results [2], we transformed this problem into finding small modular roots to reduce lattice dimensions. Table 30.1 compares our results against the previous results [1–3].

What is more is that our theoretical results are supported by experiments. We have written the corresponding programs in VC++6.0 over WINXP SP3 on computer with Intel(R) Pentium(R) 4 3.00 GHz CPU, 512M RAM.

The paper is organized as follows. In Sect. 30.2, we briefly discuss some background information. Next, in Sect. 30.3 we present our strategy in the case when  $p_1, p_2$  share a contiguous portion of bits at the middle and describe the complete experimental details with comparison of existing works. Then, in Sect. 30.4 we extend the method to the arbitrary contiguous portions of bits. In Sect. 30.5 we conclude the paper.

## 30.2 Preliminaries

In this section, we briefly present some basics on basis reduction in lattice [5–7].

**Definition 1** Let  $b_1, b_2, \dots, b_m \in R^n$  be linearly independent (row) vectors, where  $n$  and  $m$  are integers such that  $n \geq m$ . A lattice  $L$  is described as the set of vectors in  $R^n$  that are integer linear combinations of the basis vectors  $b_1, b_2, \dots, b_m$ . Formally,

$$L(b_1, b_2, \dots, b_m) = \left\{ z = \sum_{i=1}^m \lambda_i b_i \mid \lambda_i \in Z \right\}.$$

When given a lattice  $L$ , finding the shortest vector in a basis of  $L$  has been proved to be an NP-hard problem. In 1982, Lenstra, Lenstra, and Lovasz proposed the famous LLL algorithm for lattice basis reduction [8], which can produce a reduced basis in polynomial time. Here is the result:



**Table 30.1** Comparison of our results against the previous results

	May's results [1]	Sarkar's results[2]	Faugere's results [3]	Our results
$k = 2$	share $r \log_2 N$ LSBs: $t \geq 2\alpha + 3$ using two-dimensional lattices	share $r \log_2 N$ LSBs or MSBs: better than $t \geq 2\alpha + 3$ when $\alpha \geq 0.266n$	share $r \log_2 N$ MSBs: $t \geq 2\alpha + 3$ using two-dimensional lattices	share $r \log_2 N$ bits in the middle: $t > 4\alpha$ using eight-dimensional lattices
		share $r \log_2 N$ bits in the middle: $t \geq 4\alpha + 7$ but depending on the position of the shared bits using 46-dimensional lattices	share $r \log_2 N$ bits in the middle: $t \geq 4\alpha + 7$ using three-dimensional lattices	share $r \log_2 N$ discrete bits in the middle: $t > 7\alpha$ using 12-dimensional lattices
$k > 2$	Can be applied	Cannot be applied	Can be applied	Cannot be applied

**Lemma 1** *Let  $L$  be a lattice of dimension  $m$ . The reduced basis vectors  $\{v_1, v_2, \dots, v_m\}$  that the LLL algorithm outputs satisfy*

$$\|v_1\| \leq \|v_2\| \leq \dots \leq \|v_i\| \leq 2^{\frac{\omega(\omega-1)+(i-1)(i-2)}{4(\omega-i+1)}} \det(L)^{\frac{1}{\omega-i+1}}, \quad \text{for all } 1 \leq i \leq \omega.$$

**Definition 2** Given a polynomial  $|f(x_1, x_2, \dots, x_n)| = \sum_{i_1, i_2, \dots, i_n} a_{i_1, i_2, \dots, i_n} x_1^{i_1} x_2^{i_2} \dots x_n^{i_n}$ ,

we define the norm as  $\|f(x_1, x_2, \dots, x_n)\| = \left( \sum_{i_1, i_2, \dots, i_n} a^2_{i_1, i_2, \dots, i_n} \right)^{1/2}$ .

In 1998, Howgrave-Graham reformulated Coppersmith’s method and proposed a theorem as follows [9]:

**Lemma 2** *Let  $h(x_1, x_2, \dots, x_n) \in \mathbb{Z}[x_1, x_2, \dots, x_n]$  be an integer polynomial that consists of at most  $w$  monomials. Suppose that*

- (1)  $h(x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)}) = 0 \pmod{N^m}$  for some  $|x_1^{(0)}| < X_1, \dots, |x_n^{(0)}| < X_n$  and
- (2)  $\|h(X_1 x_1, X_2 x_2, \dots, X_n x_n)\| < N^m / \sqrt{w}$

*then  $h(x_1^{(0)}, x_2^{(0)}, \dots, x_n^{(0)}) = 0$  holds over the integers.*

Lemma 2 plays an important role in this paper. A common root of  $i$  variables can be extracted efficiently if the  $i$  polynomials are algebraically independent. Therefore, we have encountered the following heuristic. This heuristic often works perfectly.

**Assumption 1** *The resultant computations for the multivariate polynomials constructed in our approaches yield nonzero polynomials.*

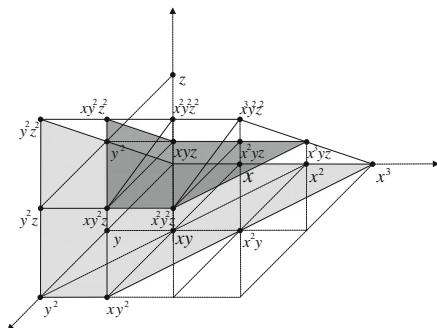
### 30.3 Primes $p_1, p_2$ Share a Contiguous Portion of Bits at the Middle

Throughout this paper, we assume  $p_1, p_2$  are primes of same bit size and  $q_1, q_2$  are primes of same bit size. Thus  $N_1 = p_1 q_1$  and  $N_2 = p_2 q_2$  are also of same bit size. We use  $N$  to represent an integer of same bit size as of  $N_1, N_2$ . Assuming that we are given two different module  $N_1 = p_1 q_1$  and  $N_2 = p_2 q_2$ , where contiguous portion of bits of  $p_1, p_2$  are same at the middle leaving the  $\gamma_1 \log_2 N$  many MSBs and  $\gamma_2 \log_2 N$  many LSBs, i.e.,  $p_1 = N^{1-\alpha-\gamma_1} p_{10} + N^{\gamma_2} p_{11} + p_{12}$  and  $p_2 = N^{1-\alpha-\gamma_1} p_{20} + N^{\gamma_2} p_{21} + p_{22}$  for some common  $p_{11}$  that is unknown to us. In this section, we will show that there is an algorithm that recovers the factorization of  $N_1$  and  $N_2$  in polynomial time provided that  $t = 1 - \gamma_1 - \gamma_2 - \alpha$  is sufficiently large.

We start with

$$p_1 - p_2 = N^{1-\alpha-\gamma_1} (p_{10} - p_{20}) + (p_{12} - p_{22}) \tag{30.1}$$

**Fig. 30.1** Example of  $M_k M_{k+1}$  when  $m = 2$ ,  $d = 1$



By reducing this equation modulo  $N^{1-\alpha-\gamma_1}$ , we can eliminate the two unknowns and get

$$p_1 - p_2 = p_{12} - p_{22} \text{mod} N^{1-\alpha-\gamma_1} \tag{30.2}$$

Since  $p_1 = N_1/q_1$  and  $p_2 = N_2/q_2$ , we have

$$N_1q_2 - N_2q_1 - N^{1-\alpha-\gamma_1}(p_{10} - p_{20})q_1q_2 - (p_{12} - p_{22})q_1q_2 = 0 \tag{30.3}$$

Then we can rewrite the equation modulo  $N^{1-\alpha-\gamma_1}$  to

$$N_1q_2 - N_2q_1 - (p_{12} - p_{22})q_1q_2 = 0 \text{mod} N^{1-\alpha-\gamma_1} \tag{30.4}$$

Thus we need to solve the following polynomial:

$$\begin{aligned} f_N(x, y, z) &= f'_N(x - 1, y, z) = N_1x - N_2y - xyz - N_1 - yz \\ &= 0 \text{mod} N^{1-\alpha-\gamma_1} \end{aligned} \tag{30.5}$$

The root  $(x_0, y_0, z_0)$  of  $f$  is  $(q_2 - 1, q_1, p_{11} - p_{22})$ . Let  $X = N^\alpha, Y = N^\alpha, Z = N^{\gamma_2}$ , then we can take  $X, Y, Z$  as the upper bound of  $x_0, y_0, z_0$  respectively.

Hence, we can get Theorem 1.

**Theorem 1** Let  $N_1 = p_1q_1$  and  $N_2 = p_2q_2$ , where  $p_1, p_2, q_1, q_2$  are primes. Let  $q_1, q_2 \approx N^\alpha$ . If a contiguous portion of bits of  $p_1, p_2$  are same at the middle leaving the  $\gamma_1 \log_2 N$  many MSBs and  $\gamma_2 \log_2 N$  many LSBs, we can factor both  $N_1, N_2$  if  $t > 4\alpha$  under Assumption 1.

*Proof* Let us discuss in general how small a root  $(x_0, y_0, z_0)$  of a polynomial

$f_N(x, y, z) = a_1x + a_2y + a_3xyz + a_4yz + a_5 = 0 \text{mod} N^{1-\alpha-\gamma_1}$  could be such that it can be found by Coppersmith method.

Jochemsz described the extended strategy [5]. So it can be prescribed as follows:

$$M_k = \bigcup_{0 \leq j \leq d} \left\{ x^{i_1+j} y^{i_2} z^{i_3} \mid x^{i_1} y^{i_2} z^{i_3} \text{ monomial of } f_N^m \text{ and } \frac{x^{i_1} y^{i_2} z^{i_3}}{l^k} \text{ monomial of } f_N^{m-k} \right\}$$

for  $l = xyz$ ,  $d = \tau m$ , and  $\tau > 0$ .

Next, we define the following shift polynomials:

$$g_{i_1 i_2 i_3 k}(x, y, z) = \frac{x^{i_1} y^{i_2} z^{i_3}}{l^k} f_N^k N^{(1-\alpha-\gamma_1)(m-k)}$$

for  $k = 0, \dots, m$ , and  $x^{i_1} y^{i_2} z^{i_3} \in M_k \setminus M_{k+1}$ .

Figure 30.1 shows that the new sets  $M_k$  when  $m = 2, d = 1$ . Points of the large shadow belong to  $M_0 \setminus M_1$ , and points of the small shadow belong to  $M_1 \setminus M_2$ .

Hence, if we reasonably order the matrix describing the lattice  $L$ , we are sure to get a lower triangular matrix.

We conclude that the condition  $\det(L) < q^{m\omega-\varepsilon}$  reduces to

$$X^{3+8\tau+6\tau^2} Y^{5+8\tau} Z^{3+4\tau} < N^{(1-\alpha-\gamma_1)(2+4\tau)} \tag{30.6}$$

Substituting the values of  $|x| \leq X = N^\alpha, |y| \leq Y = N^\alpha, |z| \leq Z = N^{\gamma_2}$ , if  $1-5\alpha-\gamma_1-\gamma_2 > 0$ , we can solve Eq. (30.6) for the maximal value of  $\tau = \frac{1-\gamma_1-\gamma_2-5\alpha}{3\alpha}$ . Let  $t = 1-\alpha-\gamma_1-\gamma_2$ , we get  $t > 4\alpha$ .

Next, we compare the conclusion of Theorem 1 in this paper and Theorem 2 in May’s paper [1] via experimental data. In this experiment, we still consider 850-bit primes  $p_1$  and  $p_2$ .

$p_1 = 60100632917456734116303555865209875275018541235077520316344103$   
 $389222613252007908441224870562785926663459589770176998171796837046$   
 $320523689341193687536324043967528101378987325022248122037708305617$   
 $873\ 9640056377459864355429485398257581885621415132927137653573.$

$p_2 = 5984825641870931585823382220962926344220670532554403933352105$   
 $675571066672703603259730323516347307682331134375921535484093155453$   
 $66310986033574699693202605043239631638906892532523449394047385276$   
 $94071\ 49342403534694186411407834893900732303380146620012528421339.$

Note that  $p_1, p_2$  share middle 504 many bits (leaving 177 bits from the least significant side).

Further,  $q_1, q_2$  are 150-bit primes

$q_1 = 1038476608131498405684472704928794724111541861$  and  
 $q_2 = 1281887704228770097092001008195142506836912053.$

Given  $N_1, N_2$  with only the implicit information, we can factorize both of them efficiently. We use lattice of dimension 8 (parameters  $m = 1, d = 1$ ) and the lattice reduction costs 15 ms.

From the test we obtain two conclusions. First, the experimental results are better than the theoretical results: the theoretical result of Theorem 1 is  $t > 4\alpha$ , that is, if  $p_1, p_2$  are 850-bit prime numbers and  $q_1, q_2$  are 150-bit prime numbers, then  $N_1, N_2$  can be factored in polynomial time when  $p_1, p_2$  share the middle of 600-bit. However, in our experiment,  $p_1, p_2$  share only the middle 504-bit. Second, because we applied the method of finding small modular roots of multivariate polynomials, the number of variables of equation in Theorem 1 of this paper was smaller than the Theorem 2 in Sarkar’s paper [2] under the same conditions. The experimental results show that compared with the 70-dimensional lattice in Sarkar’s paper [2],  $N_1, N_2$  can be factored more efficiently by constructing an eight-dimensional lattice according to Theorem 1 in this paper.

### 30.4 Sharing More Contiguous Portions of Bits at the Middle

In practice, the attacker can get the implicit information of the prime factors with shared middle discrete bits easily. In this section, we heuristically generalize the method of Theorem 1 to an arbitrary number  $n$  of shared blocks. In fact, the difference between Theorem 3 and the Theorem 2 is that the polynomials contain increased number of variables as well as monomials. For simplicity, let  $n = 2$  first.

**Theorem 2** *Let  $N_1 = p_1q_1$  and  $N_2 = p_2q_2$ , where  $p_1, p_2, q_1, q_2$  are primes. Let  $q_1, q_2 \approx N^\alpha$ . If a contiguous portion of bits of  $p_1, p_2$  is the same at the middle leaving the  $\gamma_1 \log_2 N$  many MSBs,  $\gamma_3 \log_2 N$  many LSBs, and  $\gamma_2 \log_2 N$  many bits in the middle, we assume that the first same part is  $t_1$ -bit and the second same paragraph is  $t_2$ -bit. Then under Assumption 1, we can factor both  $N_1, N_2$  if  $t = t_1 + t_2 > 7\alpha$ .*

*Proof* We can write the following formulas similarly:

$$p_1 = N^{1-\alpha-\gamma_1}p_{10} + N^{1-\alpha-\gamma_1-t_1}p_{11} + N^{\gamma_3+t_2}p_{12} + N^{\gamma_3}p_{13} + p_{14} \tag{30.7}$$

$$p_2 = N^{1-\alpha-\gamma_1}p_{20} + N^{1-\alpha-\gamma_1-t_1}p_{21} + N^{\gamma_3+t_2}p_{22} + N^{\gamma_3}p_{23} + p_{24} \tag{30.8}$$

By reducing we can get

$$N_1q_2 - N_2q_1 - N^{\gamma_3+t_2}(p_{12} - p_{22})q_1q_2 - (p_{14} - p_{24})q_1q_2 = 0 \pmod{N^{1-\alpha-\gamma_1}} \tag{30.9}$$

Therefore, our goal is to find the small root  $(q_2 - 1, q_1, p_{11} - p_{22}, p_{14} - p_{24})$  of the polynomial  $f_N = a_1x + a_2y + a_3xyz + a_4yz + a_5wxy + a_6wy + a_7 = 0 \pmod{N^{1-\alpha-\gamma_1}}$ . We can take  $X = N^\alpha, Y = N^\alpha, Z = N^{\gamma_2}, W = N^{\gamma_3}$  as the upper bound of root respectively.

By calculating, the upper bounds  $X, Y, Z,$  and  $W$  have to satisfy that

$$\begin{aligned}
 & X^{10+20\tau+10\tau^2+20\tau^4+20\tau^5} Y^{15+15\tau+20\tau^4+5\tau^5} Z^{5+5\tau+20\tau^4+15\tau^5} W^{5+5\tau-5\tau^5} \\
 & < N^{(1-\alpha-\gamma_1)(4+5\tau+20\tau^4+9\tau^5)}. \tag{30.10}
 \end{aligned}$$

Substituting the values  $|x| \leq X = N^\alpha, |y| \leq Y = N^\alpha, |z| \leq Z = N^{\gamma_2}, |w| \leq W = N^{\gamma_3}$ , if  $1 - 8\alpha - \gamma_1 - \gamma_2 - \gamma_3 > 0$ , we can solve Eq. (30.10) for the maximal value of  $\tau = \frac{1 - \gamma_1 - \gamma_2 - \gamma_3 - 8\alpha}{4\alpha}$ . Let  $t = 1 - \alpha - \gamma_1 - \gamma_2$ , we get  $t > 7\alpha$ .

Obviously, we can generalize the method of Theorem 2 to an arbitrary number  $n$  of shared blocks.

**Theorem 3** Let  $N_1 = p_1q_1$  and  $N_2 = p_2q_2$ , where  $p_1, p_2, q_1, q_2$  are primes. Let  $q_1, q_2 \approx N^\alpha$ . If a contiguous portion of bits of  $p_1, p_2$  is same at the middle leaving the  $\gamma_1 \log_2 N$  many MSBs,  $\gamma_{n+2} \log_2 N$  many LSBs, and  $\gamma_2 \log_2 N, \dots, \gamma_{n+1} \log_2 N$  many bits in the middle. We assume that the first same part is  $t_1$ -bit, the second same part is  $t_2$ -bit, and the last same part is  $t_{n+1}$ -bit. Then under Assumption 1, we can factor both  $N_1, N_2$  if  $t = \sum_i^{n+1} t_i > 7\alpha$ .

The next example considers the case of primes  $p_1, p_2$  of 900 bits and  $q_1, q_2$  of 100 bits. This is to demonstrate how our method works experimentally for discrete bits. Here we only consider 900-bit primes  $p_1$  and  $p_2$ .

$p_1 = 7505061092492112375563523501603029893103360600496947681215898$   
 $539028978880558579468898992711484001866062693594592629842369133358$   
 $102966297068584054427273763300567492499617802017459636475620491510$   
 $231406148324755843778363996138755147469460323318959606760968500389$   
 $785083814343$  and  
 $p_2 = 4226356249085321971634266775991800272784808798610270102561708$   
 $704362527040205013775291568103094901198355262363763755300814608766$   
 $443080087384559762394512580130940071069071767173778872370100880021$   
 $226567400723767205747855563912712394927148504565667243123047573407$   
 $832687510917$ .

Note that  $p_1, p_2$  share the most bits at the middle expect 95 bits which are distributed from the 1-th to the 60-th bits, from the 322-th to the 396-th bits, and from the 836-th to the 900-th bits. Further,  $q_1, q_2$  are 100-bit primes

$q_1 = 658018347648025712456491553623$  and  
 $q_2 = 684627888228678540762517278787$ , respectively.

Given  $N_1, N_2$  with only the implicit information, we can factorize both of them efficiently. We use lattice of dimensions 12 (parameter  $m = 1, d = 1$ ) and the lattice reduction costs 16 ms.

## 30.5 Conclusion

In this article we have studied polynomial time factorization strategy when  $t \log_2 N$  bits shared in one consecutive middle block of  $p_1$  and  $p_2$ . We transformed the problem into finding the small modular roots of multivariate polynomials, and then the bound  $t > 4\alpha$  is obtained by using eight-dimensional lattices. Moreover, we heuristically generalize the method to an arbitrary number  $n$  of shared blocks and obtain a bound of  $t > 7\alpha$ . Our results generalize the idea presented in May and Sarkar's papers. Furthermore, for the former case, we obtain better results than May's result under same conditions.

However, the techniques presented here cannot immediately be extended to the generalized problem in Sarkar's paper where  $N_1, N_2, \dots, N_k$  are considered. In this case, the problem has much more variables and our method cannot be directly applied. This problem would be remained as an open issue.

## References

1. May, A., & Ritzenhofen M. (2009). Implicit factoring: On polynomial time factoring given only an implicit hint. *LNCS*, 5443, 1–14.
2. Sarkar, S., & Maitra S. (2009). Further results on implicit factoring in polynomial time. *Mathematics of Communications*, 3(2), 205–217.
3. Faugère J.-C., Marinier R., & Renault G. (2010). Implicit factoring with shared most significant and middle bits. *LNCS*, 6056, 70–87.
4. Herrman, M., & May, A. (2008). Solving linear equations modulo divisors: On factoring given any bits. *LNCS*, 5350, 406–424.
5. Jochemsz, E. (2007). *Cryptanalysis of RSA variants using small roots of polynomials*. Netherlands: Technische Universiteit Eindhoven.
6. Blömer, J., & May, A. (2003). New partial key exposure attacks on RSA. *LNCS*, 2729, 27–43.
7. Jochemsz, E., et al. (2006). A strategy for finding roots of multivariate polynomials with new applications in attacking RSA variants. *LNCS*, 4284, 267–282.
8. Lenstra, A. K., Lenstra Jr. H. W., & Lovász L. (1982). Factoring polynomials with rational coefficients. *Mathematische Annalen*, 261(4), 515–534.
9. Howgrave-Graham, N. (1997). Finding small roots of univariate modular equations revisited. *LNCS*, 1355, 131–142.

# Chapter 31

## Loading Data into HBase

Juan Yang and Xiaopu Feng

**Abstract** HBase is a top Apache open-source project that separated from Hadoop. As it has most of the features of Google's BigTable system and is implemented in Java, it is very popular in days of massive data. HBase's advantages are reflected in the massive data read and query. Loading huge amounts of data into HBase is the first step to use HBase. HBase itself has several methods to load data, and different methods have different application scenarios. This article made an exhaustive study and a performance testing of them. Also, this article achieved the custom loading data, and experiments show that it has good efficiency.

### 31.1 Introduction

With the increase of massive data, a large number of business scenarios began to consider expanding the level of data storage, so that storage services can be added or deleted easily. But the current relational databases focus on a machine. The storage of massive amount of data becomes the bottleneck, and a single machine is unable to load large amounts of data. The IO read and write of a single machine have become the bottleneck when it comes to the storage of massive data. When the amount of data is increasing with linear growth, how to solve the consistency of the data IO and how to combine MapReduce computing framework to analyze massive data have become big problems of traditional databases [1]. Traditional databases are unable to meet these needs, and the emergence of HBase solves these problems well. HBase is a top open-source project separated from Hadoop Apache. As it achieves most features of Google's BigTable system with Java, it is welcome in days of massive data [2]. The scenarios of HBase are mainly massive data queries and related operations. Therefore, the first step to use HBase is loading massive data

---

J. Yang (✉) • X. Feng  
Beijing University of Posts and Telecommunications, Beijing 100876, China  
e-mail: [yangjuan@bupt.edu.cn](mailto:yangjuan@bupt.edu.cn)



into HBase. HBase itself has several methods to load data, and each method has its own characteristics and application scenarios. However, the methods of HBase are lack of flexibility, and cannot fully meet the personalized needs of users. Based on this feature, the author achieved custom loading data with the combination of HBase interface and MapReduce framework through learning the methods of HBase. The experiments show that the effect of this method is good [3].

## 31.2 HBase

### 31.2.1 HBase Architecture

HBase is the open-source version of BigTable and is based on HDFS (Hadoop Distributed File System). It is a real-time read and write database system with high reliability and high performance. It is scalable and column storage. HBase is between NoSQL (Not Only SQL) and RDBMS (Relational Database Management System), and it can only retrieve the data by the primary key (row key) and the range of the main keys. It supports single-line transaction (via Hive support to achieve complex operations such as multi-table join). HBase is mainly used to store loose of unstructured and semi-structured data. Just like Hadoop, HBase target mainly relies on horizontal expansion. With the increase of cheap commercial server, it can increase the computing and storage capacity. HBase table generally has the following characteristics:

1. Large: a table can have hundreds of millions of lines and millions of columns.
2. Column-oriented: column-oriented (family) storage and access control, independent, and retrieve of column (family).
3. Sparse: for the empty (null) column, it does not occupy storage space. Therefore, the table can be designed very sparse.

Figure 31.1 is a diagram that shows the HBase Hadoop Ecosystem [4].

HBase's data is stored in the form of table. Table consists of rows and columns. Column is divided into several column families (row family). Figure 31.2 is the logical view of HBase's storage.

Same with NoSQL databases, row key is the primary key used to retrieve the record. There are only three ways to access rows in HBase's table:

1. Key access through a single row
2. Through the row key range
3. Full table scan

Row key can be any string (maximum length is 64 KB and the length of the practical application is generally 10–100 bytes). In the internal of HBase, row key is saved as byte array. When the data is stored, it is sorted in accordance with the row key lexicographic sorting (byte order). When designing the key, it is better to

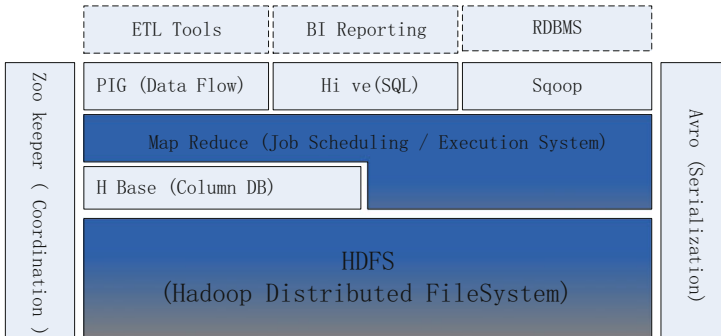


Fig. 31.1 HBase Hadoop ecosystem

Row Key	Time Stamp	Column "Contentst"	Column "anchor"		Column "mime"
"com.cnn. www"	t9		"anchor:cnnsi.com"	"CNN"	
	t8		"anchor:my.look.ca"	"CNN.com"	
	t6	"<html>..."			
	t5	"<html>..."			"text/html"
	t3	"<html>..."			

Fig. 31.2 HBase logical view

consider the characteristic of sorting storage and put the data together which are often read. The reading and writing of the line are atomic operations (regardless the number of columns to read and write). This design enables users to easily understand the program behavior when concurrent update operations are in the same line [5].

### 31.2.2 The Methods to Load Data

For the latest version of HBase (HBase-0.94.1), there are three methods to load data. In the following, we will explore and compare these three methods.

Method one: In the Hadoop environment, run the HBase own jar package to load specified data on HDFS into the specified data table. This method just needs Map function to complete the loading of data, without Reduce function. It uses the MapReduce interface of HBase to complete this operation. HBase has detailed instructions for using this method and the meaning of the parameters. It can be directly viewed in Hadoop environment with the related command.

Method two: In the Hadoop environment, run the HBase's related jar package. There are two steps to complete it. First, it will translate data in HDFS to HFile format. Then, run the related jar package to load HFile format data into specified

**Table 31.1** Test data information

Data	pk1.data	pk2.data	pk3.data	pk4.data	pk5.data	pk6.data	pk7.data
Rows	50,000	600,000	3,000,000	6,000,000	20,000,000	40,000,000	300,000,000
Size	1.10 mb	13.74 mb	71.04 mb	143.14 mb	561.61 mb	1,133.81 mb	7,860.25 mb

data table. During the implementation process in first step, both Map function and Reduce function will be executed. The execution time of Reducing/reduction function is significantly longer than the Map function. The Reducing/reduction function consumes more system resources. The second step, it is just a simple data transfer, and it does not execute MapReduce function.

Method three: In the HBase environment, run the HBase's own method. It will load the sequential file on HDFS directly into the specified data table. This method is the same as method one, completing the data loading in the Map function and did not execute the Reduce function [6].

### 31.2.3 Performance Test Comparison

#### 31.2.3.1 Test Environment and Data Description

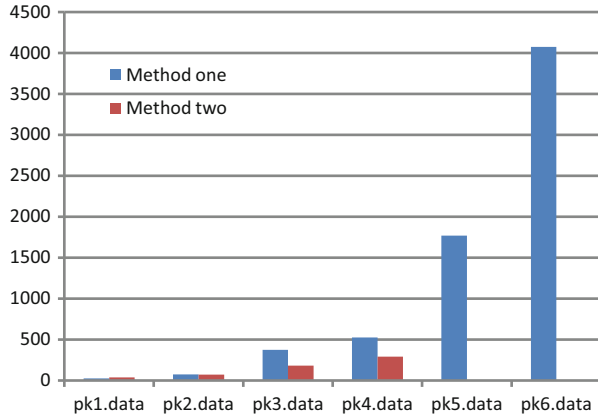
The test cluster comprises three nodes, including a master node and two slave nodes. Each single node contains four hard disks, and each hard disk has 750G storage space, four CUP with 2.00 GHz frequency, and 8G memory. The Hadoop version is 1.0.3 and the HBase version is 0.94.1. The test data is a text file and each file has five columns. The data is randomly generated. Details of the data are shown in Table 31.1.

#### 31.2.3.2 Test Result Contrast

The performance testing of method one and method two is shown in Fig. 31.3. The ordinate is the time to load data and the time unit is seconds. While using pk5.data and pk6.data to test method two, due to consuming too many resources, the hardware environment cannot meet the demand to complete.

As can be seen from Fig. 31.3, when the test data volume is small, method one is faster than method two. But when the data volume is big enough, method two is much faster than method one. However, method two consumes more system resources, and when the hardware system cannot meet the demand, it could not be completed. Method one is easy to operate, for it just needs one step to be completed. Therefore, when the hardware environment is not very well and the amount of data is small, it is better to use method one to complete loading data. When the amount of data is very big, method two is better to save time, but it needs more system resources.

**Fig. 31.3** HBase's loading data performance testing



## 31.3 Custom Loading Data

### 31.3.1 Design Description

HBase's interface allows users to customize program to load data into HBase. HBase's encapsulated MapReduce interface allows Hadoop and HBase to communicate with each other. It can define the HBase's interface in the Map function to load data into HBase. With the help of Hadoop MapReduce parallel framework, it will be easy to load massive data into HBase. Loading operation can be completed directly in the Map function, without writing the Reduce function. Firstly, set up connection with HBase in the run method of MapReduce and then in the Mapfunction calling HBase's interface to load data into HBase.

Additionally, we can optimize settings of HBase's relevant parameters for specific data and experimental environment. This chapter focuses on the following parameters for experimental cluster to optimize settings.

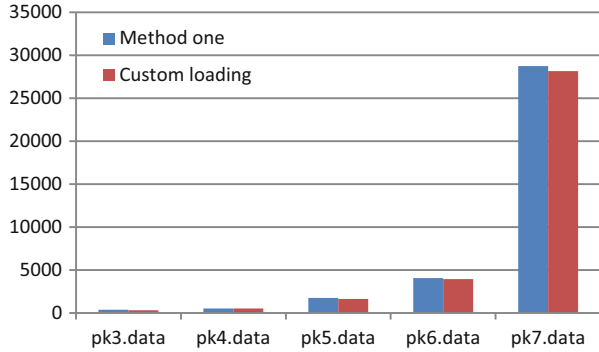
#### 1. Delay log flushes

HBase's mechanism is to refresh the log once you put a write. If we use a delay mechanism, log will be stored in memory until the log refresh cycle is over. It can delay the refresh cycle properly, thus each refresh can load more data into the data table.

#### 2. Automatic reflash

Once HBase executes a Put operation, it will send a request to the Region Server. Through `htable.add (Put)`, it will add Put to the write buffer. If the `autoFlush` is set to `false`, it will have to wait until the write buffer is full to initiate the request. In order to explicitly initiate the request, you can also call the `flushCommits`.

**Fig. 31.4** Custom loading data performance testing



3. Close Puts on WAL

When executing the Put operation, the data is firstly written to the log. Setting `put.setWriteToWAL` to false means that it will not be written to the log but directly write to the memory. So when the data is loaded, it will eliminate the time to write the log.

4. Cache size

When the data size in the cache reaches a certain percentage, HBase will automatically refresh, and the data in the cache will be loaded into the data table. We can adjust the cache size for the experimental environment and take full advantage of the cache to load as much data into the cache as possible.

In addition, we can adjust and optimize the number of regions, region size, and intermediate result storage format according to the experimental environment to achieve the best results [7].

### 31.3.2 Performance Testing

According to the characteristics of the test cluster, custom loading data optimizes the relevant parameters and, based on this, executing a performance test. Test environment and test data are the same as before. Figure 31.4 is a performance testing comparison of custom loading data and method one.

As we can see from the above results, custom loading method is flat or slightly better than method one. Additionally, you can specify a particular column and make preprocessing and other operations with custom loading method. Especially when the amount of data is large and the data contains illegal data, you can use custom loading method to filter data and do some other operations. So, custom loading method has a lot of flexibility and has advantages in terms of performance compared with the methods of HBase itself [8].

## 31.4 Conclusion

HBase itself has different methods to load data, and each method has its own application scenario. Developers can select the appropriate loading method for a specific application scenario. In addition, developers can also implement custom parallel loading data by calling the interface of HBase. Custom loading method is flexible and its performance is good. It is more suitable for programmers to develop based on HBase.

## References

1. Dean, J., & Ghemawat, S. (2008). MapReduce: Simplified data processing on large clusters. *Communications of the ACM*, 51(1), 107–113.
2. Chang, F., Dean, J., Ghemawat, S., Hsieh, W. C., Wallach, D. A., Burrows, M., et al. (2008). Bigtable: A distributed storage system for structured data. *ACM Transactions on Computer Systems (TOCS)*, 26(2), 4.
3. White, T. *Hadoop: The definitive guide [M]*. O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA95472, 2012.
4. George, L. *HBase: The definitive guide [M]*. O'Reilly Media, Inc., 1005 Gravenstein Highway North, Sebastopol, CA95472, 2011.
5. Huang, J., Ouyang, X., Jose, J., Wasi-ur-Rahman Md., Wang, H., Luo, M., et al. (2012). High-performance design of HBase with RDMA over infiniband. In *Proceedings of the 2012 I.E. 26th International Parallel and Distributed Processing Symposium, IPDPS 2012* (pp. 774–778). Washington, DC: IEEE Computer Society.
6. Li, C. (2010). Transforming relational database into HBase: A case study. In *Proceedings 2010 I.E. International Conference on Software Engineering and Service Sciences, ICSESS 2010* (pp. 683–687). Piscataway, NJ: IEEE Computer Society.
7. Vora, M. N. (2011). Hadoop-HBase for large-scale data. In *Proceedings of 2011 International Conference on Computer Science and Network Technology, ICCSNT 2011* (pp. 601–605). Piscataway, NJ: IEEE Computer Society.
8. Carstou, D., Cernian, A., & Olteanu, A. (2010). Hadoop hbase-0.20. 2 performance evaluation. In *NISS2010 – 4th International Conference on New Trends in Information Science and Service Science* (pp. 84–87). Piscataway, NJ: IEEE Computer Society.

# Chapter 32

## Incomplete Decision-Theoretic Rough Set Model Based on Improved Complete Tolerance Relation

Xia Wang

**Abstract** Recently, the decision tables used in decision-theoretic rough sets are the most complete decision tables and less for incomplete decision tables. Some authors use the filling method to deal with incomplete decision table of DTRS, which is filling the unknown values with all possible values. Its disadvantages are large calculated amount and noise. Therefore, incomplete decision-theoretic rough set model based on improved complete tolerance relation is proposed. First, improved complete tolerance relation and tolerance class were constructed. Then, the decision degree between object and target concept were also computed. Finally, the decision degree was served as conditional probability; the probability positive region, negative region, and boundary region were computed, and it would make three-way decision. The example shows that this model has less calculated amount and noise, and it more accords with practical application.

### 32.1 Introduction

Decision-theoretic rough set model (DTRS) was proposed by Yao in the 1990s, which is a kind of probability rough set model [1]. Based on Pawlak rough set model, DTRS imports the theory of minimum-risk Bayesian decision and finds least expected risk through kinds of risk costing value of classification decisions. So the least expected risk is as basic argument of partition universe into positive region, negative region, and boundary region [1–6]. In brief, based on equivalence class, the universe is partitioned into three regions, positive region, negative region, and boundary region, respectively, and this is similar to Pawlak rough sets. So under this circumstance, disposed decision table must be complete. But in real application, the decision tables are almost incomplete.

---

X. Wang (✉)

College of Computer Science and Technology, Jiangsu Normal University,  
Jiangsu 221119, China  
e-mail: [lgzwx@163.com](mailto:lgzwx@163.com)

Incomplete decision table means a table with missing data (null value). At present, there are two kinds of ways for disposing incomplete decision tables. One is completion means, which characteristic is polishing the decision table, and then using the means of completing table to deal with the incomplete table. The other is direct disposing means, which characteristic is appropriate expanding the concepts of classics rough sets, and then transforms the strict equivalence relation to broad binary relation [7].

In this chapter, an incomplete decision-theoretic rough set model is proposed, which is based on improved tolerance relation and decision-theoretic rough set. This model sufficiently absorbs the advantages of improved tolerance relation, which will be appropriate for three principal decision of the real world.

### 32.2 Preliminaries

**Definition 1[7]** Let  $S = \{U, A, V, f\}$  is a decision table, where  $U = \{x_1, x_2, \dots, x_n\}$  is nonempty finite object set,  $A = C \cup D$  is nonempty attribute set,  $C = \{c_1, c_2, \dots, c_m\}$  is set of conditional attribute,  $D = \{d_1, d_2, \dots, d_k\}$  is set of decision attribute, and  $C \cap D = \emptyset$ ,  $V$  is the domain of attribute.  $f : U \rightarrow V_i (i = 1, 2, \dots, m + k)$  is information function. When the precise values of some attributes are not known, i.e., missing or known partially, (write for  $f(x, c) = *$ ,  $x \in U$   $c \in C$ ), such  $S$  is called an incomplete decision table.

**Definition 2[8]** Let  $S = \{U, C \cup D, V, f\}$  is an incomplete decision table. The tolerance relation is defined as:

$$T_C(x, y) = \{(x, y) \in U \times U \mid \forall c \in C (c(x) = c(y) \vee c(x) = * \vee c(y) = *)\} \quad (32.1)$$

So, tolerance relation has reflexivity and symmetry, but not transitivity.

**Definition 3[8]** Let  $S = \{U, C \cup D, V, f\}$  is an incomplete decision table; the tolerance class of object  $x$  is defined as

$$I_C^T(x) = \{y \mid y \in U \wedge T_C(x, y)\} \quad (32.2)$$

**Definition 4[8]** Let  $S = \{U, C \cup D, V, f\}$  is an incomplete decision table. Improved complete tolerance relation of is  $S$  defined as

$$NS_C(x, y) = \left\{ (x, y) \in U \times U \mid \forall c \in C (c(x) = c(y) \vee c(x) = * \vee c(y) = *) \wedge \min \left( \frac{|P(x) \cap P(y)|}{\min(|P(x)|, |P(y)|)}, \frac{\gamma(x) + \gamma(y)}{2} \geq \gamma \right) \right\} \quad (32.3)$$



where  $P(x_i) = \{c|c \in C \wedge c(x_i) \neq *\}$  and assuming  $0/0 = 0$ ,  $\gamma(x)$  is complete degree of object  $x$ .

**Definition 5[9]** Let  $S = \{U, C \cup D, V, f\}$  is an incomplete decision table. The complete degree of object  $x_i$  is defined as

$$\gamma(x_i) = \frac{|P(x_i)|}{|C|} \tag{32.4}$$

where  $P(x_i) = \{c|c \in C \wedge c(x_i) \neq *\}$ ,  $x_i \in U$ ,  $|\bullet|$  is base number of set  $\bullet$ . So the complete degree of whole decision table  $S$  is defined as

$$\gamma = \frac{\sum_{i=1}^{|U|} \gamma(x_i)}{|U|} \tag{32.5}$$

**Definition 6[9]** Let  $S = \{U, C \cup D, V, f\}$  is an incomplete decision table; improved tolerance class of object  $x$  is defined as

$$I_C^{NS}(x) = \{y|y \in U \wedge NS_C(x, y)\} \cup \{x\} \tag{32.6}$$

### 32.3 Decision-Theoretic Rough Set Model

For the Bayesian decision procedure, the DTRS model is composed of two states and three actions. The set of states is given by  $\Omega = \{C, \neg C\}$  ( $C$  is target set) indicating that an object is in  $C$  and not in  $C$ , respectively. The set of actions is given by  $\varsigma = \{a_P, a_B, a_N\}$ , where  $a_P$ ,  $a_B$ , and  $a_N$  represent the three actions in classifying an object  $x$ , namely, deciding  $x \in POS(C)$ , deciding should be further investigated  $x \in BND(C)$ , and deciding  $x \in NEG(C)$ , respectively. The loss function  $\lambda$  regarding the risk or cost of actions of different states is given by the  $3 \times 2$  matrix:

	$C(P)$	$\neg C(N)$
$a_P$	$\lambda_{PP}$	$\lambda_{PN}$
$a_B$	$\lambda_{BP}$	$\lambda_{BN}$
$a_N$	$\lambda_{NP}$	$\lambda_{NN}$

In the matrix,  $\lambda_{PP}$ ,  $\lambda_{BP}$ , and  $\lambda_{NP}$  denote the losses incurred for taking actions of  $a_P$ ,  $a_B$ , and  $a_N$ , respectively, when an object belongs to  $C$ . Similarly,  $\lambda_{PN}$ ,  $\lambda_{BN}$ , and  $\lambda_{NN}$  denote the losses incurred for taking actions of  $a_P$ ,  $a_B$ , and  $a_N$  when the object belongs to  $\neg C$ .  $\Pr(C|x)$  is the conditional probability of an object  $x$  belonging to

$C$  given that the object is described by its equivalence class  $[x]$ . For an object  $x$ , the expected loss  $R(a_i|[x])$  associated with taking the individual actions can be expressed as

$$\begin{aligned} R(a_P|[x]) &= \lambda_{PP} \Pr(C|[x]) + \lambda_{PN} \Pr(\neg C|[x]) \\ R(a_B|[x]) &= \lambda_{BP} \Pr(C|[x]) + \lambda_{BN} \Pr(\neg C|[x]) \\ R(a_N|[x]) &= \lambda_{NP} \Pr(C|[x]) + \lambda_{NN} \Pr(\neg C|[x]) \end{aligned}$$

The Bayesian decision procedure suggests the following minimum-cost decision rules:

- P: If  $R(a_P|[x]) \leq R(a_B|[x])$  and  $R(a_P|[x]) \leq R(a_N|[x])$ , then decide  $x \in POS(C)$ .
- B: If  $R(a_B|[x]) \leq R(a_P|[x])$  and  $R(a_B|[x]) \leq R(a_N|[x])$ , then decide  $x \in BND(C)$ .
- N: If  $R(a_N|[x]) \leq R(a_P|[x])$  and  $R(a_N|[x]) \leq R(a_B|[x])$ , then decide  $x \in NEG(C)$ .

Since  $\Pr(C|[x]) + \Pr(\neg C|[x]) = 1$ , so we can simplify the rules based only on the probability  $\Pr(C|[x])$  and the loss function. By considering a reasonable kind of loss function with  $0 \leq \lambda_{PP} \leq \lambda_{BP} < \lambda_{NP}$  and  $0 \leq \lambda_{NN} \leq \lambda_{BN} < \lambda_{PN}$ , the decision rule P-N can be expressed concisely as:

- P: If  $\Pr(C|[x]) \geq \alpha$  and  $\Pr(C|[x]) \geq \gamma$ , then decide  $x \in POS(C)$ .
- B: If  $\Pr(C|[x]) \leq \alpha$  and  $\Pr(C|[x]) \geq \beta$ , then decide  $x \in BND(C)$ .
- N: If  $\Pr(C|[x]) \leq \beta$  and  $\Pr(C|[x]) \leq \gamma$ , then decide  $x \in NEG(C)$ .

The threshold values  $\alpha$ ,  $\beta$ , and  $\gamma$  are defined as  $\alpha = \frac{(\lambda_{PN} - \lambda_{BN})}{(\lambda_{PN} - \lambda_{BN}) + (\lambda_{BP} - \lambda_{PP})}$ ,  
 $\beta = \frac{(\lambda_{BN} - \lambda_{NN})}{(\lambda_{BN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{BP})}$ , and  $\gamma = \frac{(\lambda_{PN} - \lambda_{NN})}{(\lambda_{PN} - \lambda_{NN}) + (\lambda_{NP} - \lambda_{PP})}$ .

In addition, as a well-defined boundary region, the conditional of rule B suggests that  $\alpha > \beta$ , that is  $\frac{(\lambda_{BP} - \lambda_{PP})}{(\lambda_{PN} - \lambda_{BN})} < \frac{(\lambda_{NP} - \lambda_{NN})}{(\lambda_{BN} - \lambda_{NN})}$ . It implies  $0 \leq \beta < \gamma < \alpha \leq 1$ . In this case, after tiebreaking, the following simplified rules are obtained:

- P1: If  $\Pr(C|[x]) \geq \alpha$ , then decide  $x \in POS(C)$ .
- B1: If  $\beta < \Pr(C|[x]) < \alpha$ , then decide  $x \in BND(C)$ .
- N1: If  $\Pr(C|[x]) \leq \beta$ , then decide  $x \in NEG(C)$ .

### 32.4 Incomplete Decision-Theoretic Rough Set Model Based on Improved Complete Tolerance Relation

At present, most authors are focused on complete decision table when they research decision-theoretic rough set and less for incomplete decision table. In this chapter, the procedure of incomplete decision-theoretic rough set is proposed, based on improved complete tolerance relation. Let us see an example in Table 32.1.

Where  $U = \{x_1, x_2, \dots, x_7\}$  is object set;  $a, b$ , and  $c$  are condition attributes;  $d$  is decision attribute; and \* means possible values of attributes.

**Table 32.1** Incomplete decision table

$U$	$a$	$b$	$c$	$d$
$x_1$	1	1	1	1
$x_2$	1	*	1	1
$x_3$	2	1	1	1
$x_4$	1	2	*	1
$x_5$	1	2	*	2
$x_6$	1	*	1	2
$x_7$	1	1	1	2

**Table 32.2** Complete degree of each object

$U$	$x_1$	$x_2$	$x_3$	$x_4$	$x_5$	$x_6$	$x_7$
$\gamma(x_i)$	1	2/3	1	2/3	2/3	1	1

First, we compute complete degree of each object by using Definition 5. The result can be expressed in Table 32.2.

So the complete degree of whole decision table is  $\gamma = \frac{\sum_{i=1}^7 \gamma(x_i)}{7} = 0.8571$ .

Then, tolerance relation  $NS_C$  of table  $S$  is  $NS_C = \{(1,7),(7,1)\} \cup I_X$ , where  $I_X$  is identity relation of  $X$ .

So, tolerance class of each object  $x_i$  is

$$I_C^{NS}(x_1) = \{x_1, x_7\}, I_C^{NS}(x_2) = \{x_2\}, I_C^{NS}(x_3) = \{x_3\}, I_C^{NS}(x_4) = \{x_4\}, I_C^{NS}(x_5) = \{x_5\}, I_C^{NS}(x_6) = \{x_6\}, \text{ and } I_C^{NS}(x_7) = \{x_1, x_7\}.$$

From the result, we can find the tolerance class set is a cover of universe  $U$ .

**Definition 7** Given tolerance class  $I_C^{NS}(x_i)$  ( $i = 1, 2, \dots, |U|$ ) of object  $x_i$ , then decision degree is defined as:

$$\mu(x_i) = \frac{|I_C^{NS}(x_i) \cap X|}{|I_C^{NS}(x_i)|} \tag{32.7}$$

Replaced conditional probability of DTRS model with decision degree, three regions can be again defined as:

$$\begin{aligned} POS_{(\alpha,\beta)}(X) &= \{x | \mu(x) \geq \alpha\}. \\ NEG_{(\alpha,\beta)}(X) &= \{x | \mu(x) \leq \beta\}. \\ BND_{(\alpha,\beta)}(X) &= \{x | \alpha < \mu(x) < \beta\}. \end{aligned} \tag{32.8}$$

Correspondingly, the P, B, and N rules can be expressed as:

$$\begin{aligned}
P &: \text{if } \mu(x) \geq \alpha, \text{ then } x \in POS_{(\alpha,\beta)}(C). \\
N &: \text{if } \mu(x) \leq \beta, \text{ then } x \in NEG_{(\alpha,\beta)}(C). \\
B &: \text{if } \beta < \mu(x) < \alpha, \text{ then } x \in BND_{(\alpha,\beta)}(C).
\end{aligned} \tag{32.9}$$

Then, given the target concept set  $X = \{x_1, x_2, x_3, x_4\}$ , the decision degree of each object can be computed as:

$$\mu(x_1) = 0.5, \quad \mu(x_2) = 1, \quad \mu(x_3) = 1, \quad \mu(x_4) = 1, \quad \mu(x_5) = 0, \quad \mu(x_6) = 0, \quad \text{and} \\ \mu(x_7) = 0.5.$$

For example, when  $\alpha = 0.65$ ,  $\beta = 0.45$ , the three regions are:

$$\begin{aligned}
POS_{(\alpha,\beta)}(X) &= \{x \mid \mu(x) \geq 0.65\} = \{x_2, x_3, x_4\}. \\
NEG_{(\alpha,\beta)}(X) &= \{x \mid \mu(x) \leq 0.45\} = \{x_5, x_6\}. \\
BND_{(\alpha,\beta)}(X) &= \{x \mid 0.45 < \mu(x) < 0.65\} = \{x_1, x_7\}.
\end{aligned}$$

Different with the traditional DTRS, we use tolerance class operation and decision degree, which are foundation of building DTRS model. So this model can serve as extending of traditional DTRS model.

## 32.5 An Example of Using the New Model

Let us consider an example of incomplete decision table from the literature [8] listed in Table 32.3. We would like to make a decision when an object belongs to “ $\Phi$ ” class.

First, we can compute the complete degree of Table 32.3, which value is 0.625. Using Definitions 6 and 7, the tolerance class of each object is

$$\begin{aligned}
I_C^{NS}(a_1) &= \{a_1\}; \quad I_C^{NS}(a_2) = \{a_2, a_3\}; \quad I_C^{NS}(a_3) = \{a_2, a_3\}; \quad I_C^{NS}(a_4) = \{a_4, a_9, a_{12}\}; \\
I_C^{NS}(a_5) &= \{a_5, a_9, a_{12}\}; \quad I_C^{NS}(a_6) = \{a_6, a_7\}; \quad I_C^{NS}(a_7) = \{a_6, a_7\}; \quad I_C^{NS}(a_8) = \{a_8\}; \quad I_C^{NS} \\
(a_9) &= \{a_4, a_5, a_9, a_{11}\}; \quad I_C^{NS}(a_{10}) = \{a_{10}\}; \quad I_C^{NS}(a_{11}) = \{a_9, a_{11}, a_{12}\}; \quad \text{and} \quad I_C^{NS} \\
(a_{12}) &= \{a_4, a_5, a_{11}, a_{12}\}.
\end{aligned}$$

The second, on the base of Definition 9, we compute decision degree. Supposed target concept is  $X = \{a_2, a_4, a_6, a_7, a_8, a_9, a_{11}\}$ , and then the decision degree of each object is

$$\gamma(a_1) = 0; \quad \gamma(a_2) = \gamma(a_3) = 0.5; \quad \gamma(a_4) = 0.667; \quad \gamma(a_5) = 0.333; \quad \gamma(a_6) = \gamma(a_7) = \\ \gamma(a_8) = 1; \quad \gamma(a_9) = 0.75; \quad \gamma(a_{10}) = 0; \quad \gamma(a_{11}) = 0.667; \quad \text{and} \quad \gamma(a_{12}) = 0.5.$$

The third, suppose  $\alpha = 0.75$ ,  $\beta = 0.45$ , according to Eq. (32.8), the three regions are:

$$POS_{(\alpha,\beta)}(X) = \{a \mid \mu(a) \geq 0.75\} = \{a_6, a_7, a_8, a_9\}.$$

**Table 32.3** An example of incomplete decision table

$U$	$c_1$	$c_2$	$c_3$	$c_4$	$d$
$a_1$	*	*	1	*	$\Psi$
$a_2$	1	2	3	4	$\Phi$
$a_3$	1	2	3	4	$\Psi$
$a_4$	2	*	*	*	$\Phi$
$a_5$	2	*	*	*	$\Psi$
$a_6$	4	*	2	1	$\Phi$
$a_7$	*	3	2	1	$\Phi$
$a_8$	3	*	*	*	$\Phi$
$a_9$	2	3	4	1	$\Phi$
$a_{10}$	*	0	*	*	$\Psi$
$a_{11}$	2	3	4	*	$\Phi$
$a_{12}$	2	3	4	3	$\Psi$

$$NEG_{(\alpha,\beta)}(X) = \{a \mid \mu(a) \leq 0.45\} = \{a_1, a_5, a_{10}\}.$$

$$BND_{(\alpha,\beta)}(X) = \{a \mid 0.45 < \mu(a) < 0.75\} = \{a_2, a_3, a_4, a_{11}, a_{12}\}.$$

So in the incomplete decision table  $S$ , positive region is  $POS_{(\alpha,\beta)}(X) = \{a_6, a_7, a_8, a_9\}$ , then we can make the decision that objects  $a_6, a_7, a_8, a_9$  belong to target concept  $X$ . The negative region is  $NEG_{(\alpha,\beta)}(X) = \{a_1, a_5, a_{10}\}$ , so we can do decision that objects  $a_1, a_5, a_{10}$  do not belong to target concept  $X$  and we can reject them. This conclusion is same to original table. That is, objects of positive region and negative region are, respectively,  $\Phi$  and  $\Psi$ . With regard to each element of boundary region  $BND_{(\alpha,\beta)}(X) = \{a_2, a_3, a_4, a_{11}, a_{12}\}$ , the attribute value of object  $a_2$  is similar to object  $a_3$ , and they have no missing element. When decision attribute value of object  $a_2$  is different from object  $a_3$ , they are incomplete decision, so they need further research.

Since the other objects are similar to objects  $a_2$  and  $a_3$ , I won't say more about them here.

From building procedure, we can find this model has litter calculated amount and litter noise than filling means, so this model is more line with the need of practical use.

### 32.6 Conclusion

The decision tables of DTRS model are most complete. There are a few researches on incomplete DTRS model. In this chapter, we absorb the advantages of improved tolerance relation, and disposing means of incomplete decision table are applied to DTRS model. We propose incomplete decision-theoretic rough set model based on improved tolerance relation. This model has many advantages, such as litter calculated amount and noise. So it will be suited to three tree making decisions of the real world.

## References

1. Yao, Y. Y., & Wong, S. K. M. (1992). A decision theoretic framework for approximating concept. *International Journal of Man–Machine Studies*, 37(6), 793–809.
2. Yao, Y. Y., Wong, S. K. M., & Lingras, P. (1990). A decision-theoretic rough set model. In *Methodologies for Intelligent Systems* (Vol. 5, pp. 17–24). New York: North-Holland.
3. Yao, Y. Y. (2003). Probabilistic approaches to rough sets. *Expert Systems*, 20(5), 287–297.
4. Yao, Y. Y. (2007). *Decision-theoretic rough set models*. RSKT 2007, LNAI 4481, 1–12.
5. Yao, Y. Y., & Zhao, Y. (2008). Attribute reduction in decision-theoretic rough set models. *Information Sciences*, 178(17), 3356–3373.
6. Yao, Y. Y. (2012). *An outline of a theory of three-way decisions*. RSCTC 2012, LNAI 7413, 1–17.
7. Zhou, X., Huang, B., Li, H., & Wei, D. (2010). *Knowledge acquisition of incomplete information system and rough set theory and method* (pp. 33–52). Nanjing: Nanjing University Press (in chinese).
8. Kryszkiewicz, M. (1998). Rough set approach to incomplete information systems. *Information Sciences*, 112(1/2/3/4), 39–49.
9. Sheng, L., & Yang, H.-z. (2008). Extended rough set model based on completed tolerance relation. *Control and Decision*, 23(3), 258–262 (in chinese).

# Chapter 33

## A New Association Rule Mining Algorithm Based on Compression Matrix

Sihui Shu

**Abstract** A new association rule mining algorithm based on matrix is introduced. It mainly compresses the transaction matrix efficiently by integrating various strategies. The new algorithm optimizes the known association rule mining algorithms based on matrix given by some researchers in recent years, which greatly reduces the temporal complexity and spatial complexity, and highly promotes the efficiency of association rule mining. It is especially feasible when the degree of the frequent itemset is high.

### 33.1 Introduction

Association rule mining is one of the most important and well-researched techniques of data mining, which is firstly introduced by Agrawal et al. [1]. They presented well-known Apriori algorithm in 1993, since many methods have been involved in the improvement and optimization of Apriori algorithm, such as binary code technology, genetic algorithm, and algorithms based on matrix [2, 3]. The algorithm based on matrix could only scan the database for one time to convert the transactions into matrix and could be reordered by item support count in non-descending order to reduce the number of candidate itemsets and could highly promote Apriori algorithm efficiency in temporal complexity and spatial complexity.

A great deal of work on Apriori algorithms based on matrix has been done [4, 5]. In this chapter, a new improvement of Apriori algorithm based on compression matrix is proposed and could achieve better performance.

---

S. Shu (✉)

College of Mathematics and Computer Science, Jiangxi Science and Technology Normal University, Nanchang 330000, China  
e-mail: [sihuishu@163.com](mailto:sihuishu@163.com)

### 33.2 Preliminaries

Some basic preliminaries used in association rule mining are introduced in this section. Let  $T = \{T_1, T_2, \dots, T_m\}$  be a database of transactions and  $T_k (k = 1, 2, \dots, m)$  denotes a transaction. Let  $I = \{I_1, I_2, \dots, I_n\}$  be a set of binary attributes, called Items.  $I_k (k = 1, 2, \dots, n)$  denotes an Item. Each transaction  $T_k$  in  $T$  contains a subset of items in  $I$ . The number of items contained in  $T_k$  is called the length of transaction  $T_k$ , which is symbolized  $|T_k|$ .

An association rule is defined as an implication of the form  $X \Rightarrow Y$ , where  $X, Y \in I$  and  $X \cap Y = \phi$ . The support (min-sup) of the association rule  $X \Rightarrow Y$  is the support (resp. frequency) of the itemset  $X \cup Y$ . If support (min-sup) of an itemset  $X$  is greater than or equal to a user-specified support threshold, then  $X$  is called frequent itemsets.

In the process of the association rule mining, we find frequent itemsets firstly and produce association rule by these frequent itemsets secondly. So the key procedure of the association rule mining is to find frequent itemsets; some properties of frequent itemset are given as the following:

**Property 1 [1]** Every nonempty subset of a frequent itemset is also a frequent itemsets.

By the definition of frequent k-itemset, the conclusion below is easily obtained.

**Property 2** If the length  $|T_i|$  of a transaction  $T_i$  is less than k, then  $T_i$  is valueless for generating the frequent k-itemset.

### 33.3 An Improvement on Apriori Algorithm Based on Compression Matrix

A new improvement on Apriori algorithm based on compression matrix is introduced. The process of our new algorithm is described as follows:

1. Generate the transaction matrix.

For a given database with  $n$  transactions and  $m$  items, the  $m \times n$  transaction matrix  $D = (d_{ij})$  is determined, in which  $d_{ij}$  sets 1 if item  $I_i$  is contained in transaction  $T_j$  or otherwise sets 0.

$$D = \begin{matrix} I_1 \\ I_2 \\ \vdots \\ I_m \end{matrix} \begin{pmatrix} T_1 & T_2 & \dots & T_n \\ d_{11} & d_{12} & \dots & d_{1n} \\ d_{21} & d_{22} & \dots & d_{2n} \\ \vdots & \vdots & \dots & \vdots \\ d_{m1} & d_{m2} & \dots & d_{mn} \end{pmatrix}$$

where  $d_{ij} = \begin{cases} 1, I_i \in T_j \\ 0, I_i \notin T_j \end{cases} \cdot i = 1, 2, \dots, m, j = 1, 2, \dots, n.$



For each  $I_k, T_j, v_k = \sum_{i=1}^n d_{ij}, k = 1, 2, \dots, m; h_j = \sum_{i=1}^m d_{ij}, j = 1, 2, \dots, n.$

2. Produce frequent 1-itemset  $L_1$  and frequent 2-itemset support matrix  $D_1$ .  
The frequent 1-itemset  $L_1$  is  $L_1 = \{I_k | v_k \geq \text{min - sup}\}.$

### 33.3.1 Matrix Compression Procedure

In order to reduce the storage space and computation complexity, useless rows and columns should be discovered and removed in “matrix compression procedure,” which will be reused frequently in subsequent processes. Useless rows and columns can be classified into two classes, so the compression procedure is separated into two steps:

- (i) A row  $I_k$  is considered as worthless when the corresponding  $v_k$  is less than the support min-sup; a column  $T_j$  is considered as worthless when the corresponding  $h_j$  is less than 2 according to Property 2. Thus, we drop these rows or columns one by one and update  $v_k$  and  $h_j$  immediately after each drop operation. Subsequently, repeat the procedure (i) until there is no such row or column.
- (ii) Let’s consider the second class of useless rows and store their frequent itemsets for being used in the next procedure. Every row  $I_l$  whose corresponding  $v_l$  is less than  $\lceil \sqrt{n} \rceil$  ( $\lceil x \rceil$  is the largest integer which is no greater than  $x$ ) would be removed after its frequent itemsets are calculated as below:

Let  $\text{min-sup} = b$ . For a satisfied item  $I_l$ , let  $S_l = \{T_j | d_{lj} = 1\}$  and  $S_l'$  be the  $b$ -combinations set of elements in  $S_l$ :  $S_l' = \{(T_{j_1}, T_{j_2}, \dots, T_{j_m}) | T_{j_1}, T_{j_2}, \dots, T_{j_m} \in S_l\}$ . Each  $b$ -tuple  $(T_{j_1}, T_{j_2}, \dots, T_{j_m})$  from  $S_l'$  would be scanned in turn, if there exist items  $I_1, I_2, \dots, I_k$  except  $I_l$  that let  $d_{lj_1} = d_{lj_2} = \dots = d_{lj_m} = 1$  ( $i = 1, 2, \dots, k$ ), the collection  $(I_1, I_2, \dots, I_k, I_l)$  is one frequent itemset containing  $I_l$ . All the frequent itemsets containing  $I_l$  can be obtained though handling every  $b$ -tuple element from  $S_l'$ . After repeating step (i) and (ii) until there is no such useless row or column in the compressed matrix of  $D$ , the frequent 2-itemset support matrix  $D_1$  is produced:

$$D_1 = \begin{matrix} & T_{j_1} & T_{j_2} & \dots & T_{j_q} \\ \begin{matrix} I_{i_1} \\ I_{i_2} \\ \vdots \\ I_{i_p} \end{matrix} & \begin{pmatrix} d_{i_1j_1} & d_{i_1j_2} & \dots & d_{i_1j_q} \\ d_{i_2j_1} & d_{i_2j_2} & \dots & d_{i_2j_q} \\ \vdots & \vdots & \dots & \vdots \\ d_{i_pj_1} & d_{i_pj_2} & \dots & d_{i_pj_q} \end{pmatrix} \end{matrix}$$

where  $1 \leq i_1 < i_2 < \dots < i_p \leq m, 1 \leq j_1 < j_2 < \dots < j_q \leq n.$

3. Produce the frequent 2-itemset  $L_2$  and the frequent 3-itemset support matrix  $D_2$ .

The frequent 2-itemset  $L_2$  is the union of the 2-itemset subsets produced by frequent itemsets in step (ii) of procedure (2) and a set  $L'_2$  determined by comparing the inner product of each two row vectors of matrix  $D_1$  with the support min-sup

$$L'_2 = \{(I_{i_h}, I_{i_r}) \mid \sum_{k=1}^q d_{i_h k} d_{i_r k} \geq \text{min-sup}, h < r, h, r = 1, 2, \dots, p\}.$$

Matrix  $D'_2$  is obtained by calculating “and” operation of the two corresponding row vectors of every element  $(I_{i_h}, I_{i_r})$  in  $L'_2$ , that is:

$$D'_2 = \begin{pmatrix} (I_{i_{h_1}}, I_{i_{r_1}}) \\ (I_{i_{h_1}}, I_{i_{r_2}}) \\ \vdots \\ (I_{i_{h_s}}, I_{i_{r_t}}) \end{pmatrix} \begin{matrix} T_{j_1} & T_{j_2} & \dots & T_{j_q} \\ \begin{pmatrix} d_{i_{h_1}j_1} d_{i_{r_1}j_1} & d_{i_{h_1}j_2} d_{i_{r_1}j_2} & \dots & d_{i_{h_1}j_q} d_{i_{r_1}j_q} \\ d_{i_{h_1}j_1} d_{i_{r_2}j_1} & d_{i_{h_1}j_2} d_{i_{r_2}j_2} & \dots & d_{i_{h_1}j_q} d_{i_{r_2}j_q} \\ \vdots & \vdots & \ddots & \vdots \\ d_{i_{h_s}j_1} d_{i_{r_t}j_1} & d_{i_{h_s}j_2} d_{i_{r_t}j_2} & \dots & d_{i_{h_s}j_q} d_{i_{r_t}j_q} \end{pmatrix} \end{matrix}$$

where  $1 \leq h_1 < h_2 < \dots < h_s \leq p$ ,  $1 \leq r_1 < r_2 < \dots < r_t \leq p$ ,  $n_1$  is called row numbers of matrix  $D'_2$ .

- (i) Remove rows or columns in  $D'_2$  using the same approach in step (i) of (2), while column  $T_{j_k}$  is considered as useless when  $(h_{j_k} < 2)$  its length is less than 3 according to Property 2, we drop these columns. Update  $v_k$  and  $h_j$  immediately, and we drop these rows which the corresponding  $v_k$  is less than the support min-sup. Subsequently, repeat the procedure (i) until there is no such row or column.
- (ii) Similarly with step (ii) of (2), every row  $(I_{i_s}, I_{i_t})$  whose corresponding  $v_s$  is less than  $\lfloor \sqrt{n_1} \rfloor$  would be removed after finding and storing its frequent itemsets.

Then, the frequent 3-itemset support matrix  $D_2$  is produced by repeating the matrix compression procedure (i) and (ii) until no more row or column which is considered as a useless element could be found. That is,

$$D_2 = \begin{pmatrix} (I_{i_{h_{s_1}}}, I_{i_{r_{t_1}}}) \\ (I_{i_{h_{s_1}}}, I_{i_{r_{t_2}}}) \\ \vdots \\ (I_{i_{h_{s_u}}}, I_{i_{r_{t_v}}}) \end{pmatrix} \begin{matrix} T_{j_{p_1}} & T_{j_{p_2}} & \dots & T_{j_{p_w}} \\ \begin{pmatrix} d_{i_{h_{s_1}j_{p_1}}} d_{i_{r_{t_1}j_{p_1}}} & d_{i_{h_{s_1}j_{p_2}}} d_{i_{r_{t_1}j_{p_2}}} & \dots & d_{i_{h_{s_1}j_{p_w}}} d_{i_{r_{t_1}j_{p_w}}} \\ d_{i_{h_{s_1}j_{p_1}}} d_{i_{r_{t_2}j_{p_1}}} & d_{i_{h_{s_1}j_{p_2}}} d_{i_{r_{t_2}j_{p_2}}} & \dots & d_{i_{h_{s_1}j_{p_w}}} d_{i_{r_{t_2}j_{p_w}}} \\ \vdots & \vdots & \ddots & \vdots \\ d_{i_{h_{s_u}j_{p_1}}} d_{i_{r_{t_v}j_{p_1}}} & d_{i_{h_{s_u}j_{p_2}}} d_{i_{r_{t_v}j_{p_2}}} & \dots & d_{i_{h_{s_u}j_{p_w}}} d_{i_{r_{t_v}j_{p_w}}} \end{pmatrix} \end{matrix}$$

where  $j_1 \leq j_{p_1} < j_{p_2} < \dots < j_{p_w} \leq j_q$ ,  $(I_{i_{h_{s_y}}, I_{i_{r_{t_z}}}) \in \{(I_{i_{h_m}}, I_{i_{r_n}}) \mid m = 1, 2, \dots, s; n = 1, 2, \dots, t\}$ .

Let  $L''_2 = \{(I_{i_{h_{s_y}}, I_{i_{r_{t_z}}})\}$  be the compressed frequent 2-itemset of  $D_2$ .

4. Produce the frequent 3-itemset  $L_3$  and the frequent 4-itemset support matrix  $D_3$ . The frequent 3-itemset is the union of all 3-itemset subsets of the frequent itemsets generated in step (ii) of procedures (2) and (3), and a set defined as  $\{(I_{i_{h_{sm}}}, I_{i_{r_{tn}}}, I_{i_{r_{tk}}}) | (I_{i_{h_{sm}}}, I_{i_{r_{tn}}}), (I_{i_{h_{sm}}}, I_{i_{r_{tk}}}), (I_{i_{h_{sn}}}, I_{i_{r_{tk}}}) \in L'_2$  and inner product of corresponding row vectors of  $(I_{i_{h_{sm}}}, I_{i_{r_{tn}}})$  and  $(I_{i_{h_{sm}}}, I_{i_{r_{tk}}})$  in  $D_2$  is not less than min-sup}.

Similarly with previous steps, the intermediate matrix  $D'_3$  is produced by calculating “and” operation of the corresponding row vector of  $(I_{i_{h_{sm}}}, I_{i_{r_{tn}}})$  and  $(I_{i_{h_{sm}}}, I_{i_{r_{tk}}})$  in  $L'_2$ , which are derived from the element  $(I_{i_{h_{sm}}}, I_{i_{r_{tn}}}, I_{i_{r_{tk}}})$  in  $L_3$ .  $n_2$  is called row numbers of matrix  $D'_3$ .

- (i) Remove rows or columns using the same approach in step (i) of (2) or (3) and execute the following procedure.
  - (ii) When the sum of the corresponding row of  $(I_{i_{h_{sm}}}, I_{i_{r_{tn}}}, I_{i_{r_{tk}}})$  is less than and equal to  $\lceil \sqrt{n_2} \rceil$ , we find and store frequent itemsets containing items  $(I_{i_{h_{sm}}}, I_{i_{r_{tn}}}, I_{i_{r_{tk}}})$  by the same approach in step (ii) of (2), (3) again remove the corresponding row of  $(I_{i_{h_{sm}}}, I_{i_{r_{tn}}}, I_{i_{r_{tk}}})$ . Then the matrix compression procedure is repeated until no more row or column which is considered as a useless element could be found.
5. Analogously, the frequent 4-itemset, . . . , the frequent k-itemset is produced by step (2) to step (5) until the frequent k-itemset support matrix  $D_k$  is empty.

### 33.4 Algorithm Example Experiment Studying

Suppose that a transaction database is listed as Table 33.1 which is simulated for the number of min-sup is 2.

- (1) Generate the transaction matrix, and calculate the sum of each row  $v_s$  and the sum of each column  $h_s$  as described in Table 33.2.
- (2) Produce the frequent 1-itemset.  $L_1 = \{I_k | v_k \geq 2\} = \{I_1, I_2, I_3, I_4, I_5, I_6, I_7, I_8, I_9, I_{10}\}$ .
  - (i) It is obvious that the corresponding columns of  $T_5$  should be dropped since the sum of which is less than 2 ( $h_s < 2$ ). After updating  $v_s$ , the corresponding row of  $I_7$  is removed with regard to its  $v_s < 2$ . Then recalculate  $h_s$  and accordingly remove the corresponding columns of  $T_9$ . Finally, the new compression matrix is shown in Table 33.3.

**Table 33.1** Transaction database

ITD	Itemset
$T_1$	$I_1 I_2 I_3 I_8 I_9$
$T_2$	$I_1 I_3 I_4 I_5$
$T_3$	$I_2 I_3 I_4 I_5$
$T_4$	$I_1 I_3 I_4 I_5 I_8$
$T_5$	$I_7$
$T_6$	$I_3 I_4 I_5$
$T_7$	$I_2 I_6 I_9$
$T_8$	$I_2 I_3 I_5 I_6$
$T_9$	$I_1 I_7$
$T_{10}$	$I_2 I_3 I_4 I_6 I_9$

**Table 33.2** Transaction matrix

	$T_1$	$T_2$	$T_3$	$T_4$	$T_5$	$T_6$	$T_7$	$T_8$	$T_9$	$T_{10}$	$v_s$
$I_1$	1	1	0	1	0	0	0	0	1	0	4
$I_2$	1	0	1	0	0	0	1	1	0	1	5
$I_3$	1	1	1	1	0	1	0	1	0	1	7
$I_4$	0	1	1	1	0	1	0	0	0	1	5
$I_5$	0	1	1	1	0	1	0	1	0	0	5
$I_6$	0	0	0	0	0	0	1	1	0	1	3
$I_7$	0	0	0	0	1	0	0	0	1	0	2
$I_8$	1	0	0	1	0	0	0	0	0	0	2
$I_9$	1	0	0	0	0	0	1	0	0	1	3
$h_s$	5	4	4	5	1	3	3	4	2	5	

**Table 33.3** Compression matrix 1

	$T_1$	$T_2$	$T_3$	$T_4$	$T_6$	$T_7$	$T_8$	$T_{10}$	$v_s$
$I_1$	1	1	0	1	0	0	0	0	3
$I_2$	1	0	1	0	0	1	1	1	5
$I_3$	1	1	1	1	1	0	1	1	7
$I_4$	0	1	1	1	1	0	0	1	5
$I_5$	0	1	1	1	1	0	1	0	5
$I_6$	0	0	0	0	0	1	1	1	3
$I_8$	1	0	0	1	0	0	0	0	2
$I_9$	1	0	0	0	0	1	0	1	3
$h_s$	5	4	4	5	3	3	4	5	

(ii) Because  $\lceil \sqrt{n} \rceil = \lceil \sqrt{9} \rceil = 3$  and the corresponding  $v_s$  of  $I_1 I_6 I_8 I_9$  is less than and equal to 3, we need to find all the frequent itemsets containing items  $I_l$  ( $l = 1, 6, 8, 9$ ), then remove  $I_l$  ( $l = 1, 6, 8, 9$ ).

The given min-sup being 2, find frequent itemsets containing  $I_8$  firstly since  $v_8 = 2$ .  $S_8 = \{T_1, T_4 | d_{8j} = 1\}$  and the 2-combinations set of elements in  $S_8$  is  $S_8' = \{(T_1, T_4)\}$ . It is obvious that  $I_1$  and  $I_3$  are the rows whose matrix element with column  $T_1$  and  $T_4$  are both 1. So  $(I_1 I_3 I_8)$  is the only frequent itemset

**Table 33.4** Compression matrix<sup>2</sup>

	$T_1$	$T_2$	$T_3$	$T_4$	$T_6$	$T_7$	$T_8$	$T_{10}$	$v_s$
$I_2$	1	0	1	0	0	1	1	1	5
$I_3$	1	1	1	1	1	0	1	1	7
$I_4$	0	1	1	1	1	0	0	1	5
$I_5$	0	1	1	1	1	0	1	0	5
$h_s$	2	3	4	3	3	1	3	3	

**Table 33.5** Support matrix of the frequent 2-itemset

	$T_1$	$T_2$	$T_3$	$T_4$	$T_6$	$T_8$	$T_{10}$	$v_s$
$I_2$	1	0	1	0	0	1	1	4
$I_3$	1	1	1	1	1	1	1	7
$I_4$	0	1	1	1	1	0	1	5
$I_5$	0	1	1	1	1	1	0	5
$h_s$	2	3	4	3	3	3	3	

containing  $I_8$ , thus we store  $(I_1 I_3 I_8)$  and drop row  $I_8$ . Then another item  $I_1$  is considered,  $S_1 = \{T_1, T_2, T_4 | d_{1j} = 1\}$  and the 2-combinations set of elements in  $S_1$  is  $S_1' = \{(T_1, T_2), (T_1, T_4), (T_2, T_4)\}$ . Frequent itemsets containing items  $I_1$  are obtained by dealing with three 2-tuples in  $S_1'$  successively. Collection  $(I_1 I_3)$  is the frequent itemset determined by  $(T_1, T_2)$  using the similar approach in finding the frequent itemset containing  $I_8$ . Similarly,  $(T_1, T_4)$  determines collection  $(I_1 I_3)$  and  $(T_2, T_4)$  determines collection  $(I_1 I_3 I_4 I_5)$ . From the above, all the frequent itemsets containing items  $I_1$  are  $(I_1 I_3)$  and  $(I_1 I_3 I_4 I_5)$ . Continuing scanning other satisfied items accordingly, all the frequent itemsets containing items  $I_l$  ( $l = 1, 6, 8, 9$ ) are found:  $L'_1 = \{(I_1 I_3 I_8), (I_1 I_3 I_4 I_5), (I_6 I_2 I_3), (I_9 I_2 I_3)\}$ . After removing rows  $I_l$  ( $l = 1, 6, 8, 9$ ), the newly compressed matrix is shown in Table 33.4.

We drop the corresponding columns of  $T_7$  since the sum of which is less than 2 ( $h_s < 2$ ) and recalculate  $v_s$  again. Then the support matrix of the frequent 2-itemset is listed in Table 33.5. Regarding each row and column again, there is no useless element. In other words, the support matrix in Table 33.5 is fully compressed.

- (3) The frequent 2-itemset  $L_2$  is the union of the 2-itemset subsets produced by  $L'_1$  in step (ii) of procedure (2) and a set  $L'_2$  obtained from the support matrix in Table 33.5

$$\begin{aligned}
 L'_2 &= \{(I_i, I_j) \mid \sum_{k \in \{2,3,4,5\}} d_{ik} d_{kj} \geq 2, i < j, i, j = 1, 2, 3, 4, 6, 8, 10\} \\
 &= \{(I_2, I_3), (I_2, I_4), (I_2, I_5), (I_3, I_4), (I_3, I_5), (I_4, I_5)\}.
 \end{aligned}$$

That is  $L_2 = \{(I_1 I_3), (I_1 I_8), (I_3 I_8), (I_1 I_4), (I_1 I_5), (I_3 I_4), (I_4 I_5), (I_3 I_5), (I_2 I_3), (I_2 I_6), (I_6 I_3), (I_9 I_2), (I_9 I_3), (I_9 I_6)\}$ .

**Table 33.6** Uncompressed support matrix of the frequent 3-itemset

	$T_1$	$T_2$	$T_3$	$T_4$	$T_6$	$T_8$	$T_{10}$	$v_s$
$(I_2 I_3)$	1	0	1	0	0	1	1	4
$(I_2 I_4)$	0	0	1	0	0	0	1	2
$(I_2 I_5)$	0	0	1	0	0	1	0	2
$(I_3 I_4)$	0	1	1	1	1	0	1	5
$(I_3 I_5)$	0	1	1	1	1	1	0	5
$(I_4 I_5)$	0	1	1	1	1	0	0	4
$h_s$	1	3	6	3	3	3	3	

**Table 33.7** Support matrix of the frequent 3-itemset

	$T_2$	$T_3$	$T_4$	$T_6$	$T_8$	$T_{10}$	$v_s$
$(I_2 I_3)$	0	1	0	0	1	1	3
$(I_3 I_4)$	1	1	1	1	0	1	5
$(I_3 I_5)$	1	1	1	1	1	0	5
$(I_4 I_5)$	1	1	1	1	0	0	4
$h_s$	3	4	3	3	2	2	

Subsequently, the uncompressed support matrix of the frequent 3-itemset is constructed as listed in Table 33.6.

Firstly, we remove the corresponding columns of  $T_1$  by considering its  $h_s < 2$ . Where  $n_1 = 6$ ,  $\lceil \sqrt{n_1} \rceil = \lceil \sqrt{6} \rceil = 2$ . Secondly, because the corresponding  $v_s$  of  $(I_2 I_4)$   $(I_2 I_5)$  is equal to 2, we work out all the frequent itemsets containing  $(I_2 I_4)$  or  $(I_2 I_5)$ :  $L'_3 = \{(I_2 I_3 I_4), (I_2 I_3 I_5)\}$  and drop those corresponding rows. That is Table 33.7.

(4) Produce the frequent 3-itemset.

A frequent 3-itemset  $(I_3 I_4 I_5)$  is obtained from Table 33.7. And the frequent 3-itemset is  $L_3 = \{(I_3 I_4 I_5)\} \cup \{(I_1 I_3 I_8), (I_1 I_3 I_4) (I_1 I_3 I_5) (I_1 I_4 I_5), (I_3 I_4 I_5), (I_9 I_2 I_3), (I_9 I_2 I_6), (I_6 I_2 I_3)\}$  of  $L'_1 \cup \{(I_2 I_3 I_4) (I_2 I_3 I_5)\}$  of  $L'_3$ .

(5) Produce a frequent 4-itemset from  $(I_1 I_3 I_4 I_5)$  in  $L'_1$ . While no frequent 4-itemset could be found in Table 33.7, so our algorithm ends.

### 33.5 Conclusion

An algorithm of mining association rule based on matrix is able to discover all the frequent item sets only by searching the database once and not generating the candidate itemsets, but generating the frequent itemsets directly, which is more efficient. Many researchers have done a great deal of work on it. Here, a new algorithm for generating association rules based on matrix is proposed. It compresses the transaction matrix efficiently by integrating various strategies and achieves better performance than the known algorithms based on matrix. Some new strategies of compressing the transaction matrix are worthy of further research.

**Acknowledgment** This work is financially supported by the Natural Science Foundation of the Jiangxi Province of China under Grant No. 20122BAB201004.

## References

1. Agrawal, R., Imielinski, T., & Wami, A. S. (1993). *Mining association rules between sets of items in large databases* (pp. 207–216). Proceeding of the ACM SIGMOD Conference on Management of Data, Washington, DC.
2. Lv, T. X., & Liu, P. Y. (2011). Algorithm for generating strong association rules based on matrix. *Application Research of Computers*, 28(4), 1301–1303.
3. Cao, F. H. (2012). Improved association rule mining algorithm based on two matrixes. *Electronic Science and Technology*, 25(5), 126–128.
4. Xu, H. Z. (2012). The research of association rules data mining algorithms. *Science Technology and Engineering*, 12(1), 60–63.
5. He, B., & Xue, F. (2012). An improved algorithm for mining association rules. *Computer Knowledge and Technology*, 8(5), 1015–1017.

# **Retraction: Decoupling Interrupts from the Internet in Markov Models**

**Jinwen Ma, Jingchun Zhang, and Jinrong Guo**

School of Information Science and Engineering, Lanzhou University,  
Lanzhou 730000, China  
e-mail: guojr11@lzu.edu.cn

Several conference proceedings have been infiltrated by fake submissions generated by the SCIGen computer program. Due to the fictional content the chapter “Decoupling Interrupts from the Internet in Markov Models” by “Jinwen Ma, Jingchun Zhang, and Jinrong Guo” has been retracted by the publisher. Measures are being taken to avoid similar breaches in the future.



# Chapter 35

## Parallel Feature Selection Based on MapReduce

Zhanquan Sun

**Abstract** Feature selection is an important research topic in machine learning and pattern recognition. It is effective in reducing dimensionality, removing irrelevant data, increasing learning accuracy, and improving result comprehensibility. However, in recent years, data has become increasingly larger in both number of instances and number of features in many applications. Classical feature selection method is out of work in processing large-scale dataset because of expensive computational cost. For improving computational speed, parallel feature selection is taken as the efficient method. MapReduce is an efficient distributional computing model to process large-scale data mining problems. In this paper, a parallel feature selection method based on MapReduce model is proposed. Large-scale dataset is partitioned into sub-datasets. Feature selection is operated on each computational node. Selected feature variables are combined into one feature vector in Reduce job. The parallel feature selection method is scalable. The efficiency of the method is illustrated through example analysis.

### 35.1 Introduction

In recent years, data has become increasingly larger in both number of instances and number of features in many applications such as genome projects, text categorization, image retrieval, and customer relationship management [1, 2]. It may cause serious problems to many machine learning algorithms with respect to scalability and learning performance. How to select the most informative variable combination is a crucial problem. Feature selection is a process of choosing a subset of original features so that the feature space is optimally reduced according to a certain

---

Z. Sun (✉)

Shandong Provincial Key Laboratory of Computer Network, Shandong Computer Science Center, Jinan, Shandong 250014, China  
e-mail: [sunzhq@sdas.org](mailto:sunzhq@sdas.org)

evaluation criterion [3]. It is effective in reducing dimensionality, removing irrelevant data, increasing learning accuracy, and improving result comprehensibility. Therefore, feature selection becomes very necessary for machine learning tasks when facing high-dimensional data nowadays.

Correlation analysis is the basis of feature selection. Commonly used correlation metric is correlation coefficient, which can only measure linear correlations between variables. Another commonly used method is stepwise regression. It is mostly used to linear regression problems. Entropy is a metric that can measure uncertainty of random variables. Mutual information based on entropy can measure arbitrary statistic dependences between variables. Feature selection based on mutual information has been widely used [4, 5]. But with the development of electronic and computer technology, the quantity of electronic data is in exponential growth [6]. Data deluge has become a salient problem to be solved. Scientists are overwhelmed with the increasing amount of data processing needs arising from the storm of data that is flowing through virtually every science field, such as bioinformatics [7], biomedical [8, 9], cheminformatics [10], and web [11]. Classic feature selection method is out of work when the dataset is very big.

Feature selection method based on parallel algorithm will be the mainly choice for dealing with large-scale data. Many parallel algorithms are implemented using different parallelization techniques such as threads, MPI, MapReduce, and mash-up or workflow technologies yielding different performance and usability characteristics [12]. MPI model is efficient in computation intensive problems, especially in simulation. But it is not easy to be used in practical. MapReduce is a cloud technology developed from the data analysis model of the information retrieval field. Hadoop is the most popular open source MapReduce software. The MapReduce architecture of Hadoop doesn't support iterative Map and Reduce tasks, which is required in many data mining algorithms. Professor Fox developed an iterative MapReduce architecture software Twister. The manner of Twister MapReduce is "configure once, and run many time" [13, 14]. In this paper, a parallel feature selection method based on MapReduce model is proposed. Firstly, large-scale dataset are partitioned into small sub-datasets. Feature selection is applied on each sub-dataset. The analysis result is combined into a global result. For simplicity, the feature selection is used to analyze binary variables. The selected feature variables are input to SVM for classification. The efficiency of the proposed method is illustrated through an example analysis.

## 35.2 Mutual Information Based on Shannon Entropy

Mutual information can measure any kind of statistical dependence between variables. It has been widely applied in pattern recognition field. Probability is the basis of entropy. Many kinds of definition of entropy had been developed. The widely used one is Shannon entropy. We will introduce Shannon entropy as follows.

Feature variables are denoted by a vector  $\mathbf{X} = (X_1, X_2, \dots, X_p)^T$ , where  $X_i = (x_{ij})$ ,  $i = 1, 2, \dots, p$ ,  $j = 1, 2, \dots, q$  denotes the  $i$ th feature variable and each variable has  $q$  different values. Class variable is denoted by  $C$ ,  $C = (c_i)$ ,  $i = 1, 2, \dots, k$  that means all features are projected to  $k$  different classes. In this paper, both the feature variables and the class variable are supposed to be discrete.  $p_{X_i}$  is the probability distribution of the feature variable  $X_i$ ,  $p_C$  is the probability distribution of the class variable  $C$ , and  $p_{X_i C}$  is the joint probability distribution of  $X_i$  and  $C$ . All probability distributions are obtained through statistics of samples. The Shannon entropy  $H$  of the feature variable  $X_i$  can be described as

$$H(X_i) = -\sum_{j=1}^q p_{x_{ij}} \log p_{x_{ij}} \quad (35.1)$$

Shannon entropy of the class variable  $C$  can be described as

$$H(C) = -\sum_{i=1}^k p_{c_i} \log p_{c_i} \quad (35.2)$$

Joint entropy between the feature variables and the class variable is

$$H(X_i, C) = -\sum_{j=1}^q \sum_{l=1}^k p_{x_{ij}c_l} \log p_{x_{ij}c_l} \quad (35.3)$$

where  $X_i$  can be substituted by a subset of feature set  $X$ , i.e., the joint entropy can be generalized to the condition of  $p$  variables.

Mutual information between feature variables and class variable based on Shannon entropy can be defined as follows.

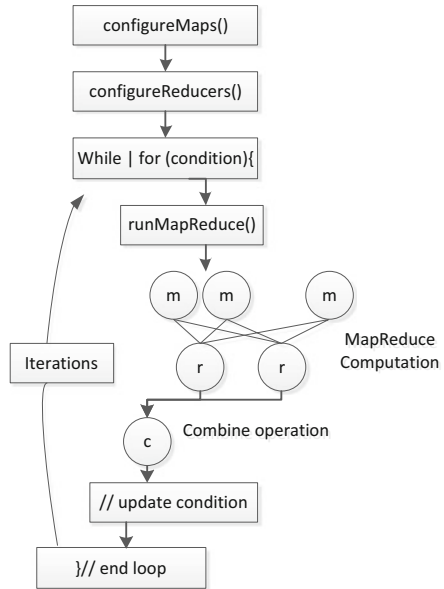
$$I(X_i; C) = H(X_i) + H(C) - H(X_i, C) = \sum_{j=1}^q \sum_{l=1}^k p_{x_{ij}c_l} \log \frac{p_{x_{ij}c_l}}{p_{x_{ij}}q_{c_l}} \quad (35.4)$$

where  $X_i$  can be substituted by subset of feature set  $X$ .

### 35.3 MapReduce Model Based on Twister

There are many parallel algorithms with simple iterative structures. Most of them can be found in the domains such as data clustering, dimension reduction, link analysis, machine learning, and computer vision. These algorithms can be implemented with iterative MapReduce computation. Professor Fox developed

**Fig. 35.1** Twister's programming model



the first iterative MapReduce computation model Twister. Twister's programming model can be described as in Fig. 35.1.

MapReduce jobs are controlled by the client program. During configuration, the client assigns MapReduce methods to the job, prepares Key/Value pairs, and prepares static data for MapReduce tasks through the partition file if required. Between iterations, the client receives results collected by the Combine method and, when the job is done, exits gracefully.

Map daemons operate on computation nodes, loading the Map classes and starting them as Map workers. During initialization, Map workers load static data from the local disk according to records in the partition file and cache the data into memory. Most computation tasks defined by the users are executed in the Map workers. Twister uses static scheduling for workers in order to take advantage of the local data cache.

Reduce daemons operate on computation nodes. The number of reducers is prescribed in client configuration step. The Reduce jobs depend on the computation results of Map jobs. The communication between daemons is through messages.

Combine job is to collect MapReduce results. Twister uses scripts to operate on static input data and some output data on local disks in order to simulate some characteristics of distributed file systems. In these scripts, Twister parallel distributes static data to compute nodes and create partition file by invoking Java classes.

## 35.4 Parallel Feature Selection Based on MapReduce

Feature selection is usually used to select the most informative feature combination with least information loss for classification problems. If we can select the most informative variables for classification, it will save lots of computation cost and reduce the effect of noise. The information between class variable and feature variables is measured with mutual information. Feature selection of classification is to select the feature variable combination that has the largest mutual information value with class variable. The parallel feature selection of classification based on mutual information metric can be formulized as follows.

Step 1: Initial dataset  $D$  is divided into  $n$  sections  $D_1, D_2, \dots, D_n$ . Each sub-dataset is deployed to each computational node. The number of selected features  $k$  is prescribed.

Step 2: Suppose  $S$  and  $V$  are the two vectors and set  $S = \Phi$  and  $V = \{X_1, X_2, \dots, X_N\}$ .  $S$  denotes selected features and  $V$  denotes unselected features.

Step 3: In computational node  $i, i \in \{1, 2, \dots, m\}$ , mutual information between  $\{S, X_i\}, i = 1, 2, \dots, N$  and  $Y$  is calculated. The feature variable  $X_j, j \in \{1, 2, \dots, N\}$  that maximizes  $I(\{S, X_j\}; Y)$  is selected. The selected variable's serial number  $j$  and corresponding mutual information  $I(\{S, X_j\}; Y)$  are collected to Reduce program.

Step 4: In Reduce program, the feature variable  $X_j, j \in \{1, 2, \dots, N\}$  with maximum count is selected. If the counts of two feature variables are equal, select the one with bigger mutual information value. Set  $S \leftarrow \{S, X_j\}$  and  $V \leftarrow V \setminus \{X_j\}$ .

Step 5: The changed  $S$  and  $V$  are feedbacks to step 3. Iterate the process until the selected feature variable's number reaches  $k$ .

The selection process based on MapReduce is shown in Fig. 35.2.

## 35.5 Example

### 35.5.1 Data Source

The source data are downloaded from NEC Laboratories America, Inc. website <http://ml.nec-labs.com/download/data/milde/>. In the adult database, 123 attributes are labeled two classes. Each attribute denoted by binary variable, i.e., 0 or 1. Labels are denoted by +1 or -1. It is a binary classification problem. The database includes two files. One is used for training and the other is used for testing. The training file includes 32,562 samples. The testing file includes 16,282 samples. In this example, four computational nodes are used. Training data are partitioned into  $m$  sections randomly. Each section has roughly equal number data.

**Fig. 35.2** Feature selection process based on MapReduce

```



---


Preparation
  Computation environment configuration
  Data partition and distribution to the computation nodes
  Create partition file


---


Main class
  JobConf; //configure the MapReduce parameters and classnames
  TwisterDriver; //to initiate the MapReduce tasks
  While(condition) //selected feature variable number doesn't reach
  prescribed value
    TwisterDriver; //Map the selected feature variable vector to Map
    job
    Get feedback results;
  If(condition) break; //the number of selected feature variables reach
  prescribed value


---


End main class


---


Map class
  Read data broadcasted by Main class


---


  Load data from local file system;
  CalMutualInformation(); //calculate mutual information between
  each combination {S,Xi}
  Collector; //sent the serial number i who has the maximum mutual
  information value and corresponding mutual information value to
  Reduce job


---


End Map class


---


Reduce class
  Read data transmitted from Map job;
  Select the final serial number with biggest count number;
  Collect; //feedback the selected final serial number;


---


End Reduce class


---



```

All examples are analyzed in India cluster node of FutureGrid. Twister0.9 software is deployed in each computation nodes. ActiveMQ is used as message broker. The configuration of each virtual machine is as follows. Each node is installed Ubuntu Linux OS. The processor is 3GHz Intel Xeon with 10GB RAM.

### 35.5.2 Feature Selection Based on the Proposed Method

Apply the proposed parallel feature selection method on the training samples. The number of feature selection is prescribed 20. Data is partitioned into four, two, and one sections, respectively. Parallel SVM introduced in reference [15] is used as the classifier. It is operated on four computational nodes also. The feature selection results and classification correct rates are listed in Table 35.1.

### 35.5.3 Feature Selection with Correlation Coefficient

For comparison, the feature variables are selected according to correlation coefficient. It is used to measure the correlation between class variable  $Y$ , and attribute variable  $X$ . The correlation coefficient is calculated as the following equation.

**Table 35.1** Feature selection and classification results

Number of nodes	Selected features	Training time (s)	Correct rate
1	39,38,0,63,72,34,79,77,1,48,13,74,5,66,51,17,50,3,82,76	1,882.873	84.30
2	39,38,74,0,81,34,1,50,5,76,72,13,66,79,2,21,15,51,14,3	951.303	83.99
4	39,38,74,0,34,81,1,50,6,76,66,15,72,5,79,4,13,51,16,21	475.657	84.32

**Table 35.2** Analysis result based on feature selection with correlation coefficient

Selected features	Correct rate
39,62,38,41,74,73,0,61,81,72,71,50,63,51,77,18,28,34,48,3	81.32

$$\rho_{X,Y} = \frac{\text{cov}(X,Y)}{\sigma_X \sigma_Y} = \frac{E[(x - \mu_X)(y - \mu_Y)]}{\sigma_X \sigma_Y} \tag{35.5}$$

where  $\text{cov}(X, Y)$  is the covariance of the two variables,  $\sigma_X, \sigma_Y$  are the standard deviations of  $X, Y$ . After calculating the correlation coefficient, 20 feature variables are selected. The selected variables are taken as the input of parallel SVM. The classification correct rate is listed in Table 35.2.

### 35.5.4 Results Analysis

From the analysis results of Table 35.1, we can find that the computation speed of feature selection is improved markedly. The accelerate ratio is approximate linear. The classification results show that classification correct rates of different partition plan are similar. It illustrates that the parallel feature election method is effective and efficiency. The analysis results of Tables 35.1 and 35.2 show that the feature selection result based on mutual information is good than that of classical feature selection method.

## 35.6 Conclusion

Feature selection is an important task of machine learning and pattern recognition. Feature selection based on mutual information is taken as one of the most efficient methods. For improving the computation speed, a novel parallel feature selection method based on MapReduce is proposed. It can accelerate the computation speed almost linearly. The example analysis results show that the proposed method is efficient in reducing computational cost. The classification correct rate based on parallel feature selection method is similar to that of feature selection without partition.

Though the method is efficient in dealing with large-scale feature selection problem, it is only useful to binary feature and class variables. How to process feature selection problem of multivalued or continuous variables is still to be studied. Furthermore, the number of selected variables is prescribed subjectively in the proposed method. How to determine the number of selected variable is our further study.

**Acknowledgments** This work is partially supported by National Youth Science Foundation (No. 61004115), National Science Foundation (No. 61272433), and Provincial Fund for Nature project (No. ZR2010FQ018).

## References

1. Yu, L., & Liu, H. (2003). *Feature selection for high-dimensional data: A fast correlation-based filter solution* (pp. 856–863). Twentieth International Conference on Machine Learning. American Association for Artificial Intelligence.
2. Dash, M., & Liu, H. (2009). Dimensionality reduction. In *Encyclopedia of database systems* (pp. 843–846).
3. Liu, H., & Motoda, H. (1998). *Feature selection for knowledge discovery and data mining* (pp. 23–45). Boston: Kluwer.
4. Kwak, N., & Choi, C. H. (2002). Input feature selection for classification problems. *IEEE Transactions on Neural Networks*, 13(1), 143–159.
5. Kari, T. (2003). Feature extraction by non-parametric mutual information maximization. *Journal of Machine Learning Research*, 3, 1415–1438.
6. Swedlow, J. R., Zanetti, G., & Best, C. (2011). Channeling the data deluge. *Nature Methods*, 8, 463–465.
7. Fox, G. C. Qiu, X. H. Beason, S. Choi, J. Y. Ekanayake, J. Gunarathne., T. et al. (2009). Biomedical case studies in data intensive computing. *Lecture Notes in Computer Science*, 5931, 2–18.
8. Blake, J. A. & Bult C. J. (2006). Beyond the data deluge: Data integration and bio-ontologies. *Journal of Biomedical Informatics*, 39(3), 314–320.
9. Qiu J. (2010). Scalable programming and algorithms for data intensive life science. *Journal of Integrative Biology*, 15(4), 1–3.
10. Guha, R. Gilbert, K. Fox, G. C. Pierce, M. Wild, D. & Yuan H. (2010). Advances in cheminformatics methodologies and infrastructure to support the data mining of large, heterogeneous chemical datasets. *Current Computer-Aided Drug Design*, 6(1), 50–67.
11. Chang, C. C. He, B. & Zhang Z. (2004). Mining semantics for large scale integration on the web: evidences, insights, and challenges. *SIGKDD Explorations*, 6(2), 67–76.
12. Fox, G. C. Bae, S. H. Ekanayake, J. Qiu, X. H. & Yuan H. P. (2008). *Parallel data mining from multicore to cloudy grids* (pp. 311–340). High Performance Computing and Grids Workshop. IOS Press.
13. Zhang, B. J. Ruan, Y. Wu, T. L. Qiu, J. Hughes, A. & Fox G. (2010). *Applying twister to scientific applications* (pp. 25–32). Proceedings of CloudCom. IEEE CS Press.
14. Ekanayake, J. Li, H. Zhang, B. J. Gunarathne, Bae, S. H. Qiu, J. et al. (2010). *Twister: A runtime for iterative MapReduce* (pp. 810–818). The First International Workshop on MapReduce and Its Applications of ACM HPDC. ACM Press.
15. Sun, Z. Q. & Fox G. C. (2012). *Study on parallel SVM based on MapReduce* (pp. 495–501). International Conference on Parallel and Distributed Processing Techniques and Applications. CSREA Press.



# Chapter 36

## Initial State Modeling of Interlocking System Using Maude

Rui Ma, Zhongwei Xu, Zuxi Chen, and Shuqing Zhang

**Abstract** In order to do formal verification of interlocking system, which is complicated but safety critical, we choose formal specification language Maude for modeling and verification based on membership equational logic and rewriting logic. In this chapter, a method is proposed to show how the initial state can be modeled and contains important information of specific interlocking system. And a case of Tongji Test Line is reported to illustrate this method in detail. The verification results show that Maude can be applied to formal object-oriented specification and model checking of railway interlocking system successfully using the proposed modeling method.

### 36.1 Introduction

Railway interlocking system is a safety-critical system and is becoming increasingly more complex. There is therefore a clear need for formal description and verification of interlocking system to prove the correctness and ensure safety [1].

Maude is not only a high-level language but also a high-performance system supporting executable specification and declarative programming [2, 3]. And it supports a range of formal analysis methods, including rewriting for simulation, search for reach ability analysis, and linear temporal logic model checking. By using Maude, a formal language which is characterized by rewriting logic, the model of static and dynamic attribute, circuit, and behavior for the interlocking system can be formed. And they constitute a general formal framework. The detail activities can be described by rules and equations of rewriting logic.

---

R. Ma (✉) • Z. Xu • Z. Chen • S. Zhang  
School of Electronics and Information Engineering, Tongji University,  
Shanghai 201804, China  
e-mail: [marymary1988@163.com](mailto:marymary1988@163.com)

Given these features and applications of Maude, it seems reasonable to assume that the tool should be a good choice for the formal modeling and verification of interlocking systems.

The main contributions of this chapter are the following: (1) shows a case in which Maude has been applied to formally model and verify the interlocking system successfully, (2) focuses on and proposes a method for initial state modeling, and (3) illustrates the method with a case of Tongji Test Line in details.

The rest of the chapter is organized as follows. Section 36.2 describes Maude. Section 36.3 gives a general method for modeling and verification of interlocking system. Section 36.4 shows how to specify the initial state model in Maude. Section 36.5 illustrates the process of modeling the initial state of an interlocking system, Tongji Test Line. Section 36.6 finally concludes the chapter.

## 36.2 Maude

Maude is particularly suitable for distributed and concurrent systems. Besides the features mentioned in last section, it is based on the membership equational logic and rewriting logic. With this logic, the behavior and the state transition relation of complicated systems can be simplified to some extent, thus simplifying the process of formal modeling and verification of the interlocking system.

### 36.2.1 Membership Equational Logic

Maude's functional modules are theories in membership equational logic, a Horn logic whose atomic sentences are equalities and membership assertions of the form, stating that a term  $t$  has sort  $s$ . Such logic extends order-sorted equational logic and supports sorts, subsort relations, subsort polymorphic overloading of operators, and definition of partial functions with equationally defined domains. Maude's functional modules are assumed to be Church–Rosser and terminating; they are executed by the Maude engine according to the rewriting techniques and operational semantics [4, 5].

A Maude program containing only equations is called a functional module. It is a functional program defining one or more functions by means of equations, used as simplification rules. Such equations (specified in Maude with the keyword `eq` and ended with a period) are used from left to right as equational simplification rules [6].

In Maude, equations can also be conditional, that is, they may only be applied if a certain condition holds, where `ceq` is the Maude keyword introducing conditional equations.

### 36.2.2 *Rewriting Logic*

Membership equational logic is a sublogic of rewriting logic. A rewrite theory is a pair of  $T$ , a membership equational theory, and  $R$ , a collection of labeled and possibly conditional rewrite rules involving terms in the signature of  $T$ . A rewrite theory has both rules and equations. Hence a Maude program containing rules and possibly equations is called a system module. The rewrite rules in  $R$  are not equations. Computationally, they are interpreted as local transition rules in a possibly concurrent system. Logically, they are interpreted as inference rules in a logical system. Rules are introduced with the keyword `r1` while conditional rules with the keyword `cr1` [6].

The essential idea of rewriting logic is that the semantics of rewriting can be drastically changed in a very fruitful way [7]. We no longer interpret a term  $t$  as a functional expression, but as a state of a system; and we no longer interpret a rewrite rule as an equality, but as a local state transition, stating that if a portion of a system's state exhibits the pattern described, then that portion of the system can change to the corresponding instance of  $t$ . Furthermore, such a local state change can take place independently from, and therefore concurrently with, any other non-overlapping local state changes. Rewriting logic is therefore a logic of concurrent state change [8, 9].

## 36.3 Modeling and Verification

The whole method for modeling and verification contains three steps:

1. *Modeling*: specify a model of this system using Maude.
2. *Deriving safety properties*: the safety properties related to the specific interlocking system should be generically derived from the model.
3. *Verification*: verify whether a given model satisfies all the safety properties.

If the methods are successfully applied and a model of an interlocking system is verified satisfying the gained safety properties, we can conclude that the interlocking system behaves as expected in relation to the documentation and is safety enough.

### 36.3.1 *Modeling*

The general idea of modeling is configurations. Maude is object-oriented. In a concurrent object-oriented system, the concurrent state, which is usually called a configuration, has typically the structure of a multiset made up of objects. And that objects evolve by concurrent rewriting modulo associativity, commutatively, and

identity, using rules that describe the effects of communication events between objects.

An object is represented in Maude as a term where  $O$  is the object's name or identifier,  $C$  is its class identifier,  $a_i$ 's are the names of the object's attribute identifiers, and  $S_i$ 's are the corresponding values.

$$\langle O : C \mid a_1 : S_1, a_2 : S_2 \cdots a_i : S_i \rangle$$

All the components of the interlocking system can be divided into classes and be modeled separately as mod in Maude. The classes are divided from top to down until it is indivisible. For example, the Class Electrical Component can be divided into fuse, button, and so on. Each mod has its own attributes and behaviors. All of the classes can use the equations and rules in Maude when modeling. And they constitute a general framework of the interlocking system model.

### 36.3.2 *Deriving Safety Properties*

The safety properties can be split up in to two categories:

1. *General safety properties.* These safety properties apply to all the railway systems and define that collision and derailling must never occur.
2. *Train route table safety properties.* These safety properties are related to specific interlocking systems. The interlocking systems ensure safety on stations by introducing the concept of train routes, which include the requirements must be satisfied, and are used to ensure the system behaves as expected. So some safety properties have to be derived from the train routes.

After all the safety properties are gained, they should be specified in Maude.

When all the safety properties have been specified, they form a mod SAFETY-PROPERTIES in Maude. It is one of the mod of the complete model and will be verified whether the model of the interlocking system satisfies all the safety properties.

### 36.3.3 *Verification*

The verification model contains three parts:

1. *Well formed:* This part is to verify the static properties of the interlocking system. For example, the lengths of the track sections are larger than zero. These properties should be specified in Maude.
2. *Confidence conditions:* Confidence conditions must be verified dynamically as they depend on the dynamic behavior of the model. For example, dropping and

drawing relays may not be cyclic. These properties should be specified in Maude.

3. *Safety properties*: This part is the safety properties mentioned in the previous section, general safety properties and train route table safety properties. And of course, they have been specified in Maude.

Only after specifying all these properties mentioned above in Maude is the modeling of the verification model finished.

Each properties check mentioned above is defined as an operator in the model. For each of these operators, a proposition is specified. We should use Maude to check each of them. Several examples are:

```
reduce in VERIFICATION : wfCircuit.
result: true
```

```
reduce in VERIFICATION : modelCheck(initState, []drawConfidenceCondition).
result: true
```

And it is important that the results of verifying the model in relation to the safety properties are only sound if the given configuration is well formed and if the confidence conditions hold.

The verification is complete after verifying all the three parts.

## 36.4 Initial State Modeling

The initial state modeling is one part of the whole formal modeling. And this part is very important since it contains the information of specific interlocking system.

When we model initial state, we form a mod named INIT-STATE in Maude. The information in this mod can be divided into two parts.

One is all the instantiated objects of the components classes. In other words, we only generate a format for the attributes of the interlocking system components in the previous modeling but haven't decided the definite value of them. Therefore, all these have to be done in the mod INIT-STATE. Then we gained instantiated objects. For example, there are some instantiate objects:

```
For relay:
< `tr2-111 : RegularRelay | drawn : false >
For signal:
< `sX2 : Signal | lamps : ('X2roe, 'X2gr, 'X2brt) >
```

(continued)

(continued)

**For lineartracksection:**

```
< 'X6G : LinearTrackSection | relay : 'X6G-111, length : 30, platform : true,
neighbour1 : outsideStation, neighbour2 : '2DG, signal1 : noSignal, signal2 :
noSignal >
```

**For trainroute:**

```
< 'tr2 : TrainRoute | lockRelay : 'tr2-111, signals : ([ 'sX3 x green], [ 'sX4 x
red]), points : ([ '1DG x minus]), trackSections : ('1DG, '5G), displayStop :
'1DG, locking : [ '1DG x '5G], conflicting : ('tr3, 'tr4, 'tr5, 'tr6, 'tr7, 'tr8, 'tr9,
'tr10, 'tr11, 'tr12) >
```

The second part is the current modeling. Wires of a circuit specify how the components are connected and whether they are conducting or not. The conducting information is very important to the control logic of the interlocking system. In this chapter the entire electrical circuit is split up into unique linear paths between the positive and negative poles. And each path should be specified in Maude according to the following format:

```
op pathX : Configuration -> Bool.
ceq pathX(..) = true if ...
```

pathX is the name of this operator. X can be the number of the path or whatever you like to name it. If the state matches the argument of the equation and the condition rewrites to true, the path is conducting. The dots in the parentheses stand for the configurations, especially the components in this path should be in the form of objects. And the dots on the right side stand for the conditions this conditional equation should satisfy. Otherwise, the configuration matches

```
eq pathX(CONF) = false [owise].
```

meaning that the path is nonconducting.

Then it will be possible to determine which components are conducting using the following operator:

```
op conducting : Qid Configuration -> Bool.
ceq conducting(C, CONF) = true
if pathX(CONF).
eq conducting(C, CONF) = false [owise].
```

The above equation matches and is rewritten to true if the object (id C) is conducting in pathX.

According to these formats, whether an object is conducting depends on the information of the other objects in the same path and the conditions. To another perspective, these can be seen as the restrictive relationships between them. Therefore, the current modeling can describe some control logic of the interlocking system.

### 36.5 Case of Initial State Modeling

Then a case of Tongji Test Line will show how the initial state of interlocking system can be specified in Maude. Figure 36.1 is the track layout of Tongji Test Line and Table 36.1 is the train route table of it.

The train route table defines some rules the trains and the track layout have to obey when train drives in this station, forming fixed routes in other words. For instance, train route 1 is from signal X2 to signal X1. When you choose this route, the button X2LA and X1LA should be pushed. This route contains no point and train drives through the track section X2G. Train can only be permitted to drive in when signal X2 is green (L for green and B for white). At last, this route has no mutually conflicting signals.

First, all the objects should be instantiated as mentioned in previous section.

And then come to the current modeling. The current modeling has eight parts in all according to the control logic. They are track section occupying logic, point locating logic, point position representing logic, train route locking logic, train route unlocking logic, signal logic, lamp logic, and lamp mutually exclusive logic. Here examples of several logic modeling are given.

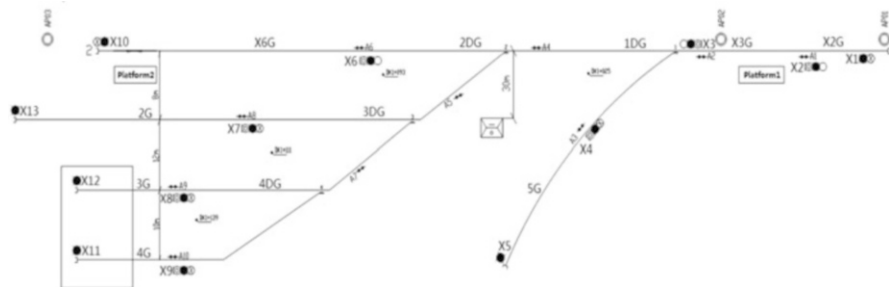


Fig. 36.1 Tongji Test Line track layout

**Table 36.1** Tongji Test Line train route table

Direction	Train routes	Signals		Points	Mutually conflicting signal	Track sections	Route number
		Route buttons	Name				
X2	To X1	X2LA, X1LA	X2	L		X2G	1
X3	To X5	X3LA, X4LA	X3	B	X4	1DG,5G	2
X3	To X10	X3LA, X6LA	X3	L	X6	1DG,2DG,X6G	3
X3	To X13	X3LA, X7LA	X3	B	X7	1DG,2DG,3DG,2G	4
X3	To X12	X3LA, X8LA	X3	B	X8	1DG,2DG,3DG,4DG,3G	5
X3	To X11	X3LA, X9LA	X3	B	X9	1DG,2DG,3DG,4DG,4G	6
X4	To X2	X4LA, X3LA	X4	B	X3	1DG,X3G	7
X6	To X2	X6LA, X3LA	X6	L	X3	2DG,1DG,X3G	8
X7	To X2	X7LA, X3LA	X7	B	X3	3DG,2DG,1DG,X3G	9
X8	To X2	X8LA, X3LA	X8	B	X3	4DG,3DG,2DG,1DG,X3G	10
X9	To X2	X9LA, X3LA	X9	B	X3	4DG,3DG,2DG,1DG,X3G	11
X6	Through X2 to X1	X6LA, X2LA	X6/X2	L/L	X3	2DG,1DG,X3G,X2G	12



### 36.5.1 Track Section Occupying Logic

A relay is used to represent whether the track section is occupied. This logic is specified as follows:

```
ceq path1(< 'X2G : V@TrackSection | ATTS > Conf) = true
if not isOccupied('X2G, < 'X2G : V@TrackSection | ATTS > Conf).
eq path1(Conf) = false [owise].
ceq conducting('X2G-111, Conf) = true if path1(Conf).
```

The track section with id 'X2G has a relay('X2G-111), V@TrackSection is a variable, and it can be matched with linear track section or point. ATTS and Conf are all variables. ATTS denotes other attribute in this object. Conf denotes other configurations besides this object. If 'X2G isn't occupied, path1 is conducting. And then 'X2G-111 is conducting, meaning drawing. Occupying leads to dropping otherwise.

### 36.5.2 Point Locating Logic

Buttons are used to decide the position of the point. And each position of the point, plus and minus, has a relay to control the locating. They are specified as follows:

```
ceq path24(< 'b11 : Button | pushed : D > Conf) = true
if isPushed('b11, < 'b11 : Button | pushed : D > Conf).
eq path24(Conf) = false [owise].
ceq conducting('1DG-110, Conf) = true if path24(Conf).
ceq path25(< 'b12 : Button | pushed : D > Conf) = true
if isPushed('b12, < 'b12 : Button | pushed : D > Conf).
eq path25(Conf) = false [owise].
ceq conducting('1DG-121, Conf) = true if path25(Conf).
```

The point(id '1DG) locating is controlled by button 'b11 to plus and button 'b12 to minus. If the button 'b11 is pushed, path24 is conducting. Then relay '1DG-110 is conducting. The point moves to the plus, to the minus otherwise as path25 specified.

### 36.5.3 Train Route Locking Logic

A train route is locked; that is to say this train route is ready for train to come in. Hence, much information has to be examined before locking a train route to ensure safety. All the track sections in this train route must not be occupied. Points must be in the correct position. All contradiction routes are unlocked. Then a train route can be locked, pressing the button of the beginning and ending. One relay is used to represent whether all the information has been examined and the button of the beginning is pushed. Another relay is used to represent whether the train route is locked.

```

ceq path13(*** track section unoccupied
<'1DG-111-11 : Contact | relay : '1DG-211, upper : true >
<'5G-111-11 : Contact | relay : '5G-111, upper : true > *** points position correct
<'1DG-210-61 : Contact | relay : '1DG-210, upper : false >
<'1DG-221-11 : Contact | relay : '1DG-221, upper : true >
<'x3la : Button | pushed : D > *** beginning
<'tr3-111-61 : Contact | relay : 'tr3-111, upper : false >
*** contradiction routes unlock
<'tr4-111-61 : Contact | relay : 'tr4-111, upper : false >
<'tr5-111-61 : Contact | relay : 'tr5-111, upper : false >
...
<'tr11-111-61 : Contact | relay : 'tr11-111, upper : false >
<'tr12-111-62 : Contact | relay : 'tr12-111, upper : false > Conf) = true
*** conditions respect to the elements in path
if conducting('5G-111, Conf)
  and-then conducting('1DG-111, Conf)
  and-then conducting('1DG-221, Conf)
  and-then not conducting('1DG-210, Conf)
  and-then isPushed('x3la, <'x3la : Button | pushed : D > Conf)
  ...
  and-then not conducting('tr11-111, Conf)
  and-then not conducting('tr12-111, Conf) .
  eq path13(Conf) = false [owise] .
  ceq conducting('tr2a-111, Conf) = true if path13(Conf) .
  eq conducting('tr2a-111, Conf) = false [owise] . *** ready for locking
  ceq path66(<'tr2a-111-11 : Contact | relay : 'tr2a-111, upper : true >
  <'x4la : Button | pushed : D > Conf) = true *** ending
  if conducting('tr2a-111, Conf)
  and-then isPushed('x4la, <'x4la : Button | pushed : D > Conf) .
  eq path66(Conf) = false [owise] .
  ceq conducting('tr2-111, Conf) = true if path66(Conf) .
  ceq conducting('tr2-111, Conf) = false [owise] . *** relay drawn, finish locking

```

### 36.5.4 *Signal Logic*

The relay represents whether a train route which is locked is used in this logic. Take lamp 'X6gr, for example, the logic is specified as follows:

```
ceq path39(< 'tr8-111-11 : Contact | relay : 'tr8-111, upper : true > Conf) = true
if conducting('tr8-111, Conf).
eq path39(Conf) = false [owise].
ceq path43(< 'tr12-111-11 : Contact | relay : 'tr12-111, upper : true > Conf) =
true
if conducting('tr12-111, Conf).
eq path43(Conf) = false [owise].
ceq conducting('X6gr, Conf) = true
if path39(Conf)
or-else path43(Conf).
eq conducting('X6gr, Conf) = false [owise].
```

Lamp 'X6gr represents two train routes, 'tr8 and 'tr12. If one of these train route's relays is drawn, the train route is locked in other words; lamp 'X6gr should be on.

### 36.5.5 *Lamp Mutually Exclusive Logic*

If a signal has more than one lamp, all the lamps cannot be on simultaneously. They have to obey some regulations. In this chapter, it is assumed that only one lamp can be on at one moment. The lamp relay is used in this logic. Take signal 'X2 for an example. The logic is specified as following:

```
ceq path45(< 'X2gr-111-61 : Contact | relay : 'X2gr-111, upper : false > Conf)
= true
if not conducting( 'X2gr-111, Conf).
eq path45(Conf) = false [owise].
ceq path77(< 'X2brt-111-61 : Contact | relay : 'X2brt-111, upper : false > Conf)
= true
if not conducting( 'X2brt-111, Conf).
eq path77(Conf) = false [owise].
ceq conducting('X2roe, Conf) = true
if path45(Conf)
and-then path77(Conf).
eq conducting('X2roe, Conf) = false [owise].
```

```

Maude> reduce in VERIFICATION : modelCheck(initState, safetyPropSignals1(confOf(initState))) .
reduce in VERIFICATION : modelCheck(initState, safetyPropSignals1(confOf(initState))) .
rewrites: 38129 in 564ms cpu (565ms real) (67600 rewrites/second)
result Bool: true
Maude> reduce in VERIFICATION : modelCheck(initState, safetyPropSignals2(confOf(initState))) .
reduce in VERIFICATION : modelCheck(initState, safetyPropSignals2(confOf(initState))) .
rewrites: 417499276 in 5594549ms cpu (5594274ms real) (74626 rewrites/second)
result Bool: true
Maude> reduce in VERIFICATION : modelCheck(initState, safetyPropPoints(confOf(initState))) .
reduce in VERIFICATION : modelCheck(initState, safetyPropPoints(confOf(initState))) .
rewrites: 1123755 in 30181ms cpu (30186ms real) (37232 rewrites/second)
result Bool: true
Maude> reduce in VERIFICATION : modelCheck(initState, safetyPropStop(confOf(initState))) .
reduce in VERIFICATION : modelCheck(initState, safetyPropStop(confOf(initState))) .
rewrites: 1375443 in 37830ms cpu (37836ms real) (36358 rewrites/second)
result Bool: true
Maude> reduce in VERIFICATION : modelCheck(initState, safetyPropReleaseInit(confOf(initState))) .
reduce in VERIFICATION : modelCheck(initState, safetyPropReleaseInit(confOf(initState))) .
rewrites: 359702 in 7172ms cpu (7174ms real) (48895 rewrites/second)
result Bool: true
Maude> reduce in VERIFICATION : modelCheck(initState, safetyPropInitToFina(confOf(initState))) .
reduce in VERIFICATION : modelCheck(initState, safetyPropInitToFina(confOf(initState))) .
rewrites: 992772 in 26749ms cpu (26759ms real) (37113 rewrites/second)
result Bool: true
Maude> reduce in VERIFICATION : modelCheck(initState, safetyPropReleaseFina(confOf(initState))) .
reduce in VERIFICATION : modelCheck(initState, safetyPropReleaseFina(confOf(initState))) .
rewrites: 346826 in 17073ms cpu (17072ms real) (20314 rewrites/second)
result Bool: true

```

**Fig. 36.2** Partial safety properties result

Signal 'X2 has three lamps, red, green, and white. The green lamp is on when all the points in locked train route are in plus. And white lamp is on when the locked train route contains minus-position point. If neither of the green and white lamps is on, the red lamp must be on.

At this point, all the logic has been specified in Maude. And the modeling of the initial state for the interlocking system, Tongji Test Line, is finished. In the further work, some possible modification could be made to the logic to make it stricter to ensure more comprehensive safety.

Figure 36.2 shows part of the verification result checked by Maude on formal model including this initial state model.

## 36.6 Conclusion

The interlocking system is typically safety-critical system. Formal modeling and verification is therefore necessary in ensuring safety. In this chapter, the interlocking system is formally modeled in an object-oriented way by Maude, which is a specification and programming language/system based on membership equational logic and rewriting logic, equipped with model checking facilities. This chapter has presented general methods for modeling and verification of the interlocking system. And it focuses on the initial state and has proposed modeling methods for initial state in a case. The case illustrated that the object instantiated

process and control logic modeling process. The control logic can be split into several parts and modeled in the way of current modeling. There is ample evidence that Maude can be successfully applied to formally modeling and verification of the interlocking system.

## References

1. Chen, B., & Wu, F. (2002). Research on formal models of railway signal interlocking logics. *Journal of The China Railway Society*, 24(6), 50–54.
2. Ji, G. (2011). *The Formal analysis of security protocols based on Maude*. Xi'an: Xidian University.
3. Eker, S., Meseguer, J., & Sridharanarayanan, A. (2004). The Maude LTL model checker. *Electronic Notes in Theoretical Computer Science*, 71, 162–187.
4. McCombs, T. (2003). *Maude 2.0 primer*. <http://maude.cs.uiuc.edu>
5. Clavel, M., Duran, F., Eker, S., Lincoln, P., Martí-Oliet, N., Meseguer, J., et al. (2011). *Maude manual (version 2.6)*. <http://maude.cs.uiuc.edu>
6. Clavel, M., Duran, F., Eker, S., Lincoln, P., Martí-Oliet, N., Meseguer, J., et al. (2002). Maude: Specification and programming in rewriting logic. *Theoretical Computer Science*, 285(2), 187–243.
7. Zhang, M. (2007). *On the model checking of SN P systems based on rewriting logic*. Shanghai: Shanghai Jiao Tong University.
8. Denker, G. (1998). From rewrite theories to temporal logic theories. *Electronic Notes in Theoretical Computer Science*, 15, 105–126.
9. Meseguer, J. (1999). Research directions in rewriting logic. *Computational Logic. NATO ASI Series*, 165, 347–398.

# Chapter 37

## Semi-supervised Learning Using Nonnegative Matrix Factorization and Harmonic Functions

Lin Li, Zhenyu Zhao, Chenping Hou, and Yi Wu

**Abstract** In order to reduce redundant information in data classification and improve classification accuracy, a novel approach based on nonnegative matrix factorization and harmonic functions (NMF–HF) is proposed for semi-supervised learning. Firstly, we extract the feature data from the original data by nonnegative matrix factorization (NMF) and then classify the original data by harmonic functions (HF) on the basis of the feature data. Empirical results show that NMF–HF can effectively reduce the redundant information and improve the classification accuracy compared with some state-of-the-art approaches.

### 37.1 Introduction

In many traditional approaches to machine learning, a target function is estimated by labeled data, which can be thought of as an example given by “age” to “stature.” But labeled examples are always time consuming and expensive to obtain, as they require the efforts of professional people, who must be very skillful. For example, getting a labeled example for the classification of giant molecule’s shape, which is a huge challenge of biological and computational science, requires months of costly calculation and analysis by specialists. So it is meaningful to combine unlabeled data with labeled data in machine learning. In recent years, an increasing amount of interest has been attracted to the semi-supervised learning problem, and many approaches have been proposed; some methods [1–3] are generally based on an assumption that similar unlabeled data should be given the same classification. For example, the approach HF proposed by Xiaojin Zhu et al. is based on Gaussian fields and harmonic functions [1], but the original data matrix may include redundancy which could influence the classification accuracy and the effectiveness.

---

L. Li (✉) • Z. Zhao • C. Hou • Y. Wu  
Department of Mathematics and System Science, National University of Defense Technology,  
Changsha 410073, China  
e-mail: [lilin110@foxmail.com](mailto:lilin110@foxmail.com)

In this chapter, we will introduce a new approach to semi-supervised learning based on nonnegative matrix factorization and harmonic functions, which can make use of the original matrix efficiently. Unlike the work based on Gaussian fields and harmonic functions, we employed nonnegative matrix factorization to HF. In the process of our approach, we extract the feature data from the original data by NMF firstly, and then classify the original data by HF on the basis of the feature data. Encouraging experimental results show that NMF–HF can effectively reduce redundant information and improve the classification accuracy compared with some state-of-the-art approaches.

In our basic approach of NMF–HF, the solution based on nonnegative matrix factorization and harmonic functions depends on the structure of the data manifold that is derived from data features. In Sect. 37.2, we will introduce the fundamental theorem of HF and NMF briefly. Then, we will give a description of NMF–HF in detail in Sect. 37.3. Encouraging experimental results for data classification will be presented in Sect. 37.4.

## 37.2 The Principle of NMF and HF

NMF is proposed by Lee D. D. et al. [4], which has been widely used in many fields, like feature extraction [5], image processing [6], and clustering [7, 8]. NMF can be formally stated as follows [9].

Given a nonnegative matrix  $X \in \mathbb{R}^{n \times m}$  and a positive integer  $r < \min(m, n)$ , find nonnegative matrices  $B \in \mathbb{R}^{n \times r}$  and  $C \in \mathbb{R}^{r \times m}$  to minimize

$$f(B, C) = \frac{1}{2} \|X - BC\|_F^2 \quad (37.1)$$

From the above NMF problem statement, it is clear that the aim of NMF is to find the approximation of  $X$  using the product of two matrices  $B$  and  $C$ . It is also easy to deduce that the rank of both matrix  $B$  and  $C$  is at most  $r$ .

HF proposed by Xiaojin Zhu et al. is based on a Gaussian random field model defined on a weighted graph over the unlabeled and labeled data, where the weights matrix  $W$  are given in terms of a similarity between instances. In particular, the most probable configuration of the random field is unique which is characterized in terms of harmonic functions and has a closed solution that can be computed using matrix methods. It can be considered as the following minimization problem:

$$\min_F \text{tr}(F^T W F) + (F_l - Y_l)^T (F_l - Y_l) \quad (37.2)$$

where  $Y_l$  is the label matrix of train data and  $F = \begin{pmatrix} F_l \\ F_u \end{pmatrix}$  is the matrix which we want to assign labels to.

### 37.3 Basic Framework of NMF-HF

Let us assume that there are  $l$  labeled points  $(x_1, y_1), (x_2, y_2), \dots, (x_l, y_l)$  and  $u$  unlabeled points  $(x_{l+1}, y_{l+1}), (x_{l+2}, y_{l+2}), \dots, (x_{l+u}, y_{l+u})$ ,  $l \ll u$ . Let  $n = l + u$ , which is the total number of data points. At first, we suppose the labels are binary:  $y \in \{0, 1\}$ . Then, considering a connected graph  $G = (V, E)$  with nodes  $V$  corresponding to the  $n$  data points among which the nodes  $L$  correspond to the labeled points with labels  $y_1, y_2, \dots, y_l$  and the nodes  $U = \{l+1, l+2, \dots, l+u\}$  correspond to the unlabeled points, our duty is to assign labels to nodes  $U$  which is unlabeled. So we can obtain a matrix  $X_{n \times m}$  where  $x_{ij}$  is the  $j$ th element of instance  $x_i$  denoted as a vector  $x_i \in R^m$ .

To realize nonnegative factorization of matrix, we need to define a loss function to describe the degree of approximation between  $X_{n \times m}$  and  $B_{n \times r}, C_{r \times m}$  at first and then figure out the solution at the nonnegative limitation.

Regard NMF as a linear mixed model including additive noise [10]:

$$X_{n \times m} = B_{n \times r} C_{r \times m} + E_{n \times m} \quad (37.3)$$

where  $B_{n \times r}$  is the base matrix,  $C_{r \times m}$  is the coefficient matrix, and  $E_{n \times m}$  is the Gaussian noise matrix. Furthermore, we can get

$$X_{ij} = (BC)_{ij} + E_{ij} \quad (37.4)$$

To figure out  $B$  and  $C$ , we consider maximum likelihood solution following:

$$\{B, C\} = \arg \max_{B, C} p(X|B, C) = \arg \min_{B, C} [-\log p(X|B, C)] \quad (37.5)$$

Because  $E_{n \times m}$  is Gaussian noise matrix, we can get

$$p(X_{ij}|B, C) = \exp\left\{-\frac{1}{2}\left[\frac{X_{ij} - (BC)_{ij}}{\sigma_{ij}}\right]^2\right\} / \left(\sqrt{2\pi}\sigma_{ij}\right) \quad (37.6)$$

where  $\sigma_{ij}$  is the weight of each observed value. Obviously, each  $\sigma_{ij}$  is equal (each  $\sigma_{ij}$  follows i.i.d); we simply define  $\sigma_{ij} = \sigma$ . The data log likelihood can be written as:

$$\log \prod_{ij} p(X_{ij}|B, C) = -\frac{1}{2\sigma^2} \sum_{ij} [X_{ij} - (BC)_{ij}]^2 - mn \log(\sqrt{2\pi}\sigma) \quad (37.7)$$

Thus, to maximize the data log likelihood is equivalent to minimize the term  $\sum_{ij} [X_{ij} - (BC)_{ij}]^2$  in (37.7). Then, we can define the loss function as follows:



$$L_{ED}(B, C) = \sum_{ij} [X_{ij} - (BC)_{ij}]^2 \tag{37.8}$$

With the traditional gradient method, we can get:

$$\begin{aligned} \frac{\partial L_{ED}}{\partial B_{ik}} &= 2 [(XC^T)_{ik} - (BCC^T)_{ik}] \\ \frac{\partial L_{ED}}{\partial C_{kj}} &= 2 [(B^T X)_{kj} - (B^T BC)_{kj}] \end{aligned} \tag{37.9}$$

So we can obtain the following additive iterative method [10]:

$$\begin{aligned} B_{ik} &\leftarrow B_{ik} + \phi_{ik} [(XC^T)_{ik} - (BCC^T)_{ik}] \\ C_{kj} &\leftarrow C_{kj} + \varphi_{kj} [(B^T X)_{kj} - (B^T BC)_{kj}] \end{aligned} \tag{37.10}$$

Let  $\phi_{ik} = \frac{B_{ik}}{(BCC^T)_{ik}}$ ,  $\varphi_{kj} = \frac{C_{kj}}{(B^T BC)_{kj}}$ , (37.10) can be written as the following multiplicative iterative method [11]:

$$B_{ik} \leftarrow B_{ik} \frac{(XC^T)_{ik}}{(BCC^T)_{ik}}, \quad C_{kj} \leftarrow C_{kj} \frac{(B^T X)_{kj}}{(B^T BC)_{kj}} \tag{37.11}$$

As mentioned above, nonnegative matrix factorization can be summed up to a problem of constrained optimization:

$$\begin{aligned} \min f(B, C) \\ \text{s.t. } B \geq 0, C \geq 0 \end{aligned} \tag{37.12}$$

where  $f(B, C)$  is a loss function. In the process of iteration, we normalize every column vector of the base matrix, namely, that we demand  $\sum_i B_{ik} = 1$  for any  $k$  [4], and then we get the following multiplicative iterative method:

$$\begin{aligned} B_{ik} &\leftarrow B_{ik} \sum_j C_{kj} X_{ij} / (BC)_{ij} \\ B_{ik} &\leftarrow \frac{B_{ik}}{\sum_l B_{lk}} \\ C_{kj} &\leftarrow C_{kj} \sum_i B_{ik} X_{ij} / (BC)_{ij} \end{aligned} \tag{37.13}$$

We assume a  $n \times n$  symmetric matrix  $W$  on the edges of the graph  $G$  according to  $B_n \times r$ . Consequently, the weight matrix can be denoted as

$$w_{ij} = \exp\left(-\sum_{d=1}^r \frac{(x_{id} - x_{jd})^2}{\sigma_d}\right), x_i \in R^r \tag{37.14}$$

where  $x_{id}$  is the  $d$ th element of instance  $x_i$  denoted as a vector  $x_i \in R^r$  and  $\sigma_1, \sigma_2, \dots, \sigma_r$  are the length scale parameters for each dimension [1]. The matrix  $W$  completely specifies the data manifold structure for our purposes.

Then, our program is to compute a real-valued function  $f: V \rightarrow R$  on  $G$  with some nice properties, and then, we assign labels to the nodes  $L$  according to  $f$ . We impose restrictions upon  $f$  to take values  $f(i) = f_i(i) \equiv y_i$  on the labeled data  $i = 1, 2, \dots, l$ . That is to say, we want unlabeled points to have similar labels with labeled points that are near to the unlabeled ones in the graph. This idea promotes us to choose the quadratic energy function:

$$E(f) = \frac{1}{2} \sum_{i,j} w_{ij} (f(i) - f(j))^2 \tag{37.15}$$

In order to make functions  $f$  assigned with a probability distribution, we construct Gaussian field  $p_\beta(f) = \frac{e^{-\beta E(f)}}{Z_\beta}$ , where  $Z_\beta$  is the partition function denoted as  $Z_\beta = \int_{f|_{L=f_i}} \exp(-\beta E(f)) df$ , which normalizes over all functions constrained to  $f_i$  on the labeled data, and  $\beta$  is an inverse temperature parameter.

It is very simple to prove that the minimum energy function  $f = \operatorname{argmin}_{f|_{L=f_i}} E(f)$  is harmonic. That is to say, it satisfies  $\Delta f = 0$  on unlabeled data points  $U$  just like  $f_i$  on the labeled data points  $L$ . This  $\Delta$  is the combinatorial Laplacian here, denoted as  $\Delta = D - W$  in matrix form while  $D = \operatorname{diag}(d_i)$  with entries  $d_i = \sum_j w_{ij}$ , and  $W = (w_{ij})$  is the weight matrix.

The harmonic property means that the value of  $f$  which is defined at each unlabeled data point is the average of  $f$  at neighboring points:

$$f(j) = \frac{1}{d_j} \sum_{i,j} w_{ij} f(i), j = l + 1, l + 2, \dots, l + u \tag{37.16}$$

Denoted as  $f = Df$ , expressed slightly differently, where  $P = D^{-1}W$ . Because of the maximum principle of harmonic functions,  $f$  is unique and is either a constant or it satisfies  $0 < f(j) < 1$  for  $j \in U$ .

In terms of matrix operations, we split the weight matrix  $W$  (and the same to  $D, P$ ) into four blocks after the  $l$ th row and column to calculate the harmonic solution explicitly:

$$W = \begin{pmatrix} W_{ll} & W_{lu} \\ W_{ul} & W_{uu} \end{pmatrix} \quad (37.17)$$

Letting  $f = \begin{pmatrix} f_l \\ f_u \end{pmatrix}$ , where  $f_u$  assigns the values to the unlabeled data points, the harmonic solution  $\Delta f = 0$  subject to  $f_L = f_l$  is captured by:

$$f_u = (D_{uu} - W_{uu})^{-1} W_{ul} f_l = (I - P_{uu})^{-1} P_{ul} f_l \quad (37.18)$$

As outlined briefly in this section, we mainly introduce NMF to HF while the HF may be inefficient and more storage space required for classification accuracy. This new viewpoint provides a rich and complementary technique for this semi-supervised learning problem.

## 37.4 Experimental Results

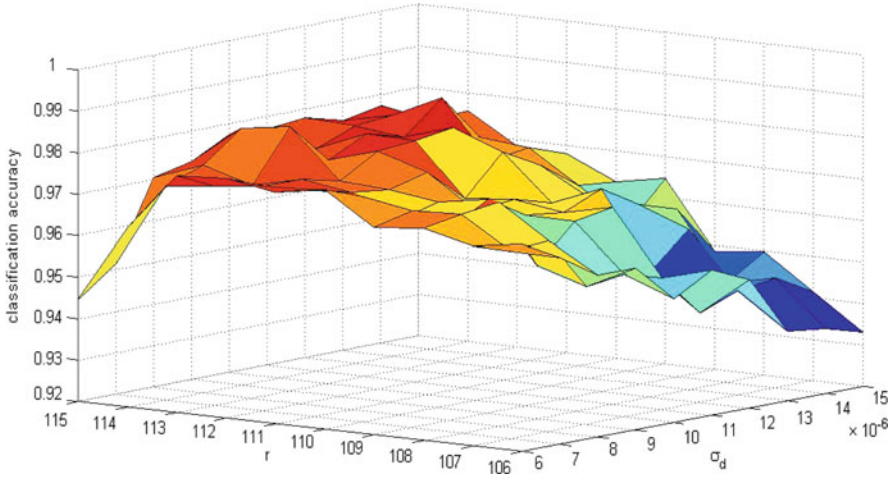
Now, we report the results of data classification experiments by using the umist dataset. Regard this dataset as a nonnegative matrix denoted as  $X_{n \times m}$ , where  $x_{ij}$  is the  $j$ th component of instance  $x_i$  denoted as a vector  $x_i \in R^m$ . In this case,  $n = 380$ ,  $m = 644$ , it means that there are 380 data points while every data point has 644 dimensions.

In our experiments, we perform several groups of trials. In each trial, let every  $\sigma_d = 1.2 \times 10^{-5}$  and  $r = 115$ . Firstly, sample some labeled data (40, 60, 80, ..., 200) randomly as train data from the entire dataset; secondly, classify the rest data (denote as test data) into 20 classes based on the approach (NMF–HF); then, repeat each trial 50 times; and finally, calculate the average classification accuracy for each trial. Compared with KNN and HF, we get the following table (Table 37.1):

According to the above table, we draw a conclusion that NMF–HF is more efficient than KNN and HF in terms of classification accuracy.

**Table 37.1** Classification accuracy

Train data	KNN	HF	NMF–HF
40	0.7198	0.8964	0.9055
60	0.8048	0.9233	0.9303
80	0.8671	0.9497	0.9546
100	0.8998	0.9561	0.9603
120	0.9295	0.9660	0.9685
140	0.9460	0.9732	0.9742
160	0.9570	0.9775	0.9786
180	0.9710	0.9851	0.9846
200	0.9747	0.9839	0.9862



**Fig. 37.1** The sensitivity analysis about parameters

We also study the stability about classification accuracy of NMF–HF while  $r$ ,  $\sigma_d$  vary in a certain range. The following figure is given (Fig. 37.1):

From the figure, we conclude that the classification accuracy of NMF–HF is stable when  $r$ ,  $\sigma_d$  vary in a reasonable range.

## 37.5 Conclusion

A new approach to semi-supervised learning is proposed that is based on nonnegative matrix factorization and harmonic functions. Promising experimental results have been presented for nonnegative data classification, demonstrating that the framework of NMF–HF has the potential to effectively exploit the structure of unlabeled data to improve efficiency and classification accuracy. Our work in this direction will be reported in a future publication.

**Acknowledgments** We gratefully acknowledge the supports from National Natural Science Foundation of China, under Grant No 61005003.

## References

1. Zhu, X., Ghahramani, Z., & Lafferty, J. (2003). Semi-supervised learning using gaussian fields and harmonic functions. In *Proceedings of the 20th International Conference on Machine Learning* (Vol. 20, Issue 2, p. 912). Washington, DC: AAAI Press.
2. Belkin, M., & Niyogi, P. (2002). Using manifold structure for partially labelled classification. *Advances in Neural Information Processing Systems*, 15, 929–936.

3. Blum, A., & Chawla, S. (2001). Learning from labeled and unlabeled data using graph mincuts. In *Proceedings of the 18th International Conference on Machine Learning* (pp. 19–26). Williams College: Morgan Kaufmann.
4. Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects with non-negative matrix factorization. *Nature*, *401*(6755), 788–791.
5. Student, M., & Eswar, K. (2012). Graph regularized non-negative matrix factorization for data representation. *International Journal of Computer Application*, *3*(2), 171–191.
6. Cichocki, A., & Amari, S. (2002). Adaptive blind signal and image processing. In *Learning algorithms and applications* (pp. 88–93). New York, NY: John Wiley Press.
7. Ding, C., Li, T., & Jordan, M. (2008). Nonnegative matrix factorization for combinatorial optimization: Spectral clustering, graph matching, and clique finding. In *ICDM* (pp. 183–192). Pisa, Italy: IEEE Computer Society.
8. He, Z., Xie, S., Zdunek, R., Zhou, G., & Cichocki, A. (2011). Symmetric nonnegative matrix factorization: Algorithms and applications to probabilistic clustering. *IEEE Transactions on Neural Networks*, *22*(12), 2117–2131.
9. Berry, M. W., Browne, M., Langville, A. N., Pauca, V. P., & Plemmons, R. J. (2007). Algorithms and applications for approximate nonnegative matrix factorization. *Computational Statistics and Data Analysis*, *52*(1), 155–173.
10. Sajda, P., Du, S., & Parra, L. C. (2003). Recovery of constituent spectra using nonnegative matrix factorization. Optical Science and Technology, SPIE's 48th Annual Meeting. Proceedings of SPIE, San Diego. pp. 321–331.
11. Lee, D. D., & Seung, H. S. (2001). Algorithms for nonnegative matrix factorization. *Advances in Neural Information Processing Systems*, *13*, 556–562.

# Chapter 38

## Exploring Data Communication at System Level Through Reverse Engineering: A Case Study on USB Device Driver

Leela Sedaghat, Brad Duerling, Xiaoxi Huang, and Ziyang Tang

**Abstract** Interactions among operating system, drivers, and peripheral devices are important for users to understand data communication at low system level, system architecture, and hardware programming. In this chapter, we study low-level data communication and resource management by conducting the development of a USB device driver. A reverse engineering approach has been adopted in this study, and we focus on exploring the USB protocol and developing a device driver for the Linux operating system. We have performed various experiments to evaluate the device driver from different aspects, and all testing results are remarkably good. We believe this work can provide users a clear practical understanding of data communication from the hardware level to user space applications as well as theoretical foundations to reproduce any unsupported peripheral hardware devices.

### 38.1 Introduction

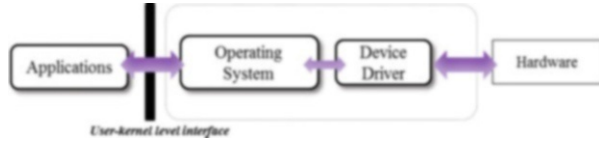
The Linux kernel is initially designed to provide a freely available, openly maintained, and modified operating system. Since its first release in 1991, the Linux kernel has been ported to countless different systems and processor architectures. It has successfully transitioned from primarily academic patronage to users in the mainstream consumer market. Today, Linux is a popular alternative to commercial operating systems, such as Windows and Mac OS, and the demand for supported applications and hardware drivers is growing at an extraordinary

---

L. Sedaghat • B. Duerling • Z. Tang (✉)  
Department of Computer and Information Sciences, Towson University,  
7800 York Road, Towson, MD 21204, USA  
e-mail: [lsedagl@students.towson.edu](mailto:lsedagl@students.towson.edu); [miamisbiggest@gmail.com](mailto:miamisbiggest@gmail.com); [ztang@towson.edu](mailto:ztang@towson.edu)

X. Huang  
Institute of Cognitive and Intelligent Computing, Hangzhou Dianzi University,  
Hangzhou 310018, China  
e-mail: [huangxix@hdu.edu.cn](mailto:huangxix@hdu.edu.cn)

**Fig. 38.1** Data communication from hardware to applications



**Fig. 38.2** Brookstone USB missile launcher and its movements



pace. From 2007, the Linux Driver Project [1] has been launched to meet this need by enlisting a group of developers to create, submit, and maintain Open Source Linux kernel drivers for different types of devices. Meanwhile, we believe that research analysis on data communication between devices and operating systems is very important to provide theoretical foundations for supporting any unsupported peripheral hardware devices. In order to successfully develop a fully functioning Linux device driver, it is essential not only to study and understand the interactions between the Linux operating system and device drivers in general but also to study and understand the specific technical details underlying the interactions between the operating system, driver, and selected target peripheral device, which also requires a thorough understanding of the data communication protocol such as USB.

Figure 38.1 shows the data communication among user application, operating system, and devices with the help of device driver and controller. Developing a Linux device driver presents a challenging and rewarding experience to understand this communication and explore the relationship between operating systems, drivers, and peripheral devices. As such, we present herein a case study to develop a Linux driver for a USB 2.0 device. We selected a peripheral device that is currently unsupported by the Linux operating system: the USB Desktop Missile Launcher by Brookstone, as shown in Fig. 38.2 [2]. The device is currently supported only by Windows and Mac OS X. Consequently, any Linux user who purchases the device would be unable to use it. We believe this study can provide general users a practical understanding of how data communication happens from low-level hardware to high-level user space applications, thereby allowing users to reproduce any device on their own.

## 38.2 USB and Device Driver

### 38.2.1 USB Data Communication Protocol

As an I/O bus, USB transfers data between a host computer and a peripheral device. The host manages the bus, and the device can only use the bus when responding to a request from the host (except in the case of remote wake-up). As such, USB forms a master/slave configuration [3, 4]. In this section, we discuss the USB protocol in detail and illustrate the key ideas through Fig. 38.3.

The host contains USB host-controller hardware, a root hub with USB ports as well as software to manage bus communication. The USB host-controller hardware is the interface between the operating system, drivers, and all of the attached peripheral devices. The USB host controller identifies each attached peripheral device with a unique address or identification number. This, in turn, permits many devices to communicate with the host over the same USB connection. The USB 2.0 driver belongs to the host-controller hardware and is needed to interpret the USB protocol and to identify any and all attached devices. It is then the responsibility of system software to find the appropriate device driver and to load it into the kernel.

The responsibilities of the host include detecting devices when they are plugged into the USB port, providing power to the devices, managing traffic on the bus, handling error checking, and exchanging data (using a message-passing system) with the peripheral devices. The device contains USB device-controller hardware, a microcontroller and usually firmware to manage bus communications (though it may contain software instead of firmware). The responsibilities of the device include detecting bus voltage, managing power, responding to requests from the host, handling error checking, and implementing the device’s functions.

Data flow on the bus is bidirectional. Transfers on the bus can be classified into four types: control, bulk, interrupt, or isochronous transfers. Furthermore, unlike

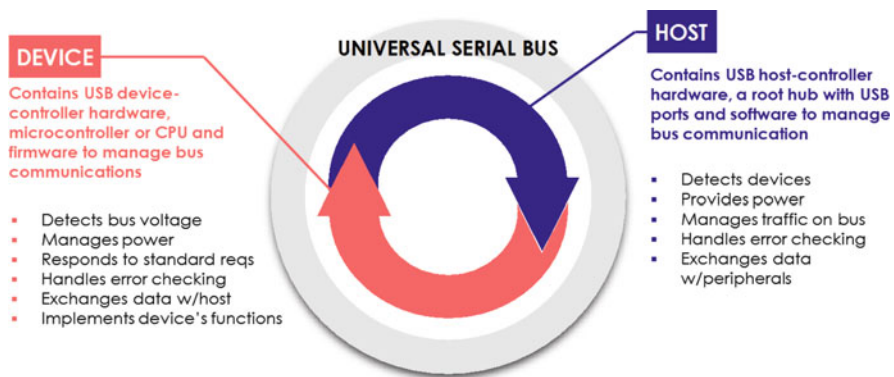


Fig. 38.3 Universal serial bus diagram



parallel and serial ports that transfer data in bits, USB encapsulates and sends data in packets called USB request blocks (URBs), which are the basic transaction unit of the message-passing system used for communication over the bus. Each URB contains all of the necessary information (e.g., data and addressing information) to perform a USB transaction. Generally, in order to initialize a URB within a device driver, the following information is required: a device pointer, the pipe (or endpoint information), the request type and value, the transfer buffer, and the desired transfer length. For example, a URB for a control-type transfer can be built and sent to a specific endpoint by invoking the following method:

```
int usb_control_msg (struct usb_device * dev,
                    unsigned int pipe,
                    __u8 request,
                    __u8 requesttype,
                    __u16 value,
                    __u16 index,
                    void * data,
                    __u16 size,
                    int timeout)
```

When a system is booted, the USB hub lets the USB host controller know of any attached devices. If a system has already been booted up, the USB host controller periodically polls the hub to learn of any newly attached or removed devices. When an attached device has been detected, whether at boot-up or afterwards, the host controller assigns the device a unique address and sends requests for its descriptors through the enumeration process. The USB descriptors requested by the host provide information about the device, allowing the host to load the appropriate device driver. Descriptors are data structures that store information about a device's capabilities and requirements. The host uses these descriptors to load the appropriate driver for the device. After the driver has been loaded, the host and device can communicate on the USB using URBs, as described above.

### ***38.2.2 Device Communication and Device File***

A device driver, which is loaded to the kernel to provide an interface for user applications, is in charge of interactions between an operating system and a device. As an interface between the hardware and the software, it provides the mechanism and not the policy in order to fulfill different requests of the OS. Hence, separation of the mechanism and the policy should be taken into consideration [5].

As pointed out by Kadav and Swift [6], there are three interfaces in the system which may be used to implement a device driver: (1) the interface between the driver and the kernel, for communicating requests and accessing OS services; (2) the interface between the driver and the device, for executing operations; and

(3) the interface between the driver and the bus, for managing communication with the device. It is clear that these interfaces might be different regarding how they are implemented in various platforms.

Generally, there are three classes of device drivers in Linux, namely, character drivers, block drivers, and network drivers. Besides this classification, there are some devices that cannot be classified into one of these classes. For example, a USB device can have a character driver (like USB serial port) or block driver (USB memory card reader) or even network driver (USB Ethernet interface) [5].

When the device is plugged in, the kernel recognizes the device and sets up all the configurations and interfaces. The kernel then checks if it can handle the USB device by searching all the drivers that it has registered and calling their probe functions. This function considers the vendorID and productID of the device against its device table. If the driver can handle the device, it pairs itself with the device and informs the kernel. Moreover, a device file is also created at this time to allow a user space program to communicate with the device through the driver. User space knows the device file by its name, while the driver knows the device file by a number. Note that the device file is not a file in the traditional sense. Instead, it is a mapping for user space functions. Each function in the user space is mapped by the device file to a function in the driver code. For example, below is the part of the driver code that tells the device file which function should be mapped to for each call.

```
static const struct file_operations skel_fops = {
    .owner = THIS_MODULE,
    .write = skel_write,
    .open = skel_open,
    .release = skel_release,
};
```

### 38.3 Reverse Engineering

In order to develop the driver for any unsupported peripheral device, we need to reverse engineer the USB device, which is the focus of our work. By reverse engineering the device, we are able to obtain the necessary information the driver needs to communicate with the device. The primary and essential information that is needed for proper driver implementation is contained in the USB descriptors for the device as well as in the payload of the URBs. Since the manufacturer provides a Windows device driver, a PC running Windows can be used as the host computer to reverse engineer the device. As such, we first install the device on a Windows system and, subsequently, we install software that would allow us to monitor the USB traffic sent and received by the host. In order to do so, USBlyzer [7], a USB traffic sniffer software, is adopted in our system. The device is plugged into an

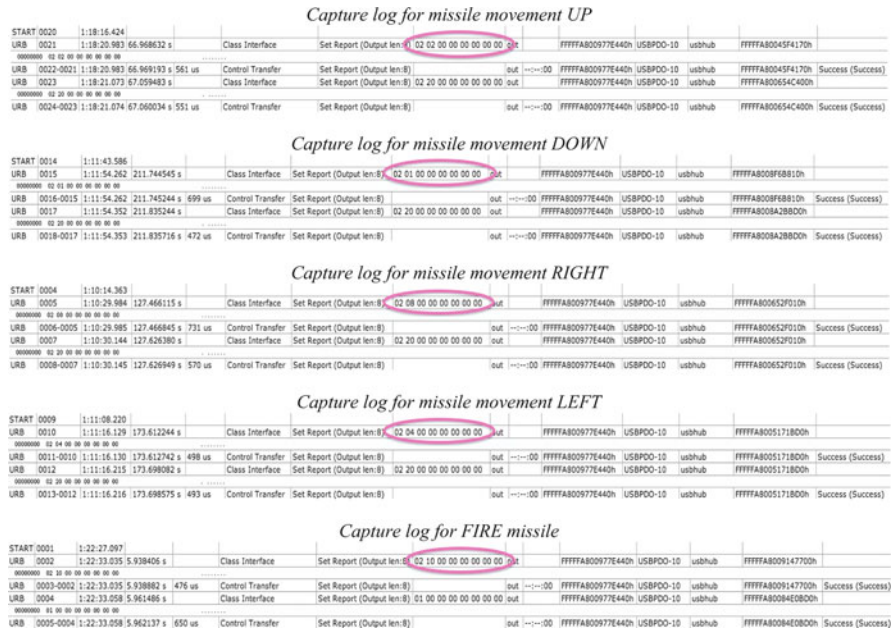


Fig. 38.4 Capture logs for different missile launcher functions

available USB port and USBlyzer is launched. USBlyzer displays various USB descriptors that are obtained by the host computer during the enumeration process in one of the user interface panels. For example, the productID and vendorID needed to implement the device driver are among the descriptors obtained from the device.

After successfully obtaining USB descriptors, we can start capturing the actual raw data that is passed in the URBs between the host and the device. In order to do so, it is necessary to have the device perform all of its possible functions (e.g., move right) and observe the values in the payload of the captured packets. Since the missile launcher device is able to move left, right, up, and down to the desired angle, we need to trigger all of these motions in individual captures so that there would be no confusion as to which packets corresponded to which functions. The results of each capture are shown in Fig. 38.4. A closer look reveals that each of the functions is comprised of a series of requests, with the majority of payload values being the same across all functions. The data of only one URB is different across all of the functions. This unique packet is circled in Fig. 38.4 for each capture. We hypothesize that it is this unique payload that drives each different function. As such, we use only these data values when implementing the driver.

## 38.4 Experimental Results

In order to provide an interaction with the USB missile launcher, we have implemented both the device driver and a user space application based on our previous studies. The device driver has to be compiled and loaded into the Linux kernel, while the user application is a typical C code which must be compiled and run in the user space and interacts with the driver.

Our device driver is implemented through USB-Skeleton where a number of *structs* are provided for the customized device driver. One of the most important ones is presented below:

```
static struct usb_driver skel_driver = {
    .name = "missile",
    .probe = skel_probe,
    .disconnect = skel_disconnect,
    .id_table = skel_table,
};
```

When we register our driver to the kernel, it will be registered by the value that we assigned to name field. As mentioned before, when a USB device is plugged in, the kernel tries to find the appropriate driver for the device based on productID and vendorID. We define these two values in *skel\_table*, and as you can see, *id\_table* is pointing to this table. So, *id\_table* field actually helps the kernel to find a driver for the plugged-in device. After finding the driver, kernel runs the Probe function of that driver which is responsible for extracting and initializing needed endpoints from the device descriptor. In the represented *struct*, Probe contains the pointer to the probe function which we have defined in our driver code.

The user application is intended to allow users to test and demonstrate the functionality of the device in the user space. To conduct testing, we have implemented both console-driven program and shell-based user-friendly dialog boxes. In both applications, we can successfully control all the missile movements. We have also experimented from system resource management aspects. In Fig. 38.5, we show process monitoring and system resource monitoring for our missile launcher process. Specifically, we trace system calls and signals and monitor CPU, memory, and I/O resources.

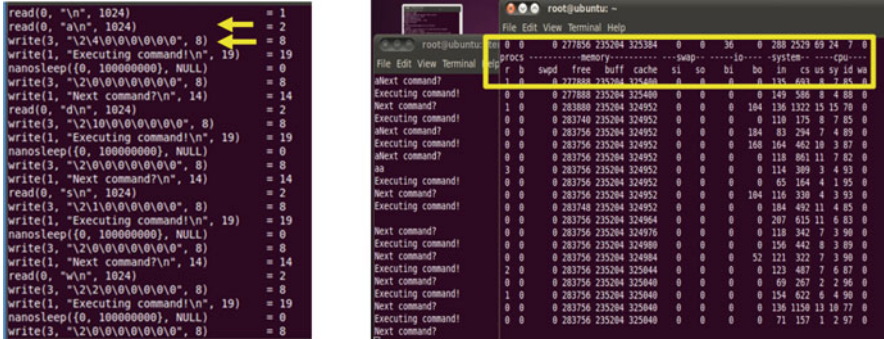


Fig. 38.5 System call trace for the missile launcher process (left). System resource monitoring (right)

### 38.5 Conclusion

In this project, we have studied the interactions between an operating system and device drivers at the system level. It provides general users a practical understanding of how data communication happens from the low-level hardware to high-level user space applications. A reverse engineering technique has been adopted in this study, and we focus on exploring the USB protocol and developing a device driver for the Linux operating system. We have conducted a detailed case study for a particular device called Desktop Missile Launcher, which is only supported by Windows and Mac OS before our work. In addition, we have performed various experiments to evaluate the device driver from different aspects, and all testing results are remarkably good.

### References

1. Linux Driver Project. <http://www.linuxdriverproject.org>
2. USB desktop Missile Launcher. <http://www.brookstone.com/usb-desktop-missile-launcher>
3. Anderson, D., & Dzatko, D. (2001). *Universal serial bus system architecture* (2nd ed.). Addison-Wesley Professional.
4. Axelson, J. (2009). *USB complete: The developer's guide* (4th ed.). Lakeview Research.
5. Corbet, J., Rubini, A., & Kroah-Hartman, G. (2005). *Linux device drivers* (3rd ed.). Sebastopol, CA: O'Reilly.
6. Kadav, A., & Swift, M. (2012). Understanding modern device drivers. *ACM SIGARCH Computer Architecture News*, 40(1), 87–98.
7. USBlyzer. <http://www.usblyzer.com>

# Chapter 39

## Using Spatial Analysis to Identify Tuberculosis Transmission and Surveillance

Jinrong Bai, Guozhong Zou, Shiguang Mu, and Yu Ma

**Abstract** Tuberculosis is a chronic infectious disease which can make serious hazard to human health and cause large social and economic burden on a country. So for experts and researchers, tuberculosis is one of the biggest public health challenges. The cause of this disease can be effectively studied by precise analysis of the spatial distribution of the disease. This chapter demonstrates that using existing health data, spatial analysis and GIS in conjunction with epidemiological analysis can identify tuberculosis transmission. This chapter also demonstrates some of the valuable results of GIS in disease surveillance and mapping. The decision-makers could master the epidemic of tuberculosis dynamically and then take better measures to control tuberculosis. Moreover, this study may add some value to traditional and molecular epidemiology and provides an alternative method that may give insight into the transmission of tuberculosis.

### 39.1 Introduction

Tuberculosis (TB), known as the “white plague,” is a chronic infectious disease which can make serious hazard to human health. The 1990 World Health Organization (WHO) reports that tuberculosis is ranked as the seventh most morbidity-causing disease in the world and predicts it to continue in the same place up to 2020. An estimated two billion people are infected with tuberculosis [1] and approximately two million people die from this disease annually [2]. China is one of the 22 countries in the world highly burdened by tuberculosis, ranking second with respect to the number of tuberculosis patients. Tuberculosis not only is an important public health problem but also is a complex socioeconomic problem. The prevalence of tuberculosis hinders social and economic development. Unless tuberculosis is properly treated, an average

---

J. Bai (✉) • G. Zou • S. Mu • Y. Ma  
School of Information Technology and Engineering, Yuxi Normal University,  
Yuxi 653100, China  
e-mail: [baijr223@163.com](mailto:baijr223@163.com)

of 10–15 people are infected by an infectious tuberculosis patient in a year. More adults are killed by tuberculosis than any other infectious disease worldwide, accounting for about 400,000 deaths annually. It mainly affects those in the economic production years of their life (equal to 15–54 years). An adult tuberculosis patient loses on average 3–4 months of working time annually. This shows that large social and economic burden is caused by tuberculosis on a country. So for experts and researchers, tuberculosis is one of the biggest public health challenges.

Previous studies on tuberculosis and other infectious diseases ignored the geographic correlation, did not perform a study on the quantitative level spatial distribution of diseases, and were limited to simple analysis of incidence of the disease. We all know that medicine is a system of micro and macro comprehensive disciplines and large amounts of data have the spatial distribution characteristics.

Epidemiological topic is the distribution of time, space, and population about diseases and about 80 % of the epidemiological data have spatial attributes. Human or animals are always in the prevalence of certain spatial location, but geographical or social factors in certain space may also affect the incidence of the diseases. The cause of the disease can be effectively studied by accurate analysis of the spatial distribution of the disease, so we can identify high-risk regions and populations and develop preventive measures.

Moreover, as an infectious disease, tuberculosis is related to population, climate, and the local environment. Because of its infectivity and universality, the occurrence, development, and prevalence of tuberculosis is a spatial phenomenon with interaction and diffusing phenomenon. Therefore, the relevant data should be based on spatial property, considering the location information and non-location information. This reflects dynamic distribution characteristics on space and its influential factors in order to meet the needs of tuberculosis control and prevention work. New methods, such as spatial analysis and mapping, may be of valuable contribution to basic elements of tuberculosis control. Such tools are used by public health professionals to visualize and explore disease patterns for guiding disease control strategies. On the basis of spatial analysis, a new way is supported by the emerging spatial statistics for the spatial autocorrelation. Its core is the understanding of space-dependent data between location-related areas.

In this study, we perform spatial statistical analysis in quantitative level to explore one of the hot spots of tuberculosis incidence; also, time, space, and spatial-temporal cluster is very necessary. The main purpose of this research was to examine the spatial distribution of tuberculosis cases in the Anxian County, Sichuan Province, China, over a 3-year period, from 2007 to 2009, using geographic information system (GIS) software and spatial analysis technology. In doing this, it was expected that some of the valuable assets of GIS in disease mapping and surveillance are demonstrated in this study. It is anticipated that the collected information by this research will help public health workers to identify and provide effective examples of using epidemiological data, public statistical software, and GIS to formulate the research problem, generate and test hypotheses, and critically evaluate mapping that is prepared using spatial statistical approaches

and GIS software. It is hoped that public health officials and workers in the future will see the added value which is brought by GIS to an already well-established disease surveillance group. Furthermore, GIS can offer data that is in a form that can be more easily communicated to the general public and community groups if necessary. An effective method for tuberculosis control programs is provided by geographically based screening and treatment to identify high-risk populations. In this research, we also sought to determine whether or not we can identify the geographical regions with ongoing tuberculosis transmission by using GIS technology.

## 39.2 Related Works

In our opinion, using GIS and spatial-temporal statistical analysis to identify geographical areas with ongoing disease transmission has become indispensable. Space-time clustering method involves the identification of greater density of a phenomenon that appears in certain places at certain times. The researchers have intensively applied these techniques in several areas such as toxicology, criminology, demography, and others.

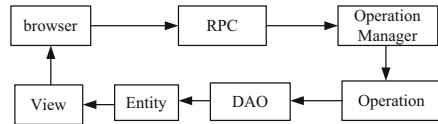
Disease mapping has a long history and it is no surprise that this descriptive analysis method is used to describe rates of spread and to identify sources of infections. It is not a new phenomenon that the geography is used in epidemiological studies. John Snow's analysis of local water pumps and their relationship with the transmission of cholera in London during the 1850s is probably the most widely cited study for first incorporating geographical analysis and the combination of field epidemiology. Epidemiology stresses the understanding of three important components of disease distribution: the time of disease outbreak or transmission, the people involved, and the location of spread. In the past decade, researchers have published several studies on geographical epidemiology all over the world. A very interesting study about the tuberculosis DOTS strategy and GIS is published by Porter [3]. GIS technology is used to identify incidence and areas of tuberculosis transmission in the USA from 1993 to 2000 [4]. Spatial distribution of *M. tuberculosis*/HIV coinfection is studied by Rodrigues in São Paulo State, Brazil, from 1991 to 2001 [5]. GIS and spatial scan statistics are used to investigate geo-spatial hot spots for the occurrence of tuberculosis in Almora district [6].

## 39.3 System Architecture

In order to allow for mapping of individual tuberculosis cases and spatial analysis, the address of each case in the study was geocoded. It needs to utilize specialized GIS software to assign a latitude and longitude for the address of the case. In this work, disease mapping was performed utilizing Google Maps. Google Maps is a web mapping service provided by Google Company, and it contains landmarks,



**Fig. 39.1** System architecture



lines, shapes, and other information, with three kinds of view: the vector maps, satellite images, and topographic maps. High-resolution satellite or aerial images for most urban areas are provided by Google Maps, which powers many map-based services. Maps embed on third-party websites via the Google Maps API. The Google Maps API is publicly accessible, does not charge for access, and is free for commercial use provided that the site is not generating more than 25,000 map accesses a day.

The system is developed by Java language and MVC framework. The framework employs DWR (Direct Web Remoting) with Ajax technology to achieve the web program interaction between foreground and background. DWR is a Java open source library and includes Ajax technology. DWR helps developers write websites, and Java functions running on a web server are called by JavaScript code in a web browser as if those functions were within the browser.

Specifically, the framework encapsulates the common functions and employs Ajax technology to access the server-side object through the operation callback interface. Therefore, the front-end operation of the system mainly uses JavaScript, whereas the back end of the system uses lightweight MVC framework. The business logic is divided into three layers (Operation, Entity, DAO). System framework is showed in Fig. 39.1.

## 39.4 Methods

Face-to-face patient survey, institution-based survey, key informant interview, and literature reviews were the main data sources. Relevant databases were set up after the collection of data from surveys on tuberculosis in the Anxian County. Spatial analysis was undertaken after the databases were linked to the software Google Maps. Correlation analysis was performed to understand the relationship among the prevalence rate of active pulmonary tuberculosis, the mortality rate of pulmonary tuberculosis, and socioeconomic factors.

### 39.4.1 Study Area and Population

This significant research was conducted in the Anxian County in Sichuan province, China. This county, which covers a total surface area of 1,404 km<sup>2</sup>, has a population of 500,000. This study was in the southwest most densely populated area of China. This county has 23 government tuberculosis diagnostic centers. All forms of

tuberculosis patients who usually reside in the Anxian County were eligible to join the study. This study obtained informed consent from all research cases.

### **39.4.2 Data Collection**

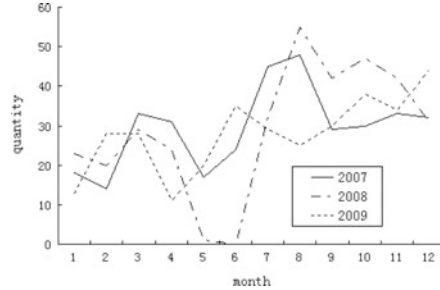
All consenting tuberculosis cases were collected between January 1, 2007, and December 31, 2009. This study included a tuberculosis case if the tuberculosis control program had undergone a full course of antituberculosis treatment for the patient. At recruitment, a structured questionnaire was answered by all study participants. This study obtained demographic data (sex, age, place of residence, occupation, and ethnicity), clinical information (type of tuberculosis, date of registration, and date of diagnosis), and past history of tuberculosis. This study also collected the residential addresses of all study cases and the geographical location of diagnostic and treatment facilities in the study area. A team of field workers monitored regularly all study participants and categorized their treatment result based on standard WHO-recommended definitions [7] transferred out, failed treatment, died, completed treatment, cured, and defaulted. This study investigated the residential status of each patient with tuberculosis and categorized the residential status of each patient as permanent or temporary residents. A permanent tuberculosis patient who is usually impossible to move out on completion of treatment was defined as a case normally resident there.

Residential address and zip code were geocoded using Google Maps at the time of diagnosis of tuberculosis. After the automatic and interactive geocoding, the longitude and latitude of most tuberculosis cases were correctly matched on the map.

### **39.4.3 Statistical Analysis**

During the 3 years of the study, 1,035 patients were diagnosed with tuberculosis and started treatment in the clinics involved in the study. The overall incidence rate of tuberculosis in the study area was 69/100,000 for the year of the study, and the male to female ratio was 2.2. There was a positive correlation between the average case load and crowding. A significant positive correlation was also found between unemployment and the average tuberculosis case load, with codependence between unemployment and crowding. The statistics of tuberculosis cases between 2007 and 2009 is showed in Fig. 39.2. As can be seen in Fig. 39.2, there was a correlation between season and the average case load and the tuberculosis cases in summer dominated other seasons.

**Fig. 39.2** The statistics of tuberculosis cases from 2007 to 2009



### 39.4.4 Results

GIS techniques were used to perform geographical analysis. All cases and diagnostic and treatment facilities in the study area were mapped using Google Maps. While the different sizes of the background population have not been taken into account in map, it appears that cases tended to concentrate in the densely populated place. The numbers of cases were then aggregated by zip code. This study aggregates data on tuberculosis distribution in the study area to the level of these settlements as with the census data for spatial analysis.

The analysis results show that global spatial autocorrelation indications are 0.2936, 0.3958, and 0.2271 in these 3 years. Global spatial autocorrelation indications are all above 0 in these 3 years. The Moran's I of 3 years all have significance. It shows that the registration rates of tuberculosis are positively correlated.

Through Spatial Lag Model and Spatial Error Model, we find that the number of tuberculosis suspects consulting and the per capita net income of rural population have influence on the registration rate of tuberculosis. The registration rate is higher in the town with more tuberculosis suspects consulting. The more per capita net income of rural population, the lower the registration rate is. Moreover, the density of population has influence on the registration rate of tuberculosis. The analytical results indicated that migrant population was responsible for the transmission of tuberculosis in the Anxian County and resulted in the huge number of drug-resistant tuberculosis cases. Based on the huge field epidemiology data, migrant population, drug-resistant tuberculosis, and TB/HIV coinfection are the three main challenges for tuberculosis control in this study.

## 39.5 Discussion

While GIS and spatial analysis have been used to study various infectious diseases by several studies, only a few have focused on tuberculosis. The spread of tuberculosis is influenced by many factors and they are often interrelated, such as unemployment, crowding, and other complex socioeconomic factors. The most elusive aspect is to determine the locations and time at which the tuberculosis

took place. Migrants are associated with the epidemic of tuberculosis in the Anxian County. The possibility of significant geographic variation about tuberculosis in the study area has been highlighted in this study. This would become the background to a series of researches whether environmental differences, organism, and particular host may be explanatory. In this respect, it is important to identify areas of those with significantly low rates of disease as well as significantly high rates. This information will assist to guide the provision and optimization of tuberculosis control strategies in China.

Disease mapping in time or space may play a good role in guiding public health policy. Exploring where and when the new cases happen, that is to say, exploring the spatial-temporal hot spot to know the spatial-temporal cluster of tuberculosis, may help the tuberculosis program in disease control activities. The systematic use of spatial-temporal analysis for the surveillance of tuberculosis incidence in the Anxian County aims to supply with the theory evidence for the policy of tuberculosis prevention and control and provide the reference for the similar study. Spatial epidemiology adds to the value of traditional and molecular epidemiology and suggests an alternative method that may provide insight into the transmission of tuberculosis.

While the usefulness of GIS and spatial analysis has been demonstrated in our study, it has some limitations. Firstly, the analysis of time and space has some defects; this study is mainly based on spatial analysis, so it cannot have accurate judgment on high-risk trend over time. Secondly, this study did not consider religious, social and economic, environmental, and lifestyle factors of monitored cases. Thirdly, the data sample size in this study is small, and the credibility and stability of the results need further validation. Therefore, future research should undertake broader study area and obtain more tuberculosis case data.

## 39.6 Conclusion

This chapter mainly discussed how to use GIS techniques on tuberculosis management and also discussed the methods and the steps of building the information system of tuberculosis management. The system is able to manage the spatial and attribute data at the same time and query information using SQL; zoom in, zoom out, and pan the layers; and display raw data, do some analysis, as well as manage report, and so on. Yet, some functions of the system are under exploitation or consideration.

GIS is able to work together with the spatial data and the attribute data of tuberculosis, track its spread, and display the outcomes of complicated raw data, as well as their analysis, by visual map. In this way the decision-makers could master the epidemic of tuberculosis dynamically and then take better measures to control tuberculosis. Using molecular epidemiological surveillance combined with GIS analysis may be an effective method for identifying tuberculosis transmission and surveillance. Targeted screening and control efforts can be strengthened

by using these methods, with the goal of incidence reduction and ultimately interruption of disease transmission.

This study demonstrates that previously undetected tuberculosis transmission can be identified if using GIS, spatial analysis, and existing health data in conjunction with epidemiological analysis. These results were used to design new targeted screening efforts. Studies of these efforts have utility in reducing tuberculosis transmission. This study may provide researchers with a meaningful picture of the disease patterns from epidemiological studies. This work will support to establish effective tuberculosis control measures and provide new thought and method to other infectious diseases.

## References

1. CDC. (2006). *Morbidity and mortality weekly report*. Retrieved April 20, 2013, from <http://www.cdc.gov/mmwr/pdf/wk/mm5511.pdf>
2. Maher, D., & Raviglione, M. (2005). Global epidemiology of tuberculosis. *Clinics in Chest Medicine*, 26(2), 631–633.
3. Porter, J. D. (1999). Editorial: Geographical information systems (GIS) and the tuberculosis DOTS strategy. *Tropical Medicine and International Health*, 4(10), 631–633.
4. Moonan, P. K., Bayona, M., Quitugua, T. N., et al. (2004). Using GIS technology to identify areas of tuberculosis transmission and incidence. *International Journal of Health Geographics*, 3(1), 23–33.
5. Rodrigues, A. L., Jr. (2006). Distribuição espacial da co-infecção M. tuberculosis/HIV no Estado de São Paulo, 1991. *Revista de Saude Publica*, 40(2), 265–270.
6. Tiwari, N., Adhikari, C. M. S., Tewari, A., et al. (2006). Investigation of geo-spatial hotspots for the occurrence of tuberculosis in Almora district, India, using GIS and spatial scan statistic. *International Journal of Health Geographics*, 5(1), 33–44.
7. World Health Organization (WHO) (2003). Global Tuberculosis Program. Treatment of tuberculosis: guidelines for national programs [WHO/CDS/TB/ 2003.313]. 3rd ed. Geneva: WHO.

# Chapter 40

## Construction Method of Exception Control Flow Graph for Business Process Execution Language Process

Caoqing Jiang, Shi Ying, Shanming Hu, and Hua Guan

**Abstract** Traditional control flow graph of exception handling lacks an explicit description of exception handling and propagation and cannot be used to well analyze the exception situations and exception handling error. To solve these problems, this chapter presents a construction method of exception control flow graph (ECFG) for BPEL process. This method uses a label that is marked exception and is of power for collection computing to describe exception information of BPEL process in building the ECFG. Moreover, the experiment shows that the ECFG generated can clearly express exception information and propagation process in BPEL process.

### 40.1 Introduction

Business process execution language (BPEL) is different from other languages such as Java at exception handling mechanism [1] which makes exception control flow graph (ECFG) in BPEL different from the construction method of traditional ECFG, and thus the study of construction method of ECFG for BPEL is a necessity [2].

There exist distinctness in the ingredients and mechanisms of exception handling of programming language; thus, researchers often propose different construction method of ECFG toward different languages [3], such as Java and C++

---

C. Jiang (✉)

The State Key Lab of Software Engineering, Wuhan University, Wuhan 430072, China

Guangxi University of Financial and Economics, Nanning 530003, China

e-mail: [jcqing@163.com](mailto:jcqing@163.com)

S. Ying • H. Guan

The State Key Lab of Software Engineering, Wuhan University, Wuhan 430072, China

S. Hu

Guangxi University of Financial and Economics, Nanning 530003, China

language, and it has become the focus of research. However, these traditional ECFG do not usually explicitly describe the exception handling and does not reflect the exception propagation and thus cannot effectively analyze exception situations and exception handling error. To solve this problem, based on a Signed-TypeSet domain compact notation, Prabhu et al. adopted inter-procedure data stream algorithm to obtain the C++ inter-procedural ECFG (IECFG) [4]. The study identifies the exception type, so it can produce a more detailed CFG and can also show which exceptions can be thrown out or spread from the method call. Based on this idea, the chapter proposes a construction method of ECFG for BPEL. Firstly, we introduce an exception label domain which is of power for collection computing and in which label elements not only are able to describe the exception information of nodes in ECFG but also reflect the exception throwing, propagation, and result after exception elimination or rethrow. Then we present a set of construction algorithms of ECFG for BPEL. Finally, an example demonstrates the feasibility and effectiveness of this method. We summarize our contributions below: (1) We propose a construction method of ECFG for BPEL process. The method is the first approach used to construct ECFG for BPEL, and the method adopts an exception label to describe throwing exception by which we facilitate to analyze exception propagation. (2) We design and implement a set of ECFG construction algorithm to model BPEL process including exception handling structure, and we can use the tools to build ECFG for a given arbitrary BPEL process.

## 40.2 Exception Control Flow Graph

In order to accurately describe and analyze ECFG, this chapter in turn gives formal definitions of exception labels and ECFG.

Exception label is used to describe the needed exception handling at exception edges and nodes. Exception label is described as exception collection, i.e., the information contained in exception label is various exceptions to be treated.

**Definition 1** The exception label set domain  $\Gamma$  is defined as  $\Gamma = \{\Gamma_{pro} \mid \Gamma_{pro} \subseteq \{e \mid e \text{ is an exception type in BPEL process}\}\}$ .

**Definition 2** The exception label  $\tau$  is defined as  $\tau = \{e \mid e \text{ is exception type thrown to be treated in BPEL process}\}$ .

We define the set operations on  $\Gamma$ . Given two elements  $\tau_a$  and  $\tau_b$  in  $\Gamma$ , element  $\tau_c$  is the result of the operation of these two elements in  $\Gamma$ :

Union( $\cup_{\Gamma}$ ): if  $\tau_a \in \Gamma$ ,  $\tau_b \in \Gamma$ , then  $\tau_c = \tau_a \cup_{\Gamma} \tau_b = \{t \mid t \in \tau_a \vee t \in \tau_b\}$ ;  
 Intersection( $\cap_{\Gamma}$ ): if  $\tau_a \in \Gamma$ ,  $\tau_b \in \Gamma$ , then  $\tau_c = \tau_a \cap_{\Gamma} \tau_b = \{t \mid t \in \tau_a \wedge t \in \tau_b\}$ ;  
 Difference( $-_{\Gamma}$ ): if  $\tau_a \in \Gamma$ ,  $\tau_b \in \Gamma$ , then  $\tau_c = \tau_a -_{\Gamma} \tau_b = \{t \mid t \in \tau_a \wedge t \notin \tau_b\}$ ;  
 Equality( $=_{\Gamma}$ ): if  $\tau_a \in \Gamma$ ,  $\tau_b \in \Gamma$ , then  $\tau_c = \tau_a =_{\Gamma} \tau_b = \{t \mid \text{if } \tau_a \subseteq \tau_b \wedge \tau_b \subseteq \tau_a \text{ then } t = \text{true else } t = \text{false}\}$ ;

ECFG for main program of process (ECFG4MP) is used to describe execution of the various activities in the main program of BPEL process. Its formal definition is as follows:

**Definition 3** ECFG4MP denoted by  $G_M$  is a 7-tuple  $(N, E_{reg}, E_{excep}, E_{exceps}, E_c, n_{ps}, n_{pe})$ , where:

- (1)  $N$  is the set of nodes of the graph which is composed of the following subset:  $N = N_{reg} \cup N_c \cup N_{cret} \cup N_{ecret} \cup N_{throw} \cup N_{catch} \cup \{n_{ps}, n_{pe}\}$ , where ①  $N_{reg}$  is a set of normal nodes, ②  $N_c$  is a set of call nodes, ③  $N_{cret}$  is a set of call-return nodes, ④  $N_{ecret}$  is a set of exception-call-return nodes, ⑤  $N_{throw}$  is a set of throw nodes, ⑥  $N_{catch}$  is a set of header nodes of catch blocks, and ⑦  $n_{ps}, n_{pe}$  are start and exit nodes, respectively.
- (2)  $E_{reg}$  is the set of normal control flow edges:  $E_{reg} \subseteq (N_{reg} \cup N_{cret} \cup N_{catch}) \times N$ .
- (3)  $E_{excep}$  is the set of exception control flow edges:  $E_{excep} \subseteq ((N_{ecret} \cup N_{throw}) \times N_{catch} \times \Gamma)$ .
- (4)  $E_{exceps}$  is the set of exception-call-summary edges:  $E_{exceps} \subseteq (N_c \times N_{ecret} \times \Gamma)$ .
- (5)  $E_c$  is the set of normal call-summary edges:  $E_c \subseteq N_c \times N_{cret}$ .

**Definition 4** Service invocation exception graph (SIEG) is defined as 3-tuple  $G_C = (\tau_e, N, E)$ , where:

- (1)  $\tau_e$  is possible exception thrown of the service.
- (2)  $N = \{n_s, n_e, n_{excepe}, n_{idea}\}$ , where,  $n_s$  are start nodes of service invocation,  $n_e$  are normal end nodes of service invocation,  $n_{excepe}$  are exception end nodes of service invocation, and  $n_{idea}$  are concept nodes of service invocation which is used to describe inner abstract logic of the service invocation.
- (3)  $E = \{(n_s, n_{idea}), (n_{idea}, n_e), (n_{idea}, n_{excepe}, \tau_e)\}$ , where edge  $(n_s, n_{idea})$  denotes that control logic flows of called service, edge  $(n_{idea}, n_e)$  denotes normal end of called service, and edge  $(n_{idea}, n_{excepe}, \tau_e)$  denotes exception end of called service and is marked with label  $\tau_e$  which describe possible exception thrown in the called service.

**Definition 5** ECFG for BPEL process (ECFG4BP), abbreviated as ECFG. It is defined as 5-tuple  $G_S = (n_s, n_e, E_c, E_{excep}, G_U)$ , where:

- (1)  $n_s = n_{ps}$  is the start node of ECFG for BPEL process.
- (2)  $n_e = n_{pe}$  is the end node of ECFG for BPEL process.
- (3)  $E_c = \{(n_c, n_s) \mid n_c \in G_M, n_s \in \cup_{C \in B(c)} G_C\} \cup \{(n_e, n_{cret}) \mid n_e \in \cup_{C \in B(c)} G_C, n_{cret} \in G_M\}$  is the set of normal call edges, which are edges between the call node in ECFG4MP and the start node in SIEG of called service and between the normal-exit node in SIEG of called service and the call-return node in ECFG4MP.
- (4)  $E_{excep} = \{(n_{excepe}, n_{ecret}, \tau_{exit}) \mid n_{excepe} \in \cup_{C \in B(c)} G_C, n_{ecret} \in G_M\}$  is the set of return edges of called service, which is edges between the exception-exit node in SIEG of called service and the service-call-exception-return node in ECFG4MP.
- (5)  $G_U = \cup_{C \in B(c)} G_C \cup G_M$  is the union of ECFG4MP and SIEG in BPEL process.



### 40.3 Construction Method for Exception Control Flow Graph

BPEL process is the application program which composes various web services to complete certain functions. Thus, when ECFG for BPEL process is to be built, BPEL process combines corresponding SIEG of the call point, i.e., we may form ECFG for BPEL process by call edges and return edges of called service.

Construction method of ECFG4MP is obtained by performing post-order traversal on abstract syntax tree (AST) for BPEL code, i.e., when analyzing  $\langle \text{process} \rangle$ , you must first analyze the various components active in the  $\langle \text{process} \rangle$ ; when analyzing activity, we need take a different approach toward the type of activity; when analyzing  $\langle \text{invoke} \rangle$  activity, build up not only its call node and normal return node but also exception-return node according to the possible exception thrown. When analyzing basic activity which is not  $\langle \text{invoke} \rangle$  activity, we need to build the corresponding node; when analyzing structure activity, we need to analyze its composition activities according to the semantics of the structured activity.

BPEL process is a special kind of scope, so its construction algorithm of CFG is similar to scope. When visiting AST's scope node, we establish corresponding a set of nodes and edges on ECFG. For every call, the *VisitScope* algorithm returns a 4-tuple  $\langle N_b, N_e, N_{unresn}, N_{unrese} \rangle$ , where  $N_b$  and  $N_e$ , respectively, are the set of nodes corresponding to start and normal exit of the ECFG region corresponding the current AST node.  $N_{unresn}$  is the set of nodes in an ECFG region that has some unresolved incoming or outgoing edges, which is resolved by an ancestor's visitor. For example, if a throw node is enclosed within a scope and has no catch node to match, then out edges of throw are given by its upper catch node of the enclosed scope, that is, to establish an exception edge for the catch node.

The ideas of *VisitScope* algorithm are as follows: First, construct the ECFG nodes and edges for scope block and all catch handling programs by visiting them recursively. It then divides the exception-related ECFG nodes in the scope block into three sets: (1) throw node, (2) the exception-call-return node, and (3) the exception node thrown from lower scope. For throw node, build an exception edge from the throw node to the corresponding head node of catch handling programs, which is marked with exception label. The information in throw node is always a single exception type which is a throw expression type. For exception-call-return node  $n_{ecret}$ , a map *ECRF* is used to obtain the exception type set for node  $n_{ecret}$ . When there is a matching catch block, *ECRF* ( $n_{ecret}$ ) decreasingly update the rest of the exception type (using the difference operator  $-\Gamma$ ). If these exceptions have no matching catch block, they can be thrown from this call node to the upper scope. And we mark  $n_{ecret}$  with the final value of *ECRF* ( $n_{ecret}$ ). If these exceptions have no matching head node of catch block, we build an exception edge from exception-call node to the head node of catch block and mark the appropriate exception information. For the exception nodes  $n_{rethrowUp}$  thrown from lower scope, construction method for ECFG is similar to exception-call-return node.

Algorithm 1: *VisitCall* algorithm

Input:  $i$ : A invoke active;

Output:  $(N_b, N_e, N_{unres}, N_{unrese})$ ;

- 1:  $N_c = N_c \cup \{n_{ci}\}$ ;  $N_{cret} = N_{cret} \cup \{n_{cret}\}$ ;  $N_{ecret} = N_{ecret} \cup \{n_{ecret}\}$ ;
- 2:  $op = CallTargets(typeOf(n_{ci}))$  where  $calltriple(n_{ci}, n_{cret}, n_{ecret})$ ;
- 3: let  $G_{Oop} = \langle \tau_{eop}, n_{sop}, n_{eop}, n_{excepops}, n_{ideaop}, E_{op} \rangle$ ;  $\tau_e = \{n_{ci}, n_{cret}, \tau_{eop}\}$ ;
- 4:  $E_c = E_c \cup \{(n_{ci}, n_{cret})\}$ ;  $E_{except} = E_{except} \cup \{(n_{ci}, n_{cret}, \tau_e)\}$ ;
- 5:  $N_b = \{n_{ci}\}$ ;  $N_e = \{n_{cret}\}$ ;  $N_{unres} = \{n_{cret}\}$ ;  $N_{unrese} = \{n_{ecret}\}$ ;
- 6: return  $(N_b, N_e, N_{unres}, N_{unrese})$ .

**Fig. 40.1** Construction algorithm of ECFG for <invoke> activity

**Fig. 40.2** Construction algorithm of ECFG for <throw> activity

Algorithm 2: *VisitThrow* algorithm

Input:  $a$ : A throw /rethrow active;

Output:  $(N_b, N_e, N_{unres}, N_{unrese})$ ;

- 1:  $N_{throw} = N_{throw} \cup \{n_a\}$ ;
- 2:  $N_b = N_{unrese} = \{n_a\}$ ;  $N_e = N_{unres} = \{\}$ ;
- 3: return  $(N_b, N_e, N_{unres}, N_{unrese})$ .

**Fig. 40.3** Construction algorithm of ECFG for exception handling program

Algorithm 3: *VisitHandler* algorithm

Input:  $eh$ : A handler where  $h=(t \ v) \ bI$ ;

Output:  $(N_b, N_e, N_{unres}, N_{unrese})$ ;

- 1: let  $(N_b1, N_e1, N_{unres1}, N_{unrese1}) = VisitBlock(bI)$ ;
- 2:  $N_{catch} = \{n_{eh}\}$ ;  $E_{reg} = E_{reg} \cup \{(n_{eh}, n_s) | n_s \in N_b1\}$ ;
- 3:  $N_b = \{n_{eh}\}$ ;  $N_e = N_e1$ ;  $N_{unres} = N_{unres1}$ ;  $N_{unrese} = N_{unrese1}$ ;
- 4: return  $(N_b, N_e, N_{unres}, N_{unrese})$ .

**Fig. 40.4** Construction algorithm of ECFG for BPEL process

Algorithm 4: *BuildSECFG* algorithm

Input:  $ss$ : Service software;

Output:  $G_S = (n_s, n_e, E_c, E_{except}, G_U)$ ;

- 1:  $G_U = \bigcup_{c \in ss} G_C \cup G_M$ ;
- 2: foreach  $calltriple(n_{call}, n_{cret}, n_{excepcret})$  do
- 3:  $op = CallTargets(typeOf(n_{call}))$ ;
- 4: let  $G_{Oop} = \langle \tau_{eop}, n_{sop}, n_{eop}, n_{excepops}, n_{ideaop}, E_{op} \rangle$ ;
- 5: let  $E_c = E_c \cup \{(n_{call}, n_{sop}), (n_{eop}, n_{cret})\}$ ;
- 6:  $E_{except} = E_{except} \cup \{(n_{excepops}, n_{excepcret}, \tau_{eop})\}$ ;
- 7:  $E_c = E_c - \{(n_{call}, n_{cret})\}$ ;  $E_{except} = E_{except} - \{(n_{call}, n_{excepcret}, \tau_{eop})\}$ ;
- 8: return  $G_S = (n_s, n_e, E_c, E_{except}, G_U)$ .

VisitScope algorithm also establishes a head node ( $n_{scope}$ ) and connection node ( $n_{join}$ ).

*VisitCall* algorithm (see Fig. 40.1) establishes three nodes and two edges for invoke activity. *VisitThrow* algorithm (see Fig. 40.2) establishes a determined node for throw activity. *VisitHandler* algorithm (see Fig. 40.3) establishes the appropriate head node for catch block and connects this node to active nodes in the block:

The *BuildSECFG* algorithm (see Fig. 40.4) shows construction method for ECFG of BPEL process based on ECFG4MP and SIEG. ECFG is the union of ECFG4MP and SIEG; it is formed by adding call edge and return edges between ECFG4MP and SIEG. Specific construction method is for each call point in the

main program of process, first determine invocation target for each call point. Second, add the following three edges: (1) call edge, from a call node to the start node of SIEG for the target service; (2) call-return edge, from normal-exit node of SIEG for the target service to call-return node; and (3) the exception edge, from exception-exit node of SIEG for the target service to exception-call-return node. And at this exception edge marks the exception information  $\tau_{\text{exit}}$  of the incoming edge of exception-exit node, which is the initial data flow of exception analysis between main programs of BPEL process and called service. Finally, delete the summary edges connecting call node with call-return node and the summary edge connecting call node with exception-call-return node:

## 40.4 Case Study

The following will give the example of “data collection vehicle assembly station BPEL process” to illustrate the effectiveness of the construction method of ECFG.

### 1. Design of exception label

- (a) Exception label domain  $\Gamma$  of this process is defined as follows:  $\Gamma = \{\Gamma_{\text{pro}} \mid \Gamma_{\text{pro}} \subseteq \{\text{Collecting}::\text{UnRegistered}, \text{Collecting}::\text{InvalidSN}, \text{Collecting}::\text{InvalidSchema}, \text{Collecting}::\text{IncompleteBasic}, \text{Collecting}::\text{Timeout}, \text{Collecting}::\text{UnavailService}, \text{Collecting}::\text{InvalidVerification}, \text{Collecting}::\text{SLA}, \text{Collecting}::\text{MissingPieces}, \text{Collecting}::\text{ThanPieces}, \text{Collecting}::\text{ProviderNoMatch}, \text{Collecting}::\text{DataRecollection}, \text{Down}::\text{InvalidCollecting}, \text{Down}::\text{InvalidUpCollecting}, \text{Down}::\text{UnavailService}\}\}$ .
- (b) According to possible exception which is thrown, use an exception label to mark corresponding edges. If calling collection service occurs the collection service unavailable exception, data collected appearing QoS exception, then exception label  $\tau$  is defined as  $\tau = \{\text{Collecting}::\text{UnavailService}, \text{Collecting}::\text{SLA}\}$ .

### 2. Construction of ECFG4MP

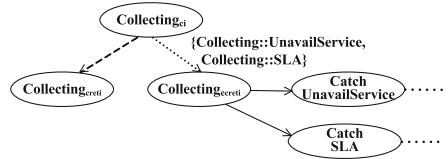
Using the method described, AST is traversed post-order and layer by layer to build ECFG for BPEL process. We use examples of invocation of collection service and <faulthandler> to illustrate the construction method for ECFG4MP.

Figure 40.5 is ECFG for collection service. In this figure,  $\text{Collecting}_{\text{ci}}$  is call node,  $\text{Collecting}_{\text{creti}}$  are normal return nodes after the collection service is called,  $\text{Collecting}_{\text{ecreti}}$  are exception-return nodes, edge  $(\text{Collecting}_{\text{ci}}, \text{Collecting}_{\text{creti}})$  is normal return edge, edge  $(\text{Collecting}_{\text{ci}}, \text{Collecting}_{\text{ecreti}})$  is exception-return edge, and the label  $\{\text{Collecting}::\text{UnavailService}, \text{Collecting}::\text{SLA}\}$  of edge  $(\text{Collecting}_{\text{ci}}, \text{Collecting}_{\text{ecreti}})$  is the possible return exception. Edge  $(\text{Collecting}_{\text{ci}}, \text{Collecting}_{\text{creti}})$  and edge  $(\text{Collecting}_{\text{ci}}, \text{Collecting}_{\text{ecreti}})$  are indicated by a dashed line, which means they are the temporary presence, and after SIEG of collection service is combined, they need to be deleted.

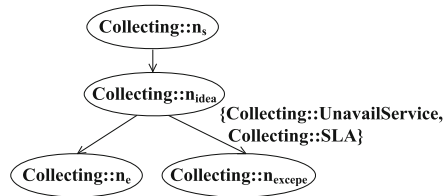
**Fig. 40.5** ECFG for collection service invocation



**Fig. 40.6** ECFG for collection service invocation and its corresponding exception handler



**Fig. 40.7** SIEG for the collection service



**Fig. 40.8** ECFG for “data collection vehicle assembly station BPEL process” (partial)

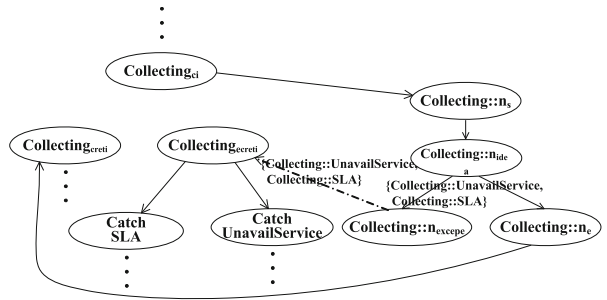


Figure 40.6 is ECFG for collection service invocation and its corresponding exception handler. After the collection service has been called, it will produce two exception, i.e., Collecting :: UnavailService and Collecting :: SLA, and they are in turn captured by exception handler CatchUnavailService and CatchSLA, where the ellipsis represents specific activities of exception handling.

3. Construction of SIEG

For the collection service, learn the possible exception thrown through the service’s WSDL document: Collecting :: UnavailService, Collecting :: SLA anomalies. Figure 40.7 is SIEG for the service.

4. Construction of ECFG for BPEL process

ECFG for BPEL process is formed by combining ECFG4MP with SIEG. Figure 40.8 is ECFG for “data collection vehicle assembly station BPEL process” (partial).

## 40.5 Related Work

There are many differences between exceptions in two-program language, thus construction of ECFG requiring different approaches. Sinha and Harrold [1] incorporate the control flow effects due to explicit Java exceptions in an inter-procedural control flow graph (ICFG) using a flow-sensitive-type analysis. In contrast, our ECFG has an exception-call-return node for every service-call point. Based on byte intermediate representation (BIR) for Java, Demange et al. present construction method for CFG [5]. Following Demange works, Amighi et al. propose a construction method of ECFG for Java byte code, in which the construction procedure divides into two steps: Step 1, Java program is transformed to BIR, and, Step 2, CFG is built through extracting information in BIR [3, 6]. In contrast to all these approaches, our analysis is based on AST.

Jiang et al. propose a construction method of ECFG for C++, in which the throw point of try block is directly conjunct to all corresponding catch block; furthermore, implicit control flow for exception and exception propagation in their ECFG is described [7]. However, construction method is different between AST in our method and source code. Based on a Signed-TypeSet domain compact notation, Prabhu et al. adopted inter-procedure data stream algorithm to obtain the C++ IECFG [4]. Our works follow some idea in constructing IECFG; in contrast to the approach, our exception label set domain  $\Gamma$  has no sign, which makes it easy to understand.

## 40.6 Conclusion

This chapter has proposed a construction method of ECFG is for BPEL process. The difference in this method from the past construction method of ECFG: When building ECFG, one use label one to describe clearly the exception information in system and its propagation, and one can use this information to analyze exception propagation and uncaught exception.

## References

1. Sinha, S., & Harrold, M. J. (2000). Analysis and testing of programs with exception-handling constructs. *IEEE Transactions on Software Engineering*, 26(9), 849–871.
2. Jo, J. W., & Chang, B. M. (2004). Constructing control flow graph for java by decoupling exception flow from normal flow. In *International Conference on Computable Science and Its Applications* (pp. 106–113). Heidelberg: Springer.
3. Amighi, A., de Gomes, P. C., Gurov, D., & Huisman, M. (2012). Sound control-flow graph extraction for Java programs with exceptions. In *Software Engineering and Formal Methods* (pp. 33–47). Heidelberg: Springer.

4. Prabhu, P., Maeda, N., Balakrishnan, G., Ivančić, F., & Gupta, A. (2011). Interprocedural exception analysis for C++. In *ECOOP 2011* (pp. 583–608). Heidelberg: Springer.
5. Demange, D., Jensen, T., & Pichardie, D. (2009). *A provably correct stackless intermediate representation for Java bytecode* (Research Report 7021). INRIA.
6. Amighi, A., Gomes, P., & Huisman, M. (2011). *Provably correct control-flow graphs from Java programs with exceptions* (Technology Report). KTH Royal Institute of Technology.
7. Jiang, S., & Jiang, Y. (2007). An analysis approach for testing exception handling programs. *SIGPLAN Notices*, 42(4), 3–8.

# Chapter 41

## P300 Detection in Electroencephalographic Signals for Brain–Computer Interface Systems: A Neural Networks Approach

Seyed Aliakbar Mousavi, Muhammad Rafie Hj. Mohd. Arshad, Hasimah Hj. Mohamed, Putra Sumari, and Saeed Panahian Fard

**Abstract** Brain–computer interface systems are communicative mediums between human brain and external device. One of the applications of these systems is P300 speller. This application provides the ability to spell the characters on the screen for disabled people. In this study, we review the character recognition and its relation to P300 detection. Then, we used three neural networks models with flexible activation functions to detect P300 patterns from electroencephalographic signals more accurately. The obtained results have shown the accuracy of the character recognition based on the precision and recall measures.

### 41.1 Introduction

Brain–computer interface (BCI) is a direct communicative pathway between brain and external devices [1–3]. BCI allows people to communicate without having to move by measuring brain activities. Disabled people are mostly benefited from BCI systems. BCI systems are commonly composed of four components as follows: signal acquisition from brain, signal preprocessing, signal feature extraction, and classification. BCI is using noninvasive electroencephalographic (EEG) method to read brainwaves [4]. EEG provides a direct measure of brain activities in milliseconds with temporal resolution. BCI systems use feature extraction techniques such as pattern recognition for classification of specific brain’s event-related potentials. One of the well-known BCI applications is the P300 speller to detect the characters from screen.

---

S.A. Mousavi • M.R.H.M. Arshad • H.H. Mohamed • P. Sumari  
School of Computer Sciences, Universiti Sains Malaysia, Pulau Pinang 11800, Malaysia  
e-mail: [pouyaye@gmail.com](mailto:pouyaye@gmail.com); [rafie@cs.usm.my](mailto:rafie@cs.usm.my); [hasimah@cs.usm.my](mailto:hasimah@cs.usm.my); [putras@cs.usm.my](mailto:putras@cs.usm.my)

S.P. Fard (✉)  
School of Mathematical Sciences, Universiti Sains Malaysia, Pulau Pinang 11800, Malaysia  
e-mail: [saeedpanahian@yahoo.com](mailto:saeedpanahian@yahoo.com)

The main task in P300 speller paradigm is to detect the P300 peak patterns in EEG accurately and instantly. Detecting P300 patterns lead to identification of corresponding characters in P300 speller paradigm. The accuracy of P300 pattern detection will ensure better character recognition in the classification step. The detection of P300 patterns is a challenging task due to presence of noise and artifacts in EEG signals.

The motivation of this study is to detect P300 patterns in EEG signals more accurately and instantly than the previous researches. The approach of this study is to use backpropagation feedforward neural networks with flexible activation functions for detection of P300 signals. These networks use different channel sets and data sizes. Each network has different activation functions with few free parameters. The significance of flexible activation functions can be expressed in networks' size reduction, and the training becomes faster [5–8]. Moreover, we try to evaluate the character recognition performance by direct measurements in P300 classification.

This study is organized as follows: in Sect. 41.2, BCI P300 speller paradigm is introduced. And the P300 detection is explained. Moreover, the P300 speller database is studied. In Sect. 41.3, the process of proposed methods is given. And data preprocessing and feature extraction is explained. Furthermore, three neural networks models with flexible activation functions are used to detect P300 signals and character recognition. In Sect. 41.4, the result of P300 classification and character recognition is given. In Sect. 41.5, a comparison of results with current methods is depicted. In Sect. 41.6, conclusions are given.

## 41.2 P300 Speller Paradigm

In BCI, P300 speller is a benchmark paradigm. P300 in BCI was initially introduced by Donchin et al. [9]. P300 signals are reflection of brain perceptive response to outside stimulus. By receiving the P300 signal on specific stimulus, we can assume that brain has perceived that event. P300 is a positive signal occurring about 300 ms after flashing light mostly in parietal lobe and occipital sites of brain [10]. Receiving P300 signal is equivalent to detection where user was looking about 300 ms before receiving the signal. The advantages of P300 in BCI are being straightforward and fast and requiring no practical training. The BCI P300 speller application assists in spelling words using P300 signals. The  $6 \times 6$  matrix of characters is composed of 26 alphabetical (A–Z) and 10 numerical characters (0–9). This matrix is shown to the user on the computer screen by randomly flashing the rows and columns. Once the user focuses on the character during the flashing light, the signals corresponding to those row and column of depicted character have P300 signal. Detection of P300 signal makes it possible to recognize the character by matching the responses to one of the rows and one of columns. In Fig. 41.1, A shows the brain region that P300 peak is occurring, B shows the P300 signal peak, and C represents the BCI P300 Speller row and column paradigm.



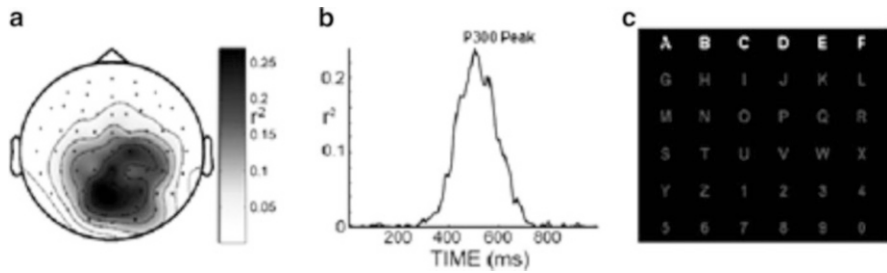


Fig. 41.1 BCI P300 speller paradigm [9]

### 41.2.1 P300 Detection

The P300 signal is an event-related potential (ERP) in EEG recording [11]. That means P300 signal is a time-locked and averaged EEG pattern. Having accurate detection of P300 signals on the P300 speller paradigm would result in better character recognition. There are two types of classification in P300 speller paradigm. In the first classification type, the P300 signals can be detected in the EEG recordings by having two sets of classes: P300 or none-P300. Due to artifacts or noises, however, there is uncertainty in P300 classification attempt. The second classification type is character identification. The output of P300 detection is applied to classify the application characters and symbols. The character identification process has strong certainty as the characters are given to user clearly.

### 41.2.2 Database

BCI P300 speller dataset II from the third BCI competition was applied in different methods [12, 13]. This dataset has records of P300 evoked potentials from two subjects A and B. In the experiment, the characters from the presented words are focused sequentially by the subjects. Character selection is done by intensifying a column and a row, and then consequently 2 of the 12 intensifications should contain P300 signals in the samples. Arbitrarily, the intensifications of columns and rows are intersected in blocks of 12. For each character trial which contains the rows/columns, the sets of 12 intensifications are repeated 15 times. Therefore, 30 possible P300 responses should exist for each character. The sampling rate is 240 Hertz (Hz) and is band-pass filtered from 0.160 Hz. The test dataset has 100 characters and training dataset has 85 characters. The number of P300 signals in training dataset is  $85 \times 2 \times 5$  and in test dataset is  $100 \times 2 \times 15$ .

## 41.3 Methods

### 41.3.1 Framework

The framework in Fig. 41.2 illustrated the process of the P300 detection and the character recognition. The process starts with preprocessing of the P300 training datasets and extracting the features with building the classifiers. Then, three neural networks models with flexible activation functions are used to train the data. Moreover, The P300 classification results are extracted with the character vectors to apply in test datasets. Finally, the results from character recognition are analyzed and compared with current methods in P300 speller paradigm.

### 41.3.2 Preprocessing and Feature Extraction

The EEG signals contain the P300 pattern, and those values from the channels are the inputs during the  $T$  time. The inputs of neural networks are  $I_{i,j}$ . The number of channels is defined by  $i$  ( $1 \leq i \leq 64$ ) and the timing of the signals is denoted by  $j$ . Although the signals are captured in time by number of samples per second, the total number of samples taken for the analysis is determined by multiplying the sampling rate (SR) to the desired time window  $T$  in second ( $1 \leq j \leq SR \times T$ ). The dataset divided half by subsampling EEG signal in sampling rate of 120 Hz [13]. Those EEG patterns are then band-pass filtered from 0.5 to 18 Hz only to keep relevant signals. A matrix of  $i \times j$  is the input of the neural networks [13]. The three classifiers are presented for feature extraction from EEG signals. These classifiers vary in terms of data size and the number of channels. The classifier NC1 uses the whole training dataset with all channels, whereas the classifier NC2 employs only eight channels. However, the classifier MultiNc uses a multiclassifier strategy. This classifier consist of five classifiers which are trained on different sets of pairs of P300 patterns. In training dataset, the signal patterns without P300 are five times more than the patterns with P300. Each classifier is trained with equal number of P300 and none-P300 patterns [13].

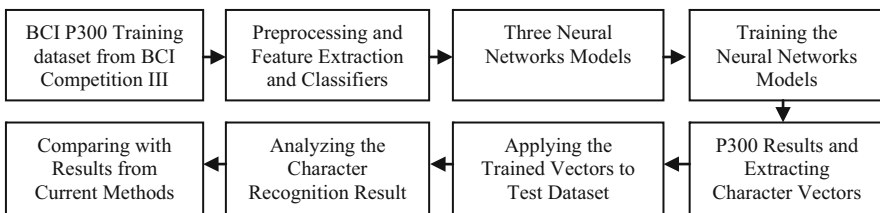


Fig. 41.2 The diagram for the proposed methods

### 41.3.3 Neural Networks Models

In this section, as a new approach in P300 signal detection and character identification, the following three neural networks models will be used:

- NN1: 35 nodes, 0.9 learning rate,  $F_1(x) = A/(1 + e^{-Bx})$  as flexible activation function which was extension of sigmoid function [8]
- NN2: 20 nodes, 0.9 learning rate,  $F_2(x) = A_1 \sin(B_1x) + A_2/(1 + e^{-B_2x})$  as flexible activation function and was used in financial modeling [6]
- NN3: 25 nodes, 0.2 learning rate,  $F_3(x) = e^{-B_1x^2}(\cos(B_2x))$  as flexible activation function was used in previous studies [5, 7]

The three neural networks models are chosen to provide a benchmark for the models and classifiers. The functions are chosen due to their capability of being flexible in the networks, and they also have not been used before in BCI P300 speller paradigm. The activation functions are improved using free parameters [5], and the number of nodes in hidden layer and the learning rate  $\beta$  are determined empirically for every model. Moreover, the free parameters are adjusted at the end of each network's epoch. The adjustment of free parameters is like the weights throughout the networks. This adjustment is based on steepest descent rule. This learning algorithm has two main parts as feedforward and error backpropagation. Furthermore, the input and output functions of each neuron in the networks are defined, respectively, in Eqs. (41.1) and (41.2):

$$I_{m,L}(x) = \sum_n [W_{m,n,L} O_{m,L-1}(x)] + \theta_{m,k} \quad (41.1)$$

$$O_{m,L}(x) = F_1(I_{m,L}(x)) = \frac{A_{m,L}}{1 + e^{-B_{m,L} \cdot I_{m,L}(x)}} \quad (41.2)$$

The input of  $m$ th neuron in the  $L$ th layer is represented by  $n$ . Apart from this, the  $O_{m,L}(x)$  is used to calculate the value of output from  $m$ th neuron in the  $L$ th layer for NN1 flexible activation function. Finally, the output of the neural networks models are character vectors. These character vectors are applied to the test dataset to evaluate the models in character recognition.

## 41.4 Results

### 41.4.1 P300 Classification

The important approach from classification of P300 signals is to find out which measurement is in direct relation to character recognition in character identification problem. The recall or the precision is the most directly related measurement in

**Table 41.1** P300 classification result

Subj	Model	Class	TP	TN	FP	FN	Recall	Precision	Recog	F-score
A	NN1	NC1	2,216	10,524	4,297	963	69.7	34.0	70.8	45.7
		NC2	1,935	10,355	4,570	1,140	62.9	29.7	68.3	40.4
		MultiNc	2,362	10,160	4,567	911	72.2	34.1	69.6	46.3
	NN2	NC1	2,241	10,465	4,326	968	69.8	34.1	70.6	45.8
		NC2	1,949	10,325	4,580	1,146	63.0	29.9	68.2	40.5
		MultiNc	2,393	10,112	4,575	920	72.2	34.3	69.5	46.6
	NN3	NC1	2,276	10,458	4,296	970	70.1	34.6	70.7	46.4
		NC2	1,964	10,316	4,578	1,142	63.2	30.0	68.2	40.7
		MultiNc	2,461	10,057	4,569	913	72.9	35.0	69.5	47.3
B	NN1	NC1	2,230	11,918	2,903	949	70.1	43.4	78.6	53.7
		NC2	2,001	11,467	3,458	1,074	65.1	36.7	74.8	46.9
		MultiNc	2,493	11,265	3,462	780	76.2	41.9	76.4	54.0
	NN2	NC1	2,255	11,859	2,932	954	70.3	43.5	78.4	53.7
		NC2	2,015	11,437	3,468	1,080	65.1	36.7	74.7	47.0
		MultiNc	2,524	11,217	3,470	789	76.2	42.1	76.3	54.2
	NN3	NC1	2,290	11,852	2,902	956	70.5	44.1	78.6	54.3
		NC2	2,030	11,428	3,466	1,076	65.4	36.9	74.8	47.2
		MultiNc	2,592	11,162	3,464	782	76.8	42.8	76.4	55.0

order to predict the target character, after running the neural networks, the following information is generated (Table 41.1): the number of false negative (FN), true negative (TN), true positive (TP), and false positive (FP) in the test database. The P300 recognition rate is defined as  $(TP + TN)/(TP + TN + FN + FP)$ ; the precision is defined as  $(TP)/(TP + FP)$ ; the recall is defined as  $TP/(TP + FN)$ ; and F-score is defined as  $2 \text{ (recall.precision)}/(\text{precision} + \text{recall})$ . In Table 41.1, the neural networks model NN3 shows better results in both subjects. And, Subject B is having better results in comparing to Subject A. Furthermore, the result of P300 recognition rate in Subject B is higher than Subject A. And NN1 and NN3 models show very close P300 recognition rate. The best recognition rates for both subjects are achieved in models NN3 and NN1 by classifier NC1. The results of Subject B indicate best precision is achieved by NN3-NC1 with 44.1 %. However, the best recall is achieved by NN3-MultiNc with 76.8 %. And in Subject A, the best precision and recall are achieved by NN3-MultiNc with, respectively, 35 % and 72.9 %. Therefore, based on the results, NN3 model shows an overall improvement in classification of P300 signals.

The important approach from classification of P300 signals is to find out which measurement is in direct relation to character recognition in character identification problem. Either recall or precision is the most directly related measurement in order to predict the target character.

**Table 41.2** Character recognition percentage in NN3 model

Classifiers	Subjects	Number of trials			
		1	5	10	15
NC1	A	14	64	70	91
	B	29	81	86	89
	Mean	21.5	72.5	78	90
NC2	A	12	49	66	79
	B	22	74	82	90
	Mean	17	61.5	74	84.5
MultiNc	A	13	65	76	91
	B	22	81	84	95
	Mean	17.5	73	80	93

### 41.4.2 Character Recognition

The results from P300 classification are taken as index to identify the target characters in this section. Based on the results in Table 41.1, the neural networks model NN3 has produced better results in P300 detection. Table 41.2 shows the character recognition rate in percentage based on number of trials for each classifier in model NN3.

In Table 41.2, the accuracy of character recognition is developed by increasing the number of trials. On the first observation, more than 65 % of target characters are recognized by having 10 trials. The MultiNc classifier has a mean rate of 93 % for 15 trials. It can be considered a good recognition rate in P300 speller paradigm. Subject B shows better character recognition accuracy in comparing to Subject A. By comparing character recognition result with P300 classification result in NN3 model, recall measurement is better representing the models and classifiers ranking in character recognition problem.

## 41.5 Comparison of Results with Current Methods

In order to provide a benchmark for this research, we intend to compare the results of proposed neural networks with previous methods [13–16]. Table 41.3 shows the character identification results from other methods including the result from this study. The MultiNc and NC1 classifier from NN3 models are representing our method. ESVM method has achieved the best result for 15 trials [16]. However, our classifiers in proposed NN3 model achieved comparatively better recognition rate when there is only five trials. Unlike the previous methods, the advantage of proposed model is to not consider any channels set nor feature selection before starting. The networks training in our model is done on raw data with some small preprocessing and it is close to real BCI implementation.

**Table 41.3** Comparison of results with current methods in BCI

Subjects	Trials	Methods				
		LDA [15]	MCNN [13]	ESVM [16]	NC1	MultiNc
A	5	45	61	72	64	65
	10	78	82	83	70	76
	15	88	97	97	91	91
B	5	76	77	75	81	81
	10	92	92	91	86	84
	15	96	94	96	89	95
Mean	5	60.5	69	72.5	72.5	73
	10	85	87	87	78	80
	15	92	95.5	96.5	90	93

## 41.6 Conclusion

In this study, we have reviewed BCI P300 speller paradigm as the well-known BCI application. Then we have used BCI P300 speller dataset II from the third BCI competition. A preprocessing and feature extraction was used to extract the P300 signals and none-P300 signals. NC1, NC2, and MultiNc were built from the extracted features. Then the classifiers were input into the three neural networks models with flexible activation functions. Based on the output of these networks, we found out MultiNc classifier in model NN3 showed high character recognition rate by having 93 % in mean for 15 trials. Furthermore, we have derived that four pivots analyze the performance of character recognition. Finally, we have compared our results with current methods in Table 41.3. The obtained results show higher accuracy of P300 signals detection and consequently increasing the rate of characters recognition.

## References

- Allison, B. Z., Wolpaw, E. W., & Wolpaw, J. R. (2007). Brain-computer interface systems: Progress and prospects. *Expert Review of Medical Devices*, 4(4), 463–474.
- Birbaumer, N., & Cohen, L. G. (2007). Brain-computer interfaces: Communication and restoration of movement in paralysis. *Physiology London*, 579(3), 621–636.
- Kostov, A., & Polak, M. (2000). Parallel man–machine training in development of EEG-based cursor control. *IEEE Transactions on Rehabilitation Engineering*, 8(2), 203–205.
- Niedermeyer, E., & Lopes, F. d. S. (2004). *Electroencephalography: Basic principles, clinical applications and related fields*. Lippincott Williams & Wilkins.
- Özbay, Y., & Tezel, G. (2010). A new method for classification of ECG arrhythmias using neural network with adaptive activation function. *Digital Signal Processing*, 20(4), 1040–1049.

6. Zhang, M., Fulcher, J., & Xu, S. (2002). Neuron-adaptive higher order neural network models for automated financial data modeling. *IEEE Transactions on Neural Networks*, 13(1), 188–204.
7. Subasi, A., Alkan, A., Koklukaya, E., & Kiyimik, M. K. (2005). Wavelet neural network classification of EEG signals by using AR model with MLE preprocessing. *Neural Networks*, 18(7), 985–997.
8. Liu, T. I. (1993). On-line sensing of drill wear using neural network approach. *IEEE International Conference on Neural Networks*, 2(1), 690–694.
9. Donchin, E., Spencer, K. M., & Wijesinghe, R. (2000). The mental prosthesis: Assessing the speed of a P300-based brain–computer interface. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 8(2), 174–179.
10. Krusienski, D. J., Sellers, E. W., McFarland, D., Vaughan, T. M., & Wolpaw, J. R. (2008). Toward enhanced p300 speller performance. *Neuroscience Methods*, 167(1), 15–21.
11. Coles, M. G. H., & Rugg, M. D. (1996). Event-related brain potentials: An introduction. *Electrophysiology of mind* (pp. 1–27). Oxford Scholarship Online Monographs.
12. Blankertz, B., Muller, K. R., Krusienski, D. J., Schalk, G., Wolpaw, J. R., Schlogl, A., et al. (2006). The BCI competition. III: Validating alternative approaches to actual BCI problems. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 14(2), 153–159.
13. Cecotti, H., & Graser, A. (2011). Convolutional neural networks for p300 detection with application to brain-computer interfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(3), 433–445.
14. Blankertz, B. (2008). *BCI competition III final results*. <http://ida.first.fraunhofer.de>
15. Liang, N., & Bougrain, L. (2008). *Averaging techniques for single-trial analysis of oddball event-related potentials* (pp. 44–49). Proceedings of the Fourth International BCI Workshop and Training Course (INRIA-00337070, version 1).
16. Rakotomamonjy, A., & Guigue, V. (2008). BCI competition III: Data set II ensemble of SVMs for BCI p300 speller. *IEEE Transactions on Biomedical Engineering*, 55(3), 1147–1154.

# Chapter 42

## Web Content Extraction Technology

Zhenyu Jiao, Xiaoben Yan, Jinjin Sun, Yuchen Wang, and Jiangbin Chen

**Abstract** In this information era, we are facing the knowledge explosion, and the information on the Internet is multifarious. It is not convenient enough for us to access to information directly on cell phones due to their limitation. Based on parsing a web page with regarding it as a DOM (document object model) tree, we extract the valuable information with considering three factors: structure, content, and programming habits. For illustration, 28 websites are utilized to show the feasibility of the method in web information extraction, and we design the mobile client to present the web content on the cell phones. The practice has proved that using the web page extraction technology related to this article to browse the corresponding news websites only consumed 8 % of cell phone traffic of the existing mobile phone browser. And the user experience is improved. This method can help people to get rid of costing too much on the cell phone traffic, redundant information, complicated operations, and so on.

### 42.1 Introduction

The twenty-first century is the era of information and the era of data rapid expansion. How fast and efficient accessing to information is becomes an important issue to us. Network is the most commonly used way of obtaining information for us in today's society. Browsing the web with smartphone is the best choice to meet the requirements of our access to information anytime and anywhere.

Comparing with computer, portability and easy access are the advantages of a cell phone, but its unavoidable disadvantage is that it has a smaller screen than a regular computer, and the cost of cell phone traffic is expensive. When using cell phones to browse the webs, which are designed by the size of the computers'

---

Z. Jiao (✉) • X. Yan • J. Sun • Y. Wang • J. Chen  
The College of Information Technology, Nankai University, Tianjin 300071, China  
e-mail: [j.shower@163.com](mailto:j.shower@163.com)



screen, we need to keep zooming in and zooming out to browse the whole picture of the web and read the specific information we need, which will not offer us a delightful experience. What is more, all of us have almost been through the counting and controlling cell phone traffic month-end. Therefore, every cell phone user hopes to be able to browse the web matching their screen with less traffic cost. At present, there are many websites specially developed for the mobile version of their site to solve this problem. However, this “sweep before your own door” approach is a large consumption on resource, which is not really effective to solve this problem. This chapter puts forward an effective solution to solve this problem and verifies the feasibility of this solution through extracting information from 28 websites.

## 42.2 Related Work

The existing web information extraction methods can be divided into five main categories: (1) The method of information extraction based on natural language processing. This method regards the web page as plain text, has a slow processing speed, and needs a lot of samples. The text in the web page does not contain complete sentences, so the using range of this method is small. (2) The method of information extraction based on wrapper induction approach has strong pointedness, low expansibility, and poor reusability and only can be applicable to a kind of web page. For different web pages, a lot of the wrappers are needed. (3) The method of information extraction based on the ontology. This way it has less dependence on the structure of web pages, but it needs domain experts to create a clear ontology in this field in detail, and its workload is big. (4) The method of information extraction based on HTML structure. Its flexibility is strong, but it is only suitable for the page which contains obvious regional structure. (5) The method of information extraction based on web query. It has good versatility and scalability, but there is no guarantee for its accuracy. Our goal is to extract information from web pages and present it on smartphones; this technology must have the following characteristics: fast processing speed, strong scalability, and high accuracy. So the above methods are not suitable for solving our problem.

At present, most of the research papers in the field of web information extraction are aimed at a particular type of website. Such as MDR [1] and Depta [2] are for commodity list or form information, these particular websites often carry with formatting information, such as record information and commodity price information. In order to get a general method, recently, some scholars put forward: This problem can be solved by establishing a template database [3]. However, due to the time of template matching being costly, this method is not suitable for browsing the web. As a result, the existing cell phone browser reduces the network traffic just by reducing the image resolution and does not deal with the content for webs which do not have phone version. Our method takes the structure, the content, and the programming habits into consideration. Its processing speed is good, because it

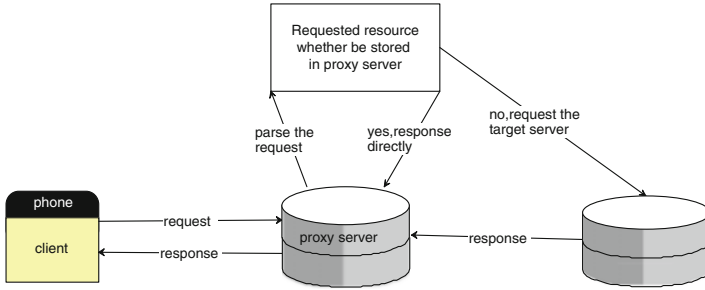


Fig. 42.1 The overall structure of the method

does not rely on a lot of templates. Its scalability and accuracy is good, because it contains a lot of rules, which is confirmed with each other. Our method can handle both strong-structured and loose-structured web pages with a reasonable time, so it meets the need of using mobile phones to browse the web.

### 42.3 Overall Architecture

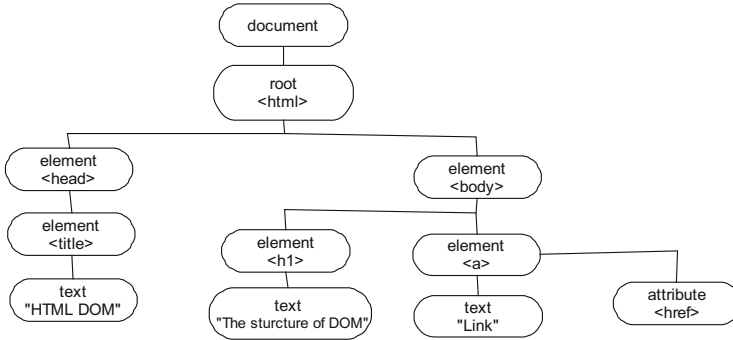
We will add a server with the function of filtering (hereinafter referred to as proxy server) between the cell phone and the server of the goal page (hereinafter referred to as the target server); therefore, the users can get useful web content by accessing the server we added instead of accessing the target server directly. We can see the flow chart below (Fig. 42.1).

The proxy server will parse the request from the client and identify the requested content, then respond directly with its resources if contents have been stored. If the requested resource is not in the proxy server, it will put forward a request to the target server, wait for the target server’s response, parse the related resource, extract valuable information, establish the image file and store it locally, and then respond the client’s request accordingly. When other users request the same page next time, it can be responded directly by the proxy server. In order to prevent storing too much content in the proxy server, we also delete the web image files, which are requested with low frequency.

The proxy server defined in this chapter is not the same as the traditional one. It does not only use traditional caching technology, but also its cached information, which is extracted from web pages, is only part of the original web page.

### 42.4 Web Content Extraction

Web can be seen as a DOM tree (shown in Fig. 42.2); the images and text in the web are nodes in the tree. Extracting content from a web page can be regarded as picking some nodes from the tree. The core work of the content extraction is how to select useful node and eliminate useless nodes at the same time.



**Fig. 42.2** An example of the DOM tree

We need to choose a parser to complete the work of extracting content from a web. There are many kinds of parser and we finally chose the Jsoup [4]. Jsoup is a Java-based parser, which can parse a web both with its URL and with its HTML text content. It provides a set of efficient APIs that can be used to extract and operate data through the DOM, CSS, and a method similar to the operation of the jQuery. We can directly load a document object through the URL of a website, such as using the following code:

```
Document doc = Jsoup. Connect (“http://example.com/”). Get ();
```

We will get the content of <http://example.com/> and store it in the document object. Then, we can find the elements through the DOM method or the selector method [5]. Through the two methods and the combination of two methods above, we can locate any position in the document. A website generally can be divided into two categories: list page and articles page; we will show you how to deal with them respectively.

#### **42.4.1 The Process of List Page**

For list page, what we need includes three items: the titles of news, the release time of news (sometimes may not exist), and the link addresses of news. Because the news headlines are some links (these links are used to jump to the content pages), we will pick out all nodes, which are links with the following code:

```
Elements chooseA = doc. Select (“a”);
```

After getting these links, we need to find which of them are useful news and which are not what we need. For this discrimination, we are going to carry out a comprehensive approach.



Fig. 42.3 An example of extracting information from a web page

For the beauty of the website, we need to set the CSS file to decorate it. CSS file uses different styles for different categories of nodes, while the same style for the same category of nodes. For news in the list, the nodes often have the same “class” attribute value. Therefore, we can select similar nodes according to its “class” attribute value with using the method <select (“className”)>.

We can locate the news nodes according to the literals. The literals of the news list node are more than the other nodes and have a gathered state. Therefore, we can use the method <text ()> to obtain the text of a node; if a node’s literal is significantly higher than others, we can mark it as a news node.

All link nodes have “href” attributes, which are used to determine the URL of the target links. Similar news should be stored under the same path on the server, so the first half of their URL is the same, different from the other kind of nodes at the same time. Since people use meaningful strings to name the directory, the strings such as “News” and “list” often appear in the “href” path. We can use the method <attr (“AttributeKey”)> to obtain the attribute values of a node.

Based on the above, we can locate the news list of each website. Then, we need to get the corresponding link address and the release time of the news. URL can be directly gotten from the “href” attributes (sometimes we should change the relative path to the absolute path by adding a prefix), with simply calling the Jsoup method <attr (“href”)>. Due to the news release time has a close relationship to the news; it is often near to the news node, appearing in the DOM tree as the child node or the sibling node of the news node. Then we pick out right ones among these nodes according to the format of release time (numbers separated with space mark). After this we can conveniently get the release time. As shown in figure, we have finished the first step in the conversion (Fig. 42.3):

### 42.4.2 The Process of Article Page

For the article page, what we need from it includes the title of the article, the released information of the article (including released time, author information, which sometimes may not exist), the text of the article, and the pictures in the article.



Fig. 4.2.4 An example of tagging the text

How to locate the body of the article: Article body content is generally in the core of the entire site and has the characteristics of big literals. So the parent node's literals in the position should be significantly higher than that of the other nodes. First of all, you can use the method `<parent().text(>>` to get the text of the parent node; if its literals are significantly higher than the other nodes, this parent node can be thought as at the core of the text. And in order not to regard the entire page as a core, a node should be compared with its sibling nodes layer by layer with a bottom-up approach. If its literals are significantly higher than its sibling nodes, it can be regarded as the core. Once the core node is identified, its sub-nodes are often the main body of the article. Regarding the parent node as the root of a DOM tree, as the body of the article in general would not be center-aligned, so we use the method `<select (" p[align! = center]">` to select all nodes which are tagged as paragraph from the parent node. It is the body of the essay (sometimes may contain the title or the release information).

The title of the article is often located above the core node and inside its parent node. Under normal circumstances, the article title is set tag `<class = title>` and is center-aligned, so we can get the title with the method `<select ("p [class = title], div [class = title]">`, or `<select (" p[align=center], span [class=title]">` based on recursion. If the returning value of these methods is empty, we will locate to the grandfather node, call again until the returning values are not empty. If there is more than one returning value, the first one shall be the title, the others, and subheadings.

As the release information of the article is center-aligned, we deal with it with the similar method: `<select ("align=center").text(>`, but we should pay attention to a case that the returning value is the title. We need to compare the first returning value with the title of the article, if they are the same, and then abandon the former one.

The layout of image is difficult to locate, but it is usually below the title. So we can select images with the method `<select("img">` after locating the article title.

We will package the transferred text to get XML (Extensible Markup Language) files for the client to recognize them. We tag the labels according to our needs of transmission and processing. For example, we use the following form for list page (Fig. 4.2.4):

### 42.4.3 Image Compression

The size of the image is much larger than the text's, so images processing is very important. We get rid of the useless images directly in the parsing process, so we will only deal with the useful images. In order to adapt to the cell phone screen display, reduce traffic burden, we will compress the images. And we will compress the images to different levels for different cell phones according to the specific screen pixel information we get from them, so cell phone users can save their mobile traffic and get the most suitable image size when they view the pictures.

## 42.5 Experiment Evaluation

According to the above methods, we designed the client on android mobile phones and tested the home pages of Nankai University and its department's websites (Fig. 42.5).

Seen from the diagram above, using our solution has improved our reading experience; the original websites are difficult to read directly on cell phones; after treatment by our system, they become concise and easy to read. We can see a record of actual test results (a small part) in detail in Table 42.1 below.

We can see the data in Table 42.1; after using our solution, the average consumption of the traffic on visiting the listed pages is only 8 % of the cost of UC Browser; the maximum consumption of traffic is less than 16.3 % of UC Browser's way.

For the article page, the test results are as follows (Table 42.2).

We can see the data in Table 42.2; after using our solution, the average consumption of the traffic on visiting the article pages is only 56% of the cost of UC Browser. And the more pictures a web page contains, the more traffic will be saved.



Fig. 42.5 The presentation of using our method to show the web page on phones

**Table 42.1** Actual test results of listed page in detail

Tested websites	Traffic consumed by UC Browser	Traffic consumed by our solution
Nankai University	236.36 KB	6 KB
Nankai News	1.26 MB	25 KB
College of Arts	40.96 KB	6 KB
School of History	92.16 KB	4 KB
Philosophy School	0.99 MB	10 KB
Law School	30.72 KB	5 KB
Zhou Enlai School of Government	71.68 KB	11 KB

**Table 42.2** Article page results

The sample pages	Traffic consumed by direct access	Traffic consumed by our solution
1	37.8 KB	15.1 KB
2	30 KB	13.7 KB
3	4.3 KB	3.1 KB
4	191.7 KB	45.3 KB
5	11.4 KB	10.0 KB
6	2.6 KB	1.6 KB
7	2.8 KB	1.9 KB

We also measured the consumed time of the two methods to discover that they have little difference, and our method is slightly better.

The causes are as follows: For the user's request, only the first user's request needs to be transmitted to the target server by the proxy server to obtain corresponding returns for making the corresponding image documents on the proxy server. Image document returned by the proxy server will transfer more quickly since it has less content compared to the original pages.

## 42.6 Conclusion

This article provides the technology of extracting content from the web and displaying it on the cell phones, and it has a certain universal applicability. Because of typifying of web pages on the overall structure and the similarity of the writing among coders, our method can be simply extended to more other types of web pages. It provides a good idea on how to show web page on a phone better in the practical work. We are facing the challenge of big data in this information era; how to extract the most valuable information from numerous and disorderly data for the users is one of the most important challenges to each operator. We believe that our approach would give operators a beneficial inspiration, help every cell phone user to get better, more high-quality service.

**Acknowledgements** The article is supported by the Nankai University undergraduates innovation research fund (Grant No. BX10-258).

## References

1. Liu, B., Grossman, R., & Zhai, Y. (2003). Mining data records in web pages. In *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 601–606). New York, NY: ACM.
2. Zhai, Y., & Liu, B. (2005). Web data extraction based on partial tree alignment. In *Proceedings of the 14th International Conference on World Wide Web* (pp. 76–85). New York, NY: ACM.
3. Qiao, F. (2012). *Web information extraction technology based on templated web crawler*. Chengdu: University of Electronic Science and Technology of China (in Chinese).
4. Jonathan Hedley. (2013). *Jsoup cookbook (EB/OL)*. <http://jsoup.org/cookbook/>
5. Andrew, W. (2005). *Beginning regular expressions* (pp. 34–60). Hungry Minds.



# Chapter 43

## A New Data-Intensive Parallel Processing Framework for Spatial Data

Dong Zhao, Yang Gu, and Zhenchun Huang

**Abstract** The explosive increase of scientific data brings in the “Fourth Paradigm” research method by Jim Gray. In order to accelerate the processing speed for these big data, parallel distributed processing is needed. As the data-intensive computing requires high throughput of IO, the data transfer from different node should be cut down as much as possible. Current technologies focus more on the framework for local reliable network with homogeneous resources, but the parallel processing framework for scientific data-intensive problems such as spatial data shared with the Internet and queried by semantics is not fully studied. In this article, we proposed a new data-intensive parallel processing framework for spatial data—Robinia DSSSD (Distributed Storage and Service for Spatial Data), which provides the flexible ability to support data distribution and allocation across the Internet, and semantics query. Experiments shows that Robinia DSSSD can achieve good acceleration with low overhead, and it can well support data-intensive computing.

### 43.1 Introduction

Scientific experiments nowadays produce large volumes of data, and researchers need to analyze and deal with these big data. This is called “Fourth Paradigm” research method [1, 2]. In order to make this procedure easier to use and accelerate the processing speed, parallel processing is necessary. As scientific data is often generated by different institutes and individuals all around the world in different locations, how to handle the distributed heterogeneous data is of great importance. Previous work focuses more on the local intranet which has reliable network topology and homogeneous nodes for storing and processing. As many valuable data has been shared on the Internet, how to do the data-intensive computing with

---

D. Zhao (✉) • Y. Gu • Z. Huang

Department of Computer Science and Technology, Tsinghua University, Beijing 10084, China

e-mail: [zhaodong8701@gmail.com](mailto:zhaodong8701@gmail.com)

Internet-shared scientific data is a challenging problem. Data grid can be viewed as the first try to do the data-intensive computing over the WAN. Since the data volume is huge, how to schedule the task running on the data node is important to decrease the network transfer. In the meantime, parallel processing is needed to accelerate the speed.

Currently many distributed parallel processing frameworks and programming methodologies have been invented to solve this problem, but the differences of scientific problems make the common framework not work well with the actual use. MapReduce [3] receives great attention recent years for simplifying the programming model for distributed data processing. Along with it are the Hadoop [4] and HDFS [5] which provide the platform and distributed file storage supporting the programming model. They have good application and performance in dealing with text processing and online stream analysis for solving big-data problems with the Internet. But when referring to scientific data such as remote sensing satellite images processing, they do not fit quite well. HDFS will automatically split the files into chunks of 64 MB block and store them distributed. But this operation is done without the meaning of metadata and other attributes. The semantics query and header metadata management is also not natively supported by these systems. Furthermore, some algorithm procedures in spatial data could not be easily converted into MapReduce job style, which makes the Hadoop difficult to handle them.

Semantics query can be done by traditional relational database with SQL query. But this has limitation when dealing with parallel computing and emerging big-data technologies, especially toward fully distributed software architecture. These databases run well on a single hardware. However, when required to scale horizontally, commonly recommended scalability approaches like “sharding” (partitioning data to run on multiple machines) often cause issues within data normalization. Therefore, deployment which does not require ACID (atomicity, consistency, isolation, and duration) makes a good candidate to store data in distributed environment with good scalability. NoSQL database is promised to be a good solution for the situation that provides a mechanism for storage and retrieval of data that use looser consistency models compared with traditional relational databases [6]. They mainly offer the storage based on the simple key-value pairs, which gain benefits in scalability and performance. Column-based products include Google’s BigTable [7], HyperTable [8], and HBase [9]. HyperTable and HBase are both the open-source implementation of BigTable. The difference is that HyperTable is written in C++ and HBase is written in Java. Performance evaluation [10] shows that HyperTable usually runs twice faster than HBase. Column-based products have advantages in computing aggregation over large numbers of similar data items. Another type of NoSQL database is document-based or row-based, which stores data that share some common properties but may have different ones. As no table is declared, no rigid structure is required to store data on it. These products include MongoDB [11], CouchDB [12], and Redis [13]. MongoDB is good at dynamic queries, frequent data changes. CouchDB is better for data accumulation. Redis has advantages in rapidly changing data with a foreseeable database size. In order to

support various semantics query with good scalability, we choose MongoDB to develop. It is an open-source document database and the leading NoSQL database written in C++ which gives good performance.

In the article, we propose a new data-intensive parallel processing framework for spatial data—Robinia DSSSD. The framework uses MongoDB as the database to support metadata query. Robinia platform is a basic component for building large-scale distributed parallel processing framework. Its main features include easy deployment and being light weighted, user friendly, and compatible for old applications running which does not require source code. In short, Robinia is a powerful servlet running on Tomcat, with the ability to do the global registry, event handling, executor management, request and response processing, web UI interface, etc. We will discuss in detail how to build parallel processing framework upon this platform and evaluate its performance.

### 43.2 Distributed Data Models for Robinia

In order to better support the data-intensive distributed parallel computing, a good data model for this should be well prepared. For the scientific data such as satellite image, it often has large size of file content, together with its metadata info, including the date created and the location record. For Robinia DSSSD, the data model is represented as document-attachment type, which means each data item is made up of one document and zero or more attachments. Documents part consists of key-value pairs. Meta and key-value pairs are structural data which can be indexed and searched. Attachment is nonstructural binary data which is suitable for store files like images or other large-size file. The basic model is shown in Fig. 43.1.

Meta data can describe the metadata of data item, mainly used by querying and indexing the data. Different types of data have different metadata. Metadata can help Robinia DSSSD to query and manipulate data (split or merge) by semantics, which makes the system more useful for spatial data that has lots of additional attributes for reference. This also makes our system different from HDFS which

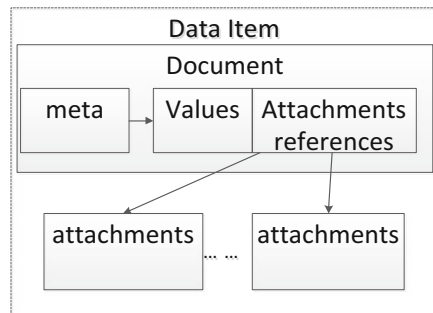


Fig. 43.1 Basic data model for Robinia

often split data by block size and lack support of query by semantics. Values serve as the main body for the data item. It contains all the fields from metadata and has its own unique ID which will be referenced as foreign key in attachments. Attachments store the binary data file with absolute path or remote URL of http or ftp. With this data model, Robinia DSSSD is capable to share and retrieve data from the Internet. In the meantime, query the required data by various semantics, which is quite useful in spatial data area.

### 43.3 System Overview and Architecture

The base component Robinia can be viewed as a powerful agent or proxy to support upper level application transparently communicates with each other, with the guaranteed services of resource discovering, status monitoring, job running, etc. Service-oriented architecture [14] and RESTful [15] technologies have been used in the system. The framework consists of two parts: data part and parallel processing part. For the data part, a storage cluster is made up of header node, backup header node, and data node. In the data node, we do not store the data as block file but use its original file item in operating system. This makes the data processing easier to get and save files. The architecture figure is shown in Fig. 43.2.

In this architecture, all the search and index operations are done by header node. Each node will be installed the supporting libraries beforehand. Header node is the central node of storage cluster and responsible for all the metadata, providing indexing and querying services. Data node is responsible for storing the real data content, including the values and attachments.

For the parallel processing framework part, master-worker style based on the data allocation is the basic idea for distributed processing. One master executor and a group of worker executors make up of this. Master processes the serial part of the algorithm, does the global schedule, and loads balance for all the workers. Workers only focus on the parallel part of the algorithm. Schedule algorithm can be performed on master node. Data layout, network latency, and node status should be taken into consideration when scheduling task together with data. In order to

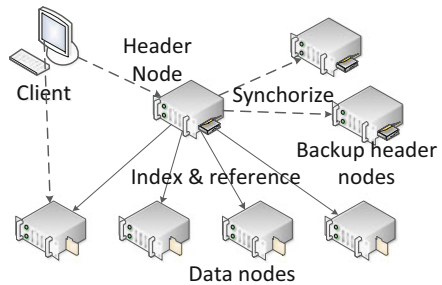


Fig. 43.2 Robinia storage architecture

achieve the minimum time cost for the whole jobs, the auto-adaptive schedule is selected. In this schedule, each processing node will retrieve the input data when it finished the previous task. This can make the system adaptive with heterogeneous node. Powerful node will process more data. Node with good data locality can do more tasks, etc.

### 43.4 Design and Implementation

In this part we will discuss the details about Robinia DSSSD design and implementation. For the data part, we have several components to do the job. The design diagram is shown in Fig. 43.3. DSSSD provides several interfaces to communicate with MongoDB for CRUD operations (create, read, update, and delete). Large binary file content is stored as attachment, which URL and absolute path is in database and binary content is stored as file in operating system. On the client side, DSSSDClient is the entry point for Java client use. Basic methods for accessing the data in Robinia platform are provided. Processing part provides corresponding framework with the specific data structure. Currently master-worker structure for parallel processing is the default structure. The worker will use the data as the input and output parameter. And the worker will try to choose the local data to reduce the network transfer time cost.

Scheduler will be defined separately to parallel select the data items for workers. When selecting the data item, several criteria should be taken into consideration: local data priority first, data transfer time (related to the data volume and network capability of node), and data processing time (related to the worker algorithm and CPU capability of node). The scheduler for jobs should follow the rules that lower network transfer priority first. The fault-tolerant is guaranteed by the status monitoring and exception catching. Failed test data will be tried until the maximum retry count is reached.

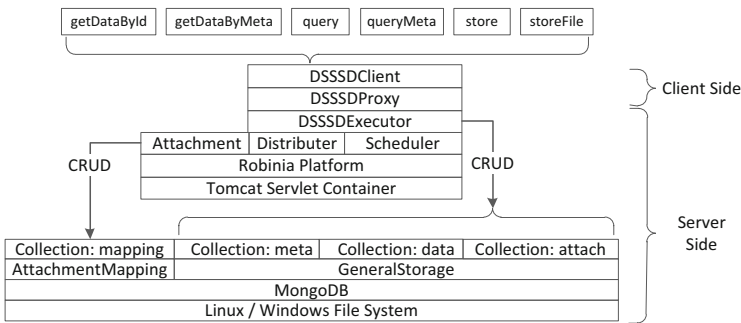


Fig. 43.3 DSSSD data design diagram



**Fig. 43.4** Screenshot of WebUI. (a) Input executor parameters, (b) view the detailed status of instance, and (c) the final result of drought detect algorithm

Figure 43.4 is the snapshot of the web UI for Robinia DSSSD. The user can select the executor to run with parameter and see the detailed status and logs on the web page. The part c is the final result of drought detection global map which we are testing in the evaluation part. The last picture is the final result of drought detection global map which we are testing in the evaluation part.

## 43.5 Evaluation and Discussion

We use the drought detect algorithm as the test case to measure the performance of Robinia DSSSD. NDWI is short for Normal Differential Water Index, which was brought up by Gao in 1996 [16]. AWI is short for Anomaly Water Index. The calculating formula is

$$\text{NDWI} = (\rho_2 - \rho_5) / (\rho_2 + \rho_5)$$

$$\text{AWI} = \text{NDWI}_i - \text{AvgNDWI}$$

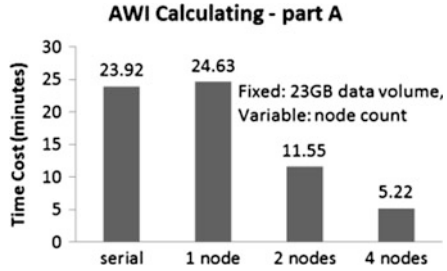
$\rho_2$  and  $\rho_5$  are the reflection rate for band 5 and band 2 in MODIS file. The main time cost of processing is the IO part of reading and writing files, and the calculating is simple as described in the above formula.

Hardware environment includes four pc nodes; each of them has Intel Core i3 @ 2.93 GHz CPU, 4 GB memory, 1 TB 7,200 rpm hard disk, and windows 7 64-bit system.

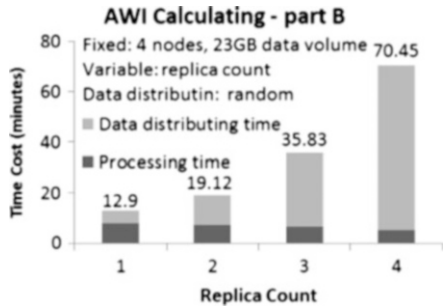
We first did the experiment with all the data can be gotten locally and compared the processing time cost with different nodes and serial part. The result is shown in Fig. 43.5.

We can see that Robinia DSSSD only put little overhead when processing with one node compared with processing the original sequential program. With the node count increasing, the processing time drops down linearly. It proves that our framework does well with the parallel processing. The auto-adaptive schedule algorithm is used by default, which means all the nodes will retrieve metadata

**Fig. 43.5** AWI calculating with different node count



**Fig. 43.6** AWI calculating with different replica count

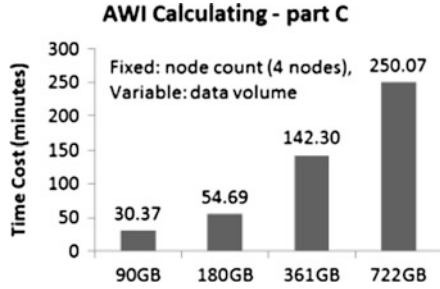


from master node and do the processing work. This schedule can make the total time cost the minimum because fast nodes will do more jobs. Another experiment was done with fixed four nodes and change replica count, which means not all the processing data be gotten locally. The original data source is from an ftp server connected by Gigabyte network. Data distribution for this is random. For example, if the replica is 1, then the whole set of data will be randomly distributed and stored on four nodes. The result shows in Fig. 43.6.

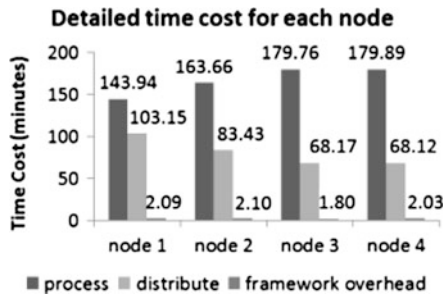
From the figure we can find the processing time differs little, while the data distributing time changes a lot. The reason is when replica is 1, data is almost distributed with each node has 1/4 part of whole set. When processing data, the probability of getting data locally is pretty good. Another experiment was done with fixed four nodes and let the input data volume change. The original data source is from an ftp server connected by Gigabyte network. The result is shown in Fig. 43.7.

From Fig. 43.7, we can find that the system has good scalability when data volume increases. This proves that Robinia DSSSD is capable for the TB level data-intensive computing. Figure 43.8 is the detailed time cost for each node of part C. We can see the auto-adaptive schedule algorithm has good load balance for the overall workload of nodes. And the overhead from Robinia DSSSD is small. Distribution time is mainly limited by the hard disk IO which is currently lower than the Gigabit Ethernet. Figure 43.9 shows the speed ratio compared with different data volume. For the small data volume 23 GB, the speed ratio is 4.58, faster than the linear speed ratio. The reason for this could be related with the disk

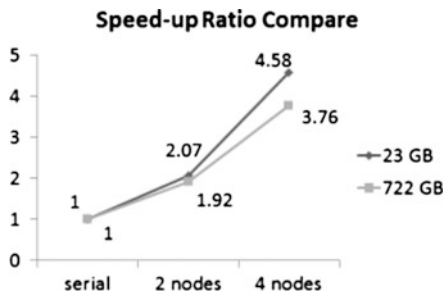
**Fig. 43.7** AWI calculating, execution time vs. data volume



**Fig. 43.8** AWI calculating with data volume



**Fig. 43.9** Speedup Ratio Compare



cache, memory limit, etc. For the large data volume 722 GB, the speed ratio is 3.76. The parallel efficiency is  $3.76 \times 361 / 722 \times 100\% = 93.75\%$ , showing Robinia DSSSD does well in parallel processing.

### 43.6 Conclusion

Data-intensive computing plays more and more important role in recent industry and academic areas. In order to better deal with big data, we propose a new data-intensive parallel processing framework—Robinia DSSSD, which is fit for the data-intensive computing over the Internet and supports semantic query well.



System overview and design implementation have been discussed in this article. Experiments show Robinia DSSSD has good performance and flexibility. It can support the data-intensive computing in TB level data. Parallel efficiency can reach 93.75 % with large data volume.

Robinia DSSSD has just implemented the basic functions for supporting data-intensive parallel processing. Lots of work can be done in the future, including the global computing over wide area network, which can take full use of all computing capability of nodes, and algorithm migration, which can better support the task running on data nodes. Security enhancement is to be refined to enhance the system authentication mechanism and resource access control.

## References

1. Hey, T., Tansley, S. and Tolle, K. (2010). The fourth paradigm of science research – a brief introduction to Jim Gray on eScience: A transformed scientific method. *e-Science Technology and Application*, *VI*(2), 92–94.
2. Hey, T., Tansley, S., & Tolle, K. (2009). *The fourth paradigm: Data-intensive scientific discovery* (pp. xvii–xxxi). Redmond, WA: Microsoft Research.
3. Dean, J., & Ghemawat, S. (2008). MapReduce: Simplified data processing on large clusters. *Communications of the ACM*, *51*(1), 107–113.
4. Hadoop. <http://hadoop.apache.org/>
5. Shvachko, K., Kuang, H., Radia, S. & Chansler, R. (2010). The hadoop distributed file system. In *2010 I.E. 26th Symposium on Mass Storage Systems and Technologies* (pp. 1–10). Piscataway, NJ: IEEE.
6. Han, J., Haihong, E., Le, G. & Du, J. (2011). Survey on NoSQL database. In *2011 6th International Conference on Pervasive Computing and Applications (ICPCA)* (pp. 363–366). Piscataway, NJ: IEEE.
7. Chang, F., Dean, J., Ghemawat, S., Hsieh, W. C., Wallach, D. A., Burrows, M., & Gruber, R. E. (2008). Bigtable: A distributed storage system for structured data. *ACM Transactions on Computer Systems (TOCS)*, *26*(2), 4.
8. HyperTable. <http://hypertable.org/>
9. HBase. <http://hbase.apache.org/>
10. Bunch, C., Chohan, N., Krintz, C., Chohan, J., Kupferman, J., Lakhina, P., & Nomura, Y. (2010). An evaluation of distributed datastores using the AppScale cloud platform. In *2010 I.E. 3rd International Conference on Cloud Computing (CLOUD)* (pp. 305–312). Piscataway, NJ: IEEE.
11. MongoDB. <http://www.mongodb.org/>
12. CouchDB. <http://couchdb.apache.org/>
13. Redis. <http://redis.io/>
14. Perrey, R., & Lycett, M. (2003). Service-oriented architecture. In *2003 Symposium on Applications and the Internet Workshops* (pp. 116–119). Piscataway, NJ: IEEE.
15. Richardson, L., & Ruby, S. (2008). *RESTful web services* (pp. 49–79). Sebastopol, CA: O’Reilly Media.
16. Gao, B. C. (1996). NDWI-a normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, *58*(3), 257–266.

# Chapter 44

## The Approach of Graphical User Interface Testing Guided by Bayesian Model

Zhifang Yang, Zhongxing Yu, and Chenggang Bai

**Abstract** GUI (graphical user interface) is becoming increasingly important in the software field for the reason that it is a friendly way for the users to interact with the software through GUI. Testing in GUI, however, is faced with many challenges, due to the immense number of event interactions. Testing all possible event interactions is impossible, since the number of required test case is huge in numbers. GUI testing mainly serves two goals: First, to establish confidence in assessment of GUI; Second, to find that more software defects in GUI testing while limiting the number of test cases. For this purpose, any testing method must be better at detecting defects. This article proposed a new technique that can be used for GUI testing, which can guide the GUI testing and find more defects as soon as possible. In this chapter, it introduces an approach of GUI testing guided by Bayesian model optimization scheme, discusses the Bayesian model topology and its issues encountered in the modeling process. It presents solutions in connection with the parameters problem. In the end, a simple case verifies the validity of the model during the GUI testing.

### 44.1 Introduction

Graphical user interface (GUI) is becoming increasingly important in the software field, while GUI testing problem has never been resolved absolutely. GUI testing is becoming the key issues restricting GUI rapid development.

Many researchers have already done a lot of meaningful work on GUI. White et al. [1, 2] and Belli [3] pointed that various responsibilities of the user can be specified as a complete interaction sequence (CIS) between the user and the GUI application under testing. Then they proposed a concept of finite-state machine for

---

Z. Yang (✉) • Z. Yu • C. Bai  
School of Automation Science and Electrical Engineering, Beihang University,  
Beijing 100191, China  
e-mail: [qwyzf@163.com](mailto:qwyzf@163.com)

CIS. A.T. Memon et al. [4] proposed that the GUI under testing is a hierarchical structure. They proposed that the test case generation problem of GUI testing is an AI planning problem. Then, they proposed several approaches guiding the GUI testing, such as usage profiles [5], dynamic adaptive automated test generation [6], and incorporating event context [7]. In the work of H. Hu, K.Y. Cai, etc. [8, 9], in order to shorten the testing time and reduce cost, a mechanism is introduced to handle multiple testing tasks on the same or different software under tests (SUTs) simultaneously. On the other hand, advanced testing techniques, such as Adaptive Testing (AT) have been proposed to improve the efficiency of traditional random/partition testing. Then, they proposed an extended adaptive testing strategy, namely Modified Adaptive Testing (MAT). The use of test history information allows the resulting test process to be adaptive in the selection of tests under a limited test budget.

In the above-mentioned work, one issue that find more software defects is not discussed absolutely in GUI testing, while not mention some very important questions, such as how can find more and important defects under the same test conditions, how to take advantage of the first-line testers testing experience to guide the testing process, and how to assess the GUI when the test is not sufficient. In this chapter, we introduced a Bayesian of model which can use test data on line and experiences of tester, guide the test process of GUI testing.

Bayesian model theory has been used to solve the complex problems of many complex scientific issues [10–12], especially the considerable uncertainties. Due to the complexity of GUI testing, we try to establish a Bayesian Model of GUI testing, which is characterized by historical data from tests, combined with the experience of the testers, established a Bayes network for analyzing of the GUI test results, and continue to guiding the further testing of GUI. The reason we use Bayesian methods rather than the other mathematical model is as follows:

First, Bayesian approach uses empirical knowledge of the testers as a priori parameter, while the tester of GUI is able to rely on their experience in finding defects faster.

Second, Bayesian approach has learning function can be given a wealth of information based on prior knowledge of the posterior distribution, which can be adopted to guide the testing process or establish confidence assessment of GUI.

Third, compared with traditional neural network method, the network structure of the Bayesian approach is more flexible, it is more important for GUI testing while GUI intricate interaction of the testing process.

The last point is the most important, Bayesian methods show higher sensitivity on a small sample of data processing than the other models. This is extremely useful for GUI testing according that sample nature and not fully tested for GUI testing.

Since Bayesian networks have advantages above with GUI testing, we introduced a Bayesian model to GUI testing. A simple case proved that GUI testing with Bayesian model can be found fault faster in GUI testing than traditional methods.

## 44.2 Background

A GUI is a graphical user interface to a program which is composed by certain number of windows and GUI components. The basic input is an event in GUI, which normally generated by action of the user, external environment of GUI or itself, it sends to GUI by the operating system, GUI respond to these events in order to accomplish a specific function [13].

Due to the various components of the GUI within a certain hierarchical relationships, it asks that a certain event can only accept after other events are applied to GUI. A.T. Memon proposed that Event-Flow Graph (*EFG*) instead of all the events interactions in the GUI, which defined as follows:

$$\langle V, E, B \rangle \quad (44.1)$$

In this model,  $V$  is the set of all vertices in the EFG, each vertex instead of one event;  $E$  is the set of edges which means event interaction;  $B$  is a set of initial event, which means that it can be immediately executed after GUI start.

When an event is input, all possible response to the event code will be named Event Handler (*EH*) and Event Handler Interaction Graph (*HIG*) is defined as follows [14]:

$$\langle H, RHI \rangle \quad (44.2)$$

In this model,  $H$  is the set of all vertices in the EFG, each vertex instead of one Event Handler;  $RHI$  is the set of edges which means Event Handler interaction.

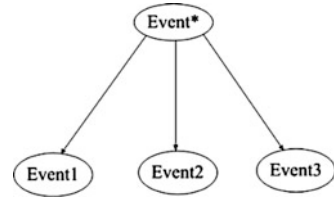
The test case generation is the most important and difficult task of the GUI testing. Due to the presence of a response to interaction, independent testing of each event cannot meet the GUI software testing requirements. Therefore, the sequence of events in the GUI test needs to be used as a test input. The sequence of events must be legitimate and efficient. In this chapter, we will utilize EFG for path searching and HIG interaction as the coverage criteria, get the test case of GUI.

## 44.3 The Topology of Bayesian Model

In this chapter, supposing that the software contains no faults before three event interactions, we considered the three events, Event1, Event2, and Event3, which stay in the same state, and the Event\*, which take place before three events above. To get a Bayesian model of GUI, we draw a graph to represent the qualitative relationships between the nodes. We then quantify all of the probabilities of the events in the graph as follows:

1. We determine the probability of occurrence of each event represented by a root node.

**Fig. 44.1** A Bayesian model of GUI



2. For each child node, we determine the probability of occurrence of the event. These specifications determine the probabilities for all combinations of events represented on the graph.
3. Because the full joint probability distribution has been specified, each time we observe an event, we may update all of the probabilities for all of the remaining events in the model by probabilistic conditioning. Further, because of the structure of the graphical model, it is straightforward to carry out such belief updating.

Our model is given in Fig. 44.1. There are four independent nodes, Event\*, Event1, Event2, and Event3. First, assuming that there was not fault before Event interaction, so every nodes of the parent of Event interaction set N have two states: 0 (this event was not be carried out) and 1 (this event was carried out), and the last event, Event1, Event2, and Event3 also have two states: 0 (this event was be executed and failed) and 1 (this event was executed correctly).

Second, we should specify probabilities for the root nodes and the conditional probability of each child node and check the rationality.

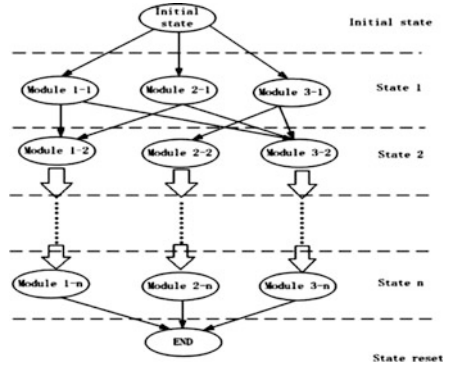
To complete the Bayesian model, we need to link test outcomes to the appropriate node, not only gaining the outcomes testers concerned but also both topology and state of nodes are not too sophisticated. There are three essential simplifications as follows:

First, we do not need to describe the test nodes so extensively. All that is important is that there is a test process which is connected to a single node, and the nodes representing the event that cannot be clicked will be removed in a particular state.

Second, any node, have not just two states, while considering all the states is likely to be too complicated for analysis and gain the goal of the designer, furthermore some states is not my concerned.

The last and most important, there were so many nodes have same or similar characteristics, taking into account the function of GUI complexity and the time constraints, a simple Module Bayesian Model of GUI testing showing in Fig. 44.2; this model is also suitable for length of GUI test case is n. Supposing that the GUI in case has three modules, and that three are one of the three events that belong to the different module can be executed in every state, Module<sub>m-n</sub> means that an event belonged to module m stay in state n. The first node Initial state has two states 0 and 1 (0 represents that GUI initially failed and 1 represents that GUI initial successfully). Then the event nodes stayed in state 1 to state n have two states 0 and

**Fig. 44.2** Module Bayesian Model of GUI testing



1 (0 means that an event of this module was not executed in the state and 1 means that an event of this was executed in the state).The last node END also has 2 states 0 and 1 (0 is in the name of that GUI resets state with fault in last event interaction recording this failure and 1 is in the name of that GUI resets state successfully).

### 44.4 Mathematical Analysis of the Bayesian Model

A basic Bayesian model contains both topology and parameter analysis; the topology has been briefly discussed above and then mathematical analysis of the Bayesian will be mentioned showing in Fig. 44.3.

First, we can give an empirical conditional probability distribution according to the testers, the test manager, or user feedback experience. In order to obtain an objective prior distribution, we define  $N_{m-n}$  is the number of Event handler which  $Module_{m-n}$  contains.

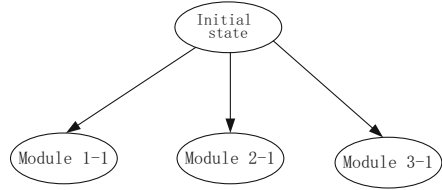
Second,  $I$  and  $M_{m-n}$  are two-dimensional discrete random variables,  $I = 0$  represents that GUI initial failed and  $I = 1$  represents that GUI initial successfully;  $M_{m-n} = 0$  means that an event of  $Module_{m-n}$  was not executed in the state and  $M_{m-n} = 1$  means that an event of  $Module_{m-n}$  was executed in the state.

Third, we consider the probability of occurrence of event  $\{M_{m-n} = m_i\}$  under the conditions of occurrence of event  $\{I = i_j\}$ , that is to find the probability of event as follows:

$$\{ M_{m-n} = m_i | I = i_j \}, i, j = 0, 1 \tag{44.3}$$

According to the conditional probability formula, it can be deduced as follows:

**Fig. 44.3** A simple case of conditional probability distribution



$$P\{M_{m-n} = m_i | I = i_j\} = \frac{P\{M_{m-n} = m_i, I = i_j\}}{P\{I = i_j\}} \tag{44.4}$$

Fourth, the experience of the conditions given in the probability distribution will be according to the formula (44.5).

$$\begin{aligned}
 P\{M_{m-n} = m_i | I = i_j\} &= \frac{P\{M_{m-n} = m_i, I = i_j\}}{P\{I = i_j\}} = \frac{N_{m-i}}{\sum_{i=1}^n N_{m-i}}, i, j \\
 &= 0, 1 \tag{44.5}
 \end{aligned}$$

In actual testing process, we define  $N_{m-n}^*$  is the number of Event handler which Module<sub>m-n</sub> contains in actual test procedure, so the updated of the conditions given in the probability distribution will be according to the formula (44.6).

$$\begin{aligned}
 P\{M_{m-n} = m_i | I = i_j\} &= \frac{P\{M_{m-n} = m_i, I = i_j\}}{P\{I = i_j\}} \\
 &= \frac{N_{m-i} + N_{m-i}^*}{\sum_{i=1}^n (N_{m-i} + N_{m-i}^*)}, i, j = 0, 1 \tag{44.6}
 \end{aligned}$$

## 44.5 Experiment Studying

### 44.5.1 Experimental Designs

In this chapter, we use open source software TerpPaint belonged to TerpOffice series, which developed by Professor A.T. Memon and his students. It has 3,922 test cases and contain 45 faults. The test case is divided into three groups on the nature of GUI. Experimental method included Randomly selected, Dynamic feedback, and Bayesian method. Bayesian model testing techniques algorithm as follows:

**Step 1** The original test suite is divided into m subdomain  $\{C_1, C_2, \dots, C_m\}$ , which contains  $k_1, k_2, \dots, k_m$  test cases;

**Table 44.1** The typical posterior probability calculation of the Bayesian approach

Node	Module <sub>1-2</sub>	Module <sub>2-2</sub>	Module <sub>3-2</sub>
Without node	0.00274	0.00639	0.01777
Module <sub>1-1</sub>	0.00260	0.00649	0.01712
Module <sub>2-1</sub>	0.00270	0.00636	0.01777
Module <sub>3-1</sub>	0.00273	0.00635	0.01752

**Step 2** Initialization parameter  $\varepsilon, 0 < \varepsilon < 1$ ;

**Step 3** According to the profile  $\{P_1, P_2, \dots, P_m\}$  to select a number of test cases, the probability of test cases being selected in the subdomain  $C_i$  is  $P_i$ ;

**Step 4** Perform steps above and record the results of testing, the number of defects in each class is cumulative;

**Step 5** If no defects are found back to step 3;

**Step 6** If the test case defects are found, recording the defects discovered what kind of test cases, adjustment of the profile parameter guided by the posterior probability of the Bayesian network;

**Step 7** Verify that it meets the test termination condition, if not, processing returns to step 3, test case input domain partition remains unchanged during the test;

**Step 8** Meet the test termination conditions, terminate the test.

Randomly selected and Dynamic feedback method is similar to the above algorithm, while Randomly selected does not contain the feedback process and Dynamic feedback does not contain the Bayesian model process.

### 44.5.2 Markov Property of the Bayesian Approach

In this chapter, the method based on Bayesian model relies on the posterior probability calculation to redefine the test path profile. During the experiment, the posterior probability showed obvious relationship and the last node, but the situation is not obvious and more upper node relations. The typical posterior probability calculation was shown in the following Table 44.1.

In this table, we can see that in our Bayesian model showing in Fig. 44.4. The posterior probability of the defect rate closely related to the third layer of nodes, and the second layer is also associated, however it is not obvious. This is evidenced by the Markov property of Bayesian networks, while it is more important issue that the posterior probability of the defect rate closely related to both the second and third layers in future work.

### 44.5.3 The Experimental Results

In this chapter, we used three experimental methods based on the same GUI and test cases. The experimental results are shown in Fig. 44.5.



Fig. 44.4 A real Bayesian model of GUI testing

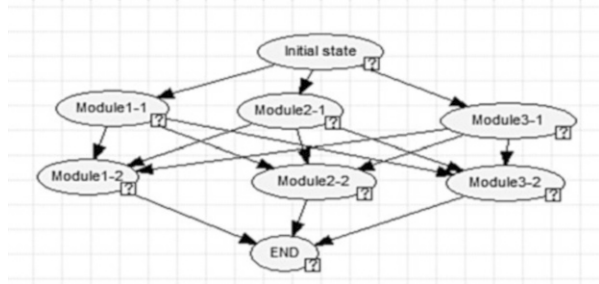
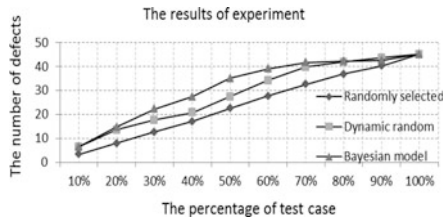


Fig. 44.5 Experimental results of three methods



In Fig. 44.5, the horizontal axis represents the percentage of test cases, and the vertical axis represents found that the number of defects. We can see that Dynamic random testing techniques and Bayesian model will find more defects as soon as possible, while Bayesian model has the best performance in the medium term.

### 44.6 Conclusion

In summary, this chapter describes an effective method for GUI testing guided by Bayesian model, discussed the Bayesian model topology and its issues encountered in the modeling process. It presents two solutions in connection with the parameters problem. Facing the problem of missing data, we use Bayesian inference method to fill all the data. During the experiment, we dynamically updated Bayesian network and calculated the posterior probability to change the test path profile. In the end, a simple case verifies the validity of the model during the GUI test. Coefficient with vector hydrophone, which measures the pressure and particle velocity instantaneously, it costs less time, more accurate in measuring, handier in practice, and had good performance both in vertical and oblique incidence.

## References

1. White, L., & Almezen, H. (2000). Generating test cases for GUI responsibilities using complete interaction sequences. In *Proceedings of the 11th International Symposium on Software Reliability Engineering* (pp. 110–121). San Jose, CA: IEEE Computer Society Press.
2. White, L., Almezen, H., & Alzeidi, N. (2001). User-based testing of GUI sequences and their interactions. In *Proceedings of the 12th International Symposium on Software Reliability Engineering* (pp. 54–63). Hong Kong: IEEE Computer Society Press.
3. Belli, F. (2001). Finite state testing and analysis of graphical user interfaces. In *Proceedings of the 12th International Symposium on Software Reliability Engineering* (pp. 34–43). Hong Kong: IEEE Computer Society.
4. Memon, A. M., Pollack, M. E., & Soffa, M. L. (2001). Hierarchical GUI test case generation using automated planning. *IEEE Transactions on Software Engineering*, 27(2), 144–155.
5. Brooks, P. A., & Memon, A. M. (2007). Automated GUI testing guided by usage profiles. In *Proceedings of the Twenty-Second IEEE/ACM International Conference on Automated Software Engineering* (pp. 333–342). New York: ACM.
6. Yuan, X., Cohen, M. B., & Memon, A. M. (2009). Towards dynamic adaptive automated test generation for graphical user interfaces. In *IEEE International Conference on Software Testing, Verification, and Validation Workshops* (pp. 263–266). Washington, DC: IEEE Computer Society Press.
7. Yuan, X., Cohen, M. B., & Memon, A. M. (2011). GUI interaction testing: Incorporating event context. *IEEE Transactions on Software Engineering*, 37(4), 559–574.
8. Hu, H., Jiang, C.H., Ye, F., Cai, K.Y., Huang, D.Z., & Stephen, S.Y. (2010). A parallel implementation strategy of adaptive testing. In *Computer Software and Applications Conference Workshops* (pp. 214–219). Seoul: IEEE Computer Society Press.
9. Hu, H., Jiang, C.H., Cai, K.Y., Kai-Yuan Cai, Wong, W.E., & Mathur, A.P. (2013). Enhancing software reliability estimates using modified adaptive testing. *Special Section: Component-Based Software Engineering (CBSE)*, 55(2), 288–300.
10. JUDEA PEARL. (2000). *Causality: Models, reasoning, and inference* (pp. 1–6). New York, NY: Cambridge University Press.
11. Bai, C. G., Jiang, C. H., & Cai, K. Y. (2010). A reliability improvement predictive approach to software testing with bayesian method. In *Proceedings of the 29th Chinese Control Conference* (pp. 6031–6036). Beijing: IEEE Computer Society Press.
12. Wooff, D.A., Goldstein, M., & Coolen, F. P. A. (2002). Bayesian graphical models for software testing. *IEEE Transactions on Software Engineering*, 28(5), 510–525.
13. Memon, A. T. (2001). *A comprehensive framework for testing graphical user interfaces* [D]. USA: University of Pittsburgh
14. Zhao, L. (2010). *GUI software testing based on event handlers* [D]. Beijing: Beihang University (In Chinese)

# Chapter 45

## A Model for Reverse Logistics with Collection Sites Based on Heuristic Algorithm

Xiaoqing Geng and Yu Wang

**Abstract** Reverse distribution has received growing attention throughout this decade. Built on the concept of green supply chain management (GSCM), this chapter presents a mathematical programming and distribution model for reverse logistics with collection sites. Due to the complexity of the GSCM model, a heuristic solution is given and improved. The solution adds a heuristic concentration procedure, where subproblems with reduced sets of decision variables are iteratively solved to improve the optimality, and it improves the capacitated plant location problem (CPLP). Computational test demonstrates that high-quality solution is obtained with the improved model.

### 45.1 Introduction

With the growing environmental concerns, green supply chain management (GSCM) is getting more interest among practitioners of operations and researchers in the fields of supply chain management. GSCM can be defined as “Integrating environmental thinking into supply chain management including product innovation and design, material sourcing and selection, manufacturing processes, delivery of the final product to the consumers as well as end-of-life (EOL) management of the product after its useful life” [1]. Environmental management devotes to reduce waste volume by moving away from one-time use and disposal to having management of the product’s recovery primarily. It includes reuse, remanufacturing, and materials recycling in a product life cycle. Thus, manufacturers are forced to consider the environment impact by their products. They have to extend and

---

X. Geng (✉)

Department of Management Information System, Tianjin University of Finance  
and Economics, Tianjin 300222, China  
e-mail: [gengxq@gmail.com](mailto:gengxq@gmail.com)

Y. Wang

School of Humanities, Tianjin University of Finance and Economics, Tianjin 300222, China

improve the traditional forward logistic distribution chain and consider the environmental effects of all products including returning process. It is a process of goods moving from final destination to another one, for the purpose of value-added or for the proper disposal of the products.

Caruso and Colorni proposed a location-allocation model for planning and designing urban solid waste management systems [2]. The results were the location and number of waste disposal plants, the amount of waste processed, specifying the technology adopted, and the service basin of each plant. Bloemhof-Ruwaard focused on the problem of by-product flows and coordinating product in a two-level distribution network. The model solution suggested the first attempt at studying the coordinated control of product and by-product flows within distribution networks [3]. Carter and Ellram noticed that the international nature and complex environment made it almost impossible to put forward decisions based on intuition. They suggested some important and critical factors in the reverse logistics process [4]. Srivastava reviewed the literature on reverse logistics and addressed the recognition of the strategic importance of reverse logistics as evident from classification and categorization of the existing GSCM literature [1]. Vaidyanathan designed a model and gave a solution procedure for reverse distribution network with GSCM [5].

Our work is to develop a mathematical programming and distribution model for one version of this problem in GSCM—reverse logistics with collection sites, which is based on a heuristic solution methodology. And the test shows that high-quality solution is achieved while improving model and expending modest computational effort.

## 45.2 Assumption

Firstly, some assumptions are given below in order to formulate a mathematical model for the reverse distribution problem [6]:

1. The recalled products are located at some source outlets (origination sites). The model in this chapter is to find an efficient method and process to return these products from a set of origination sites to specific collection sites.
2. The wholesaler or retailer is considered to be an initial site. This may be a realistic assumption because the client would be inclined to return the product to the closest collection site to get a refund or to purchase some other products from there.
3. There is a fixed cost to opening refurbishing sites and collection sites. There is a limit to the number of refurbishing sites and collection sites that can be opened, but the choice of which refurbishing sites and which collection sites to be opened should be decided by the proposed model.

In order to describe the model further, the following notation will be used in the model:

I—{i/here i is some an origination site}. This is a shop, a store, a retail outlet, or a client collection site.

J—{j/here j is some a collection site}. Collection sites are some intermediate transshipment spots. A collection site receives the recalled products.

K—{k/here k is some a refurbishing facility site}. This site may be a refurbishing spot, or a recycling plant, or the original manufacturing station.

$C_{ijk}$ —Total variable cost of transporting a single unit of recalled product from origination site i through collection site j and onto refurbishing site k.

$F_j$ —Total cost of opening a collection site j.

$G_k$ —Total cost of opening a refurbishing site k.

$a_i$ —Number of recalled products residing at origination site i.

$B_j$ —Maximum capacity of collection site j.

$D_k$ —Maximum capacity of refurbishing site k.

$P_{min}$ —Minimum number of collection sites opening and operation.

$P_{max}$ —Maximum number of collection sites opening and operation.

$Q_{min}$ —Minimum number of refurbishing sites operation.

$Q_{max}$ —Maximum number of refurbishing sites operation.

All the decision variables in this model are:

$X_{ijk}$ —Fractions of units at origination site i that is from origination site i transported through collection site j and onto refurbishing site k.

$$P_j = \begin{cases} 1 & \text{if collection site } j \text{ is open,} \\ 0 & \text{otherwise.} \end{cases} \quad Q_k = \begin{cases} 1 & \text{if refurbishing facility } k \text{ is open,} \\ 0 & \text{otherwise.} \end{cases}$$

### 45.3 Model Designing and Construction

With the assumptions, a strong formulation with collection sites can be described as:

$$\text{Min}Z = \sum_i \sum_j \sum_k C_{ijk} a_i X_{ijk} + \sum_j F_j P_j + \sum_k G_k Q_k$$

Subject to:

$$\sum_j \sum_k X_{ijk} = 1 \text{ for all } i, \tag{45.1}$$

$$\sum_i \sum_k a_i X_{ijk} \leq B_j \text{ for all } j, \tag{45.2}$$

$$\sum_i \sum_j a_i X_{ijk} \leq D_k \text{ for all } k, \tag{45.3}$$

$$X_{ijk} \leq P_j \text{ for all } i, j, k, \tag{45.4}$$

$$X_{ijk} \leq Q_k \text{ for all } i, j, k. \tag{45.5}$$

Here the objective function aims to minimize the sum of costs to transfer products from origination sites through collection sites to the destination stations and the fixed cost of opening the collection sites and destination sites.

Constraint set analysis:

- (a) All the recalled products are transported from the origination sites to destination sites by way of constraint set (45.1) (either directly or via collection sites in the reverse logistics network);
- (b) Constraint set (45.2) limits the units sent through collection site  $j$  to site  $j$ ;
- (c) Constraint set (45.3) limits the units ending up at destination site  $k$  to site  $k$ ;
- (d) Constraint set (45.4) forbids units through collection site  $j$  from being routed unless this site is opened;
- (e) Constraint set (45.5) forbids units from ending up at destination site  $k$  from being routed unless this site is opened.

Compare the strong formulation of this problem, a weak formulation can be given from aggregating the demand in constraint sets (45.4) and (45.5), obtaining the following alternative constraint sets:

$$\sum_j \sum_k X_{ijk} \leq |I|^* P_j \quad \text{for all } j, \quad (45.4')$$

$$\sum_i \sum_j X_{ijk} \leq |I|^* Q_k \quad \text{for all } k. \quad (45.5')$$

With constraint sets (45.4') and (45.5'), the weak formulation will significantly reduce computation burden with finding proper and optimal solutions.

## 45.4 Solution Method

### 45.4.1 Observations

This model is a typical zero-one mixed integer-linear programming (MIP) problem (the variables are restricted to be either 0 or 1). Assuming that there are no collection sites in this model, it will reduce to a capacitated plant location problem (CPLP). The CPLP is NP-complete problem (solutions can be verified in polynomial time), and as such, the model is also NP-hard [7]. The traditional MIP tools for solving such problem as NP-hard is limited. We should choose proper process and improve solution.

### 45.4.2 Heuristics Procedure

Since integer-linear programming is NP-complete, many problem instances are difficult and so heuristic methods must be used instead. Based on running an efficient heuristic numerous times, heuristic concentration identifies a subset of the potential collection sites that may warrant further investigation [8]. After repeatedly running a randomized greedy heuristic, the information about which sites are in the “p” set is collected, and a final subset of sites is used when solving the subset problem to optimality using integer-linear programming. Thus, using a subset of the potential collection sites results in an optimal solution if all sites that are in the optimal “p” set are in the subset.

MIP tools provide some useful module for this subproblem. In this chapter we can use AMPL (A Mathematical Programming Language) module as a front-end interface to the mathematical programming solver-CPLEX. The algorithm for this problem includes three steps: the first step is the random selection for potential collection sites and refurbishing sites, the second step is the heuristic concentration section, and the third step is the heuristic expansion (HE).

*Random selection:*

At the beginning, set  $\text{MAXIterations} = \beta$  (where  $\beta = 10, 40, \text{ or } 80$ )

1. While  $\text{Iterations} < \text{MAXIterations}$ , then do:
2. Select a subset of size  $P_{max}$  from the collection sites, and meanwhile  $Q_{max}$  from the destination sites randomly (all sites should have an equal probability to be selected).
3. Add the AMPL module file and data file in the way that only the sites selected in step 2 are considered as potential collection sites.
4. Optimize the current problem using AMPL solver.
5. Save the solution and its configuration. If the current solution is better than the best previously exist solution, then update the best solution.
6. End while.

*Heuristic concentration:*

7. Add the AMPL module file with all collection sites and destination sites chosen in the best previously solution. Collecting the information in steps 2–5 from the top 5 % best solutions, use the additional  $[P_{max} + 2$ —the number of sites used in the first best solution] most frequently collection sites, and  $[Q_{max} + 2$ —the number of sites used in the first best solution] most frequently destination sites, and change the AMPL module file.
8. Optimize the current problem using AMPL solver.
9. If the current solution is better than the best previously exist solution, then update the solution and its configuration as the best solution.
10. Report the best solution found.

*Heuristic expansion:*

11. Add the AMPL module file with all collection sites and destination sites in the best solution found. Then add one site not chosen in the best solution found.
12. Optimize the current problem using AMPL solver.
13. If the current solution is better than the best previously solution, then remember the solution and its configuration, but not change the best solution for now.
14. Repeat steps 11–13; make sure all sites have been checked.
15. If improvements are found, save this solution as the latest best solution.
16. Repeat from step 11; make sure no improvements are found.
17. Report the latest best solution.

**45.4.3 A Deterministic Heuristic**

An alternative deterministic heuristic to be proposed is designed to improve the random selection and heuristic concentration process of the previous algorithm. Compare an exact algorithm, a deterministic heuristic designing may lead a serious problem here: if we use CPLEX to select the sites and find an optimal result of demand flows (given the open sites chosen from the available sites sheet), we do not make sure whether a heuristic that would show us meaningful routing values short of using an optimal solver.

Based on the truth above, we developed procedure CC algorithm (a greedy algorithm) that uses CPLEX to define the routing planning schemes (as similar with the random selection with heuristic concentration methods we mentioned above):

1. Rank order all collection sites (all P's found) and rank order all refurbishment sites (all Q's) by the ratio of Cost = Capacity from least to greatest.
2. Select the  $P_{max} + 4$  and  $Q_{max} + 4$  cheapest Cost = Capacity site alternatives and solve with CPLEX solver.

**45.5 Computational Results**

Vaidyanathan collected some real data through research in some reverse distribution networks. This chapter chooses 5 sets of 20 randomly generated problems from reference [5] as datasets for test problem (Table 45.1):



**Table 45.1** Datasets

Problem set	Number of origination (or retail) sites	Number of available collection sites	Maximum allowed number of collection sites	Number of available refurbishment sites	Maximum allowed number of refurbishment sites
1	35	16	5	15	3
2	45	26	7	18	4
3	55	32	7	20	5
4	75	37	8	22	5
5	100	41	10	37	8

### 45.5.1 Data Generation

From origination and collection to refurbishing site are all located in a  $100 \times 100$  square randomly. And an additional collection site was used as a virtual site to figure a direct shipment from an origination site to the destination site. This additional destination site with boundless costs and boundless capacity is used to eliminate infeasible basic solutions. The data generation includes constructing two parts costs (fixed and variable) and capacities. The costs of fixed part are generated as following:

Collection sites (CS):  $C_j := 0.1([0, 10000] + B_j [0, 10])$ ,

Refurbishing sites (RS):  $R_k := 0.1([0, 10000] + D_k [0, 10])$ .

Costs in the transportation process are computed according to the following formulae:

$$C_{ijk} := \alpha(\text{Euclidian distance from } i, j \text{ to } k)$$

Here  $\alpha$  is 0.1 when using one collection site ( $i, j$  to  $k$ ) and is 0.4 when shipping direct from the origination site to one refurbishing site ( $i, m$  to  $k$ , where the Euclidian distance is computed in a direct distance from  $i$  to  $k$ ).

The demand is generated as:  $a_i = [0, 500]$ , and the capacities are:

Collection sites (CS):  $B_j = [0, 5000]$ ,

Refurbishing sites (RS):  $D_k = [0, 25000]$ .

### 45.5.2 Experimental Results

With test problem, the experimental results show that the heuristic expansion (HE) algorithm improves significantly in the speed and solution values. Applying HE to the results of Procedure CC algorithm, 80 random selection iterations with heuristic concentration (HC-80), 40 random selection iterations with heuristic concentration (HC-40) and 10 random selection iterations with heuristic concentration (HC-10) for these problem sets clearly demonstrate the additional benefit of the HE technique. Both Procedure CC and random selection with heuristic concentration provide starting solutions that the HE algorithm.

The tentative in this method is that this starting point solution using HE was stuck in a very poor local minimum. Even though the HE algorithm cannot always substantially speed every solution, Computational test shows that optimal solution was found in a large number of cases and high-quality solution is obtained while improved model and expending modest computational effort.

## 45.6 Conclusion

In the discussion of this chapter, we added GSCM into the process of reverses logistics with collection sites. Besides that, this chapter may contribute to the reverse distribution literature by developing a strong and a weak formulation for reverse distribution logistical problems including all kinds of reverse type (product recall, product reuse, product disposal, and hazardous product return). Also, this chapter adapted the heuristic concentration procedure to solve the version of problem with collection sites and provided a new solution—heuristic expansion (HE). The complexity of the proposed model was such that a heuristic solution procedure was the only viable approach to solve very severe problems. We proposed heuristic concentration procedures combined with HE to make the problem easily. Very large problems can indeed be solved in a reasonable amount of time with the HE algorithm, where they cannot be solved with traditional MIP tools within a limited computational time.

**Acknowledgments** This work was financially supported by the foundation of Tianjin Municipal Education Commission (20122129) and Tianjin Municipal Science and Technology Program (12ZLZLZF01300).

## References

1. Srivastava, S. K. (2007). Green supply chain management: A state-of-the-art literature review. *International Journal of Management Reviews*, 9(1), 53–80.
2. Caruso, C., Colomi, A., & Paruccini, M. (2003). The regional urban solid waste management system: A modeling approach. *European Journal of Operational Research*, 70(4), 16–30.
3. Bloemhof-Ruwaard, J. M., Van Beek, P., Hordijk, L., & van Wassenhove, L. N. (2005). Interactions between operations research and environmental management. *European Journal of Operational Research*, 85(3), 229–243.
4. Carter, C., & Ellram, L. (1998). Reverse logistics: A review of the literature and framework for future investigation. *Journal of Business Logistics*, 19(1), 85–102.
5. Vaidyanathan, J. (2003). The design of reverse distribution networks: Models and solution procedures. *European Journal of Operational Research*, 100(2), 128–149.
6. Pirkul, H., & Jayaraman, V. (1996). Production, transportation, and distribution planning in a multi-commodity tri-echelon system. *Transportation Science*, 30(3), 291–303.
7. Davis, P. S., & Ray, T. L. (1969). A branch-and-bound algorithm for the capacitated facilities location problem. *Naval Research Logistics*, 16(2), 331–344.
8. Rosing, K. E., & ReVelle, C. S. (2006). Heuristic concentration: Two stage solution construction. *European Journal of Operational Research*, 97(4), 75–86.

# Chapter 46

## The Storage of Wind Turbine Mass Data Based on MongoDB

Qile Wang, Zhu Shen, Long Ma, and Shi Yin

**Abstract** With the large-scale development of wind power, the storage and analysis of wind turbine's data gradually become more and more important, there are huge amount of information about wind turbines, and the traditional relational database has been difficult to meet the demand of mass data storage and analysis. This chapter proposes the solutions that data storage is based on non-relational databases (MongoDB) and compares SQL Server with MongoDB about the storage ability and query performance. The results show that using this method can increase storage speed and query performance significantly.

### 46.1 Introduction

As a clean and renewable energy, wind energy has drawn worldwide attention. The huge amount of wind energy that can be developed is 10 times as much as the whole water energy on Earth. The biggest difference between wind turbine and thermal turbine is that wind turbine has small power capacity and wide distribution and its data management is very difficult.

In order to ensure healthy and efficient operation of wind turbines in wind farm, power companies must establish monitoring center to control hundreds or thousands of wind turbines, using SCADA system to record and store operation data about wind turbine. The real-time production data is very large. In order to improve the efficiency of data storage, most wind turbine manufacturers process data before storage, they process real-time data into 10-min average data in order to use the traditional relational databases to store information, and this method may lose numbers of important turbine operation information and pose serious obstacles to data analysis for improving wind turbine performance.

---

Q. Wang (✉) • Z. Shen • L. Ma • S. Yin  
Zhongneng Power-Tech Development Co., Ltd., Beijing 100034, China  
e-mail: [wangqile@clypg.com.cn](mailto:wangqile@clypg.com.cn)

Using traditional relational database to store real-time data is not competent in the field of wind power. This chapter poses a new method based on the non-relational database (MongoDB) massive data storage solutions and analysis and implements this theory.

## 46.2 MongoDB Database Overview

### 46.2.1 *Non-relational Databases*

In industrial field, non-relational databases have drawn collective attention as a new type of database technology. The method does not completely deny relational database. A single relational database cannot meet the demand of diverse application field. The rigorous description of non-relational databases is as follows [1]:

1. Using scalable and loosely data structure to formulate logical model
2. Designed to follow CAP theorem using across multi-node data distribution model and supporting horizontal scaling
3. The ability of keeping data permanent in disk or memory
4. Support a variety of “non-SQL” interface for data access

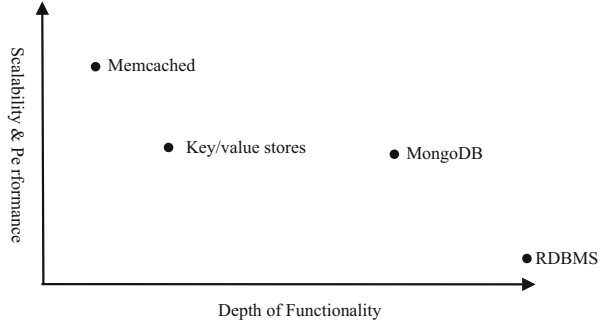
Non-relational database makes changes on two respects. First, data mode, NoSQL uses loose and extensible data model, such as key-value pairs, columns, documents, and charts, not two-dimensional table as relational database. Second, horizontal scaling, NoSQL is designed for distributed systems, supporting horizontal expansion. It is able to adapt to the rapid growth of huge amounts of data [2, 3].

### 46.2.2 *MongoDB Database*

MongoDB is designed to possess the ability of key/value storage and high-performance and scalable, in the meantime with the function of a traditional relational database to manage system [4], as shown in Fig. 46.1.

MongoDB uses data structure of loose BSON (Binary JSON) format and document-oriented storage data; it uses automatic slicing (auto-sharding) to realize mass data storage, supporting type index; it can support two kinds of data replication mechanism: master/slave and replication set. MongoDB uses document-oriented data model to store data, each document is allowed to have different settings [5, 6].

**Fig. 46.1** The magnitude calibration function



### 46.3 MongoDB Database Storage Principle

The underlying of MongoDB use BSON as the data format for data storage and network transmission. It allows data type as nested sub-documents and array. In order to solve large object storage, MongoDB uses GridFS file system to split a large object into multiple small objects.

A MongoDB system is composed of multiple databases; each database consists of a collection, each collection is made up of many documents and the document consists of a series of field, each field is a key-value pair, the key is field name, and its value corresponds to property value [7].

#### 1. Collection

MongoDB collection is a set of documents essentially, and it is similar to the table in relational database. These documents have the same field structure, but MongoDB is schema-free database, and there is no mandatory required on the field structure. Users do not need to predefine structure of a collection of fields and can store documents with different structures in the same collection, when database is running. Users can dynamically add and delete fields of the document [5].

#### 2. Document

MongoDB is a database based on document; document stores data in the form of key-value pairs. Document uses the BSON language; BSON is a binary serialization of JSON data interchange language [8]. It is binary bytes stored in the form of key-value pair, key is the string format, and the value can be any data type, like basic integer, float, string, array, and key-value pairs.

For example, a wind turbine by BSON object is expressed as follows:

```
{WT_NAME: "A-004",
  WT_AI: { {WinSpe:12.42}, {WinDir:250}, {ActPow:12.42},
    {ReaActPow:250}, {ReaActPow:12.42}, {GenSpe:250},
    {EnvTem: 12.42}},
  WT_DI: {WTstart:0, WTrun:1, WTstop:0, PlcErr:0}
}
```

Wind turbine includes wind turbine name (WT\_NAME), wind turbine analog parameter (WT\_AI), and wind turbine digital parameter (WT\_DI). Wind turbine name is the type of string, and the wind turbine analog properties are corresponding to key-value pairs [9].

## 46.4 Database Storage Design

In order to organize massive wind turbine data effectively, and in order to make data efficiently accessible, this chapter uses storage form of data in MongoDB. Convert the raw data using the quad tree method, information storage for each additional one, the information on the level  $L$  to  $L + 1$  fission into four information points, for example, decompose wind turbine into some devices, and then decomposition devices to the specific information of the device. Each sub-block coding is the sub-block that belongs to the parent block coding plus sub-block in the parent block number [10].

When querying the data, we use formulas to calculate the required data items. Then, using client-side technologies will combine information, you can get all the information by user query. The formula are as follows:

$$curXMin = XMin + j \times \frac{\Delta x}{\max^{L-1}}, curXMax = XMin + (j + 1) \times \frac{\Delta x}{\max^{L-1}} \quad (46.1)$$

$$curYmin = YMax - (i + 1) \times \frac{\Delta y}{\max^{L-1}}, curYMax = YMax - i \times \frac{\Delta y}{\max^{L-1}} \quad (46.2)$$

Calculate the number of rows and columns. The formula is as follows:

$$i = \left\lfloor \frac{YMax - y}{\Delta y} \times n \times 2^{L-1} \right\rfloor, j = \left\lfloor \frac{x - XMin}{\Delta x} \times m \times 2^{L-1} \right\rfloor \quad (46.3)$$

### 46.4.1 MongoDB Database Storage

MongoDB documentation storage can realize array types and nesting child document. The data structure of wind turbines is achieved in one collection in MongoDB [11].

Create a database “windturbineinfo”; create a collection to store fan data information statements for the following:

```
MongoServer server = MongoServer.Create("mongodb: "127.0.0.1:27017");
//Specify the database server
```

```

MongoDatabase db = sever.GetDatabase("windturbine
info");
//Obtain windturbineinfo database
MongoCollection<BsonDocument> posts = db.GetCollection
("wtcollection");
//Create a Collection, the name is wtcollection
BsonDocument wtdoc = new BsonDocument (); //Create a
document
wtdoc.put ("WT_NAME", "A-004"); //Add the wind turbines' s
name

```

Pairs of document type attribute, use BasicDBObject class to create a document first, read wind turbines' analog information, add the information to properties of the subdocument simple data types, and then insert subdocument into the properties of the document [12]:

```

BsonDocument WT_AI = new BsonDocument (); //create a doc-
ument MT_AI
WT_AI.put (WinSpe", 13.2); //add Wind speed values
WT_AI.put (WinDir", 250); //add wind direction values

```

Loop to read all the replies, repeat the above statement steps, add analog WT\_AI sub-document to document, and insert the documentation to the collection: posts. insert (doc); through the steps, achieve to store the wind turbine information [13].

#### 46.4.2 SQL Server Database Storage

In SQL Server 2008, the same type of data is stored in the same table, and an index is created for each block [14]:

Given below the key code to connect SQL Server:

```

String connectionString = "jdbc:sqlserver://172.0.0.1:
1433;
Class.forName ("com.microsoft.sqlserver.jdbc.SQLServer
Driver");
Connection con = DriverManager.getConnection (connec-
tionWT);

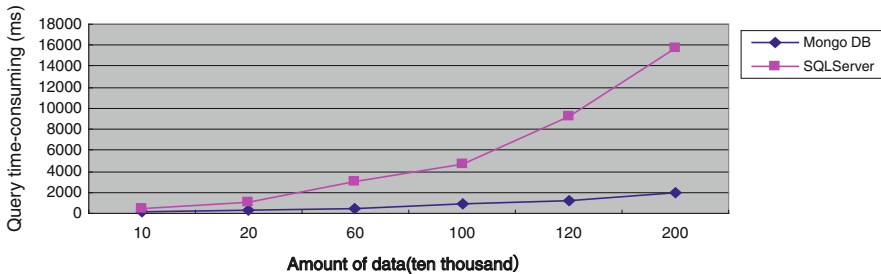
```

#### 46.5 Compared to the Performance Test

Running the SQL Server and MongoDB on the same configuration of the server, test the query performance of MongoDB and SQL Server.

**Table 46.1** The absolute error with different angle at 4 kHz

Data	Maximum (ms)		Minimum (ms)		Average (ms)	
	MongoDB	SQL	MongoDB	SQL	MongoDB	SQL
100,000	121	1,232	55	151	117	392
200,000	254	3,104	223	421	243	1,089
600,000	1,994	4,982	320	1,004	479	3,003
1,000,000	2,453	8,864	579	1,143	845	4,650
1,200,000	3,210	16,012	875	2,345	1,191	9,193
2,00,000	6,452	21,959	1,043	3,567	1,898	15,796



**Fig. 46.2** Query performance comparison about MongoDB and SQL

The server configuration is as follows: 3.0 GHz dual core CPU, 2 GB memory, Microsoft Server 2003 operating system, 2.0.3 version Professional MongoDB, and SQL Server 2008.

Using SQL Server 2008 polling analyzer and MongoDB command “explain” to query time and the use of index information to compare the performances.

In this chapter we query WT\_AI data of “A-004” database in testing system to return WT\_AI’s analog information, including the wind speed, wind direction, active power, reactive power, generator speed, and the environment temperature. Executing the query for 10 times, respectively, in the amount of data under the circumstances is 100,000, 200,000, 600,000, 1,000,000, 1,200,000, and 2,000,000. Choose its average time and compare the performance of the SQL Server and MongoDB.

In Table 46.1, we can get the query takes maximum, minimum, and average time between SQL Server and MongoDB.

Contradistinction is shown as follows (Fig. 46.2).

The experiment shows that MongoDB’s querying time is less than 1/5 of that on SQL Server, and MongoDB’s querying performance is significantly higher than SQL Server. And with the increasing of the data, MongoDB and SQL Server querying performance are both reduced.



## 46.6 Conclusion

MongoDB query takes more gentle growth, which can maintain a linear growth, and time-consuming within 2 s.

SQL Server with the amount of data grows; the query takes serious growth in data volume of 1.1 million over 10 s.

In summary, the MongoDB data storage model, in the conditions of the mass data storage, the improvement of query performance are very obvious.

## References

1. Chang, F., & Dean, J. (2008). A distributed storage system for structured data. *ACM Transactions on Computer System*, 26(2), 2–24.
2. Wang, K., & Wu, Z. X. (2008). Reducing the cluster monitoring workload by identifying application characteristics. *Grid and Cooperative Computing*, 12(8), 525–531.
3. Tai, N., & Hou, Z. (2004). New short-term load forecasting principle with the wavelet transform fuzzy neural network for the power system. *Proceedings of the CSEE*, 24(1), 24–29 (in Chinese).
4. Kassaei, H. R., Keyhani, A., & Woung, T. (1999). A hybrid fuzzy neural network bus load modeling and predication. *IEEE Transactions on Power Systems*, 1999(2), 718–724.
5. Berners-Lee, T., & Hendler, J. A. (2001). The semantic web. *Scientific American*, 284(5), 34–43.
6. Liu, C., Zhou, J., & Xie, Y. (2011). Mass data storage management technology. *Microcomputer Applications*, 8(10), 33–36 (in Chinese).
7. Xu, X., Wu, J., Yang, G., & Cheng, C. (2012). Inexpensive computing platform based on large-scale mass data processing system research. *Application Research of Computers*, 11(2), 582–585 (in Chinese).
8. Rick, C. (2010). Scalable SQL and NoSQL data stores. *SIGMOD Record*, 39(4), 12–27.
9. Peng, X., Ruichun, H., & Zhiming, Z. (2010). *Cache and consistency in NoSQL* (pp. 117–120). Computer Science and Information Technology (ICCSIT), IEEE Computer Society, USA.
10. Brewer, E. (2012). Pushing the CAP: Strategies for consistency and availability. *Computer*, 45(1), 23–29.
11. Singh, M., & Garg, D. (2009). *Choosing best hashing strategies and hash functions* (pp. 50–55). Advance Computing Conference, IEEE Computer Society, USA.
12. Sakr, S., Liu, A., Batista, D., & Alomari, M. (2010). *A survey of large scale data management approaches in cloud environments* (pp. 313–336). Communications Survey and Tutorials, IEEE Computer Society, USA.
13. Han, J., Song, M., & Song, J. (2011). *A novel solution of distributed memory NoSQL database for cloud computing* (pp. 351–355). Computer and Information Science (ICIS), IEEE Computer Society, USA.
14. Pirzadeh, P., Tatemura, J., & Hacıgumus, H. (2011). *Performance evaluation of range queries in key value stores* (pp. 1092–1101). Parallel and Distributed Processing Workshops and Phd Forum (IPDPSW), IEEE Computer Society, USA.

# Chapter 47

## Improvement of Extraction Method of Correlation Time Delay Based on Connected-Element Interferometry

Fei Wang, Zhenfei Wang, Dun Li, and Bingjie Yang

**Abstract** This study proposes a method using mean comparison to improve the accuracy of interferometry processing correlation time delay under the low signal-to-noise ratio and low residual time delay. The method uses the means of taking the average of stripe subsection, comparing threshold and eliminating outliers, which offsets the influence of channel noise on the accuracy of the signal. Using the direct method under the same SNR strike delay simulation comparison proved that the mean comparison method can get relatively high accuracy of time delay information.

### 47.1 Introduction

China has made a series of breakthroughs in satellite navigation, manned space and deep-space exploration which reflects the development of aerospace technology in recent years. As for deep space exploration projects, the Spacecraft needs accurate navigation and positioning, which requires to introduce a high-precision measurement technology to the measurement and control system as a support.

Doppler velocimetry and radar ranging technology are traditional for technologies for spacecraft navigation; these two technologies can directly measure the radial velocity and distance of spacecraft relative to the observatory. However it's difficult to measure the accurate position and velocity of the spacecraft which is perpendicular to the radial movement [1]; as the distance of spacecraft increases, the system error will increase, and measurement accuracy will decline. Interferometry technology has a high-precision angle measurement capability. By measuring the delay observables of the two tracking and measuring stations which receive the same signal that reaches the two sites, the method uses the signal processing method

---

F. Wang (✉) • Z. Wang • D. Li • B. Yang  
School of Information Engineering, Zhengzhou University, Zhengzhou 201305, China  
e-mail: [iewangfei@126.com](mailto:iewangfei@126.com)

to obtain high-precision delay observables and deduces reversely the precise angular position of the signal source, the technology been successfully applied to the tracking observation of satellites and spacecraft in deep-space, as well as positioning task. Through continuous development, it has developed a variety of processing methods in the interferometry signal based on the principle of interferometry technology, such as very long baseline interferometry (VLBI), connected-element interferometry (CEI), the same beam interference (SBI), and other high-precision deep space interferometry technologies. These methods has been successfully applied to the navigation of spacecraft in deep space in practical work [2–6]; these can adapt to the needs of future space's development, improve the accuracy of satellite orbit determination and services for the satellite in orbit and military applications [7].

Currently, for broadband signals like radio source signals, there are usually FX and XF two related processing methods [8]. According to the theory of digital signal processing, the two methods obtain the same power spectrum. Haibo Xia studies observation delay model and the model of phase stripe rotation, which makes the time delay estimation more accurate [9]; Haitao Zhang and Qinghui Liu study application and error elimination of interferometry system like VLBI [10, 11]. Longfei Hao proposes to use the first six data points to fit curves and predict the next point with the fitting curve and compares that point value with the predicted values and observes whether it meets the  $3\sigma$  principle [12]. According to the  $3\sigma$  principle, this chapter gets a phase curve and proposes a method of calculating segmented mean comparison amplitude jitter and verifies the effectiveness of the method through the simulation experiment, which is based on the analysis and introduction of interferometry principle and general signal processing model, ideological phase curve to strike.

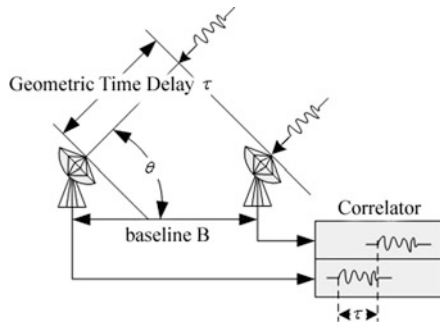
## 47.2 The Principle of Connected-Element Interferometry and Signal Processing Model

### 47.2.1 *The Principle of Connected-Element Interferometry*

The principle of connected-element interferometry is as follows: Two measuring stations receive a signal from the same signal source and perform correlation processing to obtain the time delay of the two signals, thus obtain the measurement data of one direction; the principle of connected-element interferometry is shown in Fig. 47.1.

From several relations in Fig. 47.1 we can obtain expression (47.1) as follows:

**Fig. 47.1** The basic schematic of connected-element interferometry



$$\tau = \frac{1}{c} B \cos \theta \tag{47.1}$$

where \$B\$ denotes baseline, \$\tau\$ denotes time delay, \$\theta\$ denotes the angle between target spacecraft and baseline. Taking the derivative of measuring expression, angular's accurate expression can be obtained

$$\delta\theta = \frac{c \cdot \delta\tau}{B \sin \theta} \tag{47.2}$$

From expression (47.2) we know, the accuracy of angle \$\delta\theta\$ is in inverse proportion to the length of \$B\$ and in direct proportion to the measurement accuracy of \$\tau\$, to obtain high-precision angle measurement; we can use longer baseline (\$B\$) or improve the measuring accuracy of interferometer's time delay \$\delta\tau\$ [13].

### 47.2.2 Processing Model of Connected-Element Interferometry Signal

Two measuring stations which conduct connected-element interferometry receive signals from spacecraft at the same time and have the cross-correlation processing. According to the research of Lue Chen, We need to have a time delay compensation processing before it [13]. If the signals received by measuring station 1 and measuring station 2 are \$X\_1(t)\$ and \$X\_2(t)\$, we used signal received by measurement station 1 as benchmark and analyses signal \$X\_2(t)\$ received by measuring station 2.

Suppose the signals received by two measuring stations are

$$X_1(t) = x(t)e^{j \times 2\pi f \times t} \tag{47.3}$$

$$X_2(t) = x(t + \tau_g)e^{j \times 2\pi f \times (t + \tau_g)} \tag{47.4}$$

where \$\tau\_g\$ is true geometric delay of \$X\_2(t)\$ relative to \$X\_1(t)\$, \$f\$ is the center frequency

of signal after down-conversion, and  $n1$  and  $n2$  are the noise in the signal transmission.

Before two signals conducting correlative processing, we need to have the cross-correlation processing for signal  $X_2(t)$  of measuring station 2.  $\tau_g$  is geometric delay,  $\tau_{g0}$  is compensation value of time delay model, the error between them is  $\Delta\tau_g$ , that is

$$\Delta\tau_g = \tau_g - \tau_{g0} \quad (47.5)$$

Due to adopting the digital signal for signal analysis, suppose the sampling period of signal is  $\Delta T$ , compensation value  $\tau_{g0}$  is composed of integer sampling period compensation  $\tau_{g0I} = n\Delta T$  and decimal sampling period compensation  $\tau_{g0F}$ , that is,

$$\tau_{g0} = \tau_{g0I} + \tau_{g0F} = n\Delta T + \tau_{g0F} \quad (47.6)$$

Having a integer bit time delay compensation (ISTC) to signal  $X_2(t)$  of measuring station 2, the correction is  $\tau_{g0I}$

$$\begin{aligned} X_{2D}(t) &= X_2(T - \tau_{g0I}) = X_2(t - n\Delta T) \\ &= x(t + \tau_g - \tau_{g0} + \tau_{g0F}) e^{-j2\pi f(t + \Delta\tau_g - \tau_{g0} + \tau_{g0F})} \end{aligned} \quad (47.7)$$

Having a stripe rotation to signal and eliminate the phase factor

$$\begin{aligned} X_{2FS}(t) &= X_2(t - \tau_{g0I}) \cdot e^{j2\pi f \cdot \tau_{g0I}} = x(t + \Delta\tau_g + \tau_{g0F}) e^{-j2\pi f(t + \Delta\tau_g + \tau_{g0F})} \cdot e^{j2\pi f \cdot \tau_{g0I}} \\ &= x(t + \Delta\tau_g + \tau_{g0F}) e^{j2\pi f(t + \tau_g)} \end{aligned} \quad (47.8)$$

Make Fourier transform to the signal  $X_{2FS}(t)$ ,

$$F_{2FS}(f) = F_{2D}(f) e^{j2\pi f(\Delta\tau_g + \tau_{g0F})} = F_{2D}(f) e^{j2\pi f \Delta\tau_g} \cdot e^{j2\pi f \tau_{g0F}} \quad (47.9)$$

Having a frequency decimal bit time delay compensation (FSTC), the decimal time delay of compensation is  $\tau_{g0I}$ , so

$$F_2 = F_{2D}(f) e^{j2\pi f \Delta\tau_g} \quad (47.10)$$

Perform the same processing to signal  $X_2(t)$  received by measuring station 2, since it doesn't need time delay compensation, we have

$$F_1(f) = FT(X_1(t)) \quad (47.11)$$

Calculating the cross spectrum of two measuring station signals, that is

$$P(f) = F_1(f) \cdot F_2^*(f) = A(f) \cdot e^{j2\pi f \Delta\tau_g} = A(f) \cdot e^{\varphi(f)} \quad (47.12)$$

where  $A(f)$  denotes amplitude of cross spectrum and  $\varphi(f)$  is phase difference of two signals. Thus the residual value of time delay is the following expression:

$$\Delta\tau_g = \frac{1}{2\pi} \frac{\partial\varphi(f)}{\partial f} \quad (47.13)$$

So we can easily obtain geometric time delay  $\Delta\tau_g$  [13].

### 47.3 Improvement Scheme

Now the thought of time delay extraction is to transform radio frequency down-converter into intermediate frequency first, then generate interference fringes, and use Welch method to divide the sampling data into  $L$  segments when generate interference fringes, the length of each segment is  $M$  and has part of overlap. The data window of each segment adopts window functions like rectangular window, triangular window, or Hanning window. First we calculated cross-power spectrum of each section's sampling data, then accumulated the cross-power spectrum of each section and got the average, so that we could get the average of cross-power spectrum and finally the time delay through phase fitting was calculated. This method which is based on intermediate frequency signal processing would introduce systematic errors when it transforms radio frequency down-converter into intermediate frequency and decreases accuracy. If the radio frequency signals were processed directly, stripes will vary greatly when generating interference fringes and influence the accuracy because of the increase of processing bandwidth. To this, we generally use  $3\sigma$  method to eliminate outliers of stripes, but the elimination threshold is fixed and it will shorten the length of the entire stripe while eliminate the outliers. Under the circumstance of low signal-to-noise ratio, if too many of outliers were eliminated, it would lead to the accuracy of calculating stripe slope decline.

To solve the above problems, under the circumstance of small residual time delay (No phase step from  $-\pi$  to  $\pi$ ), we can get the average data through the method of calculating subsection average before phase fitting, according to the thought of  $3\sigma$  method. Threshold of the fluctuation range was set, if the difference between the data point and the mean was greater than the threshold in scope of this segment, and then we gave the average to this point. Using this method to eliminate outliers, the shortcomings of  $3\sigma$  method which reduce data segment and inability to specify threshold can be overcome. This increases the accuracy of stripe fitting effectively and then increases the measuring accuracy.

## 47.4 Simulation Experiment

We wrote programs for signal processing algorithm according to signal processing procedure in unit 1, next the time delay performance of improved extraction signal through simulation algorithm was tested.

A satellite beacon signal  $X_0(t)$  received by antenna was simulated. Simulating signals consist of a modulation signal and Gaussian white noise  $n(t)$

$$X_0(t) = \cos(2\pi \times 10^9 \times t + \cos(2\pi \times 5 \times 10^7 \times t)) + n(t) \quad (47.14)$$

The signal from the measuring station 1 as the reference signal and the signal from two stations were

$$X_1(t) = X_0(t) + n_1(t) \quad (47.15)$$

$$X_2(t) = X_0(t + \tau_g) + n_2(t) \quad (47.16)$$

where  $n_1(t)$  and  $n_2(t)$  are transmission noise of the channel. Parameters of simulation signals were set as follows: The sampling frequency  $f_s$  was 4 GHz, the length of analytic signal was 262,144, geometric time delay of two signals  $\tau_g$  was 12.525 ns. Supposing estimated model's compensation time delay  $\tau_{g0}$  is 12.2725 ns; the signal-to-noise ratio of the two measurement signal's channel noise is  $-5$  dB; Welch segment  $L$  is 128, window function adopts Hanning window; threshold compared to the average is  $\pi/6$ .

The time domain waveform and the frequency spectrum of the spacecraft simulation signal are as follows in Fig. 47.2.

Constructing two measuring station signals to be analyzed  $X_1(t)$  and  $X_2(t)$  according to expressions (47.15) and (47.16). In the case of that channel noise is  $-10$  dB, the interference fringes and fitting images were obtained as in Fig. 47.3 after the process procedure of interferometry signals, the two images are the direct method and the signal processing image after a mean comparison. From Fig. 47.3 we can see directly that using the mean comparison method can eliminate outliers effectively.

Calculate the straight line's slope after fitting, dividing  $2\pi$  and obtaining the estimation of the residual time delay  $\overline{\Delta\tau}$ , and obtaining the estimation of geometric time delay  $\overline{\Delta\tau_g} = \tau_g + \overline{\Delta\tau}$ . Fitting images of generated stripes when SNR is  $-10$  dB in Fig. 47.3 and calculate the residual time delay estimation of the two methods  $\overline{\Delta\tau_g} = 1.831053167995637 \times 10^{-10}$ s,  $\overline{\Delta\tau_2} = 1.902373750762525 \times 10^{-10}$ s.

From it we can deduce the estimation of geometric time delay  $\overline{\tau_g} = \tau_{g0} + \overline{\Delta\tau}$  and obtain the error between ideal geometric time delay  $\tau_g$  and  $\overline{\tau_{g1}}$  and  $\overline{\tau_{g2}}$  are respectively  $-1.6894683 \times 10^{-11}$ s and  $-9.7626249 \times 10^{-12}$ s. From that we can prove the improvement method can increase the accuracy of time delay.

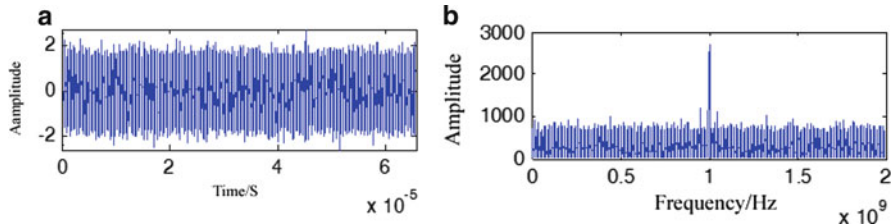


Fig. 47.2 Time-domain waveform and spectrum of spacecraft simulation signal  $X_0(t)$ . (a) Time-domain waveform and (b) spectrum

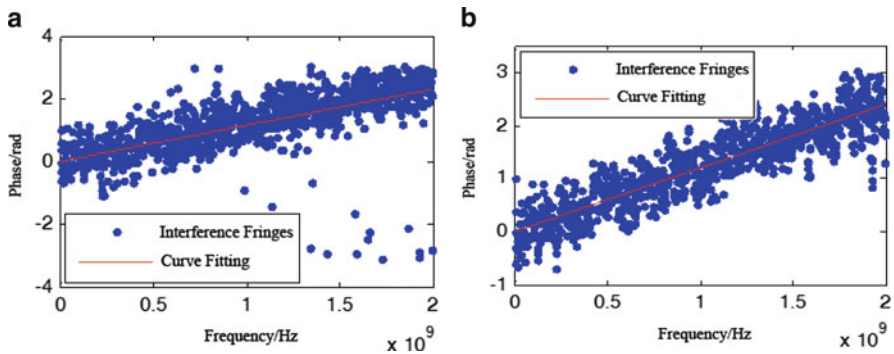
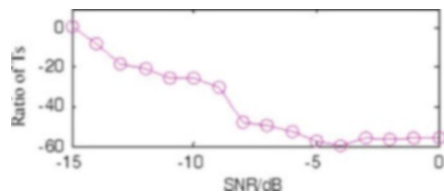


Fig. 47.3 Conduction of interference fringes and images fitting. (a) Direct method and (b) mean comparison method

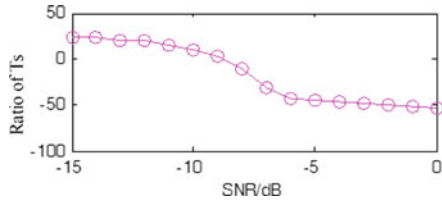
Fig. 47.4 The delay-error-mean-to-sampling-period ratio under different SNR



Using mean comparison method to simulate with different SNR (parameters unchanged). Figures 47.4 and 47.5, respectively, reflect the delay-error-mean-to-sampling-period ratio and delay-error-mean-to-standard-deviation ratio under different SNR. From the figure we can see when SNR is greater than  $-10$  dB, we can obtain the accuracy of signal time delay greater than  $1/10 T_s$  and has better performance when processing residual time delay error.



**Fig. 47.5** The delay-error-standard-deviation-to-sampling-period ratio under different SNR



## 47.5 Conclusion

This study introduced an observational method for the satellite signal time delay, elaborated an improved algorithm of the interferometry signal processing, and made the simulation analysis verification. The results showed that adopting the improved algorithm can effectively eliminate the influence of noise on the data processing when residual time delay is small, and it can obtain the respective time delay of two measuring station signals better and increase the accuracy of measurement when SNR is low. This algorithm can have a further verification in the follow-up observation of the practical measured signals and then it can be applied to the exploration tasks of high-precision navigation in the future, and it can also provide future interferometric studies with valuable experience.

## References

1. Thornton, C. L., & Border, J. S. (2003). *Radiometric tracking techniques for deep-space navigation*. Wiley-Interscience (Chaps 3 and 5).
2. Martín-Mur, T. J., Antreasian, P., Border, J., Benson, J., Dhawan, V., Fomalont, E., & Walker, C. (2006). Use of very long baseline array interferometric data for spacecraft navigation. Jet Propulsion Laboratory.
3. Hodgart, M. S., & Sarannah, G. A. (2008). A triple estimating receiver of multiplexed binary offset carrier (MBOC) modulation signals. In *ION GNSS 2008-21st International Technical Meeting of the Satellite Division* (pp. 877–886). The Institute of Navigation.
4. Chen, L., Tang, G. S., & Chen, M. (2011). Research and application of connected interferometry measurement signal processing method. In *Photonics and Optoelectronics (SOPO) Symposium* (pp. 1–4). IEEE.
5. Madde, R., Morley, T., & Abello, R. (2006). Delta-DOR—a new technique for ESA’s deep space navigation. *ESA Bulletin*, 128, 69–74.
6. Kikuchi, F., Liu, Q., Hanada, H., Kawano, N., Matsumoto, K., Iwata, T., Sasaki, S., et al. (2009). Picosecond accuracy VLBI of the two subsatellites of SELENE (KAGUYA) using multifrequency and same beam methods. *Radio Science*, 44(2), 1–7.
7. Ichikawa, T. (2010). Application of high-precision two-way ranging to the spacecraft navigation. In *Proceedings of SICE Annual Conference 2010* (pp. 817–821). IEEE.
8. Thompson, A. R., Moran, J. M., & Swenson, G. W. (2008). *Interferometry and synthesis in radio astronomy* (pp. 289–293). Wiley-Vch.

9. Xia, H. B. (2005). Satellite the VLBI observations delay and phase stripes rotational model. *Space Science*, 25(1), 52–56 (In Chinese).
10. Zhang, H. T., Qin, Y. Y., & Zhan, Y. M. (2002). Sub-inertia guide fulcrum identify and Eliminate import GPS reference system processing. *Navigation*, 2, 75–80 (In Chinese).
11. Liu, Q., Shi, X., Kikuchi, F., Huang, Q., Kamata, S. I., Matsumoto, K., Wang, N., et al. (2009). High-accuracy same-beam VLBI observations of Shanghai and Urumqi radio telescope. *Science in China, Series G*, 39(10), 1410–1418 (In Chinese).
12. Hao, L. F., Wang, M., & Yang, J. (2010). VLBI observation with the Kunming 40-meter radio telescope. *Research in Astron and Astrophysics*, 10(8), 805–814.
13. Chen, L., Tang, G. S., Wang, M., Liu, H. C., Li, L., & Han, S. T. (2011). Connection interfere with the measurement signal processing method and experimental validation. *Journal of Telemetry, Tracking and Command*, 32(6), 28–31, 43 (In Chinese).

# Chapter 48

## Modeling and Evaluation of the Performance of Parallel/Distributed File System

Tiezhu Zhao, Xin Ao, and Huaqiang Yuan

**Abstract** The mass data storage systems need to be coupled with efficient parallel/distributed file systems, such as Lustre and HDFS, which can effectively solve the problems of the mass data storage and I/O bottlenecks. This chapter systematically studies the performance factors and distribution of parallel/distributed file systems and proposes a valuation scheme for the classic parallel/distributed file system by capturing the changes in workload characteristics. The experiment results show that the proposed evaluation scheme can reach better accuracy and efficiency.

### 48.1 Introduction

Distributed file systems are the key component of any cloud-scale data processing middleware. Evaluating the performance of distributed file system is very important. To avoid the cost for late cycle performance fixes and architectural redesign, providing performance analysis before the deployment distributed file system is also particularly important. System architects can use design-time performance to evaluate the resource utilization, throughput, and timing behavior of a system prior to the deployment due to the following reasons (1) analyzing performance of the system is much less expensive than testing the performance of the system by running it, (2) it is simply infeasible to test all kinds of different configurations of the system by running it, and (3) performance analysis on models helps architects make configuration and deployment decisions to avoid costly redesign, reconfiguration, or redeployment [1].

Automated storage management within an enterprise has proven to be an effective remedy for the low utilization that plagues storage system. Storage system administrator tries to ensure that workload characteristics of application are

---

T. Zhao (✉) • X. Ao • H. Yuan  
Engineering and Technology Institute,  
Dongguan University of Technology, Dongguan 523808, China  
e-mail: [tzzhao83@163.com](mailto:tzzhao83@163.com)

appropriately matched to the performance and availability characteristics of the storage to which they are assigned. Automated storage system design, a solution proposed by many, relies on fast and accurate performance prediction. Constructing good performance evaluation scheme is a critical first step in effectively designing and optimizing the performance of storage systems.

Although the parallel/distributed file systems have been widely studied for several years, the potential impact to workload characteristics of parallel/distributed file system is not clearly understood. In this chapter, we focus on the performance study of the parallel/distributed file system. We systematically study the performance factors and distribution of parallel/distributed file system and propose a novel evaluation scheme for the classic parallel/distributed file system by capturing the changes in workload characteristics.

The remainder of this chapter is organized as follows. We begin by introducing related work in Sect. 48.2. We introduce the performance factors and distribution and present an evaluation scheme for parallel/distributed file system in Sect. 48.3. We perform a series of evaluation experiments, discuss the experiment results in Sect. 48.4, and conclude the chapter in Sect. 48.5.

## 48.2 Related Work

Performance analysis and modeling is an important concern in the distributed file system research area. The related work in the field mainly evaluates the performance of distributed file systems and computing paradigms by relying on running benchmarks or application programs, performance measurements under a variety of workloads/strategies, and comparing with other distributed file systems. The typical approaches can be divided into two categories (1) Experiment-driven performance analysis and prediction. These approaches are mainly based on the analysis of the experiment results or draws conclusion by comparing with existing distributed file systems. Yu et al. indicated excessively wide striping can cause performance. To mitigate striping overhead and benefit collective IO, authors proposed two techniques: split writing and hierarchical striping to gain better IO performance [2]. Wang et al. proposed a two-level metadata management method to achieve higher availability of the parallel file system while maintaining good performance [3]. Yu et al. presented an extensive characterization, tuning, and optimization of parallel I/O on the Cray XT supercomputer (named jaguar) and characterized the performance and scalability for different levels of the storage hierarchy [4]. (2) Model-driven performance analysis and prediction. Li et al. modeled the whole storage system's architecture based on closed Fork-Join queue model and proposed an approximate parameters analysis method to build performance models [5]. Yu et al. adopted a user-level perspective to empirically reveal the implications of storage organization to parallel programs running on Jaguar and discovered that the file distribution pattern can impact the aggregated I/O bandwidth [6]. Piernas et al. adopted a novel user-space implementation of active storage for Lustre, and

the user-space approach has proved to be faster, more flexible, portable, and readily deployable than the kernel-space version [7]. Zhang et al. developed a new mechanism named Logic Mirror Ring (LMR) to improve the reliability and availability of the parallel file system. A logic mirror ring is built over all I/O nodes to indicate the mirror relationship among the nodes [8].

## 48.3 Performance Factors and Evaluation Scheme

### 48.3.1 Performance Factors and Distribution

In parallel/distributed file system, the key problem of performance research is to discover the potential performance factors of application workload and system configuration (e.g., stripe unit size, replica number, concurrent I/O number). Figure 48.1 shows the distribution of performance factors in the classic distributed file system. The performance factors can be divided into four parts: performance factors associated with metadata server, performance factors associated with data storage server, performance factors associated with the network, and performance factors associated with client nodes and applications.

This chapter focuses on performance factors associated with applications. So, it is important to study the characteristics of application workload. Workload characteristics can be used to describe I/O. Common among them is measures of the read/write ratio, I/O request size, spatial locality, temporal locality, and concurrency. Other, more description, characteristics include the temperature of data, the inter-arrival time of I/O bursts, the phasing (overlap) of different I/O streams, measures of temporal burstiness, measures of spatial burstiness, and spatiotemporal correlations [9].

### 48.3.2 Evaluation Scheme

The parameters of our proposed evaluation scheme are defined as follows:

*W*: the workload characteristic vector, including I/O request size (*ReqSize*), read/write ratio (*RWRatio*), request queue length (*ReqQueue*), request arrival rate (*ReqRate*), I/O randomness (*Rand*), I/O inter-arrival delay, etc., and is defined as

$$W = [ReqSize, RWRatio, ReqQueue, ReqRate, Rand, \dots] \quad (48.1)$$

*PMetric*: the performance metric vector, consisted of bandwidth (*Bandwidth*), throughput (*Throughput*), and latency (*Latency*) and is defined as

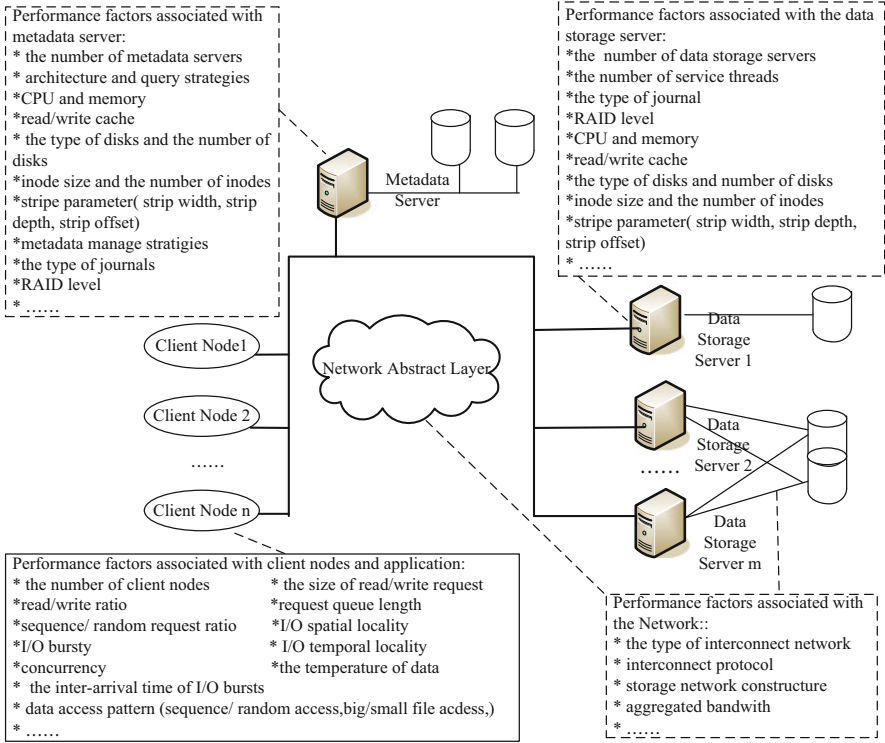


Fig. 48.1 Distribution of performance factors in classic distributed file system

$$PMetric = [Bandwidth, Throughput, Latency] \tag{48.2}$$

For a specific parallel/distributed file system  $DFS_i$ , a map function  $\phi_i$  is defined to map the application’s workload characteristics  $W_i$  to a performance prediction metric  $PMetric_i$ . This relationship can be expressed as follows:

$$PMetric_i = \phi_i(W_i) \tag{48.3}$$

where  $W_i$  is a vector of values representing workload characteristics of  $DFS_i$  and  $PMetric_i$  is any performance metric of  $DFS_i$ .

Similarly, we can define

$$PMetric_j = \phi_j(W_j) \tag{48.4}$$

The first objective of our evaluation scheme is to capture the changes in workload characteristics from  $DFS_i$  to  $DFS_j$ , that is, predict  $W_j$  given  $W_i$ . The function  $\phi_i \rightarrow_j$  is defined as

$$W_j = \phi_{i \rightarrow j}(W_i) \quad (48.5)$$

According to Eqs. (48.4) and (48.5), the composition function  $CF_{i \rightarrow j}$  of  $\phi_i$  and  $\phi_{i \rightarrow j}$  can be expressed as follows:

$$PMetric_j = \phi_j(W_j) = \phi_j(\phi_{i \rightarrow j}(W_i)) = CF_{i \rightarrow j}(W_i) \quad (48.6)$$

## 48.4 Experiment Analysis

### 48.4.1 Experiment Setup

The experiment was performed on HDFS (Hadoop Distributed File System), the primary distributed file system used in cloud computing with Hadoop. The HDFS consists of a single NameNode, a master server that manages the file system namespace and regulates access to files by clients. In addition, there are a number of DataNodes, usually one per node in the cluster, which manage storage attached to the nodes, which they run on. HDFS exposes a file system namespace and allows user data to be stored in files. Our experiments are done on Hadoop 0.20.1. The Hadoop cluster is constructed as two nodes for servers (one for NameNode, one for JobTracer), and the remaining nodes are DataNodes. Each node has 6 core Xeon X5650 processors at 2.66 GHz and 48 GB memory. The interconnect network is 10-gigabit TCP/IP Ethernet.

### 48.4.2 Experiment Result Analysis

The experiment mainly considers four workload characteristic factors: write request size, read request size, write queue depth, and read queue depth. Bandwidth, throughput, and latency are chosen as the performance measurement metrics. The configuration of workload characteristics or performance metrics is shown in Table 48.1.

The classification and regression trees algorithm is chosen as the machine learning algorithm because of their simplicity, flexibility, and interpretability. We measure the proposed evaluation scheme in term of average relative error (*Err*) and Pearson correlation coefficient (*Pearson*) [10]. Table 48.2 shows the accuracy analysis of our proposed evaluation scheme.

As shown in Table 48.2, when considering only one workload characteristics factor (Case1, Case2, Case3, and Case4), the average relative error can be controlled between 11.3 and 18.7 %, and Pearson is 0.79–0.91. When considering two factors (Case5 and Case6), the average relative error can be controlled between 20.9

**Table 48.1** Configuration of workload characteristics or performance metrics

Workload characteristics or performance metrics	Units	Variable
Read request size	KB	<i>WRS</i>
Write request size	KB	<i>RRS</i>
Read queue length	IOs	<i>WQL</i>
Write queue length	IOs	<i>RQL</i>
Bandwidth	MB/s	<i>Bandwidth</i>
Throughput	IO/s	<i>Throughput</i>
Latency	ms/IO	<i>Latency</i>

**Table 48.2** The accuracy analysis of the evaluation scheme

Case	Workload characteristics	Performance metrics					
		Bandwidth		Throughput		Latency	
		<i>Err (%)</i>	<i>Pearson</i>	<i>Err (%)</i>	<i>Pearson</i>	<i>Err (%)</i>	<i>Pearson</i>
Case1	<i>RRS</i>	11.3	0.91	12.3	0.89	13.1	0.90
Case2	<i>WRS</i>	13.1	0.86	15.6	0.83	13.8	0.85
Case3	<i>RQL</i>	12.7	0.88	16.9	0.89	14.4	0.86
Case4	<i>WQL</i>	15.3	0.80	18.7	0.79	16.2	0.80
Case5	<i>RRS, RQL</i>	21.8	0.78	20.9	0.76	23.3	0.76
Case6	<i>WRS, WQL</i>	25.5	0.83	24.9	0.73	23.9	0.80

and 25.5 %, and Pearson is 0.73–0.83. The higher the Pearson correlation and the lower the average relative error are, the more reliable is the prediction model. The results implicate our proposed evaluation scheme can obtain better prediction accuracy and efficiency.

## 48.5 Conclusion

In this chapter, we study the performance of the parallel/distributed file system. We first introduce the performance factors and distribution of the classic parallel/distributed file system and propose a performance valuation scheme by capturing the changes in workload characteristics. Then, we perform a series of experiment to investigate the evaluation scheme. The average relative error can be controlled between 11.3 and 25.5 %. The experiment results show that the proposed evaluation scheme can reach better accuracy and efficiency.

**Acknowledgments** This work is supported by the Natural Science Foundation of Guangdong Province, China (Grant No. S2012040007746), the Scientific Research Foundation for Doctors of DGUT (ZJ130604), the National Natural Science Foundation of China (Grant No. 61170216, 10805019, 61272200).



## References

1. Wu, Y., Ye, F., Chen, K., Zheng, W. (2013). Modeling of distributed file system for practical performance analysis. *IEEE Transactions on Parallel and Distributed Systems*, 99, 1–12.
2. Yu, W., Vetter, J. S., Canon, R. S., Jiang, S. (2007). Exploiting lustre file joining for effective collective IO. In *Proceeding of the Seventh IEEE International Symposium on Cluster Computing and the Grid* (pp. 267–274). Washington, DC: IEEE Computer Society.
3. Wang, F., Yue, Y. L., Feng, D., Wang, J., Xia, P. (2007). High availability storage system based on two-level metadata management. In *Proceeding of the 2007 Japan-China Joint Workshop on Frontier of Computer Science and Technology* (pp. 41–48). Washington, DC: IEEE Computer Society.
4. Yu, W., Vetter, J. S., & Oral, H. S. (2008). Performance characterization and optimization of parallel I/O on the cray XT. In *Proceeding of the 2008 I.E. International Symposium on Parallel and Distributed Processing* (pp. 1–11). Piscataway, NJ: IEEE.
5. Li, H. Y., Liu, Y., & Cao, Q. (2008). Approximate parameters analysis of a closed fork-join queue model in an object-based storage system. In *Proceeding of the Eighth International Symposium on Optical Storage and 2008 International Workshop on Information Data Storage* (pp. 1–8). Bellingham: SPIE.
6. Yu, W., Oral, H. S., Canon, R. S., Vetter, J. S., Sankaran, R. (2008). Empirical analysis of a large-scale hierarchical storage system. In *Euro-Par 2008, LNCS 5168* (pp. 130–140). Berlin: Springer.
7. Piernas, J., Nieplocha, J., & Felix, E. J. (2007). Evaluation of active storage strategies for the lustre parallel file system. In *Proceeding of the 2007 ACM/IEEE Conference on Supercomputing* (pp. 1–8). New York, NY: ACM.
8. Zhang, H., Wu, W., Dong, X., Qian, D. (2005). A high availability mechanism for parallel file system. In *APPT 2005, LNCS 3756* (pp. 194–203). Berlin: Springer.
9. Zhao, T., Verdi, M., Dong, S., & Simon, S. (2010). Evaluation of a performance model of Lustre file system. In *Proceeding of the Fifth Annual ChinaGrid Conference* (pp. 191–196). Washington, DC: IEEE Computer Society.
10. Benesty, J., Chen, J., Huang, Y., Cohen, I. (2009). Pearson correlation coefficient. *Springer Topics in Signal Processing*, 2, 1–4.

# Chapter 49

## CoCell: A Low-Diameter, High-Performance Data Center Network Architecture

Peng Wang, Huaxi Gu, Yan Zhao, and Xiaoshan Yu

**Abstract** As critical infrastructures in the Internet, data centers play an important role in supporting large-scale distributed applications as well as data-intensive computing. This chapter presents CoCell, a server-centric architecture, which uses servers to relay packets. CoCell has several nice properties for desired data center networking. The average node degree in CoCell network is close to 3, and the longest routing path length is no larger than 7. Besides, CoCell network is able to provide high network capacity to support bandwidth-intensive applications. Leveraging the multi-paths between any pairs of servers in CoCell network, we propose relative routing schemes and make a comparison among these paths. The evaluation indicates that CoCell performs well in all-to-all traffic pattern.

### 49.1 Introduction

With the prevalence of cloud computing, mega data centers have emerged as indispensable infrastructures for supporting large-scale distributed applications, such as GFS [1], BigTable [2], and MapReduce [3]. A desired DCN architecture should meet the following design goals: support of incremental expansion, commodity hardware that scales out, and abundant server to server connectivity. However, the conventional tree-based structure does not scale well to hundreds of thousands of servers and is also vulnerable to “single point failure,” thus becomes increasingly difficult to meet these design goals. Recently, researchers are actively

---

P. Wang (✉) • Y. Zhao • X. Yu  
State Key Laboratory of ISN, Xidian University, Xi’an 710100, China  
e-mail: [pengwang.xd@gmail.com](mailto:pengwang.xd@gmail.com)

H. Gu  
State Key Laboratory of ISN, Xidian University, Xi’an 710100, China

Science and Technology on Information Transmission and Dissemination in Communication Networks Laboratory, Xidian University, Xi’an 710100, China

designing new network architectures to build high-performance data centers. The first thread of research work focuses on switch-centric structure network, leveraging switches to relay packets, such as VL2 [4] and Fat-Tree [5]. Another thread puts networking intelligence on servers instead of switches. Such designs include DCell [6], BCube [7], FiConn [8], and DPillar [9]. This kind of design has many obvious advantages over switch-centric designs as listed in the above articles.

In this chapter, we propose a regular server-centric DCN architecture named CoCell, which can be incrementally expanded by adding NICs to only a part of servers. Each server in CoCell not only undertakes the task of computation but also serves as an intermediate node to relay packets. It offers high scalability to support hundreds of thousands of servers with low diameter, low average node degree, and high network capacity. As for the routing algorithm we propose, it is specially customized for the CoCell network. We find many sets of paths for arbitrary pairs of servers, and the length of alternative paths is not too much larger than that of default path.

The rest of this chapter is organized as follows. Section 49.2 describes the structure of CoCell in detail and its topological properties. Section 49.3 presents the routing algorithm. Section 49.4 evaluates the topology properties and routing protocols in CoCell. Finally, Sect. 49.5 concludes the chapter.

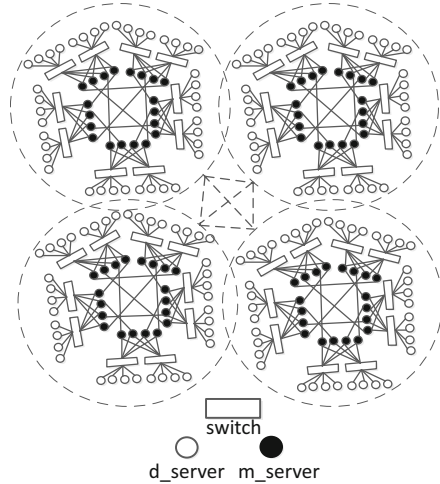
## 49.2 The CoCell Network Structure

In this section, we will first present the physical structure of CoCell and its topological properties. Then we discuss the addressing and interconnection rules of servers in CoCell.

### 49.2.1 CoCell Physical Structure

CoCell is a server-centric network structure, designed to efficiently interconnect an increasingly number of servers. The implementation leverages architecture design rules to build scalable and reliable network architecture at a low cost. All switches in this structure are of the same type that connect a constant number of servers. We use two parameters to define a CoCell network:  $n$ , the number of ports in each switch, and  $i$ , the number of ports in each server. Thus a CoCell network can be expressed as  $CoCell(n, i)$  ( $i \geq 2$ ). Before explaining how to construct  $CoCell(n, i)$ , we first introduce the smallest module,  $CoCell_0$ , and the building block,  $Cell(n, i)$ . In our design,  $CoCell_0$  consists of an  $n$ -port commodity switch;  $n/2$  multi-port servers, which are denoted as  $m\_server$ ; and  $n/2$  dual-port servers, which are denoted as  $d\_server$ . The first port of  $m\_server$  and  $d\_server$  is used to connect the commodity switch. In a  $CoCell_0$ ,  $m\_servers$  are used to connect the other  $n/2$   $CoCell_0$ s, and  $d\_servers$  are used to connect  $Cell(n, i)$ s into a  $CoCell(n, i)$ .

**Fig. 49.1** A partial *CoCell* (8, 3) structure. A *CoCell* (8, 3) structure consists of 41 *Cell*(8, 3)s



The smallest *Cell*(*n*, *i*), *Cell*(*n*, 2) is constructed by  $n/2 + 1$  *CoCell*<sub>0</sub>s by means of a complete graph. There is only one link between any pair of *CoCell*<sub>0</sub>s.

Then, each *Cell*(*n*, *i*) is treated as a virtual node, fully connected with each other homogeneous virtual node to form a *CoCell*(*n*, *i*) in the same way. To scale from *Cell*(*n*, *i* - 1) to *Cell*(*n*, *i*), we need to add a switch connecting  $n/2$  servers into each *CoCell*<sub>0</sub> and connecting another  $n/2$  ports of the switch with  $n/2$  *m\_servers* in the *CoCell*<sub>0</sub>. In this way, additional  $(n/2) * (n/2 + 1)$  *d\_servers* are added into a *Cell*(*n*, *i* - 1) to form a *Cell*(*n*, *i*). The structure of *CoCell*(8, 3) is shown in Fig. 49.1.

### 49.2.2 Topological Properties

We denote the total number of servers in a *CoCell*(*n*, *i*) as *N*(*n*, *i*). The number of servers in a *Cell*(*n*, *i*) network can be expressed as follows:

$$N(n, i) = \left(\frac{n}{2} + 1\right) * \frac{n}{2} * i + \left[\left(\frac{n}{2} + 1\right) * \frac{n}{2}\right]^2 * i * (i - 1) \tag{49.1}$$

Bisection bandwidth refers to the smallest number of links removed to divide the nodes in the network into two parts of equal size.

In a *CoCell*(*n*, *i*), the lowest bound of bisection bandwidth has been obtained:

$$Bd(n, i) = \left[\frac{n}{4}(i - 1)\right]^2 + \left(\frac{n}{4}\right)^2 + i \tag{49.2}$$

### 49.2.3 Addressing and Interconnection Rules

We denote a server in a  $CoCell(n, i)$  as a 4-tuple  $\langle p, q, a_1, a_0 \rangle$ . In this tuple,  $p$  ( $p \geq 0$ ) means the  $Cell(n, i)$  a server locates in,  $q$  indicates the switch that a server belongs to, and  $a_1 a_0$  are 2-tuple used to distinguish the multi-port servers in each  $Cell(n, i)$ . Based on the features of the structure,  $a_1$  and  $a_0$  take values from  $[0, n/2 + 1]$  and  $[0, n/2]$ . Particularly, if  $q = 0$ , it means that this server is an  $m\_server$  and connected to all the  $i-1$  switches.

In a  $CoCell(n, i)$ , two  $m\_servers$  of  $\langle l, p, a, b-1 \rangle$  and  $\langle p, q, b, a \rangle$  are connected with a link for every  $a$  and every  $b > a$ . In a  $CoCell(n, i)$ , we treat each  $Cell(n, i)$  as a virtual node. A  $d\_server$  in a  $Cell(n, i)$  is denoted as  $\langle l, u \rangle$ , where  $l$  is the  $Cell(n, i)$  the server belongs to and  $u$  is calculated from the other three tuples:

$$u = (p - 1) * \left(\frac{n}{2} + 1\right) * \frac{n}{2} + a_1 * \frac{n}{2} + a_0 \quad (49.3)$$

We illustrate the design via the following example. Server  $\langle 0, 1, 2, 1 \rangle$  should be denoted as server  $\langle 0, 5 \rangle$  through the method above. Thus we can easily find server  $\langle 6, 0 \rangle$  of another cell to connect to.

## 49.3 Routing in CoCell Network

In this section, we propose CoCell routing algorithm, which fully exploits the characteristics of CoCell structure to achieve effective resource utilization.

Leveraging the characteristics of the network structure and node addressing, we propose an efficient and simple routing scheme, called *CoCellRouting*, to find a single-path between any pairs of servers in a  $CoCell(n, i)$ . *CoCellRouting* is divided into two steps, finding path in the same  $Cell(n, i)$  and finding path in the different  $Cell(n, i)$ s.

Considering two nodes  $src$  and  $dst$  are in the same  $Cell(n, i)$ , we propose the routing algorithm as follows. *FindLink* calculates the link that interconnects two  $m\_servers$  belonging to different  $CoCell_0$ s. According to the interconnection rules of CoCell, the link is  $\langle S_l, 0, D_{a_0}, D_{a_1-1} \rangle, \langle D_l, D_p, D_{a_1}, D_{a_0} \rangle$ .

Considering two nodes  $src$  and  $dst$  in the different  $Cell(n, i)$ s, we first use Eq. (49.3) to convert the address of a  $d\_server$  into  $\langle l, u \rangle$ . Then based on the function *FindLink*, we can find the link connecting the two  $Cell(n, i)$ s.

We consider the worst-case scenario where two  $d\_servers$  are in different cells. From each server, the packet needs to be forwarded three hops to reach the cell where the other one located in. Thus, the maximum path length between any pairs of servers in  $CoCell(n, i)$  is at most 7.

---



---

**Pseudo-code of this single-path routing algorithm**


---

```

/* src and dst are denoted using the 4-tuples
src = <Sb, Sp, Sa1, Sa0> = <Sb, 0, Sa1, Sa0>
dst = <Db, Dp, Da1, Da0> = <Db, 0, Da1, Da0> */
01: CellRouting(src, dst) /* in the same Cell(n,i) */
02:   if (Sa1 == Da1)
03:     if ((Sp & Dp) & (Sp != Dp) == 1)
        /* belong to the different switches */
04:       path1 = (src, n1);
05:       FindLink(n1, n2);
06:       path2 = (n2, dst);
07:       return path1 + path2;
08:     else return(src, dst);
09:   else return(src, dst);

```

---



---

## 49.4 Evaluation

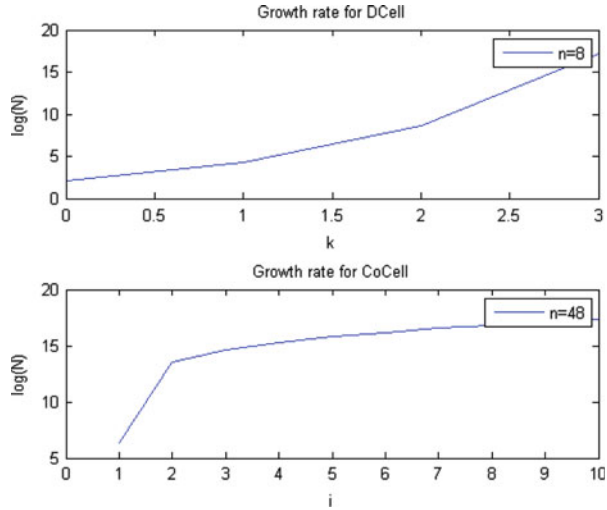
We have analyzed the basic properties of CoCell in Sect. 49.2, such as the high scalability and high bandwidth, which play an important role in the performance of data center networking. In this section, we conduct simulations to evaluate the properties of CoCell and compare it with a famous server-centric structure, DCell.

### 49.4.1 Growth Rate of Server Number

For a current server-centric architecture, such as DCell and FiConn, it can be difficult to build a complete structure when the server number in its lower structure has been very large, since the number of servers increases doubly exponentially. In our structure, the server number increases smoothly and meets the requirement of current data center. So, we do not have to build partial network and design partial routing algorithm for network anymore.

In Fig. 49.2, we can clearly find that CoCell expands more smoothly than DCell. DCell has the limitation of using mini switches ( $n \leq 8$ ) [1]. With the increase of  $n$  and basic expansion unit (symbol  $k$  for DCell and symbol  $i$  for CoCell), the number of servers will increase more acutely. We find that even *CoCell*(48,  $i$ ) also changes smoothly with  $i$ . However, DCell increases very fast even when  $n$  and  $k$  are still small integers.

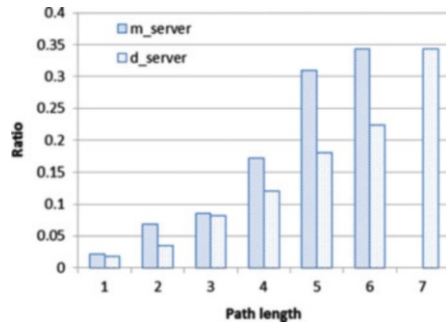
**Fig. 49.2** Growth rate of DCell and CoCell vs. expansion unit



**Table 49.1** The distribution of routing path in *CoCell*(4,3)

Path length		1	2	3	4	5	6	7
Number	<i>m_server</i>	5	16	20	40	72	80	0
	<i>d_server</i>	4	8	19	28	42	52	80

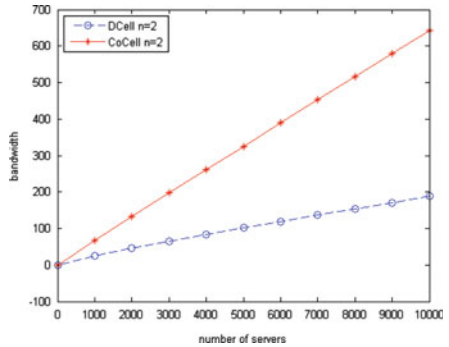
**Fig. 49.3** The distribution of routing path in *CoCell* (4, 3)



### 49.4.2 Distribution of Path Length

We run simulations on *CoCell*(4, 3) to evaluate the distribution of path length. The results are shown in Table 49.1. From these data, we find that the average length of path from *m\_server* to other servers is shorter than that of *d\_server*. Therefore, when the network scales to a large one, the number of *m\_servers* increases; thus the average path length will decrease. Figure 49.3 indicates the distribution of routing path in *CoCell*(4, 3). Such a distribution of routing path performs well in all-to-all traffic pattern [7].

**Fig. 49.4** Bandwidth of DCell and CoCell vs. server order



### 49.4.3 Evaluation of Bandwidth

Bandwidth is an important metric to measure the network throughput. Larger bisection bandwidth means that there are more possible paths between any pair of servers and thus intrinsically more fault tolerant. The results in Fig. 49.4 show that the bandwidth of CoCell performs better than that of DCell with the same type of switches.

## 49.5 Conclusion

In this chapter, a novel server-centric architecture named CoCell is proposed, which has many desirable DCN features. It is highly scalable to accommodate up to hundreds of thousands of servers with low diameter, low average node degree, and high bisection bandwidth. As a decentralized routing solution, the routing algorithm in CoCell effectively exploits the features of the structure. Finally, we evaluate the performance of this structure and find that CoCell shows nice topological properties.

## References

1. Ghemawat, S., Gobioff, H., & Leung, S. (2003). The Google file system. In *Proceedings of the Nineteenth ACM Symposium on Operating Systems Principles* (pp. 29–43). New York: ACM Press.
2. Chang, F., Dean, J., Ghemawat, S., Hsieh, WC., Wallach, DA., Burrows, M., et al. (2008). Bigtable: A distributed storage system for structured data. *ACM Transactions on Computer Systems (TOCS)*, 26(2), doi:10.1145/1365815.1365816
3. Dean, J., & Ghemawat, S. (2008). MapReduce: Simplified data processing on large clusters. *Communications of the ACM*, 51(1), 107–113.



4. Greenberg, A., Hamilton, JR., Jain, N., Kandula, S., Kim, C., Lahiri, P., et al. (2009). VL2: A scalable and flexible data center network. In *ACM SIGCOMM Computer Communication Review* (pp. 51–62). New York: ACM Press.
5. Al-Fares, M., Loukissas, A., & Vahdat, A. (2008). A scalable, commodity data center network architecture. In *ACM SIGCOMM Computer Communication Review* (pp. 63–74). New York: ACM Press.
6. Guo, C., Wu, H., Tan, K., Shi, L., Zhang, Y., & Lu, S. (2008). Dcell: A scalable and fault-tolerant network structure for data centers. In *ACM SIGCOMM Computer Communication Review* (pp. 75–86). New York: ACM Press.
7. Guo, C., Lu, G., Li, D., Wu, H., Zhang, X., Shi, Y., et al. (2009). BCube: A high performance, server-centric network architecture for modular data centers. In *ACM SIGCOMM Computer Communication Review* (pp. 63–74). New York: ACM Press.
8. Li, D., Guo, C., Wu, H., Tan, K., Zhang, Y., & Lu, S. (2009). FiConn: Using backup port for server interconnection in data centers. In *INFOCOM 2009* (pp. 2276–2285). Washington, DC: IEEE Computer Society Press.
9. Liao, Y., Yin, D., & Gao, L. (2012). DPillar: Dual-port server interconnection network for large scale data centers. *Computer Networks*, 56(8), 2132–2147.

# Chapter 50

## Simulation Investigation of Counterwork Between Anti-radiation Missile and Active Decoy System

Huaqiang Hu and Dandan Wen

**Abstract** Simulation test has provided a favorable method and platform for quantitatively evaluating impact on countering ARM by overcoming the disadvantage of high price and poor privacy in regard to outfield experiment. The essay tries to make a deep research on modeling simulation of active decoy and ARM and thus to formulate active decoy interference model and anti-radiation missile movement model. Then it carries through simulation process of active decoy's effect; the simulation result validates the model's effectiveness and accuracy, and therefore it has provided theoretical bases for designing and deploying active decoy and ARM's base station program.

### 50.1 Introduction

As a tactical offensive weapon against electromagnetic radiation source (radar, military communications equipment, etc.), ARM is mainly directed by guidance signal given out by enemy radiation source; it would lead the missile to radiation source and thus destroy it. ARM plays the “ace in the hole” role against electromagnetic radiation source and tries to threaten military use and survival of radiation source. As a result, it is significant that deep research should be made on measures against ARM and that viability of electromagnetic radiation source thus would be improved. Nowadays one of the effective methods for confronting ARM is active decoy.

It is an effective measure for confronting ARM by adopting the active decoy system constituted by active decoy. It is capable of inducing space synthetic

---

H. Hu (✉)

The First Aeronautic Institute of Air Force, Xinyang 464000, China

e-mail: [hqsnial@163.com](mailto:hqsnial@163.com)

D. Wen

Xinyang Normal University Huarui College, Xinyang 464000, China

distortion; in certain conditions, ARM would be unable to aim at any radiation source, and therefore it plays the role of protecting ground radar station and decoy station [1]. In this way, it has a practical significance for research on related weapon system development and tactical application that a deep investigation on active decoy and countering ARM would be made. Simulation test has provided a favorable method and platform for quantitatively evaluating impact on countering ARM by overcoming the disadvantage of high price and poor privacy in regard to field experiment.

## 50.2 Mathematical Modeling of Active Decoy System

One-source active decoy system adopts the system constituted by one decoy and radar. Despite its limitations, it lays the foundation for two-source active decoy system and multisource active decoy system, and therefore a detailed analysis of it is necessary.

Anti-radiation seeker actually functions as a broadband passive mono-pulse detection system. Its constitution and working principle resemble roughly that of mono-pulse radar, though anti-radiation missile itself does not emit electromagnetic waves. ARM, as passive radar, tries to ascertain azimuth and elevation of external active radiation by employing external electromagnetic wave and thus locate its targets [2]. At present, most of ARM makes use of the method of “phase comparison” to measure azimuth and elevation both at home and abroad. The mono-pulse system has a good ability of anti-interference with one-source active decoy system; however, when there exist several radiation sources within discrimination angle of mono-pulse radar, it would trace source energy center while deviating from the source. In this way it lays the theoretical foundation for interference with ARM.

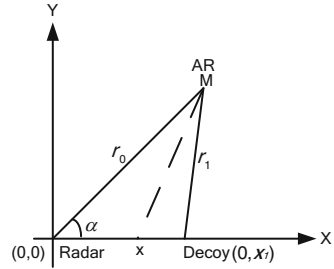
It is assumed that the plane constituted by radar, active decoy, and ARM is set within a rectangular coordinate system, as shown in Fig. 50.1.

Origin of coordinate is location of radar, active decoy being on x-axis, and coordinate of ARM should be  $(x, y)$ . Thus, the following relationship can be obtained from Fig. 50.1:

$$\begin{aligned} r_0 &= \sqrt{x^2 + y^2} \\ r_1 &= \sqrt{(x - x_1)^2 + y^2} \end{aligned} \quad (50.1)$$

It is assumed that angular frequency of radar and decoy would, respectively, be  $\omega_0$  and  $\omega_1$ , with the initial phase difference between them being  $\varphi_{10}$ , and when arriving at ARM, electric field intensity of radar and decoy would, respectively, be

**Fig. 50.1** Position relationship diagram of radar, decoy, and ARM



$$\begin{aligned}
 E_0 &= E_{00} \cos \left( \omega_0 t - \frac{2\pi}{\lambda_0} r_0 \right) \\
 E_1 &= E_{10} \cos \left( \omega_1 t - \frac{2\pi}{\lambda_1} r_1 + \varphi_{10} \right)
 \end{aligned}
 \tag{50.2}$$

Taking no account of field polarization direction difference between radar and decoy, synthetic field intensity at the position of ARM would be

$$\begin{aligned}
 E &= E_{00} [1 + \beta^2 + 2\beta \cos(\varphi_0 - \varphi_1)]^{0.5} \\
 \varphi(x, y, t) &= \arctan \left( \frac{\sin \varphi_0 + \beta \sin \varphi_1}{\cos \varphi_0 + \beta \cos \varphi_1} \right)
 \end{aligned}
 \tag{50.3}$$

Since tracking direction of mono-pulse seeker is the same as normal direction of phase line, connecting with the abovementioned formula, coordinate of x would be

$$\frac{x}{x_1} = \frac{\beta^2 + \beta \cos \Delta\varphi}{\beta^2 + K + \beta(K + 1) \cos(\Delta\varphi)}
 \tag{50.4}$$

It can be made out by formula (50.4) that when  $\beta = 1$ , formula (50.4) would be simplified as follows:

$$x = \frac{x_1}{1 + K}
 \tag{50.5}$$

It can be seen from formula (50.5) that in this case, ARM's target direction has nothing to do with its phase difference  $\Delta\varphi$ ; on the contrary, the target direction is closely related to decoy's frequency, radar's frequency, and ARM's attacking direction. In decoy process of actual combat, regulation of  $\omega_0$  and  $\omega_1$  can be achieved according to distance parameter detected by radar. ARM is thus made to direct to two radiation sources as far as possible, and in this way both radar and decoy can be assured their safety. When distinction between  $r_0$  and  $r_1$  is made quite

great by ARM’s direction (ARM is launched from the other side of layout with radar and decoy), radar and decoy should be adjusted so that distinction between  $\omega_0$  and  $\omega_1$  would be larger; however, in such a situation as for ARM which would fulfill its target recognition by using carrier frequency of radiation sources, the system concerned would not achieved its decoy effect, and single-point source decoying has its limitations [3].

Basing on this, the universal model of decoying system can be inferred. Supposing that the decoying system is composed of radar and  $n_i$  irradiation source, formulating a coordinate system with radar for origin, radar would be (0,0,0), decoy No. i would be  $(x_i, y_i, z_i)$ , and ARM is located at  $(x_A, y_A, z_A)$ , it can be concluded by employing the same method of single-point source decoying that

$$\begin{aligned}
 x &= \frac{\sum_{i=0}^n \sum_{k=0}^n E_{0i} E_{0k} C(x_k) \cos(\varphi_i - \varphi_k)}{\sum_{i=0}^n \sum_{k=0}^n E_{0i} E_{0k} D_k \cos(\varphi_i - \varphi_k)} \\
 y &= \frac{\sum_{i=0}^n \sum_{k=0}^n E_{0i} E_{0k} C(y_k) \cos(\varphi_i - \varphi_k)}{\sum_{i=0}^n \sum_{k=0}^n E_{0i} E_{0k} D_k \cos(\varphi_i - \varphi_k)}
 \end{aligned}
 \tag{50.6}$$

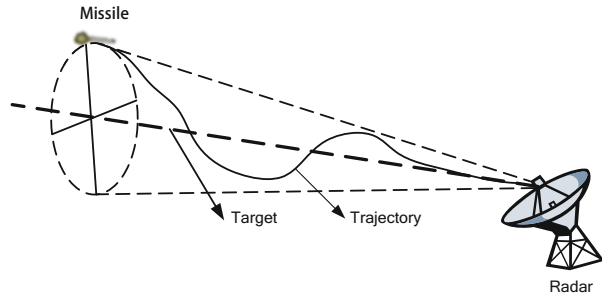
In formula (50.6)  $C(t_k) = \frac{t_k z_A - t_A z_K}{R_K \lambda_K}$ ,  $D_k = \frac{z_A - z_K}{R_K \lambda_K}$ , in the same way, since single-pulse ARM seeker tracks along with normal direction electromagnetic wave front; ARM’s tracking direction can be ascertained during its flight at any time.

### 50.3 ARM’s Mathematical Modeling of Dynamic Flight

ARM’s dynamic flight refers to its whole process of space flight after being launched. The essay mainly deals with active decoy’s degree of interference on the ARM seeker which is attacking, with regard to multipoint embattling. Before modeling, firstly ARM’s space flight line while attacking ground radar station would be introduced. Figure 50.2 displays ARM’s space flight line while attacking a static ground radar station. It can be seen from Fig. 50.2 that during the process when ARM is attacking its target (static target), its trajectory is usually operating negative feedback near the axis and finally moves towards its target. Hereunto, the thin broken line designates missile axis movement [4].

As shown in Fig. 50.2, ARM’s ballistic curve equation can be denoted as follows:

**Fig. 50.2** ARM attacking target radar trajectory



$$\frac{x - x_A}{x_A - x_0} = \frac{y - y_A}{y_A - y_0} = \frac{z - z_A}{z_A - z_0} \tag{50.7}$$

Formula (50.7)  $(x_A, y_A, z_A)$  denotes initial coordinate of ARM trajectory axis;  $(x_0, y_0, z_0)$  denotes coordinate of target radar. The ARM's trajectory orbit sways around axes by operating negative feedback. Therefore, its trajectory axis orbit can be shown in the following equation:

$$\begin{cases} x = S - Vt \\ y = k * \sin x \\ z = k * \cos x \end{cases} \tag{50.8}$$

Hereunto according to relative coordinate system formulated by  $(x_0, y_0, z_0) = (0,0,0)$  and  $(x_A, y_A, z_A) = (S,0,0)$ , where S denotes the distance between ARM and its target, V denotes the radial velocity of ARM, and k denotes the negative feedback coefficient, that is, attenuation coefficient of amplitude, it can be seen from Fig. 50.2 that trajectory axis is only an approximate simulation of ARM's flight line. For within one time step, the attachment between ARM's location from one point to next point and radar is not in the same straight line [5]. Here in consideration of the small time step, in order to simplify the modeling, it is approximately presumed that these three points are on the same line. Since it does not have an obvious effect on placement of distribution, trajectory orbit can be approximately substituted by its axis orbit.

The essay tries to formulate ARM's mathematical modeling of dynamic flight basing on time differencing method. With firing data of ARM's location at certain moment of flight as the starting point, it is presumed that ARM would be located at the next moment  $(t + \Delta t)$  on the attachment between its position at previous moment and radar energy center point. Meanwhile attenuation is introduced, and it can make ARM deviate from the straight line with a tiny distance during a time step of  $\Delta t$ . As a result, these discrete points in the space would be connected as the orbit of ARM's flight in the space. The essence of time differencing method lies in that it makes use of a step-by-step approach to obtain ARM's flight line. In this way, mathematical modeling can be thus simplified and conclusion would not be distorted.

First of all, three-dimensional coordinate system  $xyz$  should be formulated, with radar being located at the origin of coordinate, decoy coordinate being presumed as  $(x_1, y_1, z_1)$ ,  $(x_2, y_2, z_2)$ , etc. In its initial state, that is,  $t = 0$ , ARM's coordinate can be presumed as  $(x_{A0}, y_{A0}, z_{A0})$ ; supposing that time step be  $t$ , one group of discrete time variables can thus be obtained, being, respectively,  $t_1, t_2, t_3 \dots$ . ARM's coordinate would be  $(x_{Ak}, y_{Ak}, z_{Ak})$  at the point of  $t_k$ . During its flight, coordinate of ARM's location at any time can be shown as  $(x_A, y_A, z_A)$ , which is obviously a group of random variables.  $R$  is employed to denote the distance between ARM and radiation source,  $R_{A0}$  the distance between ARM and radar, and  $R_{Ai}$  the distance between ARM and decoy source  $i$ . For the purpose of highlighting the direction of missile attack, make that the azimuth angle of ARM be  $\varphi$ , pitching angle be  $\theta$ , and  $\varphi \in (0, 2\pi)$ ,  $\theta \in (0, \pi/2)$ . According to the definition of right-angle coordinate system, conversion relation between them can be shown as follows:

$$\begin{cases} x = R \cos \theta \sin \varphi \\ y = R \cos \theta \cos \varphi \\ z = R \sin \theta \end{cases} \tag{50.9}$$

Basing on the definition mentioned above, and meanwhile employing subscript method to differentiate energy center of decoy system, the following conclusion can thus be obtained:

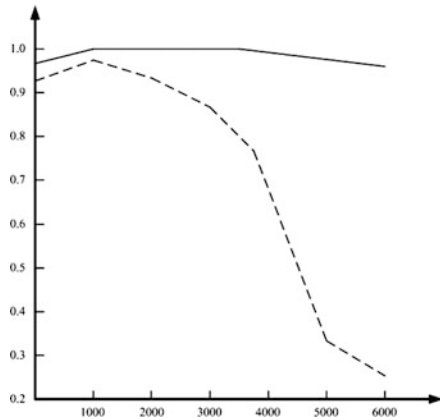
$$\begin{aligned} R_{A0} &= \sqrt{(x_A - x_0)^2 + (y_A - y_0)^2 + (z_A - z_0)^2} \\ R_{ij} &= \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2} \end{aligned} \tag{50.10}$$

### 50.4 Simulation of Decoying Effect

In order to make an analysis of modeling simulation, it is presumed that ARM's parameters are the following: radius should be 25 m, speed on the location of critical differentiating point be  $V = 3M$ , maximum overload of ARM be  $a_m = 10g$ , and resolution of missile seeker be  $\theta = 15^\circ$ . Meanwhile when it is presumed that there are three active decoys, being arranged as regular triangle, power ratio being 1:1:1, and side length of triangle being  $L$ , its simulation process would be as follows [6]:

Suppose that ARM being as far as possible, three active decoys are all located within its field of view; for each given  $L$ , several directions are to be chosen evenly from ARM's probable attacking directions, for each direction  $\varphi_i \in (0, 2\pi)$ , and values would be evenly taken. ARM's final impact point can be obtained on the basis of previous calculation, decoying hit probability would then be counted, and finally relation diagram between decoying hit probability and lengths of triangle sides is shown as in Fig. 50.3 by dotted line.

**Fig. 50.3** Decoying hit probability



It is presumed that after ARM would fly over the critical position of a three-point source, without considering ARM's mobility, it would hit the target directly along this direction; in this way, relationship diagram between decoying hit probability and lengths of triangle sides is shown in Fig. 50.3 by solid line.

From Fig. 50.3, it can be seen that without considering ARM's mobility, when  $L$  is more than 600 m, decoying hit probability is nearly 1. The reason lies in that the absolute distance between critical point and each point source would increase along with increase of  $L$ ; decoying distance concerned would also increase and decoying hit probability can thus be improved. Decoying hit probability would be the largest when  $L$  should be 600 m or so. The reason lies in that absolute distance between critical point and each point source would increase along with  $L$ 's increase. There would be much more time for ARM to modulate its direction; however, due to limitation on ARM's maximum lateral overload, its modulation is rather limited, so in this case, appropriate side length would bring about the largest decoying hit probability.

## 50.5 Conclusion

During the process of ARM's countering multipoint source active decoy, its seeker's resolution angle is one of critical factors which have an effect on its combat effectiveness. From the analysis of simulation, it can be seen that seekers with wide angle of resolution do not possess the ability of countering decoying; within the range which angle of sight would permit, its interference effect would be strengthened along with decrease of seeker's resolution angle. Meanwhile seeker's precision of measure angle can also be heightened to optimize seeker's guidance control parameters, and in this way ARM's motor overload can be increased. It has been an effective method for improving ARM's combat ability of countering decoy and for reducing towed target quantity.



## References

1. Luo, X., Ning, J., & Luan, S. (2005). Fuzzy systematic judgment of ranking multi-radiant imperilment. *Fire Control and Command Control*, 30(4), 66–68.
2. Zhang, B. (2004). Set pair analysis and multi-attribute decision making. *System Sciences and Comprehensive Studies in Agriculture*, 20(2), 123–125.
3. Pan, Q., & Zhang, H. (2008). Comprehensive modeling and simulation verification of radiation threat based on closeness degree of ideal point. *Journal of System Simulation*, 20(14), 3896–3898.
4. Zheng, M. (2005). *Study of radiant-point beguiling resisting anti-radiation missile technology, tactics and embattling mode*. Changsha: National University of Defense Technology.
5. Xie, B., Yin, J., & Song, J. (2004). Building of kill probability simulation model for the ARM against the radar targets. *Journal of System Simulation*, 16(9), 2044–2047, 2051.
6. McLendon, R., & Turner, C. (1983). Broad-band sensors for lethal defense suppression. *Micro-wave Journal (S0192–6225)*, 15(6), 85–101.

# Chapter 51

## Simulation Jamming Technique on Binary Phase-Coded Pulse Compression Radar

Yulin Yang and Lijuan Qiu

**Abstract** Binary phase-coded signal is usually used in pulse compression (PC) radar, which is mostly used for surveilling and tracking targets. Two prominent characteristics of the binary phase-coded signal are introduced in this article; according to these characteristics, some jamming forms are analyzed which can be used to jam binary phase-coded PC radar; simulation with MATLAB is done to test these jamming forms. The result indicates that noise jamming has less effectiveness on binary phase-coded PC radar and continuous and partial code replicated jamming have preferable effectiveness on it. Because binary phase-coded signal is compressed to narrow pulse when it is received by receiver, jamming signal can easily capture the range gate when the jamming side adopts range-gate pull-off (RGPO), and so the jamming effectiveness is perfect.

### 51.1 Introduction

Modern radars are required, not only detecting the targets which are far from them but also having high-range resolution [1]. Radars with high-range resolving power must have extremely narrow pulse width, which limits the increase in power of the transmitter, thus affecting the detecting range of radars. In order to resolve the contradiction between the detecting range and the range resolving power, the designers of radars adopt PC technique and design and manufacture PC radar. The radar of this type transmits wide pulse and receives narrow pulse which is compressed by the matched filter network. Binary phase-coded signal is a type of sending signal commonly used by PC radars. It may be coded pseudo-random. When the matched filter of the echo signal Doppler frequency detuning, the matched filter cannot compress the pulses, so we called that the Doppler frequency

---

Y. Yang (✉) • L. Qiu  
The First Aeronautic Institute of Air Force, Xinyang 464000, China  
e-mail: [yanyline@163.com](mailto:yanyline@163.com)

sensitive signal. It is often used in the target Doppler frequency changing within narrowband occasions. So the binary phase-coded signal has strong anti-jamming capability.

## 51.2 Characteristics of the Binary Phase-Coded PC Signal

Binary phase-coded symbols are composed of the 0, 1 sequence or +1, -1 sequence. Phase of transmitting signal is alternating between  $0^\circ$  and  $180^\circ$  in accordance with the order of symbols (0, 1, or +1, -1). Because the frequency of the transmitting signal is usually not the integral times of the sub-pulse width's reciprocal, the coded signal is generally not continuous in the contrary phase points [2].

At the receiving end, the compressed pulses are obtained through a matched filtering or correlation processing. The half-amplitude point's width of the compressed pulses is nominally equal to the width of the sub-pulses. Therefore, the range resolution is proportional to the duration of one symbol (one sub-pulse). Time width-bandwidth product and pulse compression ratio is equal to the sub-pulses' number of the waveform, i.e., the number of coded symbols [3].

The plural expression of the binary phase-coded signal is

$$s(t) = a(t)e^{j\varphi(t)}e^{j2\pi f_0 t} \quad (51.1)$$

To binary phase-coded signals,  $\varphi(t)$  only has two possible values 0 or  $\pi$ .  $\varphi(t)$  can be expressed as two-phase sequence  $\{\varphi_k = 0, \pi\}$ , and it can also be expressed as a binary sequence  $\{c_k = e^{j\varphi_k} = +1, -1\}$ . If the envelope of the binary phase-coded signal is rectangular, that

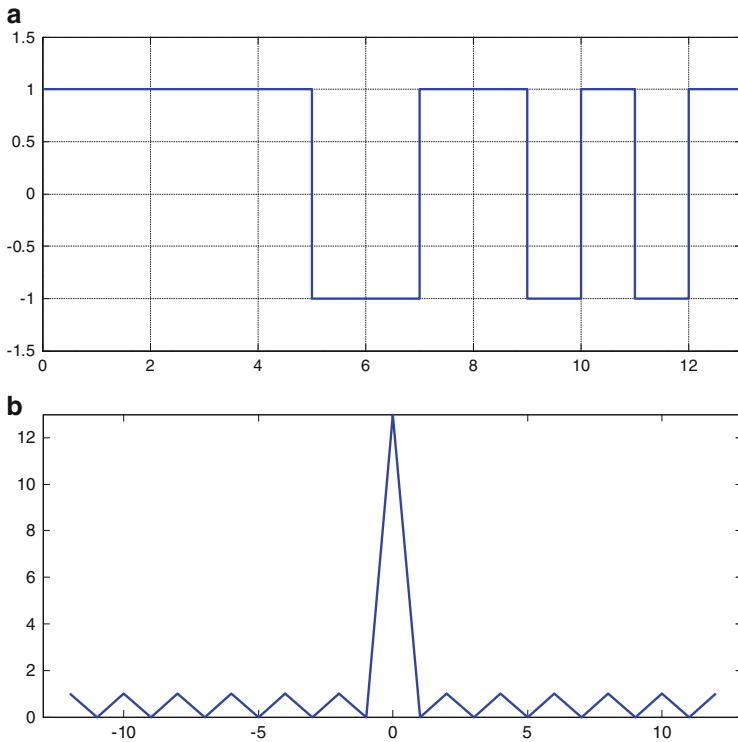
$$a(t) = \begin{cases} 1/\sqrt{PT}, & 0 < t < PT \\ 0, & \text{otherwise} \end{cases} \quad (51.2)$$

Plural envelope of the binary phase-coded signal can be written as

$$u(t) = \begin{cases} \frac{1}{P} \sum_{k=0}^{P-1} c_k v(t - kT), & 0 < t < PT \\ 0, & \text{otherwise} \end{cases} \quad (51.3)$$

Wherein  $v(t)$  is a function of the sub-pulses, T is sub-pulse's width, P is the code's length, and PT is the duration of the encoded signal.

Barker code binary sequence is usually used in binary phase-coded PC radar system. This sequence is characterized by the following: when the aperiodic self-correlation function for a given length of the code is calculated, the peak of the side lobes has a minimum value (the amplitude is 1), and the peak of the main lobe is

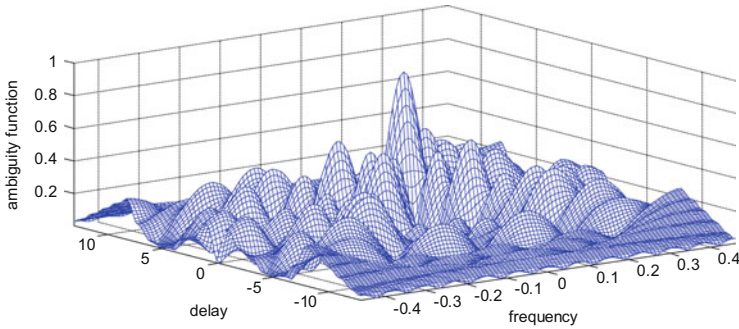


**Fig. 51.1** The waveform (a) and the self-correlation (b) function of the 13-bit Barker code sequence

N. The Barker code which is used by PC radar is usually the 13-bit Barker code sequence, that is, + 1, + 1, + 1, + 1, + 1, - 1, - 1, + 1, + 1, - 1, + 1, - 1, + 1. Figure 51.1a shows the waveform of the 13-bit Barker code sequence; the self-correlation function of it is shown in Fig. 51.1b. Figure 51.2 shows the ambiguity function of the 13-bit Barker code sequence. Seen from the figure, the 13-bit Barker code PC radar's signal has favorable self-correlation and speed and distance resolution.

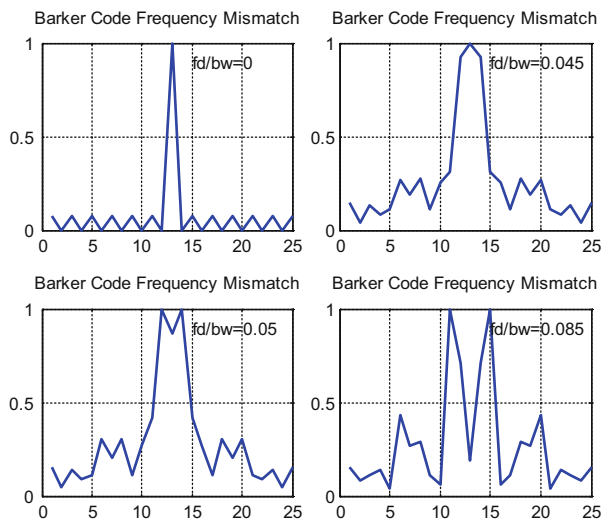
The signal of the binary phase-coded PC radars is more sensitive at the Doppler frequency shift than that of LFM signal. Figure 51.3 shows that in the case of the Doppler frequency shift to PC bandwidth, ratio is bigger than 0.05, the Barker coded PC signal's waveform which is sent out by matched filter will come into being a serious distortion, and the time side lobe is much bigger. Therefore, the binary phase-coded PC radar is not fit for detecting high-speed targets.

The binary phase-coded waveforms with low side lobes of the self-correlation function or zero Doppler frequency response are precisely needed by PC radars. Moving target's frequency response is different to zero Doppler frequency response. However, through the appropriate waveform design, the minimum of



**Fig. 51.2** The ambiguity function of the 13-bit Barker code sequence

**Fig. 51.3** The matched filter's output of different Doppler frequency shift to waveform bandwidth ratio



the Doppler frequency shift to bandwidth ratio can be obtained, and the perfect Doppler frequency response within the interested speed range can be reached. In this speed range, the Doppler frequency response corresponding to the distance between the target and radar or fuzzy function is approximately equal to the self-correlation function of the sequence.

### 51.3 Analysis of Binary Phase-Coded PC Radar Jamming

The binary phase-coded waveform is usually used for surveillance radar and tracking radar. The purpose of jamming depends on the type of radar which we want to jam.

For surveillance radar, the general jamming purpose is to produce background jamming or generate the synchronous decoys which can blanket the real target detection and masking the target. This latter process is sometimes called false-target jamming [4].

For tracking radar, its aim is to produce false targets to capture radar tracking wave gate. In this operating mode, when the jamming side receives the signal of radar, it can make the signal generating coherent frequency shift and then transmit the signal back to the radar. This jamming signal ahead of the target echo which comes from matched filter, and it makes electronic protection technology which used cutting-edge tracking useless.

In general, noise is the least effective jamming waveform in all types of jamming signal on binary phase-coded PC radar. This will be verified if the matched filter of binary phase-coded PC radar is regarded as a correlator. Noise which is completely uncorrelated with the radar's signal cannot get any gain when the noise passes through the matched filter along with the radar's signal and, when wholly related to the target's echo, will get all the processing gain. The jamming waveform will be effective than random noise even if there is little correlation to the radar's signal [5].

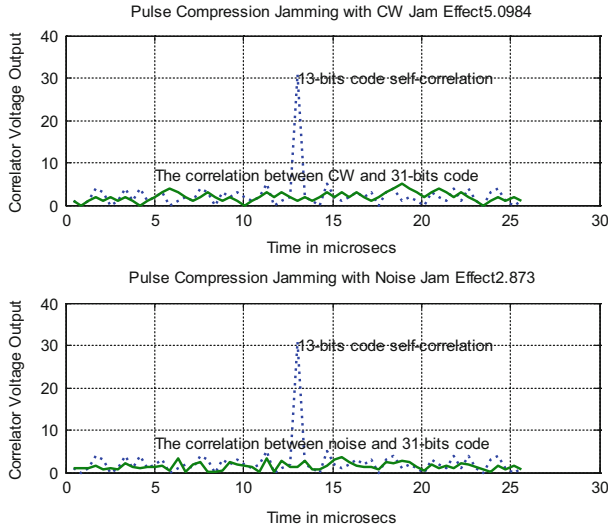
Continuous-wave signal will normally produce a certain degree of correlation with the matched filter of the binary phase-coded PC radar. For example, a continuous waveform whose frequency is same as that of carrier can get a power gain of 6 dB higher than the noise waveform if it is coherent with binary phase-coded waveform in about half of its duration.

Because the efficiency of noise jamming signal passing through the binary phase-coded PC radar's receiver is very low, the other forms of jamming signal must be considered. Replicating the radar's signal and making necessary jamming modulation is a feasible measure. However, binary phase-coded PC radar's signal is generally wide pulse or quasi-continuous-wave signal, if the pulse is used to modulate jamming signal wholly; when the modulated jamming signal compressed by the radar's receiver, its jamming effectiveness will be weakened. Therefore, the partial pulse replicated jamming form is used generally.

Binary phase-coded signal is sensitive to Doppler frequency shift, so it is very difficult to jam it by velocity deception. But the range resolution of the binary phase-coded signal is much higher than that of ordinary pulse radar; as long as its transmitted signal is attached a small delay, then send back to the radar, it will be deceived by range deception [6].

## 51.4 Simulation of Binary Phase-Coded Pulse Compression Radar Jamming

Continuous-wave and random noise signal to jam the 31-bit maximum length sequence of binary phase-coded signal is selected, simulating the jamming process by MATLAB, and the result is shown in Fig. 51.4. Assuming that continuous-wave



**Fig. 51.4** Continuous-wave and noise jamming effectiveness

signal has the same frequency as the carrier, certain phase components are come into being at the central frequency of the binary phase-coded waveform. Also assumed that the noise is also at the central frequency, but relative to the binary phase-coded signals, it produces a uniform phase modulation ( $0 - \pi$ ).

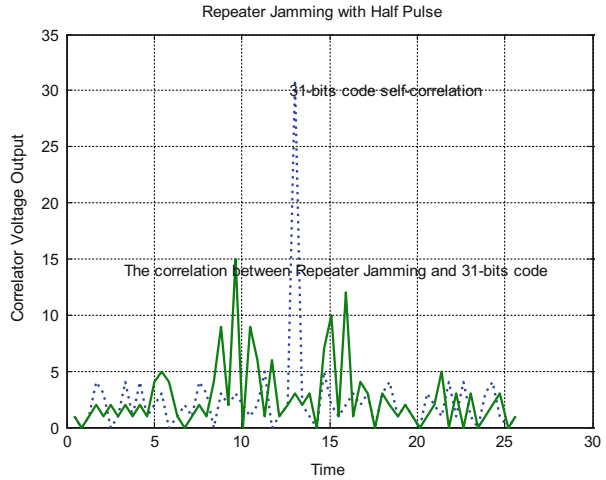
Because the phase-coded compression signal is generally a wide pulse or quasi-continuous-wave signal, better jamming effectiveness can be gained when partial code replicated jamming form is adopted. In this experiment, given the half-code, sending back jamming signal was used to interfere the 31-bit maximum length sequence signals of binary phase-coded PC radar, with the combination of the two 15-bit half-code plus 1 to fill the PC filter network. The result is shown in Fig. 51.5.

In this experiment, the RGPO jamming which is used to jam the binary phase-coded PC radar was simulated; the Digital Radio Frequency Memory (DRFM) technique was used to store radar's operating frequency and repeat the original signals of radar. The course that the range gate tracking the echo and the jamming signal has been simulated, as shown in Fig. 51.6, with the towing time increasing and the range gate gradually tracked on the jamming signal.

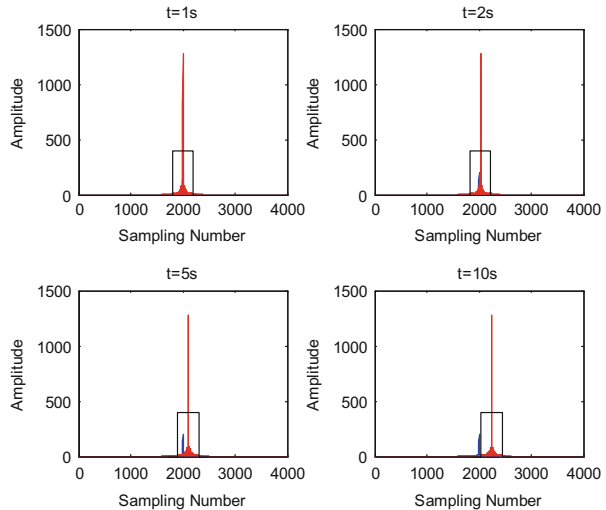
## 51.5 Conclusion

The binary phase-coded PC technique is widely used in the new system of radar, solving the contradiction between the detection range and range resolution, having perfect speed detection and anti-jamming capability. In this chapter, four forms of jamming were used to jam the binary phase-coded PC radar's signal; they are noise,

**Fig. 51.5** Partial code replicated jamming effectiveness



**Fig. 51.6** Range-gate pull-off jamming effectiveness



continuous wave, partial code replicated, and RGPO jamming; the jamming theory and effectiveness were analyzed; the jamming course and result were simulated by MATLAB. These jamming forms are dynamic, and the effectiveness is relative. “There is no radar that unable to jam, and there is no jamming form unable to counterwork.” The anti-jamming capability of the binary phase-coded PC radar will be gradually increased as science and technology development, and the challenge of the jamming side will be increasing. Studying on the jamming technique on the binary phase-coded PC radar will be going on.



## References

1. Xiang, J. (2001). *Radar system* (pp. 188–256). Beijing: Publishing House of Electronics Industry.
2. Skolnik, M. I. (1990). *Radar handbook* (2nd ed., pp. 247–289). New York, NY: The McGraw-Hill Companies.
3. Schleher, D. C. (1999). *Electronic warfare in the information age* (pp. 167–201). Norwood, MA: Artech House.
4. Chen, Z. (1995). Effects of range gate pull-off (RGPO) jamming on some pulse radars. *Aerospace Shanghai*, Vol. 12, Nr. 5, 1995, (Cama, Vol. 3, Nr. 1, 1996); pp. 23–26
5. Xie, B., Yin, J., & Song, J. (2004). Building of kill probability simulation model for the ARM against the radar targets. *Journal of System Simulation*, 16(9), 2044–2047, 2051.
6. McLendon, R., & Turner, C. (1983). Broad-band sensors for lethal defense suppression. *Micro-wave Journal (S0192–6225)*, 15(6), 85–101.

**Part III**  
**Pattern Recognition**

# Chapter 52

## Personalized Information Service Recommendation System Based on Clustering and Classification

Yu Wang

**Abstract** To solve the recommendation system, the prevalence of blindness, and low resistance, in this paper, an in-depth study on the personalized automatic recommendation system user model and automatic recommend technology. The paper first introduces the automatic recommendation system user model representation and update. Then the user modeling, clustering, classification, and automatic recommendation technology are combined to develop the automatic personalized document recommendation system based on clustering and classification. First, offline, the system form clustering points of interest to the article and build user model based on clustering interest points. Then realize the automatic recommendation online by recommendation algorithm which is based on classification. Theoretical analysis and experimental results show that the system can significantly improve the online response speed and efficiency.

### 52.1 Introduction

The Internet has become an important way for people to obtain information. But with the increase of Internet information, people have to spend a lot of time in the complex ocean of information and search the information they need. Information retrieval technology is used to meet some needs of people, but because of its universal nature, it still cannot meet the different backgrounds, different purposes, and different times of the query request. Therefore, personalized information service technology arises at the historic moment. Personalized information serve for different users system with different services to meet different needs. Automatic recommendation system is one of the main forms of the personalized information

---

Y. Wang (✉)

Jilin Business and Technology College, Changchun 130062, China

e-mail: [wang\\_zfy@sina.com](mailto:wang_zfy@sina.com)

service, and it is used to solve the problem of Internet information “get lost” and “information overload” and put forward a kind of intelligent agent system. Many large websites, such as Amazon, eBay, and Dangdang online bookstore, have varying degrees of using various forms of recommendation system.

## **52.2 Personalized Information Service**

Personalized information service is according to user’s interest characteristics and purchasing behavior, referring the information and commodities which they are interested in. Personalized information service technology can fully improve the website’s service quality and access efficiency to attract more visitors. Personalized information service has brought profound influence on the concept of information service, service mode, and content; to improve the level of information service plays a powerful role [1].

## **52.3 Research Status of User Model**

User modeling is the base of personalized information service. The accuracy of the models and timeliness directly determines the quality of service. The user model is the real users’ “virtual” technology. Virtual technology depending on the system requirements and specific functions can give the user a full range of real description and some characteristics of the description. User model research mainly includes user model representation and user model update research.

### ***52.3.1 The Representation of User Model***

The representation of user model provides a structured model storage form [2]. Recommender system user model representation method covers a very wide scope; the main user model method includes the following categories:

1. Representation based on neural network: Neural network represents after network stability. The network connection weights characterized of network state represent the user model. Network states composed by the network input state, network output state, input and output of the connection states. Neural network representation method is used information structural organization, and is not a simple keyword listed to represent user model, which reflects the relationship between the information. However, the said method based on neural network depends on the user model in the process of the neural network

algorithm; the categories and scope are narrow, and representation is not easy to be understood.

2. Representation based on user evaluation matrix: Based on the user appraisal matrix, representation method is used more based on collaborative filtering recommendation system. The user appraisal *matrix* uses a matrix  $R_{m * n}$  to indicate the user model set,  $m$  stands for the system users,  $n$  for resources/number, and  $n$  for resources/project number. In the matrix, each element  $r_{ij}$  is the representation of the user  $i$  to the project  $j$  evaluation. The greater the  $r_{ij}$  value is, the higher degree of interest the user to the corresponding project has.

The user appraisal matrix representation method is simple, intuitive, and doesn't need any learning technologies. It can be collected from the original data directly generated. But this method requires users to display the given evaluation data. This indicates the user's lack of interest and change method, and it is difficult to reflect the user's latest interest.

### 52.3.2 The User Model Update

After the establishment of the user model, according to the user, it is best to recommend dominant or recessive feedback on the user's current change of interest and to constantly update the user model in order to timely reflect the dynamic changes in interest preference and to ensure the effectiveness and practicality of the recommendation system. The existing user model updating technique is mainly divided into the following categories:

1. Genetic algorithm: Genetic algorithm is a kind of search optimization technique based on natural selection and iterative genetic mechanism. Fitness function, chromosome population and selection, and crossover and mutation are the three main operators of genetic algorithm. The genetic algorithm update technology system is usually the first to form a chromosome for the user model coding and generate the initial population. When the initial population evolves iteratively to meet the termination conditions, chromosomes with high fitness are decoded using the new user model, so as to develop the user model update [3].
2. Neural network technology: Neural network update technology is a kind of adaptive updating technology. When the user interest changes, neural network is adaptive to adjust the network connection weights and update the network output results to track changes in the user's interest. Some updates need to establish a new interest preference category and cut the old to set interest preference category. In this kind of circumstance, the network usually needs training. Neural network update technology depends on the user model representation established, usually only the user model based on neural network method establishes the representation of the system used.

## 52.4 Automatic Recommendation Technology

Automatic recommendation systems are commonly used. Mainly recommended technologies are as follows:

1. Recommendation technology based on rules

A rule-based recommendation technology is knowledge engineering methods in the application of personalized information service. Rules can be used by customer and also can be used in association rule data mining techniques for discovery. Recommended technical performance depends on the quality and quantity of rules, rules can take advantage of user static attribute setup, and it can also use the user to establish dynamic information [4].

2. Recommendation technology based on collaborative filtering

The traditional collaborative filtering-based recommendation technology is also known as user-based collaborative filtering recommendation technology or nearest-neighbor collaborative filtering recommendation technology. The basic thought on recommendation technology based on collaborative filtering is based on similar nearest-neighbor rating data from scores of recommended target users who do not score on a project which can be approximated using the weighted average approach.

But the traditional collaborative filtering-based recommendation technology has three difficult problems to solve: algorithm scalability, the sparse solution evaluation data, and recommended problem from initial resources [5].

## 52.5 Automatic Personalized Article Recommendation System Research Based on Clustering and Classification

Automatic personalized article recommendation system research is based on clustering and classification. First, by clustering, it forms the point of interest and the establishment user model is based on clustering interest points. Then the same interest points have greater interest users cluster, and form the the user group of interest point. Finally, to recommend paper classification, it realizes the active recommend according to their interest points. The automatic personalized recommendation based on clustering and classification system is presented in this paper. This system includes offline user model and user groups obtain subsystem and online personalized article recommendation subsystem of two parts.

### 52.5.1 *Offline User Model and User Group Acquisition Subsystem Based on Clustering*

Offline user model and user group acquisition subsystem based on clustering are using implicit ways to acquire user model. First of all, users collect clustering articles to form multiple interest points. Then the user model is expressed weighted interest point vector, according to the user's interest clustering to form each interest point user groups.

Specific steps of the structure based on the clustering of interest point user model are introduced as follows:

Step 1: For all users' collection of all the articles on pretreatment, the article based on semantic and statistical characteristics of the feature vector is expressed as

$$V(d_i) = (w(d_i, T_1), w(d_i, T_2), \dots, w(d_i, T_m))$$

Step 2: For the article to cluster analysis, form the cluster set  $C_n$

Step 3: Each cluster  $C_i$  is as an interest point, and construct the characteristic vector:

$$V(C_i) = (tdr(C_i, T_1), tdr(C_i, T_2), \dots, tdr(C_i, T_m)). \text{ The } tdr(C_i, T_j) \text{ (52.1) as feature item } T_j \text{ for classes } C_i \text{ of the resolution.}$$

$$tdr(C_i, T_j) = \begin{cases} \frac{N_i}{\sum_{k \neq i} N_k} & \text{if } \sum_{k \neq i} N_k \neq 0 \\ 2 \times N_i, & \text{if } \sum_{k \neq i} N_k = 0 \end{cases} \quad (52.1)$$

Among them,  $N_i$  as  $C_i$  class appear feature  $T_j$  of article number.

Step 4: According to each user  $U_i$  collection of the articles, to statistical articles belong to interest point, to construct weighted clustering interest point vector  $UM_i$  to represent the user model:  $UM_i = (w_{i1}, w_{i2}, \dots, w_{i|C|})$

The  $w_{ij}$  (52.2) for users  $U_i$  to interest points  $C_j$  of the weight,  $1 \leq i \leq |U|$ ,  $1 \leq j \leq |C|$

$$w_{ij} = DC_{ij}/DN_i \quad (52.2)$$

$DC_{ij}$  collect the article number belongs to the interest point  $C_j$  for the user  $U_i$ , and  $DN_i$  collect articles for user  $U_i$

### 52.5.2 *The Online Personalized Recommendation Subsystem Based on Classification*

The online personalized recommendation subsystem based on classification is the first to recommend paper classification, namely, calculating it with the interest point

feature vector similarity. To search their interest points, and then we can recommend all user groups to make the interest point.

This subsystem proposed the classification of the recommendation algorithm, and the specific steps are as follows.

*Algorithm 1: The recommendation algorithm based on the classification.*

Input: To recommend article  $RD$ , set the interest point feature vector as  $CF$  and the interest point main user cluster as  $CU$ .

Output: Recommend user group  $RU$ .

- 1)  $RU = \Phi$
- 2)  $Maxsim = 0$
- 3) Clustering  $ID = 0$
- 4) For every  $CF_i$  in  $CF$
- 5) If  $Sim(RD, CF_i) > Maxsim$
- 6)  $Maxsim = Sim(RD, CF_i)$
- 7) Clustering  $ID = i$
- 8) Endif
- 9) Endfor
- 10)  $RU = CU_{ClusteringID}$
- 11) Return

## 52.6 The Experimental Results and Analysis

This paper adopts a certain site after the preprocessing part of the real data set, including 714 users and 20,000 articles. From the experiment, extract 20 articles as the recommended article, use Recall and Precision as two quality indicators of recommendation system, the definition is as follows:

*Definition 1: For recommend system  $RS$ ,  $RD$  for to be recommended article, the  $RDU$  for the system recommend  $RD$  users set,  $SDU$  for collection  $RD$  users set, then recall is defined as:*

$$Recall = \frac{|RDU \cap SDU|}{|SDU|} \quad (52.3)$$

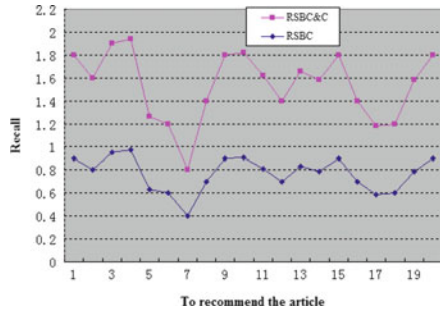
Precision is defined as

$$Precision = \frac{|RDU \cap SDU|}{|RDU|} \quad (52.4)$$

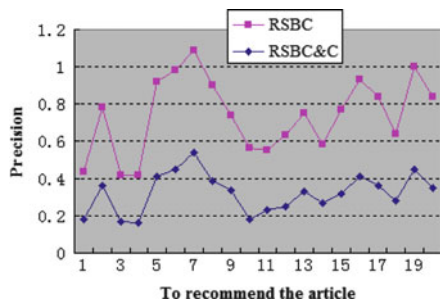
In order to test automatic personalized recommendation system (referred to as RSBC&C) based on clustering and classification performance, this paper made two



**Fig. 52.1** Recommendation system recall rate comparisons



**Fig. 52.2** Recommendation system precision rate comparisons



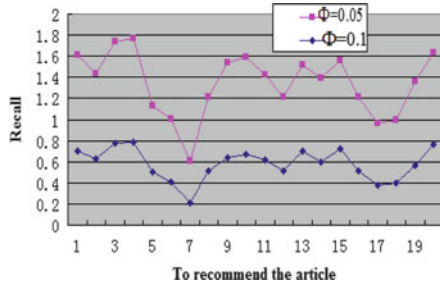
groups of experiment. The first group of experiments was based on content filter paper automatic recommendation system (referred to as RSBC system) and RSBC&C system comparison experiments. Set the interest point of value  $\Phi = 0.05$ . The experimental similarity calculation used is the cosine similarity calculation method. The experimental results were shown in Figs. 52.1 and 52.2.

The experimental results (Figs. 52.1 and 52.2) show that the RSBC&C system recall rate is higher than RSBC system; RSBC&C system article precision rate is slightly lower than RSBC system.

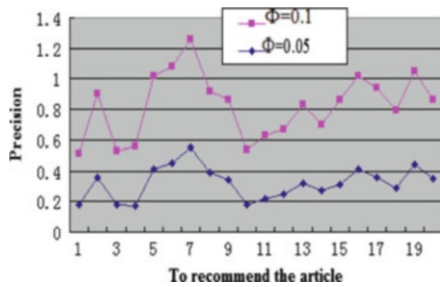
The second group tests interest degree of value  $\Phi$ , the effect of system performance. Interest degree of value  $\Phi$  is an adjustable parameter set at  $\Phi$  0.05 and 0.1, respectively. The experimental results were shown in Figs. 52.3 and 52.4.

The experimental results show (Figs. 52.3 and 52.4): when  $\Phi$  takes 0.05, the system recall rate is higher than  $\Phi$  take 0.1. But the system precision rates are lower than take 0.1 system precision accuracy. In the analysis of the experimental results, it can be concluded that the system's recall rate requires higher precision rate if it can properly lower  $\Phi$  value and the system's precision requires higher recall rate if it can appropriately increase  $\Phi$  value.

**Fig. 52.3** Different  $\Phi$  value recall rate comparison



**Fig. 52.4** Different  $\Phi$  value precision rate comparisons



### 52.7 Conclusion

This paper mainly aims at the Internet personalized information service of automatic recommendation system which has launched research. Automatic recommendation is one of most important ways of the Internet’s personalized information service. With more and more users online reading habits and collecting articles and the keyword vector user model dimension growing too fast, it is hard to reflect the user’s interest. Current recommendation techniques affect the effectiveness of the system. In view of the above questions, this paper puts forward the Internet’s automatic recommendation system based on clustering and classification of articles.

### References

1. Fan, G. (2005). Personalized information service in our country. *The Library in the New Century*, 5(5), 22–25.
2. Wu, L., & Liu, L. (2006). Personalized recommendation system user modeling technology. *Intelligence Review Journal*, 1(5), 55–62.
3. Shahabi, C., & Chen, Y. S. (2003). Recommendation system without explicit acquisition of use relevance feedback. *Distributed and Parallel Databases*, 14(3), 173–192.

4. Zeng, C., Xing, C., & Zhou, L. (2003). The personalized search algorithm based on content filtering. *Journal of Software*, 14(5), 999–1004.
5. Sawra, B., Karypis, G., Konstan, J., & Riedl, J. (2000). Analysis of recommender algorithms for e-commerce. In *Proceedings of the 2nd ACM E-Commerce Conference* (pp. 158–167). New York: ACM Press.

# Chapter 53

## Palmprint Recognition Based on Subclass Discriminant Analysis

Pengfei Yu, Haiyan Li, Hao Zhou, and Dan Xu

**Abstract** Subspace-based palmprint recognition methods, such as principal components analysis and linear discriminant analysis, assume that each class can be grouped in a single cluster. However, this assumption is not reasonable at some situations where a class is assembled in two or more clusters. In order to solve this problem, a novel palmprint recognition method based on subclass discriminant analysis is proposed in this chapter. Each palmprint class is divided into a set of subclasses that can be separated easily in the new subspace representation. After that, the Euclidean distance and nearest neighbor method are employed as the similarity measurement. Experimental results conducted on a database of 86 hands (10 impressions per hand) show that the equal error rate (EER) of the proposed method yields promising result of  $EER = 0.67\%$  for verification rate, which demonstrates that the proposed method is effective to solve the problem mentioned above.

### 53.1 Introduction

As palmprint is a relatively novel biometric trait, solutions in the field of palmprint recognition generally require extensive study involving algorithms and testing with sample images captured on different conditions with digital cameras, scanners, or other new devices. Although many algorithms of palmprint recognition have been developed during the past 10 years, the need of new method that is flexible, comprehensive, and robust is greatly urgent.

To date, there are numerous methods applied in palmprint recognition. Based on how to extract the palmprint feature, these algorithms can be broadly separated into three classes: line-based, subspace-based, and statistical methods.

---

P. Yu (✉) • H. Li • H. Zhou • D. Xu  
School of Information, Yunnan University, Kunming 650091, China  
e-mail: [pfyu@ynu.edu.cn](mailto:pfyu@ynu.edu.cn)

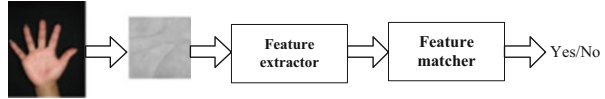
Line-based methods usually apply edge or point detectors to extract principle lines [1, 2] and crease [3, 4]. For instance, Huang et al. [2] used the modified finite Radon transform (MFRT) as the feature extraction approach to extract principle lines. Two images, direction image and energy image, were produced by the MFRT. After that, the principle lines could be extracted from the energy image and used to calculate the similarity between two palmprints. It is obvious that line is the basic feature of palmprint. Line-based methods directly extract the basic identity information contained in the texture of palmprint. However, the relatively stable texture of palmprint is difficult to be obtained because it is inconstant under different light condition or the stretching level of the hand. As a result, the recognition performance of the most line-based methods is not better than other ones.

Statistical methods usually subdivide the region of interest (ROI) of palmprint images into nonoverlapping small blocks. These blocks were subsequently transformed by wavelets [5, 6], Gabor [7], or other approaches [8, 9]. Finally, the means, variances, or energy of the small blocks were calculated and regarded as the features. For instance, Han et al. [8] used Sobel and morphological operators to process the ROI images. After that, these images were uniformly divided into several small grids whose sizes are  $32 \times 32$ ,  $16 \times 16$ , or  $8 \times 8$ . Lastly, the mean values of pixels in the grids were calculated to build the feature vectors.

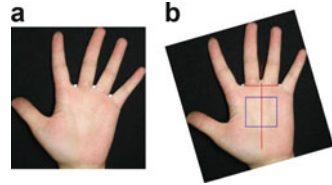
Subspace-based methods have been widely used in the literature of biometrics. Among subspace-based methods, principal components analysis (PCA) and linear discriminant analysis (LDA) are the famous methods employed in pattern recognition for data reduction and feature extraction. Consequently, they are applied in palmprint recognition. For example, Lu et al. [10] proposed a PCA-based palm recognition method named Eigenpalm. Wu et al. [11] used Fisher's linear discriminant (FLD) approach to extract palmprint features. However, these linear methods mentioned before are not effective to extract nonlinear features. In order to make these linear approaches applicable to extract nonlinear features, kernel-based methods have been proposed. The main idea of kernel-based methods is to map the input data to a feature space via nonlinear mapping, and the input data can be easily separated in the new feature space, for example, kernel principal component analysis (kernel PCA) [12] and kernel fisher discriminant analysis (KFDA) [13]. But if the data of a class is assembled into two or more clusters or cannot be represented by a single cluster, it is difficult to get a reasonable recognition performance by using previously mentioned subspace methods. To solve this problem, Zhu and Martinez [14] proposed a novel discriminant analysis method named subclass discriminant analysis (SDA).

In this chapter, we apply the SDA method to palmprint images for personal authentication. The rest of this chapter is organized as follows. In Sect. 53.2, we introduce the architecture of the proposed method. In Sect. 53.3 we briefly review the theory of SDA, and proposed SDA-based feature extraction method in Sect. 53.4. The experimental results are reported in Sect. 53.5. Conclusions are drawn in Sect. 53.6.

**Fig. 53.1** A block diagram of the proposed method



**Fig. 53.2** (a and b) the generation of palmprint ROI



## 53.2 System Overview

Figure 53.1 shows the architecture and work flow of our system.

First, we capture a colorful whole-hand image via a digital camera. To eliminate the variation caused by the rotation of the hand image, we rotate the hand image to a relative fixed angle. To do so, three key points, between the roots of index finger, middle finger, and ring finger, are located as shown in Fig. 53.2a. After the key points are detected, a coordinate system, based on the valley points, is established, as shown in Fig. 53.2b, and then a square region can be segmented from the center part of palm images as the palmprint ROI.

In order to facilitate matching with the identity template, the ROI images is further processed by a feature extractor to get a compact but expressive representation regarded as the identity of the hand's owner. If we use subspace-based methods, the feature extractor can be PCA based, LDA based, SDA based, etc. Then the representation is fed to the feature matcher, which compares it against the template of a user (verification system) or all templates of all users (identification system). It is note that Fig. 53.1 is a block diagram of verification system rather than an identification system. During the feature matching phase, simple method such as nearest neighbor method or complex classifiers such as support vector machine (SVM) and artificial neural network (ANN) can be applied.

## 53.3 Theoretical Foundation of SDA

In this section, we present the theoretical foundation of SDA and show that SDA is able to extract directly features from palmprint ROI images.

As we know, once the data distribution of each class has been approximated using a mixture of Gaussians, most of subspace-based methods are suitable to be selected as the feature extractor.

Assuming that each class has been approximated using a mixture of Gaussians, it is easy to use the following generalized eigenvalue decomposition equation to find those discriminant vectors that can best classify the data [14]:

$$\Sigma_B V = \Sigma_X V \Lambda \tag{53.1}$$

where  $V$  is a matrix whose columns correspond to the discriminant vectors,  $\Lambda$  is a diagonal matrix of the corresponding eigenvalues, and  $\Sigma_B$  is the between-subclass scatter matrix, given by

$$\Sigma_B = \sum_{i=1}^{C-1} \sum_{j=1}^{H_i} \sum_{k=i+1}^C \sum_{l=1}^{H_k} p_{ij} p_{kl} (\mu_{ij} - \mu_{kl})(\mu_{ij} - \mu_{kl})^T, \tag{53.2}$$

where  $C$  is the number of classes,  $H_i$  is the number of subclass divisions in class  $i$ ,  $p_{ij}$  and  $\mu_{ij}$  are the prior and mean of the  $j$ th subclass in class  $i$ , and  $p_{kl}$  and  $\mu_{kl}$  are the prior and mean of the  $l$ th subclass in class  $k$ . Note that  $p_{ij} = \frac{n_{ij}}{n}$ , where  $n_{ij}$  represents the number of the  $j$ th subclass in class  $i$  and  $n$  represents the number of samples.

$\Sigma_X$  is the covariance matrix of the data, defined as [15]

$$\Sigma_X = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)(x_i - \mu)^T \tag{53.3}$$

where  $x_i$  is the  $i$ th sample vectors and  $\mu$  is the sample mean of all the data.

Since  $\Sigma_B$  measures the scatter of the subclass means, the method is called SDA. Compared with LDA, SDA can be used to classify both linear and nonlinearly separable classes. Moreover, LDA can only extract  $C - 1$  features from the original feature space, but SDA can extract features larger than  $C - 1$ .

Note that the real challenge of this method is how to find the best subclass divisions which can be applied to get the optimal classification results [14]. To achieve this, Zhu and Martinez [14] proposed two approaches: leave-one-out-test (LOOT) criterion and stability criterion. Using the LOOT criterion, all possible values of the number of subclass  $H_i$  are exhaustively tested. At the same time, the recognition rate of the specified  $H_i$  is recorded. The optimal value of the number of subclass corresponds to the max recognition rate. When the optimal subclass is selected, the between-subclass scatter matrix  $\Sigma_B$  can be computed by Eq. (53.2), so the discriminant vectors can be obtained by solving Eq. (53.1). In this study, we choose LOOT criterion to find the optimal value of the number of subclass. More details about the stability criterion can be referred in Zhu and Martinez [14].

### 53.4 Feature Extraction

Feature extraction is an essential but difficult step in the proposed system. In order to effectively implement the proposed palmprint recognition, SDA is chosen to extract the features from the palmprint ROI images.

Assume that there are  $c$  persons and the number of ROI images belonging to the  $i$ th person is  $n_i$ . Let the  $j$ th ROI images of class  $i$  be  $x_{ij}$ , where  $1 \leq i \leq c$ ,  $1 \leq j \leq n_i$ , and  $n = \sum_{i=1}^c n_i$ .

The purpose of SDA is to maximize the following Fisher criterion:

$$J(W) = \frac{W^T \Sigma_B W}{W^T \Sigma_X W}, W \neq 0 \quad (53.4)$$

where  $\Sigma_B$  is the between-subclass scatter matrix defined by Eq. (53.2) and  $\Sigma_X$ , defined by Eq. (53.3), is the covariance matrix of the data in the feature space  $F$ .

The optimal discriminant vectors  $w_i$  under Fisher criterion can be achieved by solving the following eigenvalue problem defined by Eq. (53.1).

Furthermore, if the set of generalized eigenvectors  $w_i$  are composed of the optimal linear transformation matrix  $W_{opt}$ , then the ROI images can be transformed to the feature space by using  $y_{ij} = W_{opt}^T x_{ij}$ ,

where  $y_{ij}$  is the extracted ROI feature from  $x_{ij}$  by using the SDA.

During the feature matching phase, the Euclidean distance and nearest neighbor method are employed as the similarity measurement.

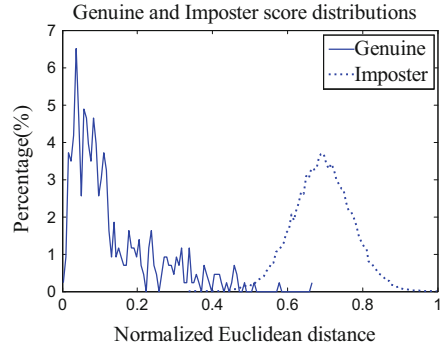
### 53.5 Experiment Results

In this study, the YNU Hand database is chosen to verify the SDA-based palmprint recognition method. The data set used in our experiments consists of 860 images, 10 per hand. The center area of the palm is obtained as the ROI during the preprocessing phase. After that, five ROI images per hand are selected for training, and the remaining five ROI images are chosen for testing. The size of palmprint ROI is  $200 \times 200$ .

Generally, the performance of a biometric system is often measured in terms of false accept rate (FAR) and false reject rate (FRR). However the FAR and FRR cannot be decreased simultaneously. For this reason, equal error rate (EER) is usually chosen as the quality indices because both of them can combine FAR and FRR into a number. In this study, we choose the EER, which refers to the point in a detection error trade-off (DET) curve where FAR equals the FRR.



**Fig. 53.3** The corresponding genuine and imposter distributions



**Fig. 53.4** The FAR and FRR of the proposed method

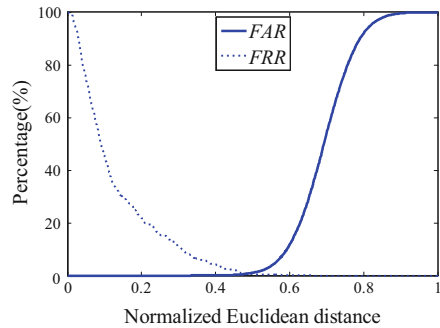
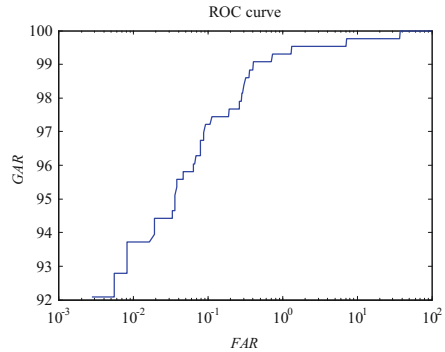


Figure 53.3 shows the performance evaluation for genuine and imposter matching distribution, which is estimated by matching scores obtained by using normalized Euclidean distance. The genuine matching score is gained by comparing palmprints from the same palm, and the imposter matching score is obtained by comparing palmprints of two different palms. Consequently, a total of 36,980 ( $5 \times 86 \times 86$ ) comparisons are performed, in which 430 ( $5 \times 86$ ) comparisons are genuine matching, and the remaining 36,550 ( $5 \times 86 \times 85$ ) comparisons are imposter matching. It is clear in Fig. 53.3 that the proposed method can separate palmprints of different persons well.

A threshold value 0.487 is determined based on the EER criteria. The FAR and FRR at different distance thresholds are plotted in Fig. 53.4, in which the EER is 0.67 %.

Figure 53.5 shows the ROC curve obtained from the proposed method. It must be noted that we plot FAR as logarithmic scales for  $x$ -axis for viewing the curve clearly.

**Fig. 53.5** ROC curve for the proposed method



## 53.6 Conclusion

Over the years, many palmprint recognition approaches have been proposed. Among these palmprint recognition approaches, subspace-based methods, such as PCA and LDA, have been widely used. Most of them assume that each class is represented by a single cluster in the feature space. However, this assumption is not reasonable at some situations in which the data of a class is assembled into two or more clusters. So the SDA is proposed to solve this dilemma. In this chapter, a new subspace-based method named SDA is applied in palmprint feature extraction. The experimental results on YNU Hand database show that the SDA-based method yields a good recognition rate.

**Acknowledgments** Supported by Science Research Foundation of Yunnan Provincial Education Department (Grant No. 2011Z029).

## References

1. Wu, X., Wang, K., Zhang, D., & Huang, B. (2004). Palmprint classification using principal lines. *Pattern Recognition*, 37(10), 987–1998.
2. Huang, D., Jia, W., & Zhang, D. (2008). Palmprint verification based on principal lines. *Pattern Recognition*, 41(4), 1316–1328.
3. Chen, J., Zhang, C., & Rong, G. (2001). Palmprint recognition using crease. *Proceeding of 2001 International Conference on Image Processing, IEEE in Piscataway, NJ* (pp. 234–237).
4. Cook, T., Sutton, R., & Buckley, K. (2010). Automated flexion crease identification using internal image seams. *Pattern Recognition*, 43(3), 630–635.
5. Liu, Y. Q., Yuan, W. Q., & Guo, J. Y. (2010). Block statistic under wavelet decomposition for palmprint recognition. *Proceedings of 2010 International Conference on Electrical and Control Engineering (ICECE), IEEE CS, Piscataway, NJ* (pp. 5073–5075).

6. Madasu, H., Gupta, H. M., Mittal, N., & Vasikarla, S. (2009). An authentication system based on palmprint. *Proceedings of 6th International Conference on Information Technology: New Generations (ITNG '09), IEEE CS, Piscataway, NJ* (pp. 399–404).
7. Kumar, A., & Shen, H. C. (2004). Palmprint identification using PalmCodes. *Proceedings of the 3rd International Conference on Image and Graphics, IEEE CS, Piscataway, NJ* (pp. 258–261).
8. Han, C. C., Cheng, H. L., Lin, C. L., & Fan, K. C. (2003). Personal authentication using palmprint features. *Pattern Recognition, 36*(2), 371–381.
9. Wu, X., Wang, K., & Zhang, D. (2002). Fuzzy directional element energy feature (FDEEF) based palmprint identification. *Proceedings of 16th International Conference on Pattern Recognition, IEEE CS, Piscataway, NJ* (pp. 95–98).
10. Lu, G., Zhang, D., & Wang, K. (2003). Palmprint recognition using eigenpalms features. *Pattern Recognition Letters, 24*(9), 1463–1467.
11. Wu, X., Zhang, D., & Wang, K. (2003). Fisherpalms based palmprint recognition. *Pattern Recognition Letters, 24*(15), 2829–2838.
12. Schölkopf, B., Smola, A., & Müller, K. (1998). Nonlinear component analysis as a Kernel eigenvalue problem. *Neural Computation, 10*(5), 1299–1319.
13. Mika, S., Rätsch, G., Weston, J., Schölkopf, B., & Müller, K. R. (1999). Fisher discriminant analysis with kernels. In *Neural networks for signal processing IX, 1999. Proceedings of the 1999 I.E. Signal Processing Society Workshop, IEEE, Piscataway, NJ* (pp. 41–48).
14. Zhu, M., & Martinez, A. M. (2006). Subclass discriminant analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 28*(8), 1274–1286.
15. Martinez, A. M., & Zhu, M. (2005). Where are linear feature extraction methods applicable? *IEEE Transactions on Pattern Analysis and Machine Intelligence, 27*(12), 1934–1944.

# Chapter 54

## A Process Quality Monitoring Approach of Automatic Aircraft Component Docking

Guowei Yang, Chengjing Zhang, and Xiaofeng Zhang

**Abstract** In order to evaluate automatic aircraft docking quality, a new method of process quality monitoring of the automatic aircraft component docking is proposed. This method is based on laser tracker measurement. The position of automatic aircraft docking component is obtained by laser tracker measurement for a plurality of feature points. By doing identification labels on the surface of docking components, component surface images replace complex docking images. Image processing technology is used to extract the features of component surface image information to evaluate docking quality so as to control the perfect match of docking component. The new method enforces automatic docking process visual monitoring and absorbs human visual and reliable monitoring advantages of manual assembly model.

### 54.1 Introduction

With the development of computer and information technology, in order to adapt to the social progress and meet its quality and efficiency, aircraft assembly has been from the traditional manual assembly mode developed gradually and quickly into today's automatic assembly model, which is fast efficient digital and integrated and has high precision and high coordination [1]. Aircraft component has the features of large-scale, complex matching surface and difficult measurement. Domestic and foreign advanced technology on component about pose measurement and accurate calibration is based on laser tracker, providing reliable basis for large parts of the pose adjustment and position driver [2, 3]. Although the automatic docking technology has made great progress, the traditional manual assembly model is not

---

G. Yang • C. Zhang (✉) • X. Zhang  
School of Information Engineering, Nanchang Hangkong University,  
Nanchang 330063, China  
e-mail: [jingshui0326@126.com](mailto:jingshui0326@126.com)

completely replaced. To a great extent, automatic docking technology did not achieve the simulation monitoring of human vision.

The method of process quality monitoring refers to the automatic docking process about large aircraft components that not only relies on the data returned by pose control system but also uses effective scheme to enforce docking process visual monitoring and evaluate whether or not the docking process is completed. If the evaluation result is up to standard, the docking process quality is good; if not, the need to adjust the position and pose is necessary.

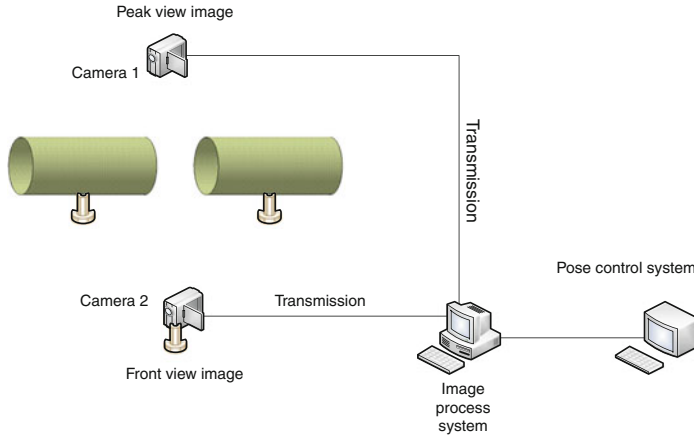
## **54.2 The Measurement Method of Laser Tracker**

In assemble process of laser tracker measuring the position, after the large aircraft component shelving, target balls should be respectively arranged, and laser tracker should be placed in front of the docking component. In adjusting process, in order to achieve the current standard point automatic tracking and positioning measurement, laser trackers provide second-development interface, through programming to set search area based on putting the standard point into search center; then discrete laser signal to the theory coordinates of standard point; and use their automatic search function to search the target balls around the space of the standard points. Then we can theoretically obtain the theory coordinates of the next standard point. So we can measure the second standard point coordinates by using the same method. In this turn, every dynamic coordinate of the standard point in adjustment progress can be measured.

A few problems will appear above. For saving the cost, assembly project usually uses a single laser tracker; this way, at the same time only a standard point (target) can be achieved, and its measure time is different. In order to solve the docking pose, standard points' measure time must be unified. Moreover, assembly parts are complex, so errors in assembly would not be avoided [2]. This chapter proposes a new docking quality assessment method, which can effectively lessen the deviation and reduce the influence on subsequent docking assembly.

## **54.3 Docking Component Surface Monitoring Method Based on Image**

The measurement method of laser tracker adjusts and positions posture with digital coordinates which is obtained from computer feedback of control system. For the assembly filed, the method of process quality monitoring of automatic component docking based on image improves a new level about perfect docking.

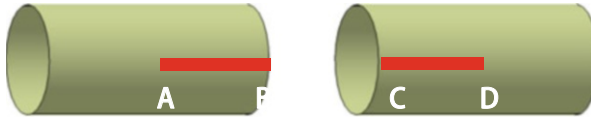


**Fig. 54.1** Process quality monitoring system diagram of automatic aircraft component docking

### 54.3.1 Image Monitoring Method

To realize process quality control based on visualization, an ideal way is to capture surface image. General docking is inserted, namely, component's several main shafts inserted into another's corresponding holes precisely. Thus, as long as capturing hold images and extracting hold center and shafting coordinates with algorithm, we can contrast these two coordinates. If the coordinates contrast result is coincident, the docking result is perfect. If it is deviation, position adjustment is necessary. But considering components' complex surface and the practice without camera condition, the project cannot be carried out successfully. In order to overcome these problems, a new method is proposed. As follows, Fig. 54.1 shows process quality monitoring system diagram of automatic component docking based on image. Using CCD cameras with high quality to capture surface images and achieving quantitative calculation through image processing algorithm can evaluate docking quality. Camera 1 placed in the peak of docking captures peak view images. Camera 2 has the same height with components' horizontal axis and captures front-view images. The two cameras transmit the images to image process system synchronously. The computer uses algorithm to achieve image information for quality evaluation. Then the evaluation results are transmitted to pose control system for guiding position adjustment.

Because aircraft components are rigid and manufacturing process is controlled strictly and every step of project by means of precise calculation, it is feasible to use component surface quality to assess the ensemble docking quality.



**Fig. 54.2** Front-view image of label location

### 54.3.2 Label Location Method

Aircraft components' surfaces are very complex. Either fuselage-to-fuselage or fuselage-to-wing docking, there are different surface characteristics. It's difficult to extract surface images' features, and feature extraction algorithm of the fuselage-to-fuselage docking is not applicable for the fuselage-to-wing docking process; conversely, it is the same. Based on these problems, this chapter proposes a new feature extraction method—label location method.

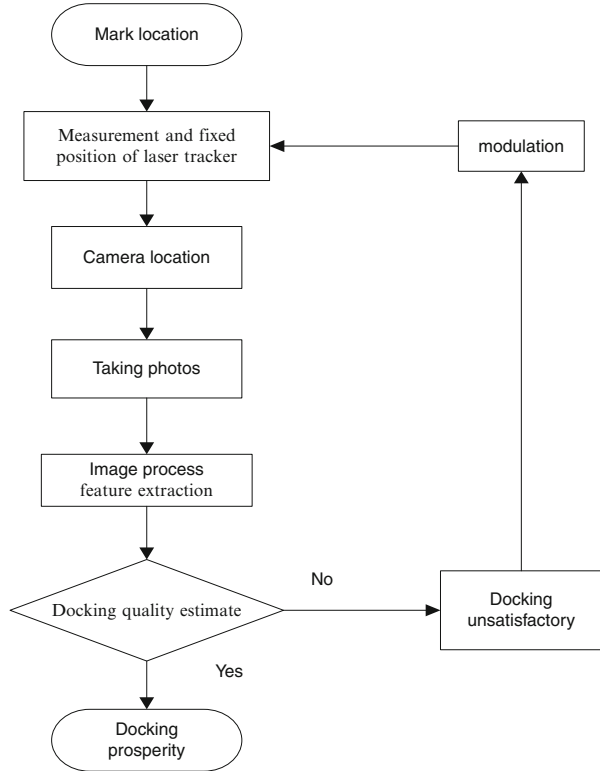
Firstly, feature extraction chooses feasibility labels—red lines which have obvious color difference between the labels and background. The labels are placed in two components' front and peak mid sides. The front images captured by camera 1 are shown in Fig. 54.2. Similarly, the camera 2 captures these peak view images. This chapter only takes the front-view image as an example.

## 54.4 The Flow Chart of the Process Quality Monitoring System

In this chapter, the quality monitoring on aircraft component docking process is based on laser tracker positioning system. Figure 54.3 shows the flow chart of the system.

Before assembly, specified labels should be placed on docking component surfaces, according to label location method. By the formulation of the program, laser tracker control system controls implementation of component docking automatically. Position control system controls the components' posture drawing on the return data from laser tracker. When the docking components get to predetermined position, the process quality monitoring system starts up. Calibrated camera 1 and 2 transmit images to process quality monitoring system in the same time. Through image processing, it extracts target image information and then evaluates whether or not the quality in docking process satisfies the standard.

**Fig. 54.3** Process quality monitoring system flow diagram



## 54.5 Experiment Studying

### 54.5.1 The Color Segmentation of the Advantages

In this chapter, in order to facilitate identification [4], we make red line for distinguishing labels and components obviously. In RGB space, color feature is used to extract target designations from images. First thing is to display images' color, namely, original images are displayed in R, G, and B three independent spaces, with gray image expression, as shown in Fig. 54.4.

Of R, G, and B three spaces grayscale, in RGB image displaying, red, green, and blue ingredients have different values, and their effects are also different. In R space, the red ingredient on the line target salience is very obvious. In G and B spaces, ingredients that are changing are not obvious. Through the experiments, only based on R spatial gray level, the extraction of target label line is not very ideal. By the methods of refs. 4, 5, the use of RGB space (2R-G-B) [4, 5] gray difference can make further processing Fig. 54.5 is the (2R-G-B) gray difference image.



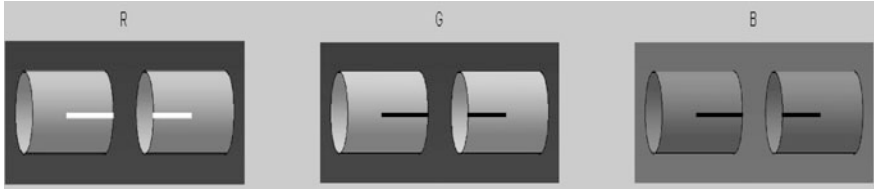
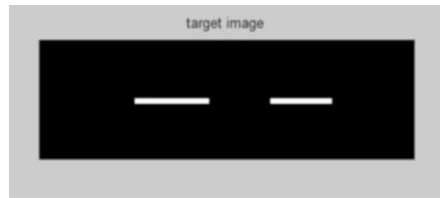


Fig. 54.4 R, G, B gray image

Fig. 54.5 (2R-G-B) color gray image



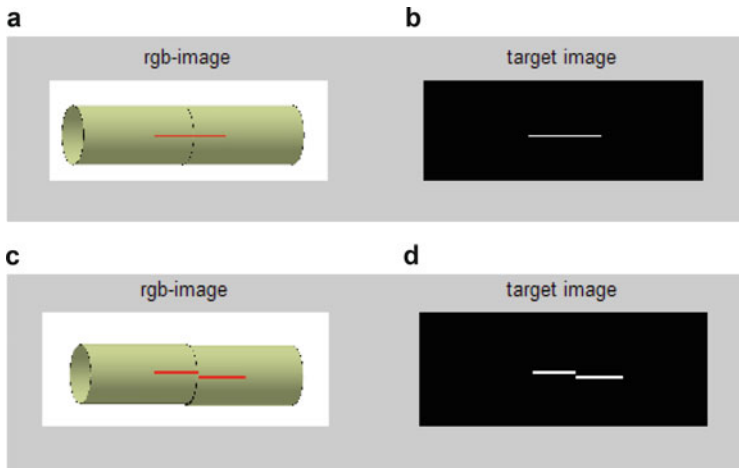
Fig. 54.6 Target binary image



### 54.5.2 Otsu Adaptive Threshold Segmentation

Otsu is a good kind of performance of automatic threshold segmentation method [6] through calculating image's target and background class within the minimum of within-class variance, the maximum of between-class variance, for automatic threshold calculation. According to grayscale of image features, if the image has larger variance between foreground and background, the two portions of the image have bigger difference. Whenever foreground ingredients are divided into wrong background or backgrounds are divided into wrong prospects, the two ingredients' difference becomes smaller, so the maximum between-class variance segmentation means that probability of wrong points is minimum variance.

The synthesis images of (2R-G-B) gray image by means of the Otsu adaptive threshold segmentation can quickly and effectively distinguish between targets and background. As Fig. 54.6 displays, getting adaptive threshold segmentation of target image can get ideal target binary image.



**Fig. 54.7** Verified experiment of simulative assembly. (a) Original image of docking prosperity (b) Target image of docking prosperity (c) Original image of docking unsatisfactory (d) Target image of docking unsatisfactory

**Table 54.1** *ABCD* coordinate points of docking target image in Fig. 54.7

Target image	Axis	A	B	C	D
b	X	379	520	521	638
	Y	197	197	197	197
d	X	367	508	509	661
	Y	185	185	198	198

### 54.5.3 Verified Experiment

In order to verify feasibility, this experiment extracts the result images among ideal and not ideal docking assembly on the CATIA simulation. By the proposed control method, we got the processing image [7] on MATLAB platform and finally got the target binary images, as shown below in Fig. 54.7.

The above experiments are hypothesized to extract the images captured by camera 1. The results have shown a good target extraction within simulated docking assembly. According to target binary images, the result image which is not ideal docking displays that the label line is obviously wrong and describes that the docking process quality has obvious problem so that we need to adjust the docking pose; result image of docking success displays a straight line, and if the image of camera 2 also has the same result, it realizes ideal docking.

From Table 54.1, we can learn that the target image (b) shows the docking prosperity, because the *ABCD* coordinate points have the same *Y* coordinates, and

$B$  and  $C$  have the regular  $X$  coordinates. Target image (d) has different  $Y$  coordinates; obviously, it shows an unsatisfactory docking. In order to modulate pose, pose control system should apply instrumentation code to control components to move  $|Y_B - Y_C|$  and to obtain same  $Y$  coordinates.

So the above experiments achieve process quality evaluation and verify the feasibility of docking process quality monitoring system.

## 54.6 Conclusion

In past studies, the research of process quality monitoring for automatic aircraft component docking is very few, and general evaluation methods examine whether or not the data returned by laser tracker meets tolerance standard. The new method has the advantages of non-contact and achieves visual docking process quality monitoring, which is based on laser tracker pose measurement and calibration. It enforces aircraft assembly automation to a higher level.

**Acknowledgments** This program is financially supported by the Natural Science Foundation of China (No. 61272077), Natural Foundation of Jiangxi Province (No.20114BAB201034), Research Foundation of Jiangxi Province Education Department (No.GJJ12413), and Key Laboratory Open Foundation of Jiangxi Province (No.TX201204003).

## References

1. Mei, Z. Y., & Fan, Y. Q. (2009). Docking flexible assembly technology based on laser tracking and locating parts. *Journal of Beijing University of Aeronautics and Astronautics*, 35(1), 68–69.
2. Zhu, Y. G., Huang, X., & Fang, W. (2012). Medium fuselage position and attitude adjustment and tracking measurement. *Journal of Mechanical Science and Technology*, 31(7), 1121–1127.
3. Renliang Yu. (2008). *CATIA V5 basic tutorial* (pp. 166–169). Beijing: China Machine Press.
4. Li, Y., Ji, C., Wang, H., & Zhao, W. (2012). The image segmentation method of apple tree branches based on Matlab. *Science Technology and Engineering*, 12(13), 55–59.
5. Cai, J. R., Zhou, X. J., Li, Y. L., & Fan, J. (2008). The mature citrus recognition based on machine vision natural scenes. *Journal of Agricultural Engineering*, 24(1), 175–178.
6. Jichao, W. (2010). *Object extraction recognition algorithm of machine vision image*. Hebei: Agricultural University of Hebei.
7. Gonzalez, R. C., & Woods, R. E. (2011). *Digital image processing* (3rd ed., pp. 285–307). Beijing: Publishing House of Electronics Industry.

# Chapter 55

## Overhead Transmission Lines Sag Measurement Based on Image Processing

Wengang Cheng and Long Chen

**Abstract** Sag is one of the important parameters for operation and maintenance of the transmission lines, and its size directly affects the safe and stable operation of the line. In recent years, in order to improve the transmission capacity, many existing transmission lines allow the temperature from 70 °C to 80 °C, and then the transmission line sag becomes a major constraint for the transmission security. This chapter presents a novel sag measurement method based on image processing. Firstly it grays the collected color images and preprocesses the images with some image-denoising methods. Secondly, special points generated by the isolation rod are extracted by the corner extraction algorithm, and the spatial coordinate values of the extracted points are identified according to the principle of binocular vision and the relationship of three-dimensional coordinate space coordinates and image coordinates. Finally via the method of the curve fitting, the actual sag of the transmission line is calculated. The experimental results show that this method is suitable for both the cases in which the height of the transmission line is equal or not, and it has good adaptability.

### 55.1 Introduction

Transmission line sag is one of the main parameters of the circuit design and operation and maintenance, and it directly affects the safety and stable operation of the transmission line. The sag will change with the transport capacity, weather condition change, and line increase of service life, and sometimes the sag will be beyond the design requirements that may lead to the security risks [1]. So it is very necessary to measure the transmission line sag regularly.

---

W. Cheng • L. Chen (✉)

School of Control and Computer Engineering, North China Electric Power University, Beijing 102206, China

e-mail: [wgcheng@ncepu.edu.cn](mailto:wgcheng@ncepu.edu.cn); [happy1988@163.com](mailto:happy1988@163.com)

At present, a lot of methods and applications have been brought up on the measurement of the power line sag by domestic and foreign research institutions and companies. It mainly contains two cases: they are the traditional manual measurement in the field and using certain devices to implement the measurement. The traditional manual measurement methods include the different length measurement method, the same length measurement method (parallelogram method), the angle measurement method, and so on. These methods that use the actual observed data and combine with mathematical formulas can calculate the sag more accurately. But the traditional sag measurement methods have low efficiency and they are sensitive to the geographical environment. Besides, it is also difficult to conduct real-time monitoring. So a lot of real-time automatic monitoring power line sag methods have been proposed and have had a very good application. Literature [2, 3] proposed monitoring power line stress and temperature-based methods to measure the sag. The literature [4, 5] proposed the sag measurement method based on the global positioning system technology. These methods have very good application in practice, such as The Group Inc. in the USA that produced the CAT-1 that calculated sag by measuring the transmission line stress.

This chapter describes the real-time measurement of sag in an overhead power transmission line by using the image-processing method. We just need to extract the feature points of the image that is captured according to the image corner extraction methods and then calculate the spatial coordinates of the curve feature points in accordance with the principle of binocular vision and the relationship of the world coordinate system and the image coordinate system. Finally, we recover the curve equation by curve fitting to calculate the sag. This method is easy to operate, and it has a good real-time. Besides, the calculated results are very accurate and meet the applied requirements.

## 55.2 The Principle of Sag Measurement

### 55.2.1 *The Process of Sag Measurement*

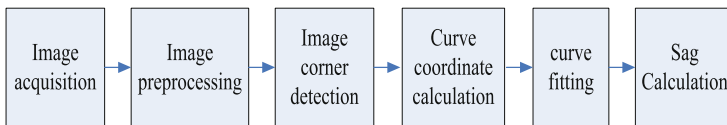
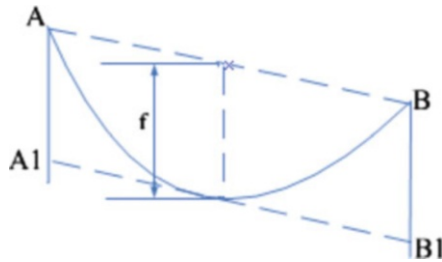
Overhead line curve refers to the arc-shaped curve that is supported by both the end towers. Any point plumb distance to both ends of the suspension point connection line on the transmission line is called the overhead line sag of the point.

Overhead line sag is denoted by  $f$  generally. When both ends of the suspension point of the transmission line have the same height, the maximum sag is in the center of the span, as shown in Fig. 55.1. When the height of the two end points of the line is not equal, as shown in Fig. 55.2, the maximum sag is the plumb distance between the cutoff point that is the straight line A1B1 that parallels the connection line of the two point tangent to the curve and the point in the connection line. That is the cutoff point sag of the parallelogram ABB1A1.

**Fig. 55.1** The model of the unequal height transmission line



**Fig. 55.2** The model of the equal height transmission line



**Fig. 55.3** The flow chart of sag calculation

As the camera shoots the same scene several times, we can get the different coordinate values of the feature point that is the same point in the different images. The space coordinates of the feature points of the curve can be calculated according to the principle of binocular vision as well as the relationship of the world coordinate system and the image coordinate system. Finally we calculate the sag according to the fitting curve equation. The flow chart of computing transmission line sag is shown in Fig. 55.3.

### 55.2.2 Harris Corner Detection Algorithm

Corner is an important local feature of the images. In the actual image, corners such as contour inflection point, the end of the segment corner, have a rich amount of information, and they are not only very easy to be measured but also able to adapt to changes in ambient light. Most corner detection methods detect the image point with specific characteristics rather than just the “corner.” These feature points often have some digital features, such as local maximum or minimum gray and some gradient feature. The corner detection algorithm can be summarized into three categories. They are corner detection based on gray-scale images, binary image corner detection, and corner detection based on the profile curve. Corner detection

based on gray-scale image can be divided into three types. They are gradient-based, template-based, and template gradient combination.

Harris corner detection algorithm is a typical corner detection based on gray-scale image. It contains the following detection steps:

1. Calculate the directional derivative of the image, save for two arrays  $I_x$  and  $I_y$ .
2. Calculate the localized autocorrelation matrix  $U(x, y) = [I_x(x, y)^2 * W \ I_y(x, y) I_x(x, y) * W; I_x(x, y) I_y(x, y) * W \ I_y(x, y)^2 * W]$  for each point; here  $*W$  refer to that the  $x, y$  as the center do convolution with Gaussian template  $W$  and the size of the template needs to be specified.
3. If the two eigenvalues of  $u$  are very small, it indicates that this region is a flat region. If one of the eigenvalues of  $u$  is larger and the other is smaller, it is the line. If both of the eigenvalues are larger, then it shows that this is a corner point. Harris provides another formula to get an evaluation that whether the current point is the corner point:  $corn = \det(u) - k * \text{trace}(u)^2$ . Here the  $corn$  is referred to the point value and  $k$  is a fixed variable, typically between 0.04 and 0.06.

Since the Harris corner operator has good robustness, rotational invariance, and many other advantages, it can extract the feature points of the transmission line curve easily. The position of the spacer rods in the transmission line will be considered the corner, and we can get all positions of the spacer rods in every image, so we can obtain the position coordinates of the curve on the same physical point for each different image.

### 55.2.3 Camera Calibration and the Three-Dimensional Coordinate Measuring Algorithm

According to the principle of camera imaging, we can conclude the relationship between the coordinate value  $(u, v, 1)$  of the point P in the pixel coordinate system and the coordinate value  $(X_w, Y_w, Z_w)$  of the point P in the world coordinate system. The conversion formula is as follows:

$$Z_c \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_1 & 0 & u_0 & 0 \\ 0 & f_2 & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0^T & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = M_1 M_2 \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} \tag{55.1}$$

$(u_0, v_0)$  stands for the coordinate value of the origin of the image coordinate system in the pixel coordinate system whose origin is the upper left corner of image plane; matrix  $R$  stands for the camera rotation matrix; matrix  $T$  stands for the camera translation matrix.

According to Eq. 55.1, we can figure out the computer calibration parameters and calculate the spatial point coordinate value by combining the binocular stereo

vision principle [6]. According to the principle of binocular vision, we need to take two photos of the same scene and calculate other information by using the relationship of the two photos. The specific applications are as follows:

### 1. Calculating the camera calibration parameters

After analyzing Eq. 55.1, we can learn that we just need to calculate the parameters  $f_1$  and  $f_2$ . In order to facilitate the calculation, we assume that the first camera coordinate system coincides with the world coordinate system and the second camera translates  $L$  distance along the  $X$ -axis. Besides, both of the two cameras are with no rotation. That is to say, the matrix  $R$  is a unit matrix and the matrix  $T = [L \ 0 \ 0]^T$ . Above all, we will be able to get the two equations about  $f_1$  and  $f_2$  and can get the following results:

$$f_1 = \frac{Z_w(u_1 - u_2)}{L} \quad (55.2)$$

$$f_2 = \frac{Z_w(v_1 - v_2)}{L} \quad (55.3)$$

$(u_1, v_1)$  and  $(u_2, v_2)$  stand for the coordinate values of the same point in the two images.

### 2. Calculating the space coordinate values of the feature points

According to Eq. 55.1, the calculated camera calibration parameters  $f_1$  and  $f_2$ , as well as the camera rotation matrix  $R$  and translation matrix  $T$  that can be obtained in the actual shooting operation, we can get four linear equations about  $X_w$ ,  $Y_w$ , and  $Z_w$  by taking two images, so we can work out the feature point  $P$  coordinate value  $(X_w, Y_w, Z_w)$  of the world coordinate system.

## 55.3 Experiments

In the process of calculation of the camera calibration parameter, we capture several images at five different locations and take 20 images in each location. Comparing the recovery coordinate value with the actual coordinate values, we can conclude the maximum error that mainly affected by the camera calibration parameter in each direction. The maximum error in  $X$ -,  $Y$ -, and  $Z$ -direction are  $-2.25$ ,  $2.98$ , and  $2.15\%$ , respectively.

The experimental result of using Harris corner extraction algorithm to detect the spacer bars of the transmission lines is shown in Fig. 55.4 (the picture has been rotated  $15^\circ$  clockwise). Experiments show that using the corner extraction algorithm to detect spacer bars is more effective than using image matching method to detect the spacer bars, especially for those blurred images that are captured on the long shooting distance.



**Fig. 55.4** The experimental results of corner detection



Comparing the measurement data that are obtained by using image-based corner detection method to measure the sag with the measurement data that is measured by the field theodolite, except that individual data error is larger, most data error were less than 2.5 %. The measurement accuracy meets the applied requirements.

## 55.4 Conclusion

Through analysis of the chapter and related experimental results, we can get the following conclusions.

1. Based on the characteristics of the transmission line and combined with image processing technology, it uses the Harris corner extraction algorithm to detect spacer bars of the transmission line and computes transmission line sag with the curve information. According to the experimental results, we can conclude that this method is simple and has a good adaptability.
2. The results of the sag calculation will be affected by the camera calibration accuracy, position of the camera, and shooting angle. Therefore, it is our future research work to improve the accuracy of the camera calibration and select the appropriate coordinate point and shooting angle. That is to say, we should choose the shooting position that can capture the image as much as possible and the shooting angle that is easy to calculate such as  $30^\circ$  and  $45^\circ$ .
3. Due to the complexity of the environment, it is very important for the accuracy of the experiment to remove the image noise. This chapter uses the median filtering denoising algorithm for image denoising, which can overcome the problem of the fuzzy edge of the low-pass filtering and the noise enhancement of the high-pass filtering. Besides, this chapter also uses the image enhancement function that is from the MATLAB function library to highlight the edge of the transmission line.

**Acknowledgements** This work is supported by “the Fundamental Research Funds for the Central Universities” of China.

## References

1. Xu, Q. S., Ji, H. X., & Wang, M. L. (2007). Monitoring sag of the transmission lines real-time. *High Voltage Engineering*, 33(7), 206–208 (In Chinese).
2. Chen, S. X., Wu, P. S., Sun, Y. T., et al. (2008). Application of on-line temperature and sags monitoring device of transmission line. *North China Electric Power*, (3), 14–15.
3. Muhr, M., Pack, S., & Jauffer, S. (2005). Sag calculation of aged overhead lines. In *Proceedings of the 16th International Symposium on High Voltage Engineering*. Tsinghua University, Beijing, China (pp. D-14).
4. Mensah-Bonsu, C., Krekeler, U. F., Heydt, G. T., et al. (2002). Application of the global positioning system to the measurement of overhead power transmission conductor sag. *IEEE Transactions on Power Delivery*, 17(1), 273–278.
5. Mahajan, S. M., & Singareddy, U. M. (2008). Real time GPS data processing for 'sag measurement' on a transmission line. In *Power System Technology and IEEE Power India Conference, New Delhi*. IEEE (pp. 1–6).
6. Deng, Z. D., Niu, J. J., & Zhang, J. D. (2007). Three-dimensional modeling approach based on stereo vision. *Journal of System Simulation*, 19(14), 3258–3262.

# Chapter 56

## Chinese Domain Ontology Learning Based on Semantic Dependency and Formal Concept Analysis

Lixin Hou, Shanhong Zheng, Haitao He, and Xinyi Peng

**Abstract** The ontology construction process is very expensive and time consuming when performed manually. In order to solve the problems of time consumption and high cost, a Chinese domain ontology learning method based on semantic dependency and formal concept analysis is proposed in this chapter. During the learning process, semantic dependency analysis technology is used for extracting formal context from unstructured domain texts, and Godin algorithm is used for constructing concept lattice. At last, the chapter takes a medical ontology construction as an example to verify this method. The experiment results show that the method we proposed can construct domain ontology automatically and reduce manual intervention. In addition, the ontology we got is a formal ontology, so it has more advantages in sharing and reusability.

### 56.1 Introduction

Ontology learning is a subtask of information extraction, and its purpose is to semiautomatically extract relevant concepts and relations from a given corpus or other kinds of data sets to form ontologies. Ontology learning can solve the problems of low efficiency and time consumption when performed manually. Therefore, ontology learning is becoming an active research area.

Ontology learning focuses on the extraction of domain concepts and relations among them automatically [1]. There are many methodologies for concept extraction. Extraction based on linguistics has advantages in terms of disambiguation and

---

L. Hou (✉) • S. Zheng • H. He  
College of Computer Science and Engineering, Changchun University of Technology,  
Changchun 130012, China  
e-mail: [houlixinmingxuan@126.com](mailto:houlixinmingxuan@126.com)

X. Peng  
Software Vocational and Technical College, Changchun University of Technology,  
Changchun 130012, China

accuracy [2, 3], but it relies on the results of the word segmentation. Extraction based on statistics does not require lexical and syntactic information and has better portability [4, 5]. Also there are many methods that mix linguistics with statistics [6, 7]. Structured vocabularies [8], linguistics [9, 10], and statistics [11, 12] have been used to extract relationships among concepts. However, methods based on traditional structured vocabularies extract only several kinds of relations, and most of them are hierarchical relations. Relationships extracted with linguistics have high accuracy, but the effects are easily affected by the completeness of the grammar rules. Extractions with statistics have nothing to do with the language and the fields. Therefore, they have a strong portability, but the effects largely depend on the quality and size of the corpus.

In the study of ontology construction, formal ontologies are better than ontologies based on natural language in sharing. Formal concept analysis (FCA) invented by Wille is a tool of identifying conceptual structures among data sets. FCA is widely used in the ontology construction. Obitko put forward that FCA can explore potential objects and attributes by constructing concept lattice and present them in a visual way automatically [13]. This chapter improves the above method by adopting the semantic dependency analysis technology to extract formal context, using Godin algorithm for building the concept lattice, and adopting the mapping rules between the concept lattice and the ontology to generate the domain ontology. Finally, this method is testified by constructing a medical ontology. The experiment results show that ontology construction based on semantic dependency and FCA not only can extract semantic information automatically and objectively but also can construct expected ontology from unstructured Chinese texts. Therefore, the method we proposed can construct domain ontology automatically and reduce manual intervention. In addition, the ontology we get is a formal ontology, so it has more advantages in sharing and reusability.

## 56.2 The Chinese Domain Ontology Learning Procedure Based on FCA

### 56.2.1 Formal Context Extraction

The formal context and concept hierarchy are basic structures of FCA. A formal context comprises two sets and a binary relation between them. The definition is as follows:

**Definition 1** *Formal context* is a triple  $K := (O, A, I)$ , where  $O$  is a set of objects,  $A$  is a set of attributes, and  $I$  is the binary relationship between  $O$  and  $A$ , that is,  $I \subseteq O \times A$ . For  $o \in O$  and  $a \in A$ ,  $(o, a) \in I$  is defined as “attribute  $a$  is an attribute of the object  $o$ .”

Formal context is the basic of the ontology learning based on FCA. This chapter adopts Philipp Cimiano method which was used in IST-Dot Kom project by AIFB

research institution [14]. The main idea is that we use a Chinese dependency syntactic parser for analyzing the sentences in texts [15], and then the parsing trees with dependence relation and semantic role labels will be produced. To build the formal context we extract the backbone of the sentence from the parsing tree and turn the subject of the sentence head into an FCA object and the objects into attributes. The algorithm for extracting the formal context based on the definition 1 and the main idea we proposed is given below.

- Step 1. Initialize HashMap for store pair of object and attribute.
- Step 2. For any sentence  $i$  (the sentence number) get word segmentation sequence filed as wordlist based on conditional random field (CRF).
- Step 3. For every word  $j$  (the word number) in wordlist, we adopt graph-based parser output, the dependency relation of word  $i$  in sentence  $j$ .
- Step 4. Judge the type of the dependency relation of word  $j$ , if the type equals "SBV," and then let word  $j$  be an object filed as  $o$ ; if the type equals "VOB," then let word  $j$  be an attribute filed as  $a$ , and then add  $(o, a)$  to HashMap.
- Step 5. Build formal context  $K = (O, A, I)$ ; use all pairs in HashMap.

## 56.2.2 Concept Lattice Construction

**Definition 2** For any  $X$  as an object subset, we define

$$X' = \{a \in A \mid \forall x \in X, \exists (x, a) \in I\}$$

For any  $Y$  as an attribute subset, we define

$$Y' = \{o \in O \mid \forall y \in Y, \exists (o, y) \in I\}$$

**Definition 3** If  $X \subseteq O$ ,  $Y \subseteq A$  satisfies  $X' = Y$  and  $Y' = X$ , then we regard  $C = (X, Y)$  as a formal concept of  $K = (O, A, I)$ ,  $X$  as the extension, and  $Y$  as the intension of the  $C$ .

**Definition 4** For a formal context  $K = (O, A, I)$  and two formal concepts  $C_1 = (x_1, y_1)$  and  $C_2 = (x_2, y_2)$  of  $K$ , a concept hierarchy relation is given by

$$(x_1, y_1) \prec (x_2, y_2) \Leftrightarrow (x_1 \subseteq x_2) \wedge (y_2 \subseteq y_1)$$

All the formal concepts in  $K$  are ordered by this concept hierarchy relation, and this ordering is called the concept lattice of the formal context  $K$ .  $C_1$  is the sub-concept of  $C_2$ , whereas  $C_2$  is the super-concept of  $C_1$ . Concept lattice is the core data structure of FCA, and it reflects the generalization and specialization of the relationship between the concepts; also it can be represented graphically by using Hasse diagram. There are two kinds of lattice construction algorithms: batch algorithms and incremental algorithms [16].

This chapter adopts Godin algorithms to construct concept lattice. Godin algorithm is an incremental algorithm introduced by Robert Godin in 1995. If we want to use Godin algorithms, two problems should be resolved [17]: (1) update the nodes in concept lattice and (2) update edges between lattice nodes. The new lattice nodes can be divided into three types, and the types can be described by the following definitions.

**Definition 5** Let  $L(K)$  be corresponding concept lattice of context  $K = (O, A, I)$  and  $M_i = (x_i, y_i)$  be any lattice node in  $L(K)$ . Let  $x_j$  be newly increased object and  $y_j$  be corresponding attribute. Meanwhile, if  $L'(K)$  is new concept lattice through adding  $x_j$ , then the lattice nodes in  $L'(K)$  can be divided into three types:

1. If  $y_j \cap y_i = \emptyset$ , then let  $M_i = (x_i, y_i) \in L(K)$ ;  $M_i$  is called unchanged node.
2. If  $y_i \subset y_j$ , then modify  $M$  to  $M' = (x_i \cup x_j, y_i)$  and add  $M'$  to  $L'(K)$ ;  $M'$  is called update node.
3. If  $(y_j \cap y_i \neq \emptyset) \wedge (\neg \exists ((y_j \cap y_i)', y_j \cap y_i) \in L(K))$ , then create a new node  $M' = (x_j \cup x_i, j_j \cap y_i)$  and add it to  $L'(K)$ ;  $M'$  is called newly increased node.

The following are the procedure for building lattice based on Godin algorithms:

Step 1. Initialize concept lattice  $L(K) \leftarrow \emptyset$ .

Step 2. Take out one pair  $(o, a)$  from  $K = (O, A, I)$ .

- (1) Take out formal concept  $M_i = (x_i, y_i)$  from  $L(K)$  in proper sequence.
- (2) Update the node in  $L(K)$  according to the result of  $(a \wedge y_i)$  and the rules in definition 5.

Step 3. If producing a new node in step (2), then add it to  $L(K)$  and update the edges between the nodes.

Step 4. Repeat steps 2 and 3 until generating a complete concept lattice.

### 56.2.3 Mapping Concept Lattice to OWL Ontology

Concept lattice can be viewed as a prototype of ontology. So the Hasse diagram should be trimmed circularly by deleting unreasonable concepts according to the domain knowledge. After that, ontology description language OWL recommended by W3C is adopted to describe the ontology mapped from concept lattice. The mapping rules between lattice elements in FCA and semantic elements in OWL are the followings:

**Definition 6** Let  $O := (C, root, < c)$  be a domain ontology,  $C$  be the set of concepts,  $root$  be the root element in  $O$ ,  $< c$  be the concept hierarchical relation,  $L(K)$  be the corresponding lattice of formal context  $K = (O, A, I)$ ,  $H_i = (x_i, y_i)$  and  $H_j = (x_j, y_j)$  be any lattice nodes in  $L(K)$ ,  $e$  be lattice element in FCA,  $C$  be the corresponding class name of  $H_i$ ,  $\sup C$  be the corresponding class name of  $H_j$ , and  $f : L(K) \rightarrow O$  be the mapping rules from concept lattice to ontology; then the rules are the following:

```

R1: IF e is Hi THEN <owl:Class rdf:about="#C"/> IN O
R2: IF e is Hi AND Hi ≤ Hj THEN <owl:Class rdf:about="#C">
<rdfs:subClassOf rdf:resource="#sup C"/><rdfs:sub-
ClassOf>IN O
R3: IF e ∈ xi THEN <owl:NamedIndividual rdf:about="#xi">
<rdf:type rdf:resource="#C"/></owl:NamedIndividual>
IN O
R4: IF e ∈ yi THEN <owl:DatatypeProperty rdf:about="#yi">
<rdfs:domain rdf:resource="#C"/></owl:DatatypeProperty>
IN O

```

Based on definition 6, we design the mapping algorithm as follows:

Step 1. Initialize an empty ontology  $O$ .

Step 2. Map the root node in  $L(K)$  to root according to  $R_1$ .

Step 3. Take out the child node  $H_i = (x_i, y_i)$  of the root node.

Step 4. Apply the mapping rules:

- (1) Apply rule  $R_1$  to lattice node  $H_i$ .
- (2) Apply rule  $R_2$  according to the hierarchical relation between the root node and the child node.
- (3) Apply rule  $R_3$  to the new objects that lattice node  $H_i$  owns, whereas the new objects do not belong to  $H'_i$  ( $H'_i$  is the child node of  $H_i$ ).
- (4) Apply rule  $R_4$  to the new attributes that  $H_i$  owns whereas the corresponding parent node of  $H_i$  does not own.

Step 5. Let  $H'_i$  be the new root node.

Step 6. Repeat steps 3–5 until mapping of the whole concept lattice is complete.

## 56.3 Experimental Verification

To verify the effectiveness of the method, we apply it to many medical domain files to learn ontology. The learning process is divided into three stages: (1) obtain the formal context from domain texts; (2) construct the concept lattice; and (3) map the concept lattice to the ontology. At last a Chinese ontology of medical domain will be obtained.

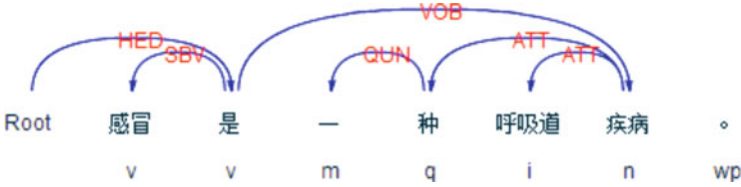
### 56.3.1 Obtaining the Formal Context Through Learning Medical Texts

The semantic dependency of every sentence in medical texts should be analyzed according to the method in Sect. 56.2.1. The process is as follows:

```

LTML ltml = ls.analyze(LTPOption.ALL, str);
for(int i = 0; i<ltml.countSentence(); ++i) {
    ArrayList<Word>wordList=ltml.getWords(i);

```



**Fig. 56.1** The semantic dependency tree of the example sentence

```

for(int j=0; j<wordList.size(); ++j){
System.out.print("\t"+wordList.get(j).getWS());
System.out.print("\t"+wordList.get(j).getParser-
Relation());
if(wordList.get(j).getParserRelation().equals(s1))
{name=wordList.get(j).getWS();}
if(wordList.get(j).getParserRelation().equals(s2))
{predicate=wordList.get(j).getWS();}
if(wordList.get(j).getParserRelation().equals(s3))
{value=wordList.get(j).getWS();}
hashmap.put(name, value);}}

```

Take “感冒是一种呼吸道疾病 (Cold is a kind of respiratory disease).”; for example, the visual result of the dependency analysis would look like as in Fig. 56.1.

In the figure, the semantic role of “是(is)” is “HED,” while the semantic role of “感冒 (cold)” is “SBV” and the semantic role of “疾病 (disease)” is “VOB.” From the analysis result we can get

```
name="感冒"; value="呼吸道疾病"; hashmap.put(name, value);
```

As a result, the pair(“感冒”, “呼吸道疾病”) will be obtained.

We learn all the domain texts and obtain all the pairs by the following method:

```

Iterator it=hashmap.entrySet().iterator();
while (it.hasNext()){Map.Entry entry=(Map.Entry) it.
next();
String
filecontents=entry.getKey()+predicate+entry.getValue
()+"\n";}

```

We place all the pairs in the table, and small part of the formal context would look like as in Table 56.1.

### 56.3.2 Constructing Medical Concept Lattice

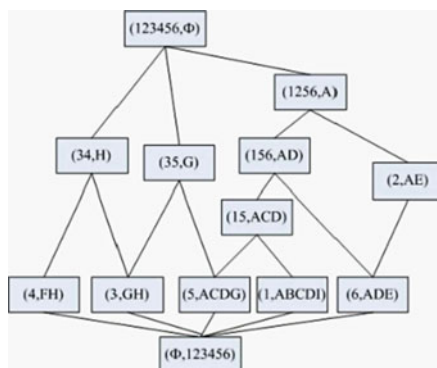
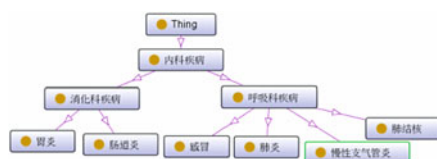
We analyze the medical formal context through the Godin algorithms described in Sect. 56.2.2. Then we adjust the location of part lattice nodes. Finally the medical concept lattice will be obtained, and part of it looks like as in Fig. 56.2.



**Table 56.1** Part of formal context of medical domain

	A	B	C	D	E	F	G	H	I
1	×	×	×	×					×
2	×				×				
3							×	×	
4						×		×	
5	×		×	×			×		
6	×			×	×				

1.感冒 (cold) 2. 慢性支气管炎 (chronic bronchitis) 3. 胃炎 (gastritis) 4. 肠道炎 (gastroenteritis) 5. 肺炎 (pneumonia) 6. 肺结核 (phthisis) A. 咳嗽 (cough) B. 鼻塞 (nasal obstruction) C. 头痛 (headache) D. 发热 (fever) E. 咳痰 (expectoration) F. 腹痛 (abdominal pain) G. 恶心呕吐 (nausea and vomiting) H. 反酸 (sour regurgitation) I. 流鼻涕 (rhinorrhea)

**Fig. 56.2** Part of the medical concept lattice**Fig. 56.3** Part of the medical ontology

Concept lattice describes the generalization and specialization relationship of the concepts where the above node is parent node whereas the below node is child node.

### 56.3.3 Generating the Chinese Ontology in Medicine

We adopt the mapping rules between lattice elements in FCA and ontology elements in OWL defined in Sect. 56.2.3 to map the concept lattice in Fig. 56.2. After the mapping, the unreasonable concepts are deleted, and finally a Chinese ontology is generated. Furthermore, one plug-in of the Protégé, called OntoGraf, is used for visualizing the Chinese ontology, and the result would look like as in Fig. 56.3.

## 56.4 Conclusion

We proposed a Chinese domain ontology learning approach based on semantic dependency and FCA. First, this study adopted the Chinese dependency parser to parse every sentence in Chinese domain texts to construct the formal context. Then, by adopting the Godin algorithms, the concept lattice was built. At last, the concept lattice was mapped to the domain ontology through the mapping rules between them, showing the ontology graphically by using Hasse diagram. In addition, we verified the effectiveness of the approach proposed in this chapter through the experiment in Sect. 56.3. The advantages are that the method can conduct the unstructured Chinese texts and can express the automatic objective extracting semantic characteristics of FCA. Shortages inevitably exist, and we will design more specific algorithms for constructing the formal context to reduce artificial participation and improve the automation degree of the ontology learning in the next phase.

## References

1. Du, X. Y., Li, M., & Wang, S. (2006). A survey on ontology learning research[J]. *Journal of Software*, 17(9), 1837–1847.
2. Shamsfard, M., & Barforoush, A. (2004). Learning ontologies from natural language texts [J]. *International Journal of Human-Computer Studies*, 60(1), 17–63.
3. Sabou, M. (2004). From software APIs to web service ontologies: A semi-automatic extraction method[C]. In *Proceedings of International Semantic Web Conference (ISWC)*, Springer, Berlin (pp. 410–424).
4. Huang, C. (2009). *Research of domain ontology construction and using in web information extraction[D]*. Jiangxi: Jiangxi University of Science and Technology.
5. Yu, J. (2010). *Learning domain ontologies from Chinese text corpora[D]*. Dalian: Dalian University of Technology.
6. Liang, J., & Wu, D. (2006). Seed concept method and its application in texts-based ontology learning[J]. *Library and Information Service*, 50(9), 18–21.
7. Wen, C., Wang, X., & Shi, Z. (2009). Automatic domain specific term extraction in Chinese domain ontology learning[J]. *Application Research of Computers*, 26(7), 2652–2655.
8. Ahlnad, K., Tariq, M., et al. (2003). Corpus-based thesaurus construction for image retrieval in specialist domains[C]. In *Proceedings of the 25th European Conference on Advances in Information Retrieval (ECIR)*, Springer, Berlin (pp. 502–510).
9. Buitelaar, P., Olejnik, D., Hutanu, M., & Schutz, A. (2004). Towards ontology engineering based on linguistic analysis[C]. In *Proceedings of LREC, Lisbon, Portugal* (pp. 7–10).
10. Wang, S. (2010). *The research on acquisition method of relationships between subject concepts in ontology construction[D]*. Beijing: The Chinese Academy of Agricultural Sciences.
11. Fu, K. (2007). *The study of ontology learning from web pages[D]*. Wuhan: Wuhan University of Technology.
12. He, L. (2009). *Research on domain ontology semi-automatic construction and retrieval[M]* (pp. 156–163). Nanjing: Southeast University Press.
13. Obitko, M., Sná, eIV., & Smid, J. (2010). *Ontology design with formal concept analysis [EB/OL]*. <http://ftp.informa-tik.rwth-aachen.de/Publications/CEUR-WS/Vol-110/paper12.Pdf>

14. Huang, M., & Liu, Z. (2006). Research on domain ontology building methods based on formal concept analysis[J]. *Computer Science*, 33(1), 210–212. 239.
15. Che, W., Li, Z., & Liu, T. (2010). LTP: A Chinese language technology platform. In *Proceedings of the Coling 2010: Demonstrations, Beijing* (pp. 13–16).
16. Baixeries, J., Szathmary, L., et al. (2009). Yet a faster algorithm for building the Hasse diagram of a concept lattice[C]. In S. Ferré & S. Rudolph (Eds.), *ICFCA 2009* (pp. 162–177). Berlin: Springer.
17. Jiang, Y., Zhang, J., & Zhang, S. (2007). Incremental construction of concept lattice based on linked list structure [J]. *Computer Engineering and Applications*, 43(11), 178–180.

# Chapter 57

## Text Classification Algorithm Based on Rough Set

Zhiyong Hong

**Abstract** In text classification community, k-nearest neighbor (kNN) and support vector machine (SVM) are all effective classifiers, but both of them have their own drawbacks. kNN involves high cost to classify a new document when training set is large; SVM is too sensitive to the noisy data when the noisy data is close to the hyperplane it suffers. So one hybrid algorithm based on variable precision rough set is proposed. It combines the strength of both KNN and SVM techniques and overcomes their weaknesses. Finally some experiments are carried out to compare the efficiency and classification accuracy with different classification algorithms. Results show that the proposed method achieves significant performance improvement.

### 57.1 Introduction

Text classification is the core technology of the research in the field of information retrieval and data mining; it has rapidly developed in recent years. Text classification or categorization is the task of automatically assigning unseen documents to suitable predefined categories. A number of well-known algorithms have been introduced to deal with text classification [1], such as k-nearest neighbor (kNN) [2], support vector machine (SVM) [3], neural network [4], naive Bayesian (NB), centroid-based classifier, decision tree, and Rocchio classifier.

The kNN is an instance-based learning algorithm, which is simple and intuitive but very effective for a variety of problem domains including text classification. It is well known to be one of the most effective methods on Fudan University classification corpus—one of the benchmark corpora used in text classification. However, kNN has a high cost of classifying new patterns; its training phase just

---

Z. Hong (✉)

School of Computer Science, Wuyi University, Jiangmen 529020, Guangdong, China  
e-mail: [hongmr@163.com](mailto:hongmr@163.com)

stores all training patterns as classifier; thus it has often been called as lazy learner since it defers the decision on how to generalize beyond the training data until each new query pattern is encountered. The efficiency of kNN prohibits it from being applied to areas where efficiency is particularly required for text classification, such as dynamically mining large-scale collection.

SVM is a machine learning method based on statistical learning theory [3]. An SVM constructs an optimal hyperplane that has the largest distance to the nearest training data point of any class to find the good support vector. Since SVM has good generalization performance, it has been used for classification, regression analysis, handwriting recognition, speech recognition, face recognition, and so on. In classification of the SVM traversal the feature space consists of support vector instead of the original input space consisting of training samples to ensure higher classification accuracy and greatly reduce the computational complexity. Although SVM has good generalization performance, SVM also has its drawbacks, such as the classification accuracy is not very high when encountering complex classification. Especially when training samples is confusing or overlapping, SVM classifier is prone to over-fitting.

In this chapter, a hybrid algorithm based on variable precision rough set (VPRS) is proposed to combine the strength of both kNN and SVM techniques and overcome their weaknesses. Firstly, feature space of training data is partitioned by using VPRS, and lower and upper approximations of each class are defined. Then kNN and SVM classifiers are built on these new subspaces, respectively. The SVM classifiers are used to classify most of the new documents effectively and efficiently. The kNN classifier is only required to classify new document which lies in the boundary region where SVM classifier suffers. And it is just required to find nearest neighbors of new document in the subset of training dataset, which can save time obviously compared with finding nearest neighbors in the whole training dataset. Experiments are carried out on Fudan University classification corpora. The experimental results indicate that the proposed hybrid algorithm achieves significant performance improvement.

## 57.2 Background

The rough set theory [5], introduced by Pawlak in the early 1980s, is a formal mathematical tool to deal with incomplete or imprecise information. As a generalized version of rough sets, VPRS [6] allows objects to be classified with an error smaller than a certain predefined level. In text classification, when the training data is confusing or overlapping, VPRS model is more suitable than the classic rough set. In this section, a brief introduction to rough set and VPRS is given.

### 57.2.1 VPRS

As a generalization of the standard inclusion relation, majority inclusion relation introduced by the VPRS is defined as follows:

**Definition 1** Majority inclusion relation

$$c(X, Y) = \begin{cases} 1 - |X \cap Y|/|X|, & |X| > 0 \\ 0, & |X| = 0 \end{cases} \quad (57.1)$$

where  $X$  and  $Y$  are subsets of the universe  $U$ . The majority inclusion relation denotes the relative degree of misclassification of the set  $X$  with respect to set  $Y$ . Based on this measure, one can define the standard set inclusion relation between  $X$  and  $Y$  as  $X \subseteq U$  if and only if  $c(X, Y) = 0$ . Consider an equivalence relation  $R$  on  $U$ , For a subset  $X \subseteq U$ , the  $\beta$ -lower approximation and  $\beta$ -upper approximation of  $X$  can be defined as follows:

**Definition 2**  $\beta$ -lower approximation and  $\beta$ -upper approximation:

$$\underline{R}_\beta X = \{x \in U : c([x]_R, X) \leq \beta\} \quad (57.2)$$

$$\overline{R}_\beta X = \{x \in U : c([x]_R, X) < 1 - \beta\} \quad (57.3)$$

**Definition 3**  $\beta$ -positive,  $\beta$ -negative, and  $\beta$ -boundary region based on VPRS:

$$POS_\beta X = \underline{R}_\beta X \quad (57.4)$$

$$NEG_\beta X = U - \overline{R}_\beta X \quad (57.5)$$

$$BND_\beta X = \overline{R}_\beta X - \underline{R}_\beta X \quad (57.6)$$

## 57.3 Hybrid Algorithm Based on VPRS

### 57.3.1 Algorithm

In this chapter, we only consider binary text classification that assigns each document  $d$  either to the positive class  $C_d$  or to its complement negative class  $C_n$ . Theoretically, binary classification is more general than the multi-class classification, and a multi-class classification can be transformed into a set of binary classifications. We can characterize the two class document sets of  $C_d$  and  $C_n$  with respect to a hidden equivalence relation  $R$  which may lead the documents belonging to the same class to have the tendency of clustering. The kNN algorithm is used to create equivalence classes for set  $C_d$  and  $C_n$  according to the

concepts of VPRS. For a document  $d$  in training set, the kNN algorithm is used to find its kNN by means of calculating the similarity of  $d$  to anyone in training set, which forms a neighborhood of  $d$ . If all neighbors are from a single class, e.g.,  $C_d$ , then there is no uncertainty in the neighborhood. However, if any neighbor belongs to another class  $C_n$ , the rough uncertainty arises in the neighborhood. This uncertainty can be captured by using the modified majority inclusion relation. For any document  $d$  and document set of class  $C_d$  in training dataset, modified majority inclusion relation is defined as follows:

**Definition 4** (Modified majority inclusion relation):

$$c(N_d, C_d) = 1 - \frac{|N_d \cap C_d|}{|N_d|} \quad (57.7)$$

where  $N_d$  is the neighborhood region around  $d$  and  $|\cdot|$  denotes the cardinality of the set. According to VPRS,  $\beta$ -positive region,  $\beta$ -negative region, and  $\beta$ -boundary region of class  $C_d$  can be obtained. The  $\beta$ -positive region of class  $C_d$  denotes the document set which lies in the positive region where documents can be certainly classified to class  $C_d$ ,  $\beta$ -boundary region of class  $C_d$  denotes the document set which lies in the boundary region where documents cannot be classified uniquely to the class  $C_d$ , and  $\beta$ -negative region of class  $C_d$  denotes the document set which cannot be surely classified to the class  $C_d$ . Similar description is also fit to class  $C_n$ .

After feature space is partitioned into three regions, i.e.,  $\beta$ -positive region,  $\beta$ -negative region, and  $\beta$ -boundary region for each class, then we can get the training set `init_set1` which consists of all the  $\beta$ -positive for each class and `init_set2` which consists of all the  $\beta$ -boundary for each class. For overcoming the problem of model over-fitting of SVM and improving the classification performance of kNN, the SVM classifier is trained with `init_set1` and kNN classifier is trained with the `init_set2`. After the SVM classifier is built, we can build one new hybrid classification algorithm. For a new unseen document  $d$ , firstly calculate the distance to SVM optimal hyperplane; if the distance is larger than the given threshold value then the algorithm will use SVM to classify the document  $d$ , or else the algorithm will use kNN to find the k-partition nearest neighbors around  $d$ ; then  $d$  is classified by a majority vote of its neighbors. The hybrid classification algorithm is described as Algorithm 1.

### 57.3.2 The Proposed Algorithm

**Algorithm 1** The hybrid algorithm

*Input:* Training set  $\chi$ , the parameter  $k$ ,  $\beta$ , new document  $d$

*Output:* The class label of  $d$

Step 1: Use kNN to partition space of training set  $\chi$  based on VPRS and definition 7, and then obtain the  $POS_{\beta}C_i$  and  $BND_{\beta}C_i$  of every class  $C_i$  ( $i = 1, 2 \dots n$ )

Step 2: Classifier training

- (1) Combine all the  $POS_{\beta}C_i$  of class  $C_i$  to constitute the training set  $init\_set1$ .
- (2) Combine all the  $BND_{\beta}C_i$  ( $i = 1, 2 \dots n$ ) of class  $C_i$  to constitute the training set  $init\_set2$ .
- (3) Train the SVM classifier with the training set  $init\_set1$  and then obtain support vector set ( $SV_i, i = 1 \dots m$ ), factor ( $\alpha_i, i = 1 \dots m$ ), and constant  $b$ .
- (4) Train the kNN classifier with training set  $init\_set2$ .

Step 3: Classifying the new document  $d$

- (1) For a new document  $d$ , firstly calculate the distance to SVM optimal hyperplane with the following formula (57.8):

$$f(x) = \sum_{i=1}^m \alpha_i y_i k(x_i, x) + b \quad (57.8)$$

if  $|f(x)| > \varepsilon$  ( $\varepsilon$  is the given threshold value) then output the class label of  $d$ , or else go to next step.

- (2) Use kNN to find the k-partition nearest neighbors around  $d$ ; the nearest neighbors can be obtained by the vector similarity calculation formula (57.9):

$$Sim(d_1, d_2) = \cos \theta = \frac{\sum_{k=1}^n W_{1k} \times W_{2k}}{\sqrt{\sum_{k=1}^n W_{1k}^2 \cdot \sum_{k=1}^n W_{2k}^2}} \quad (57.9)$$

where  $W_{ik}$  is the  $k$ th term weight of  $d_i$ ; then the class label of  $d_i$  can be obtained by methods in Sect. 57.3.1.

## 57.4 Experiment Results and Discussion

To evaluate performance (compute speed and classification accuracy) of the proposed approach, we have conducted experiments on the Fudan University classification corpus obtained from the International database center of Fudan University. The Fudan University classification corpus is a standard text



**Table 57.1** Comparison of the classification accuracy of each algorithm

Category	SVM	kNN	Hybrid algorithm
Transportation	89.12	91.04	91.98
Sports	91.35	92.61	92.93
Military	92.78	92.21	94.35
Medicine	93.89	89.23	94.50
Politics	93.00	93.56	93.99
Education	93.45	94.00	95.38
Environment	87.37	90.76	91.77
Economy	90.81	91.23	92.26
Art	88.78	89.75	91.68
Computer	89.29	86.46	90.81
Macro F1	90.98	91.09	92.97

classification benchmark which contained 1,882 training examples and 934 test examples. It is a large-scale collection, and the samples are confusing and overlapping. In our experiments, only the most populous ten categories from this corpus are used as our dataset, i.e., transportation, sports, military, medicine, politics, education, environment, economy, art, and computer.

Three commonly used indicators of the evaluation of the classification performance [7] are the precision  $p$ , recall  $r$ , and F1-Measure. The recall and precision rate reflects two different aspects of the quality of classification, but F1-Measure considers not only the precision but also recall. So we used the F1-Measure as the evaluation index in the experiment. To evaluate the performance of the algorithm on the whole dataset, we used the Macro F1 as the evaluation index. Assuming that F1-Measure value of  $i$ th category is  $F_1(i)$ , the Macro F1 can be defined as

$$Macro - F_1 = \frac{\sum_1^n F_1(i)}{m} \quad (57.10)$$

In order to test the hybrid classification algorithm, we construct two experiments on the Fudan University classification corpus as follows:

1. Comparing the classification accuracy of the hybrid algorithm with the kNN and SVM classifier
2. Testing the performance of the hybrid algorithm when selecting different  $\beta$  value

In the first experiment, we select 2,000 words as features for each algorithm. For kNN the value of the parameter  $k$  is 35. For the proposed algorithm, the value of  $k$ -partition is 20 and the value of  $\beta$  is 0.3. For SVM, we used the LIBSVM package [8], which supports both two-class and multiclass classification. The F1 value of each algorithm on each category and the corresponding Macro F1 on each algorithm are presented in Table 57.1. The compute speed of each algorithm on the dataset is presented in Table 57.2.

In the second experiment, we tested the performance of the hybrid algorithm when the number of features is 2,000 and the value of  $\beta$  varies from 0.05 to 0.45 with step 0.05. The performance of the hybrid algorithm is presented in Fig. 57.1.

**Table 57.2** Comparison of the cost time of each algorithm

SVM	kNN	Hybrid algorithm
16.28	53.45	18.72

**Fig. 57.1** The performance with the  $\beta$

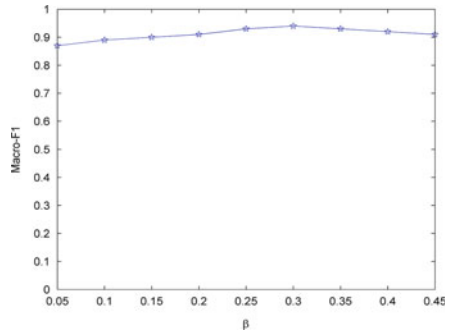


Table 57.1 shows that the F1 of proposed approach is slightly higher than that of other two algorithms on each category of the dataset; the Macro F1 of proposed approach is 92.97 %, which is approximately 1.07 % higher than that of kNN and 1.99 % higher than that of SVM. Table 57.2 shows that the time cost of kNN is the highest, that of the SVM is 16.28 which is lowest, and the cost time of hybrid algorithm is 18.72s which is about only 35 % of the cost time of kNN and close to that of SVM. So the hybrid algorithm is more efficient than kNN. From 1, we discovered a phenomenon that with the increase of  $\beta$ , the Macro F1 of proposed hybrid algorithm in the beginning increases and afterward decreases. When  $\beta$  takes 0.3, the proposed algorithm achieves the best results.

Above all, the proposed algorithm outperforms kNN and SVM in performance on the Fudan University classification corpus, and it is more efficient than kNN. Thus, the proposed approach is a good alternative for SVM algorithm and kNN in some scenarios of text classification such as dynamically mining large Web repositories where kNN is not suitable due to its lower efficiency and SVM is too sensitive to noise which easily leads to over-fitting.

## 57.5 Conclusion

In this chapter, two widely used techniques for text classification, i.e., the kNN and the SVM algorithm, are analyzed and some shortcomings of each are identified. Based on the analysis, a hybrid algorithm based on VPRS is proposed to combine the strengths of kNN and SVM classifier and overcome the problems of low efficiency of kNN and model over-fitting of SVM. Extensive experiments conducted on Fudan University classification corpus show that the hybrid algorithm achieves significant performance improvement.

**Acknowledgements** This study is supported by the National Natural Science Foundation of China (Grant No. 60474022).

## References

1. Sebastiani, F. (2002). Machine learning in automated text categorization. *ACM Computing Surveys*, 34(1), 1–47.
2. Duwairi, R. (2005). An eager k-nearest-neighbor classifier for Arabic text categorization. In *Proceedings of the International Conference on Data Mining (ICDM'05)*, IEEE Computer Society Press, Houston (pp. 187–192).
3. Vapnik, V. N. (1999). *The nature of statistical learning theory* (pp. 19–24). Berlin: Springer.
4. Xia, Y. S., & Wang, J. (2004). A one-layer recurrent neural network for support vector machine learning. *IEEE Transactions on Systems Man and Cybernetics Part B*, 34(2), 1261–1269.
5. Pawlak, Z. (1982). Rough sets. *International Journal of Information and Computing Science*, 32(11), 341–356.
6. Ziarko, W. (1993). Variable precision rough set model. *Journal of Computer System Science*, 46(1), 39–59.
7. Estévez, P. A., Tesmer, M., Perez, C. A., & Zurada, J. M. (2009). Normalized mutual information feature selection. *IEEE Transactions on Neural Networks*, 20(2), 189–201.
8. Hsu, C. W., & Lin, C. J. (2002). A comparison of methods for multi-class support vector machines. *IEEE Transactions on Neural Networks*, 13(2), 415–425.

# Chapter 58

## Robust Fragment-Based Tracking with Online Selection of Discriminative Features

Yongqiang Huang and Long Zhao

**Abstract** In order to solve the variation of target appearance and background influence to the visual tracking, we extend the robust fragment-based tracker to an adaptive tracker by selecting features with an online feature ranking mechanism, and the target model is updated according to the similarity between the initial and current models, which makes the tracker more robust. What is more, we reposition the integral histogram's bin's structure and that makes our tracker quicker. The proposed algorithm has been compared with fragment-based tracker, and the results proved that our method provides better performance.

### 58.1 Introduction

Visual tracking is an important task in computer vision and has been widely applied in traffic surveillance system [1], suspicious person monitoring system [2], etc. The problem in a visual tracker is the variation of target appearance and background. To improve the performance, an adaptive tracking mechanism is necessary. The normal idea is to adaptively select features of the object that can be discriminated from surrounding background.

In recent years, adaptive tracking algorithms have been widely studied [3–8]. Stern et al. [3] proposed an algorithm that chooses the best feature from five color spaces. Collins contrasted the foreground/background and selected the color features that can be best distinguished from the background [9]. Wang extended Collins's method by selecting reliable features from both color and shape–texture cues [10]. Chockalingam proposed an adaptive fragment-based

---

Y. Huang (✉) • L. Zhao  
Science and Technology on Aircraft Control Laboratory, Beihang University,  
Beijing 100191, China

Digital Navigation Center, Beihang University, Beijing 100191, China  
e-mail: [buaa\\_huangyongqiang@foxmail.com](mailto:buaa_huangyongqiang@foxmail.com)

tracking using level sets [11]. Leandro proposed a patch-based tracking and used simple updating scheme to cope with appearance and illumination changes; however, the model computed is not good enough for tracking [12].

In this chapter we use multiple image fragments to represent the object [13]; in addition, our fragment template is adaptively modeled by the most discriminative features, and the target model is updated according to the similarity between the initial and current models.

## 58.2 Feature Subset Selection

We choose color histograms to represent the color distributions in RGB and HSV color spaces. We omit the intensity because it is useless in our tracking.

After selecting the features, the log-likelihood ratio and variance ratio are used to select the most descriptive features. The pixel frequency can be calculated as  $\xi_f^{(bin)} = H_f^{(bin)}/n_{fg}$  and  $\xi_b^{(bin)} = H_b^{(bin)}/n_{bg}$ , where  $H_f$  and  $H_b$  are histograms and  $n_{fg}$  and  $n_{bg}$  are the pixel number of the target and background, respectively.

The feature's log-likelihood ratio can be computed through Eq. (58.1)

$$L^{(bin)} = \log \frac{\max\left(\xi_f^{(bin)}, \sigma_L\right)}{\max\left(\xi_b^{(bin)}, \sigma_L\right)} \quad (58.1)$$

where  $\sigma_L$  is a small number (here 0.001).

And the variance ratio is computed through Eq. (58.2)

$$\text{var}(L; p) = E\left[(L^{bin})^2\right] - (E[L^{bin}])^2 \quad (58.2)$$

Then the color features are ranked according to the discriminative ability by comparing the variance ratio. We assume that the most discriminative feature is corresponding to the maximum variance ratio.

## 58.3 Adaptive Robust Fragment-Based Tracking

### 58.3.1 The Fragment-Based Tracking Algorithm

In fragment-based tracking algorithm, the target template is represented by multiple histograms of multiple rectangular patches. By comparing each patch's histogram with the corresponding image patch histogram, its vote on the positions of the object in the image can be calculated, and a distance map describing the possible

positions of each patch in current image is obtained. Combine the distance maps obtained from all template patches, and get the target position.

However, like standard meanshift tracking, the basic fragment-based tracking algorithm assumes that the target representation is good enough to discriminate the foreground from the background. However, this is not always reliable in a dynamic background tracking.

### 58.3.2 Integral Histogram

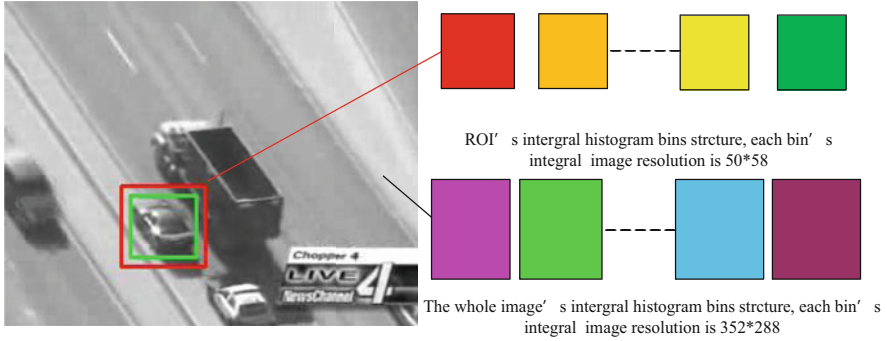
The fragment-based tracker needs to compute multiple histograms of multiple rectangular patches; the integral histogram is used to realize it in real time [13, 14].

Integral histogram extends from the integral image data structure. In the integral image, each pixel  $P_t$  holds the sum of all values of the relative image to the left and above of the pixel including the value of the pixel itself. After the image is computed, the sum of the pixels on any rectangular regions can be computed with four arithmetic operations. In order to extract histograms of any rectangular regions, each bin of the histogram (in the integral histogram) is built from an integral image counting the cumulative number of pixels falling into it. Then the number of pixels that belong to a bin in a given region can be easily computed through these integral images, and hence the histogram of that region is obtained.

Although extraction of a histogram over any patch is cheap, the experiments show that the integral histogram used in [13] is still computationally expensive. For an image of  $352 \times 288$ , the integral histogram data structure requires each bin one big image of that size in the histogram it takes about 500 ms (on a 2.0 GHz P2) to compute the integral histogram. Here we propose to develop a method which can reduce the computational cost. As the tracker searches in the neighborhood of the position estimate from the previous frame, we call this neighborhood area ROI; the region outside the ROI is useless; we just compute ROI's integral histogram to reduce the computational cost. See Fig. 58.1.

### 58.3.3 Target Localization

As mentioned in Sect. 58.2, we have ranked the features according to their discriminative ability from the background, and we will take advantage of the best two discriminative features to represent the target. The proposed tracking algorithm calculates the joint histogram ( $P_f^{(b_{in}^1, b_{in}^2)}$ ) of the target with the best two features, replaces the patch's histogram in Frag (fragment-based tracking algorithm) with the joint histogram, then computes the similarity between the corresponding patches, and gains the distance map for each template patch.



**Fig. 58.1** Each color rectangle in the *right* of the figure represents one bin's integral image of the histogram (in this chapter  $N_{bins} = 8$ ), and in the *left* of the figure, the region inside the *red* rectangle represents the ROI. Obviously, ROI's integral histogram computation complexity is much lower than the whole image's (Color online)

The distance map gives a value for every likely position of the target in the current frame, combines the distance maps obtained from all template patches, sums the distance maps, and gets the target position which has the minimal value.

### 58.3.4 Model Update

As the appearance of a target will always change during tracking, the target model needs to be updated.

We adopt the method proposed in [10] to update the target model, which involves the initial model, previous model, and current candidate. In the first frame, the initial target is labeled and its model  $M_i$  is computed. The initial model's weight is updated according to the similarity between the initial and current target. Before the  $i$ th update, the tracker searches the target in the current frame using the previous computed target model  $M_p^{i-1}$ . The new target model  $M_c^i$  can be computed by

$$M_c^i = (1 - S_{ic})M_i + S_{ic}M_{cp}^i \tag{58.3}$$

where  $M_i$  is the initial model;  $M_{cp}^i$  is the combined histogram of the current and the previous target model; and  $S_{ic}$  is the similarity between the initial and current target model:

$$M_{cp}^i = (1 - S_{cp})M_p^{i-1} + S_{cp}M_c \tag{58.4}$$

where  $M_p^{i-1}$  is the previous target model;  $M_c$  is the histogram of current target model; and  $S_{cp}$  is the similarity between the previous and the current target model, which is measured by Bhattacharyya coefficient.

### 58.4 Experimental Test and Result Analysis

In our experiment, the feature number  $K = 2$ , the histogram bin number  $N = 12$ , and the search radius  $R = 7$ ; as two adjacent frame change is small, our feature selection runs every five frames.

We tested on three sequences (“woman,” “bolt,” and “Person”). The tracking results are compared with the Frag. See Figs. 58.2–58.4, where the red rectangles are the results of Frag and the green rectangles are ours. For the sequence bolt, the position errors of our tracker and the Frag are plotted in Fig. 58.5.

In Figs. 58.2 and 58.3, as the target appearance changes constantly, the Frag loses the target, while our tracker can still get it. The next sequence “Person” shown in Fig. 58.3 is under complete occlusion, and our tracker is still robust to it, whereas the Frag has lost the target. The Frag just uses the gray feature to represent the target, while our tracker uses the best subset of features; what is more, the Frag uses the same initial model all the tracking time, while we take into account the



Fig. 58.2 Woman

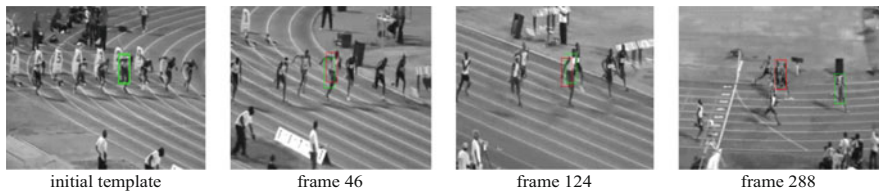


Fig. 58.3 Bolt

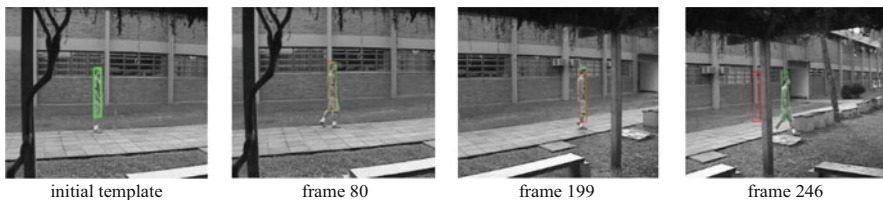
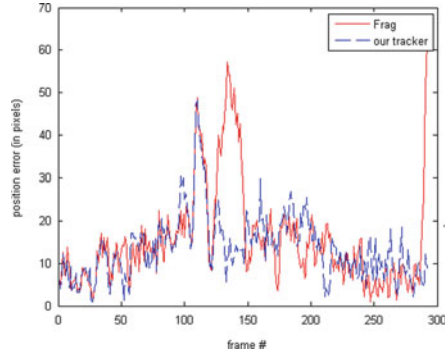


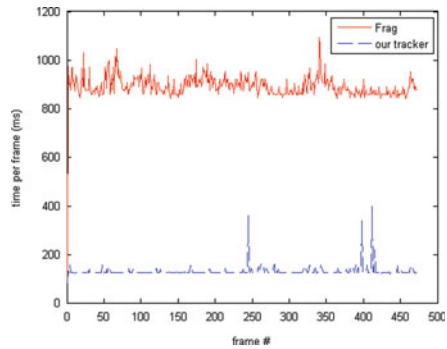
Fig. 58.4 Person



**Fig. 58.5** Bolt sequence—error with respect to ground truth. Our tracker—*green*. Frag—*red* (Color online)



**Fig. 58.6** Woman sequence—time per frame, our tracker—*red*, Frag—*green* (Color online)



similarity between the initial and current appearance of the target; this is robust to appearance variation.

Figure 58.6 shows the computing time of our tracker and the Frag during each frame, from which we can see that our tracker is less time consuming than Frag.

## 58.5 Conclusion

We extend the fragment-based tracker to an adaptive tracking algorithm by integrating color features adaptively. The target is represented by a joint histogram of the best two features. We also modify the integral histogram's structure, which can make our tracker quicker than the Frag. The proposed approach demonstrates good performance.

**Acknowledgements** Project supported by the key program of the National Natural Science Foundation of China (Grant No. 61039003), the National Natural Science Foundation of China (Grant No. 41274038), the Aeronautical Science Foundation of China (Grant No. 20100851018), and the Aerospace Innovation Foundation of China (CASC201102).

## References

1. Lee, L., Romano, R., & Stein, G. (2000). Monitoring activities from multiple video streams: Establishing a common coordinate frame[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8), 758–767.
2. Haritaoglu, I., Harwood, D., & Davis, L. S. (2000). Real-time surveillance of people and their activities[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8), 809–830.
3. Chen, H. T., Liu, T. L., & Fuh, C. S. (2008). Probabilistic tracking with adaptive feature selection[C]. In *Pattern Recognition, Proceedings of the 17th International Conference, Roma* (pp. 736–739).
4. Collins, R. T., Liu, Y. (2003). On-line selection of discriminative tracking features[C]. In *Computer Vision, 9th IEEE International Conference on IEEE, Istanbul* (pp. 346–352).
5. Collins, R., Zhou, X., & Teh, S. K. (2005). An open source tracking testbed and evaluation website[C]. In *IEEE International Workshop on Performance Evaluation of Tracking and Surveillance, Munich* (pp. 17–24).
6. Comaniciu, D., Ramesh, V., & Meer, P. (2003). Kernel-based object tracking[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(5), 564–577.
7. Wang, J., & Yagi, Y. (2006). Integrating shape and color features for adaptive real-time object tracking[C]. In *Robotics and Biomimetics, International Conference on IEEE, Sydney* (pp. 1–6).
8. Gevers, T., Smeulders, W. M., & W. A. (1999). Color based object recognition[J]. *IEEE Transactions on Pattern recognition*, 32(3), 453–464.
9. Collins, R., & Liu, Y. (2003). On-line selection of discriminative tracking features[C]. In *Computer Vision, 9th IEEE International Conference, Mamai, IEEE, 2003* (pp. 346–352).
10. Wang, J., & Yagi, Y. (2008). Integrating color and shape-texture features for adaptive real-time object tracking[J]. *IEEE Transactions on Image Processing*, 17(2), 235–240.
11. Chockalingam, P., Pradeep, N., & Birchfield, S. (2009). Adaptive fragments-based tracking of non-rigid objects using level sets[C]. In *Computer Vision, 12th International Conference on IEEE, Chengdu* (pp. 1530–1537).
12. Dohl, L., Jung, C. R., & Bins, J. (2011). Robust adaptive patch-based object tracking using weighted vector median filters[C]. In *Graphics, Patterns and Images (Sibgrapi), SIBGRAPI Conference on IEEE, Rio de Janeiro* (pp. 149–156).
13. Adam, A., Rivlin, E., & Shimshoni, I. (2006). Robust fragments-based tracking using the integral histogram[C]. In *Computer Vision and Pattern Recognition, Computer Society Conference on IEEE, New York* (pp. 798–805).
14. Porikli, F. (2005). Integral histogram: A fast way to extract histograms in Cartesian spaces [C]. In *Computer Vision and Pattern Recognition, Computer Society Conference on IEEE, San Diego* (pp. 829–836).

# Chapter 59

## Extraction Method of Gait Feature Based on Human Centroid Trajectory

Xin Chen and Tianqi Yang

**Abstract** Gait features obtained by current extraction methods are easily affected by people's walking direction, dresses, and carryings, due to which gait recognition system has not yet appeared. An extraction method based on centroid is proposed in this chapter. Segment and track the moving silhouettes of a walking figure in image sequences to calculate the silhouettes' centroid. The complex silhouette is represented by a point to avoid the influence of dresses and carryings. Divide centroid coordinate value by the height of detecting walking figure to normalize to remove the disturbance caused by walking direction relative to the camera optical axis angle. By denoizing centroid trajectory remove the noise caused by some accidental factors to obtain regular wavelet curve whose main frequency component distribution vector is the final gait feature. Experimental results show that this approach can obtain identical gait features even when experimenters change their walking directions, dresses, or carryings, tolerating noise and low resolution.

### 59.1 Introduction

Face recognition systems have been used widely around the world, but with the development of making up technology, face can no longer mark a human. Most criminals will shelter their faces from the cameras, and face recognition system's application is becoming narrower and narrower. Gait is influenced by muscle strength, length of tendon, bone density, and so on, which makes disguising gait very difficult.

---

X. Chen (✉) • T. Yang  
Department of Information Science and Technology, Jinan University,  
Guangzhou 510000, China  
e-mail: [843597029@qq.com](mailto:843597029@qq.com)

Currently, there are mainly two kinds of method about gait feature extraction: one is establishing human models to describe gait with model parameters; the other is establishing a relationship between two neighboring frames by characteristics such as position, speed, and color [1]. Zhang, e.g., uses GVF to model the walking figures [2], making description of human contour possible. He judges current gait gesture by matching the models of current frames with those of key frames, considering cyclic features of gait fully. However, it is quite difficult to model a figure's outline because the human body is irregular. Wang, e.g., uses the distance between the centroid and outline to describe figure's contour to convert complex 2D images to simple 1D array, which increases the calculation speed greatly. But in this chapter centroid is just used as a static symbol instead of dynamical movement parameters. David Cunado takes trajectory of thigh's swing angle in walking as the gait feature [3]. The method successfully describes dynamic gait habits without contour extraction. However, the approach is feasible only when the pedestrian's leg outline is visual; it will lose worth when pedestrians wear gowns which cover their legs. Although there is technology which has improved the recognition rate greatly by extracting light stream information of image sequences [4], the cost is quite high because of complex algorithms' usage.

The former surveys revealed that human motion features can be reflected on the centroid trajectory. The overall process is that knee joint muscles make concession contraction from a person's one foot touching the ground, making the angle of the knee joint reduced and the leg shortened, the centroid showing a downward trend. In the driving process, all the joints and muscles of the leg involve restrained shrinkage to make the joint angles enlarge and the leg longer, the centroid showing an upward trend. Hao Ding and Junping Jiang had demonstrated that different gait owns different centroid swing arc or fluctuate curve, which proves that centroid is able to reflect essential features.

Considering human body as a nonhomogeneous object and each divided link as the uniformity, centroid can be calculated based on Varignon theorem without weighing the body which is expressed as a summary of all links. Centroid transforms gait period process to the clear centroid trajectory to avoid complex silhouette extractions and cycle tests, providing the theoretical basis for gait recognition towards commercialization.

## 59.2 Feature Abstract

### 59.2.1 *Moving Target Detection*

In this part, the main task is to detect the moving target from the original image sequences to get the two-value image sequences of pedestrians [5].

Figure 59.1a represents the original image obtained from the video database. It is firstly transferred into a single channel image and is smoothed by Gauss method.



**Fig. 59.1** Examples of moving silhouette extraction: (a) An original image, (b) the background image constructed by three-frame difference method, (c) the extracted silhouette form

Then separate the moving target from the background (Fig. 59.1b) to get moving targets by three-frame difference method:

$$\begin{cases} |I_n(x) - I_{n-1}(x)| > T_n(x) \\ |I_n(x) - I_{n-2}(x)| > T_n(x) \end{cases} \quad (59.1)$$

where  $I_n(x)$  represents domain value describing the gray change at position  $x$  of the  $n$ th image statistically,  $I_{n-1}(x)$  represents the domain value at position  $x$  of the  $(n-1)$  th frame, and  $T_n(x)$  represents an experimental threshold value. When the gray value at position  $x$  has changed beyond the threshold value  $T_n(x)$  relative to the previous frame  $I_{n-1}(x)$  and the frame  $I_{n-2}(x)$ , the pixel is considered to belong to moving objects and is separated from the background. Repetitively operations are to obtain a series of foreground images. The pixel domain value in figure area is set to 1 and the remaining 0, by which we can get two-value images of walking figures (Fig. 59.1c).

## 59.2.2 Calculation of Human Body's Centroid

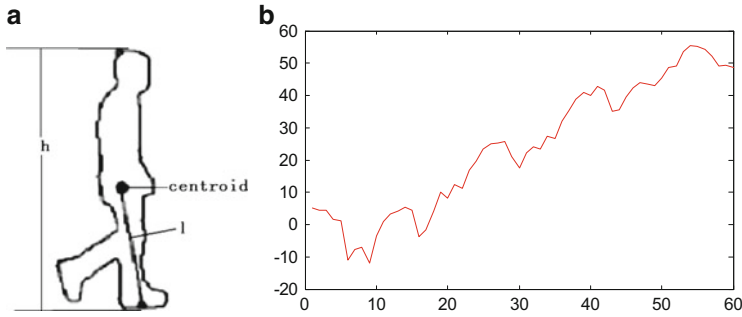
By centroid calculation formula (Eqs. (59.2) and (59.3)) these 2D silhouette changes are converted to an associated sequence of 1D signal. To eliminate the influences of spatial scale and signal length caused by walking direction, we normalize these centroid signals by dividing by the walker's silhouette height. This process is illustrated in Fig. 59.2.

The summation of target centroid can eliminate the impact of hollows. Variety laws of centroid can describe the essential gait features:

$$x_{w(i)} = \frac{1}{N} \sum_{x_i \in Area} x_i \quad (59.2)$$

$$y_{w(i)} = \frac{1}{N} \sum_{y_i \in Area} y_i \quad (59.3)$$

where  $N$  represents the number of pixels in the target area,  $x_i$  and  $y_i$  represent the X- and Y-coordinates of one pixel in the target area. Centroid coordinate is divided by the silhouette height of current frame to reduce the impact of the above factors; the formula is given by



**Fig. 59.2** Parameters used in calculation and centroid trajectory: (a) Position of centroid, distance of centroid to ankle, (b) tested centroid trajectory

$$h = y_{\max} - y_{\min} \tag{59.4}$$

$$gait_2 = \frac{gait_1}{h} \tag{59.5}$$

where  $h$  represents the current frame body height,  $y_{\max}$  represents human target’s maximum y-coordinate,  $y_{\min}$  represents human target’s minimum y-coordinate,  $gait_1$  represents the initial centroid coordinate, and  $gait_2$  represents normalized centroid coordinate. Finally construct a gait database storing the trajectories (Fig. 59.2b).

### 59.3 Wavelet Deionizing

The gait waveform is decomposed by the wavelet function to get high-frequency coefficients on every decomposition scale and low-frequency coefficients on maximum decomposition. The process can be achieved through the DWT () function in MATLAB Toolbox; the programming usage form is given by “[cAc, cDc] = Dwt (Yc, wname),” where “Yc” is the centroid trajectory vector, “wname” is the selected wavelet function’s name, “cAc” is the low-frequency coefficient vector, and “cDc” is the high-frequency coefficient vector after decomposition. Refactor along with processed low-frequency components “cAc” to get the deionized results. Adjust the wavelet function, and repeat the above steps until the ideal result is obtained.

### 59.4 Spectral Analysis

If legs are rigid, distance of the heel to the centroid remains the same. When one of the legs touches the ground, the other leg starts to swing immediately. In the frontal perspective, the trajectory of body’s centroid is a convex arc, with the heel as its center and the distance from the centroid to the heel as its radius [6]. The chart is shown in Fig. 59.3.

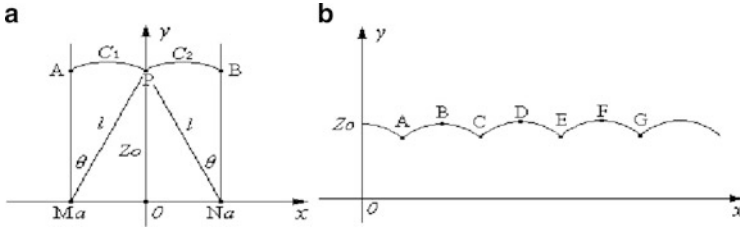


Fig. 59.3 Centroid’s mathematical model: (a) From the frontal view, (b) from the lateral view

In the lateral view, centroid’s trajectory is composed of fluctuations on the up and down direction and the walking direction. The one obtained in any perspective is composed of fluctuations on the up and down direction, the left and right direction, and the walking direction.

In Fig. 59.3a, direction of X-axis represents the left and right direction, and the positive direction of X-axis represents just the direction of people’s right hand. Y-axis represents centroid’s height. Curves C1 and C2 represent the transferring process from the state left foot is perpendicular to the ground to the state right foot does which takes point M and N as the center and distance  $l$  as the radius. In the arc, the boundaries of human body swinging on both sides is just when legs are perpendicular to the ground. Here  $\theta = \angle AMP = \angle BNP = \arcsin a/l$  is the angle of swing on left and right direction. As Fig. 59.4a shows, AN and BM represent the moment that left and right legs become perpendicular to the ground (Fig. 59.3a). In Fig. 59.3b, points A, C, E, and G represent the lowest points of centroid positions, which are the moments that two feet alternate up and down. Points B, D, and F represent the highest points of centroid positions, which are the moments that a leg becomes perpendicular to the ground [7].

When the distance of heel to centroid is long, radius formed from centroid’s swing will be relatively large; sine functions with greater cycle must be used in spectral analysis, which leads to appearance of peaks in the lower frequency band on spectrum map. However, fluctuations are mainly influenced by walking gestures. If the fluctuation amplitudes are large, amplitudes of peaks in every frequency band will be relatively large. Different human body characteristics and walking habits will show as different spectrum distributions.

Gait waveform is a continuous function  $x(t)$ , but in fact only limit values can be collected during the limit time. Take  $x(t)$  as a continuous function whose cycle is  $T$ ; then it can be expanded as a Fourier series; the exponential form is given by

$$x(t) = \sum_{K=-\infty}^{+\infty} C_K e^{j2\pi Kft} \tag{59.6}$$

$$C_K = \frac{1}{T} \int_{-T/2}^{T/2} x(t) e^{-j2\pi Kft} dt \tag{59.7}$$

where  $K = 0, \pm 1, \pm 2$ , and so on,  $f = 1/T$  is the fundamental frequency of  $x(t)$ , and  $kf$  represents  $K$  harmonic frequency.  $C_k$  is the Fourier coefficient of  $x(t)$ ; it is a complex.  $|C_k|$  is harmonic's amplitude.  $|C_0|$  is the harmonic's average amplitude, and  $|C_k|$  is known as the  $K$  harmonic amplitude.

We can just extract features from a spectrum map by segmentation to obtain data feature vector. First segment the X-axis representing frequency ranges into small bands; if there are peaks in one band, the quantization result is the number of peaks in the band; otherwise, it is 0. The combination of all the quantization values is the final quantization result.

## 59.5 Experimental Result

### 59.5.1 Result Analysis in Three Views

Experiments in this chapter are done on Dell personal computer with windows XP, MATLAB 7.0 as the simulating software, and C++ as the main programming language. Grouped by dress and carrying, the testers are divided into standard group, dress group, and burden group. Each group is measured from multi-angle. Record centroid trajectories of each group in each angle to construct a gait database.

Select 1,000 as the fundamental frequency, and simulate in the MATLAB to get results as shown in Fig. 59.4 (note: the spectrum is drawn  $f = 500$  as a symmetry axis symmetrical; analysis of the frequency spectrum is limited to the range  $f \leq 500$ ).

From Fig. 59.4a, there will be slight concussions in trajectory due to changes of walking ground, but the overall trend is the combination of fluctuations on the walking direction and the up and down direction, which on the spectrum map are two peaks on the band of 0–100 and 350–400.

In the frontal view, the trajectory is the combination of fluctuations on up and down direction and left and right direction. At this time it can be expressed by peak distribution and corresponding amplitudes on high-frequency band. As is shown in Fig. 59.5, the spectrum peaks allocate in each frequency band. By recording the peak positions we can get the walking figure's essential gait features.

In the perspective of  $45^\circ$ , measured trajectory is a superposition of all fluctuations on three directions. As is shown in Fig. 59.4, on  $45^\circ$  condition, spectrum is mainly composed of high-frequency components of fluctuations on left and right direction and low-frequency components of fluctuations on the up and down direction. The principal components distribute in the low-frequency band.

Figure 59.5 gives the frequency parameters of two persons in database in perspective of  $45^\circ$ . Select 50 as separation unit. The quantification result of Fig. 59.5a is "01,01,11,11,00" while that of Fig. 59.5b is "01,01,11,10,00." The Euclidean distance of the same person walking two times is 1; the quantification result of Fig. 59.5c is "11,01,11,10,10" while that of Fig. 59.5d is "11,01,11,10,10." The Euclidean distance that the same person walks two times is 0. In a word, the



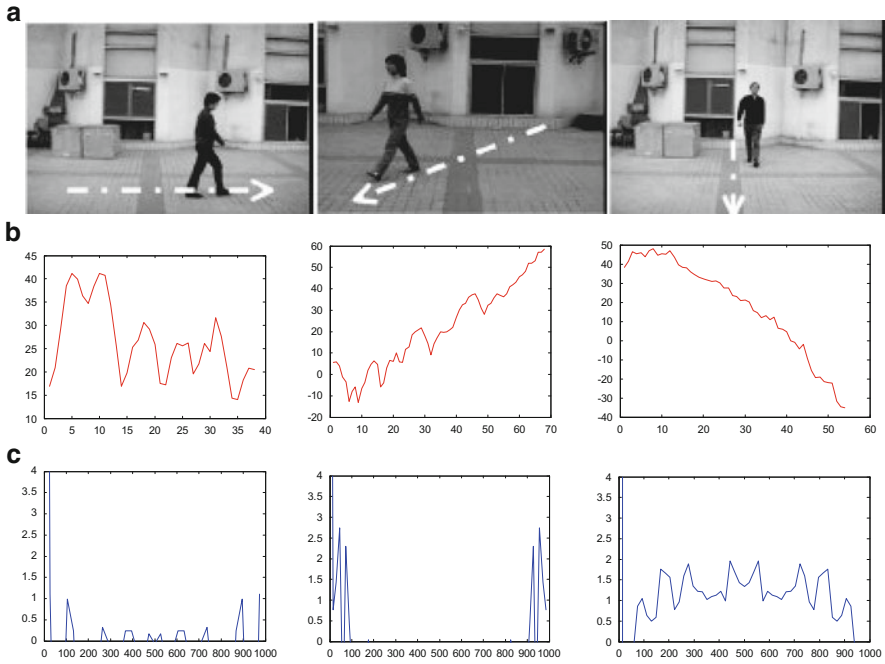


Fig. 59.4 Gait spectrum analysis: (a) Input sequences, (b) centroid trajectory, (c) spectrum map

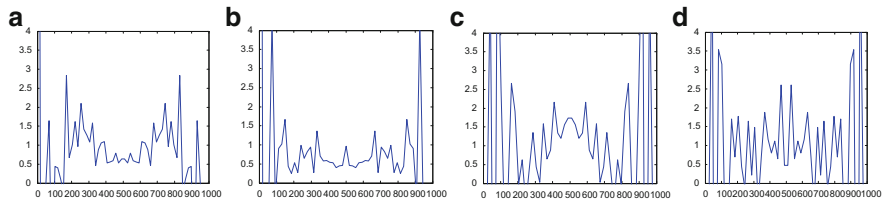
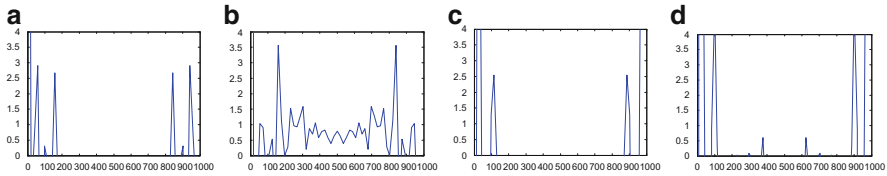


Fig. 59.5 Gait spectrum results: (a) First walking result of sixth person in NLPR database, (b) second walking result of sixth person in NLPR database, (c) first walking result of fifth person in NLPR database, (d) first walking result of fifth person in NLPR database

same person’s Euclidean distance is very small, while different person’s is large. This algorithm has a perfect distinction, and reducing the separating unit can further improve the ability of distinguishing.

### 59.5.2 Gait Analysis in Special Cases

In addition to the frontal view and lateral view, as centroid trajectory obtained from the rest angles are just cycle’s and amplitude’s overall enlargement or reduction compared with the 45° perspective, a strong unity can be detected in the spectrum



**Fig. 59.6** Gait spectrum analysis: (a) Person A walking from  $45^\circ$  perspective in the lab, (b) person A walking from  $135^\circ$  perspective in the lab, (c) person B walking from  $45^\circ$  perspective in the lab, (d) person B walking from  $135^\circ$  perspective in the lab

map obtained in  $45^\circ$  view and any other views. Individual frequency parameters obtained in  $45^\circ$  and  $135^\circ$  views are given in Fig. 59.6.

Changes of perspective mainly influence the details of the centroid trajectory, which will cause disturbance. The overall extracted characteristic can still reflect essential gait features. As the figure in one frame has been compressed into a point, carrying and dress will just influence the overall coordinate values of centroid trajectory instead of variety laws. Spectrum variety laws remain the same when pedestrians change their carrying or dress.

## 59.6 Conclusion

This chapter presents a new gait extraction method. It does not need complex contour extractions and cycle detections compared with the prior methods. The experimental results demonstrated that this extracting method can reflect essential gait characteristics of humans. It is not affected by dress, carrying, and shooting angle. Noise and low resolution will not affect the recognition result.

## References

1. Xu, D., & Huang, Y. (2012). Human gait recognition using patch distribution feature and locality-constrained group sparse representation[J]. *IEEE Transactions on Image Processing*, 21(1), 316–326.
2. Zhang, F., Zhang, X., & Cao, K. (2012). Contour extraction of gait recognition based on improved GVF snake model[J]. *Computers and Electrical Engineering*, 38(4), 882–890.
3. Cunado, D., & Nixon, M. S. (2003). Automatic extraction and description of human gait models for recognition purposes[J]. *Computer Vision and Image Understanding*, 90(1), 1–41.
4. Derawi, M. O., Ali, H., & Cheikh, F. A. (2011). Gait recognition using time-of-flight sensor[C]. In *Lecture Notes in Informatics (LNI), Proceedings-Series of the Gesellschaft für Informatik (GI). Gesellschaft für Informatik (GI), Ahrstrabe 45, 53175 Bonn, Germany* (pp. 187–194).
5. Yu, T., & Zou, J.-H. (2012). Research on gait recognition based on the combination of HMM and Bayes rules[J]. *Chinese Journal of Computers*, 35(2), 386–396 (In Chinese).

6. Ding, H., Jiang, J., & Li, Z. (2008). Establishment of human centroid trajectory mathematical model in walking. *Journal of Jiangsu Police Institute*, 23(6), 167–170. (In Chinese).
7. Peng, Z., Wu, X., & Yang, J. (2007). Multiview gait recognition algorithm based on the parameters of the limb length[J]. *Journal of Automation*, 33(2), 211–213. (In Chinese).

# Chapter 60

## An Algorithm for Bayesian Network Structure Learning Based on Simulated Annealing with Adaptive Selection Operator

Ao Lin, Bing Xiao, and Yi Zhu

**Abstract** In order to solve the problems that the intelligence algorithm falls into the local optimum easily and has a slow convergence in Bayesian networks (BN) structure learning, an algorithm based on adaptive selection operator with simulated annealing is proposed. This chapter conducts the adaptive selection rule in combination with conditional independence tests of BN nodes to guide the generation of neighbor. In order to better compare the adaptive effect, an algorithm based on selection operator with simulated annealing (SOSA) is proposed; at the same time 15 data sets in the three typical networks are accessed as learning samples. The results of the Bayesian Dirichlet (BD) score, Hamming distance (HD), and evolution time of the network after learning show that it has the quicker convergence and it searches the optimal solution more easily compared with simulated annealing (SA) and SOSA.

### 60.1 Introduction

As a graph model, BNs is an effective tool to deal with the uncertain problems in modeling and analysis, which is widely applied in many domains. Meanwhile, the learning problem of BNs is an important part of BN study. BN learning includes BN structural learning and parameter learning. Structure learning is needed to disclose the qualitative and the quantitative relationship between variables to light at the same time, while the BN structure learning is proved to be NP hard. Therefore, studying the BN structure learning problems is more challengeable and meaningful.

---

A. Lin (✉) • Y. Zhu  
Department of Graduation Management, Air Force Early Warning Academy,  
Wuhan 430019, China  
e-mail: [lin\\_ao4035@163.com](mailto:lin_ao4035@163.com)

B. Xiao  
No. 4 Department, Air Force Early Warning Academy, Wuhan 430019, China

In the investigation of BN structure learning, there are two classes of methods to deal with the structure learning problems [1]: The method of independence analysis and the method of score searching. The former determines whether there is border between the corresponding points or not by examining the conditional independence and dependence between variables, thereby establishing the skeleton of BN structure and orienting the border to get the BN structure. The latter is the method of score searching. For the network search space is of great extent generally, some BN structure learning adopts heuristic greedy algorithm, which tends to lead to the local optimum in the learning outcomes.

Nowadays, some researchers adopt the method of intelligence evolution to avoid the shortage of the heuristic algorithm [2, 3]. The SA algorithm is just the intelligence algorithm which is applied therein firstly, which is proved to be successful [4]. In contrast with the typical methods, it has a persistent evolution. In fact, examining the dependence and the independence between variables can reveal the variables' relational information which is camouflaged in data. Using this information can guide the evolution of the intelligence algorithm, thereby achieving the target of rapid convergence. ASOSA algorithm for the BN structure learning is proposed in this chapter.

## 60.2 BN Structure Learning

As a pictorial model that represents the joint probability distributions between variables, BNs include two parts: directed acyclic graph (DAG) and BN parameters. BN joint probability distributions can be decomposed as following through the independence relationship between variables contained in BN structure:

$$p(X_1, \dots, X_n) = \prod_{i=1}^n p(X_i | \mathbf{Pa}_i, G) \quad (60.1)$$

wherein  $\mathbf{Pa}_i$  refers to the father node of the variable  $X_i$  in BN structure (G). BN structure learning refers to obtaining the network structure which matches the sample data fitting best by analyzing a variety of samples, which is the focus of this chapter.

The method based on conditional independence test is efficient. But in certain cases, the conditional independence test order (the variables' number of the conditional set is the number of the conditional independence test orders) increases exponentially with respect to the number of the variables.

This chapter adopts the mutual information and conditional mutual information in the conditional independence test.  $I(X_i; X_j)$  expresses the mutual information between variables  $X_i$  and  $X_j$ . According to the information theory,  $I(X_i; X_j)$  is

$$I(X_i; X_j) = I(X_j; X_i) = \sum_{X_i, X_j} p(X_i, X_j) \log \frac{p(X_i, X_j)}{p(X_i)p(X_j)} \quad (60.2)$$

The mutual information is nonnegative. The more variables  $X_i$  and  $X_j$  incline to independence, the more  $I(X_i; X_j)$  approaches 0.

The method based on score searching is another method of the BN structure learning; this method defines grading function  $S$  for each candidate network. Generally,  $S$  is posterior probability of the network. When assuming that BNs have equal a priori probability, the comparison of the posterior probability  $S$  of different BN structures is the comparison of the structure likelihood  $p(D|G)$ . Cooper and Herskovits provided the calculating procedure of the structure likelihood:

$$p(D|G) = \prod_{i=1}^n \prod_{j=1}^{q_i} \frac{\Gamma(\alpha_{ij})}{\Gamma(\alpha_{ij} + N_{ij})} \prod_{k=1}^{r_i} \frac{\Gamma(\alpha_{ijk} + N_{ijk})}{\Gamma(\alpha_{ijk})} \quad (60.3)$$

wherein  $\Gamma$  is gamma function,  $N_{ijk}$  is the number of cases in the dataset in which the parents of  $X_i$  are in state  $j$  and  $X_i$  itself is in state  $k$ , and  $q_i$  and  $r_i$  are the number of the parents of  $X_i$  and  $X_i$  in its own state separately.  $\alpha_{ij}$  and  $\alpha_{ijk}$  are the Dirichlet prior distributions. Equation (60.3) is also the famous Bayesian Dirichlet (BD) score function; the greater the structure score, the better the network structure.

### 60.3 Adaptive Selection Operator

The traditional SA algorithm of BN structure learning gets the neighbor by randomly selecting add, delete, or reverse the directed edges. This kind of operation is purposeless and ineffective. The zero-order independence information of the BN nodes can characterize the relationship between nodes substantially. So the information can be used to guide the generation of neighbor. The main thought is constructing the initial selection matrix with the nodes' zero-order information in the independence test. With the conducting of the annealing, effect of selection matrix is weakened gradually and selection matrix will not change any more when it meets certain conditions.

The selection matrix can be classified into the dependence selection matrix and the independence selection matrix which are applied to the addition and the deletion of the BN edges, respectively.

The selection matrix can be gained in two steps:

**Step 1: Initialization.** Set the coefficient of renovation  $k$  ( $k > 1$ ). Calculate the mutual information of the two different random variable  $C_{ij}^0 = I(X_i; X_j)$ ,  $i = 1 \cdots m - 1$ , and  $j = i + 1 \cdots m$ ;  $m$  is the number of the BN nodes. The matrix  $C^0$  is the initial dependence selection matrix; the independence selection matrix and the dependence selection matrix have the opposite

effects. In order to ensure the nonnegativity of the probability,  $D_{ij}^0 = -C_{ij}^0 + \max(C_{ij}^0)$ ,  $i = 1 \cdots m - 1$ ,  $j = i + 1 \cdots m$ .  $D^0$  is the initial independence selection matrix.

Step 2: Updating.  $p = p + 1$ .  $C_{ij}^p = (C_{ij}^{p-1})^{1/k}$ ,  $D_{ij}^p = (D_{ij}^{p-1})^{1/k}$ ,  $p \geq 1$  when it meets the conditions;  $C_{ij}^p = C_{ij}^{p-1}$ ,  $D_{ij}^p = D_{ij}^{p-1}$ ,  $p \geq 1$  when it does not meet the conditions.

In order to keep the selectivity of the selection matrix, the maximum of the number of updating can be set.

## 60.4 ASOSA Algorithm

For the SA algorithm of BN structure learning, the initialized network structure is empty. It looks for the better network locally through the operation such as randomly adding, deleting the edge, and converting the edge's direction. Lower the temperature gradually to look for the locally optimal network, until the suspense condition is reached. Adaptive selection operator with simulated annealing (ASOSA) algorithm has the same main body frame as the SA algorithm, but it uses the selection probability of the selection matrix instead of the random selection to operate the edges. The algorithm proposed in this chapter is presented in Algorithm 1.

The marking criterion of the algorithm based on ASOSA adopts the BD score. For SA algorithm seeking for the minimum, the score results should maintain the negative value only. The condition of the step 13 and 14 refers to  $\Delta S \leq 0$ , and the updating time does not reach the maximal time  $\lambda$ .

In order to compare the effectiveness of the ASOSA algorithm, an algorithm based on selection operator with simulated annealing (SOSA) is designed the selection matrix of which will not change in the annealing process.

## 60.5 Experimentation

The standard way of assessing the effectiveness of a learning algorithm is to draw samples from a known BN, apply the algorithm on the artificial data, and to compare the learned structure with the original one [3]. This chapter chooses three typical networks in different domains for the experiment: Asia network [5] (8 variables, 8 edges), insurance network [6] (27 variables, 52 edges), and alarm network [7] (37 variables, 49 edges). For the three experimental networks, datasets with 100, 200, 500, 1,000, and 5,000 samples were generated separately by applying Gibbs sampling.

**Algorithm 1: ASOSA algorithm****Input:** Set of learning data**Output:** Bayesian network

1. Set the initial temperature  $T_0$ .
2. Set the minimum temperature  $T_{end}$ .
3.  $\nu = 0.99$ ; //The descent velocity of temperature.
4.  $\beta = 20$ ; //The maximal time of the outside loop.
5.  $\sigma = 20$ ; //The maximal time of the inside loop.
6. Calculate the mutual information between any two variables to get  $C^0$ .
7.  $D^0 = -C^0 + \max(C^0)$ .
8.  $k = 1.1$ ; //Coefficient of renovation.
9.  $\lambda = 15$ ; //The maximal updating time of the selection matrix.
10.  $G =$  empty graph; //The candidate graph is initialized into empty graph.
11.  $T = T_0$ .
12. **Repeat**
13. **If** the conditions are met, update the selection matrix  $C$  and  $D$ ; **end**.
14. **If** the conditions are not met, do not update the selection matrix; **end**.
15. **For**  $\sigma$  times **do**.
16. Add one edge and reduce one edge to  $G$ , according to the selection matrix  $C$  and  $D$ , and reverse the direction of one of the directed edges in  $G$  randomly.
17. Calculate the score difference  $\Delta S$  coming from the three operations above.
18. **If**  $\Delta S > 0$  or  $e^{(\Delta S/T)} > rand(0,1)$  **then** apply these actions to  $G$ ; **end**.
19. **End**.
20.  $T = T \times \nu$ .
21. **Until** the maximal time of the loop or  $T > T_{end}$  is obtained.
22. **Return**  $G$ .

**60.5.1 Qualitative Analysis of Algorithms**

Table 60.1 shows the BD score statistics to the learning results of the 15 experimental datasets from the three experimental networks by three algorithms (SA, SOSA, and ASOSA).

Hamming distance (HD) is an available approach to describe the difference between the network  $G$  after learning and the original network  $G^0$ . HD is the sum of excessive edge, deleted edge, and reversed edge. Table 60.2 shows the statistic results of the three different algorithms from the three different samples. It can be found from the statistic results of the HD that the performance of ASOSA is optimal and the posterior is SOSA.



**Table 60.1** BD score for the learned structure (normalized for correct network score)

Sample	Method	Sample size					Avg.
		100	200	500	1,000	5,000	
Asia	SA	1.008	1.010	1.009	1.004	1.004	1.007
	SOSA	0.994	0.995	0.997	1.000	1.000	0.997
	ASOSA	0.994	0.995	0.997	0.999	1.000	0.997
	TRUE	1.000	1.000	1.000	1.000	1.000	1.000
	Empty	1.271	1.384	1.333	1.336	1.322	1.329
Insurance	SA	0.876	0.932	0.970	0.990	1.003	0.954
	SOSA	0.889	0.916	0.965	0.982	1.000	0.950
	ASOSA	0.886	0.927	0.964	0.982	0.998	0.951
	TRUE	1.000	1.000	1.000	1.000	1.000	1.000
	Empty	1.109	1.251	1.403	1.492	1.588	1.369
Alarm	SA	0.994	1.009	1.006	1.005	1.005	1.004
	SOSA	0.997	1.004	1.002	1.004	1.007	1.003
	ASOSA	0.990	1.003	0.996	0.999	1.000	0.998
	TRUE	1.000	1.000	1.000	1.000	1.000	1.000
	Empty	1.571	1.682	1.768	1.836	1.888	1.749

**Table 60.2** HD for the learned structure

Sample	Method	Sample size					Avg.
		100	200	500	1,000	5,000	
Asia	SA	15	15	15	8	13	13.2
	SOSA	5	7	3	0	2	3.4
	ASOSA	5	3	3	2	2	3.0
Insurance	SA	53	43	31	37	22	37.2
	SOSA	49	29	25	24	23	30.0
	ASOSA	40	25	22	24	22	26.6
Alarm	SA	45	34	33	21	26	31.8
	SOSA	33	31	21	16	21	24.4
	ASOSA	25	20	18	13	15	18.2

Since the operation platforms of the algorithms are at variance, the scores should be normalized for the correct network score. As BD score is subtractive, the lower the normalized score, the better the network structure

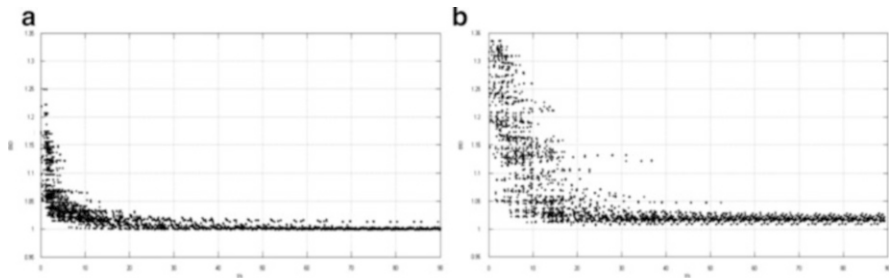
It can be found from the learning results of the three networks that the normalized score of the SA algorithm is higher than the other two improved algorithms except the 100 samples of the insurance and alarm network learning, and the score of SOSA and ASOSA are superior to SA algorithm. In the learning, the score of ASOSA is no higher than SOSA in each group or on an average

### 60.5.2 Constringent Analysis of Algorithms

The runtime of algorithm can be the leading indicator to measure the algorithm efficiency, but the final results of the different algorithms are different. For convenience in comparison, take the time consumption of the BD final score of the SA

**Table 60.3** Runtime (normalized for SA)

Sample	Method	Sample size					Avg.
		100	200	500	1,000	5,000	
Asia	SA	1.30	1.15	1.26	1.00	1.15	1.30
	SOSA	0.22	0.26	0.52	0.37	0.30	0.22
	ASOSA	0.19	0.19	0.33	0.26	0.15	0.19
Insurance	SA	1.15	1.00	1.10	1.13	1.47	1.17
	SOSA	1.04	0.63	0.55	0.67	0.71	0.72
	ASOSA	1.06	0.63	0.58	0.52	0.64	0.69
Alarm	SA	1.15	1.00	1.07	1.10	1.48	1.16
	SOSA	0.85	0.67	0.70	0.78	0.78	0.76
	ASOSA	0.62	0.55	0.60	0.72	0.78	0.65



**Fig. 60.1** (a and b) The performance record of ASOSA and SA

algorithm calculated by SOSA and ASOSA as the comparison object, and the result is normalized with the computation time of the SA, as shown in Table 60.3.

It can be found from the comparison of the runtime of the algorithms that the time of ASOSA is smaller than the other two algorithms, but there is little difference in SOSA and ASOSA.

In order to better compare the convergence of the algorithms, the performance process of the Asia network learning results from 500 samples which is obtained by ASOSA and SA is recorded. The cross axle is time, and the axis of ordinates is the BD score result normalized for the correct network score as shown in Fig. 60.1.

It can be found from the figure above that ASOSA converges rapidly. The searching directivity of the initial algorithm is conspicuous. The final convergence result is in the low level. The convergence process of the traditional SA is slow, and the final convergence result is worse than that of ASOSA.

## 60.6 Conclusion

This chapter introduces the ASOSA algorithm for BN structure learning, which fuses the independence analysis method and the score searching method of the BN structure learning, by the agency of the searching optimal network of the simulated annealing intelligence algorithm. The comparison of learning in three different

samples from multi-aspect is among the algorithms ASOSA, SOSA, and SA. The result shows that the ASOSA is superior to SOSA and SA in the learning accuracy and the time consumption.

## References

1. Singh, M., & Valtorta, M. (1995). Construction of Bayesian network structures from data: A brief survey and an efficient algorithm[J]. *International Journal of Approximate Reasoning*, 12(2), 111–131.
2. Wong, M. L., & Leung, K. S. (2004). An efficient data mining method for learning Bayesian networks using an evolutionary algorithm-based hybrid approach[J]. *IEEE Transactions on Evolutionary Computation*, 8(4), 378–404.
3. Pinto, P. C., Nagele, A., et al. (2009). Using a local discovery ant algorithm for Bayesian network structure learning[J]. *IEEE Transactions on Evolutionary Computation*, 13(4), 767–777.
4. Kirkpatrick, S., Gelatt, C. D., et al. (1983). Optimization by simulated annealing[J]. *Science*, 220(4598), 671–680.
5. Lauritzen, S. L., Spiegelhalter, D. J., et al. (1988). Local computations with probabilities on graphical structures and their application to expert systems[J]. *Journal of the Royal Statistical Society, Series B*, 50(2), 157–224.
6. Binder, J., Koller, D., et al. (1997). Adaptive probabilistic networks with hidden variables [J]. *Machine Learning*, 29(2–3), 213–244.
7. Beinlinch, I. A., Suermond, H. J., et al. (1989). The ALARM monitoring system: A case study with two probabilistic inference techniques for belief networks[C]. In *Proc. 2nd Europ. Conf. on Artificial Intelligence in Medicine Care* (pp. 247–256). Berlin: Springer-Verlag.

# Chapter 61

## Static Image Segmentation Using Polar Space Transformation Technique

Xuan Luo, Tiancai Liang, and Weifeng Wang

**Abstract** This chapter proposes a polar space-based method to segment the static image automatically. The proposed method aims at segmenting the object of interest by finding the optimal closed contour in the polar space, solving the long-term problem of scale in the Cartesian space. Experimental results further verify and demonstrate the efficacy of the proposed polar space-based method on the challenging datasets.

### 61.1 Introduction

In computer vision literature, segmentation is the process of partitioning an image into disjoint, homogeneous, and compact regions where each part constitutes connected pixels with similar properties, such as brightness, color, and texture. Over the years, many algorithms [1–3] have been proposed for segmentation. Generally speaking, all the algorithms could be classified into three categories: (1) feature space-based techniques; (2) image domain-based techniques; and (3) physics-based techniques [4].

This chapter addresses the problem of separating the object of interest from a static image. We propose a segmentation framework that takes a point as its input and outputs the region containing that point, as shown in Fig. 61.1. Essentially, segmenting this region is equivalent to find the enclosing contour, which is a

---

X. Luo (✉)

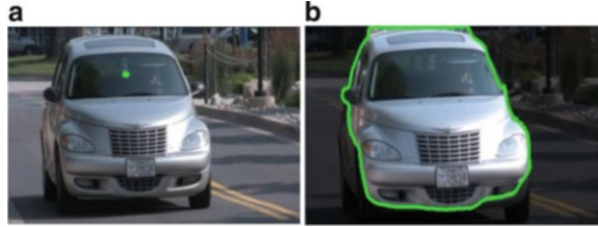
School of Information Science and Technology, Sun Yat-sen University,  
Guangzhou 510006, China

Research Institute of GRG Banking Equipment Co., Ltd., Guangzhou 510006, China  
e-mail: [xuanluo@ieee.org](mailto:xuanluo@ieee.org)

T. Liang • W. Wang

Research Institute of GRG Banking Equipment Co., Ltd., Guangzhou 510006, China

**Fig. 61.1** (a) The image with an input point. (b) The final segmentation given by the proposed approach



connected set of boundary edge fragments in the edge map, around the input point. The edge map is generated by using all available visual cues.

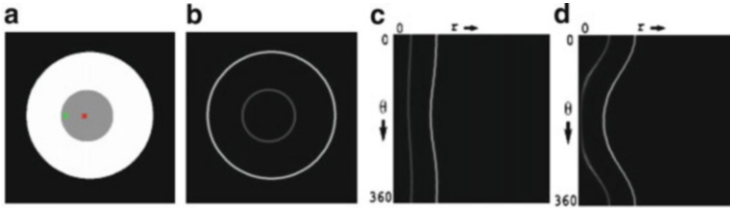
The proposed algorithm framework is a four-step process: First, the edge map of the image is generated using low-level cues; second, the edge map is transformed from Cartesian space to polar space with the input point as the pole; third, the “optimal” path through this transformed edge map is found; fourth, the path is mapped back to an enclosing contour. Figure 61.1b shows the final result.

## 61.2 The Significance of Polar Space

At present, any segmentation algorithm, which finds the “optimal” enclosing contour through edge map in the Cartesian space, usually prefers the smaller contours as the overall cost, which is defined as the product of the length of the closed contour and the average cost of tracing the edge pixel along the contour, increases with the contour size [3]. Take Fig. 61.2a for example; we intend to find the optimal contour for the red point in it. Figure 61.2b shows the gradient edge map of the disc. Apparently, the small circle is just the internal edge, while the big circle is the actual boundary of the disc. Let us make an assumption: (1) The edge map assigns the internal contour intensity 0.33 and the boundary contour 0.66; (2) the numbers of the pixels of the two circles are 100 and 250, respectively. According to the definition of the overall cost, we can get the cost of internal contour 67 and boundary contour 85. Unfortunately, the former one would be considered the “optimal” enclosing contour as a result of its lower cost.

Focus on dealing with this “the shorter, the better” problem, we propose to transform these contours from Cartesian space to polar space, where the lengths of these contours no longer depend on the area they enclose. In the polar space, the cost of tracing these contours would be independent of their scales in the Cartesian space. The transformed edge map with the red point as a pole is shown in Fig. 61.2c; both contours are changed into open curves, arranging from  $0^\circ$  to  $360^\circ$ . In this way, the costs of tracing the internal contour and the external contour become 242 and 124. Clearly, the actual boundary contour becomes the optimal contour logically.

For the new input point (the green one) in Fig. 61.2a, although both the corresponding contours have changed shape (see Fig. 61.2d), the brighter one still remains optimal.



**Fig. 61.2** (a) The mentioned disc. (b) The gradient edge map. (c) The polar edge map corresponding to the *red point*. (d) The polar edge map corresponding to the *green point*

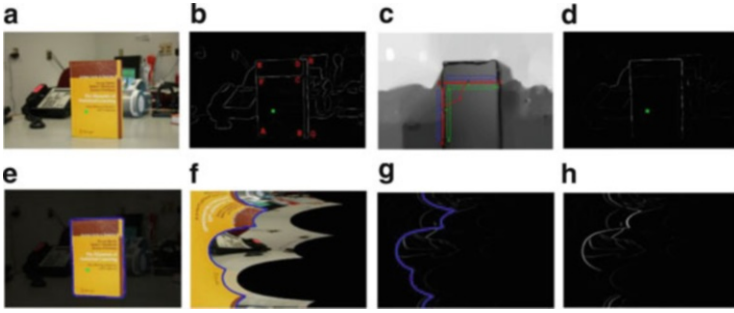
### 61.3 Generating the Edge Map by Combining Cues

In this section, we explicate the first step of the proposed segmentation algorithm: generating the edge map where the boundary edges (the actual boundary) are much brighter than the internal edges. In most scenes, the static visual cues such as color, intensity, or texture can precisely locate the edges. In our framework, we use the Berkeley edge detector [5], which learns the color and texture properties of the boundary pixels versus the internal pixels from a dataset containing human-labeled segmentations of 300 images, to generate our initial boundary edge map, as shown in Fig. 61.3b. This edge detector handles texture much better than any intensity-based edge detectors. We can see that the unauthentic texture edges have been basically removed while the boundary edges are bright.

However, some internal edges (see BC, CD, and DF in Fig. 61.3b) still remain bright. In order to separate the boundary edges from internal edges, we can use the motion cue sufficiently. It is known that the optical flow value changes significantly at the boundary of an object while substantially remaining unchanged inside an object. Based on this concept, we can modify the edge map so that the edge pixels with strong gradient of optical flow values are stronger than the ones with weak gradient. We break the initial edge map into straight line segments and select rectangular regions of width  $\alpha$  at a distance  $\beta$  on its both sides (see FC and FA in Fig. 61.3c). Then, we calculate the average flow inside these rectangles. The difference in the magnitude of the average flow on both sides is the standard measurement of the probability of the segment to be boundary edge. The brightness of an edge pixel on the segment is changed as

$$I'(x, y) = \lambda I(x, y) + (1 - \lambda) \Delta f / \max(\Delta f) \quad (61.1)$$

$\Delta f$  represents the change in optical flow;  $\lambda$  is the weight related to the relative importance of the static visual cue-based boundary estimate. Figure 61.3d shows the final boundary edge map where the internal edges are clearly fainter and the boundary edges are brighter.



**Fig. 61.3** (a) An example of the input image. (b) The Berkeley edge map. (c) The magnitude of optical flow field. (d) The boundary edge map combining the static visual cues with motion cue. (e) The final output by mapping back the optimal path to the Cartesian space. (f) The optimal cut dividing the polar image into two parts: *left* (inside the object) and *right* (outside the object). (g) The optimal contour overlapped on the polar edge map. (h) The corresponding polar edge map

### 61.4 Transforming from Cartesian Space to Polar Space

Let us say  $I^p(\cdot)$  is the corresponding polar edge map of the edge map  $I^c(\cdot)$  in the Cartesian space and  $Q(x_0, y_0)$  is chosen as the input point. We can see that a pixel  $I^p(r, \theta)$  in the polar coordinates corresponds to a pixel location  $\{I^c(x, y) : x = r \cos \theta + x_0, y = r \sin \theta + y_0\}$  in the Cartesian space. Adopting the method of bilinear interpolation, which only considers four immediate neighbors, we can figure out  $I^c(x, y)$  directly.

We propose to generate a continuous 2D function  $F(\cdot)$  by putting 2D Gaussian kernel functions on every edge pixel and aligning the major axis of those Gaussian kernel functions with the orientation of the edge pixel. Let  $S$  be the set of all edge pixels. The intensity at any pixel location  $(x, y)$  in the Cartesian coordinates is defined as

$$F(x, y) = \sum_{i \in S} \exp\left(-\frac{x_i^t}{\sigma_{x_i}^2} - \frac{y_i^t}{\sigma_{y_i}^2}\right) * I^c(x_i, y_i) \tag{61.2}$$

where

$$\begin{bmatrix} x_i^t \\ y_i^t \end{bmatrix} = \begin{bmatrix} \cos \theta_i & \sin \theta_i \\ -\sin \theta_i & \cos \theta_i \end{bmatrix} \begin{bmatrix} x_i - x \\ y_i - y \end{bmatrix} \tag{61.3}$$

$$\sigma_{x_i}^2 = \frac{A_1}{\sqrt{(x_i - x_0)^2 + (y_i - y_0)^2}} \tag{61.4}$$

$$\sigma_{y_i}^2 = A_2 \tag{61.5}$$

$\theta_i$  is the orientation of the edge pixel  $i$  and  $A_1$  and  $A_2$  are constant and are chosen to be 900 and 4, respectively, in our experiments. On account of keeping the gray values of the edge pixels in the polar edge map the same as the corresponding edge pixels in the Cartesian edge map, we set the square of variance along the major axis,  $\sigma_{x_i}^2$ , to be inversely proportional to the distance between the edge pixel  $i$  and the pole  $Q$ .

The polar edge map  $I^p(r, \theta)$  is obtained by sampling  $F(x, y)$ . The intensity values of  $I^p(r, \theta)$  are scaled to the interval ranging from 0 to 1, which can depict the probability of an edge pixel being at boundary. Figure 61.3h displays the polar edge map corresponding to Fig. 61.3d. What is more, we come to an agreement that the angle  $\theta \in [0^\circ, 360^\circ)$  is represented along the vertical axis and increases from top to bottom while the radius  $r \in [0, r_{max}]$  varies along the horizontal axis and increases from left to right.  $r_{max}$  stands for the maximum Euclidean distance between two arbitrary pixels in the image.

## 61.5 Finding the Optimal Cut Through the Polar Edge Map

Let us regard every pixel  $p \in P$  of  $I^p$  as a node in a graph. Every node (pixel) is connected with their four immediate neighbors (see Fig. 61.3). A row of the graph represents the radiation from the input point at an angle ( $\theta$ ) equal to their row number. The first and the last rows are the rays  $\theta = 0^\circ$  and  $\theta = 360^\circ$ , which are exactly the same in the polar space. So, the pairs of nodes  $\{(0^\circ, r), (360^\circ, r)\}$  and  $\forall r \in [0, r_{max}]$  could be connected by edges in the graph. Let us say  $l = \{0, 1\}$  are the two possible labels for each pixel, where  $l_p = 0$  indicates inside the object and  $l_p = 1$  indicates outside the object.  $\gamma$  denotes the set of all the edges between neighboring nodes in the graph. Evidently, we aim at assigning a label  $l_p$  to every pixel  $p$  (i.e., finding the mapping  $M(p) \rightarrow l$ ), which corresponds to the minimum energy. The energy function is

$$G(M) = \sum_{p \in P} \phi_p(l_p) + \mu \sum_{(p,q) \in \gamma} \varphi_{p,q} \cdot \delta(l_p, l_q) \quad (61.6)$$

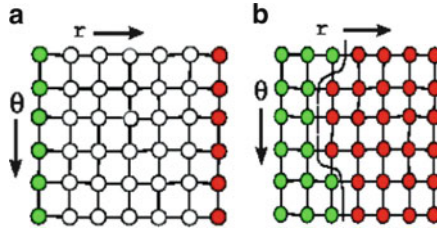
$$\varphi_{p,q} = \begin{cases} \exp(-\tau I_{pq}^p) & \text{if } I_{pq}^p \neq 0 \\ C & \text{otherwise} \end{cases} \quad (61.7)$$

$$\delta(l_p, l_q) = \begin{cases} 1 & \text{if } l_p \neq l_q \\ 0 & \text{otherwise} \end{cases} \quad (61.8)$$

$$I_{pq}^p = (I^p(r_p, \theta_p) + I^p(r_q, \theta_q)) / 2 \quad (61.9)$$

The cost of assigning a label  $l_p$  to the pixel  $p$  is denoted by  $\phi_p(l_p)$ , and the cost of assigning different labels to the neighboring pixels  $p$  and  $q$  is denoted by  $\varphi_{p,q}$ .





**Fig. 61.4** (a) The *green* nodes in the first column are initialized to be inside the object, while the *red* nodes in the last column are initialized to be outside the object. (b) The binary labeling output after minimizing the energy function by means of graph cut algorithm

At the beginning of this process, we assign value 0 to the data term  $\phi_p(l_p)$  for all the nodes except those in the first column and the last column:

$$\phi_p(l_p) = 0, \quad \forall p \in (r, \theta), \quad r \in (0, r_{max}), \quad \theta \in [0^\circ, 360^\circ) \quad (61.10)$$

Meanwhile, the nodes in the first column (corresponding to the input point in the Cartesian space) must be inside the object and are assigned label 0:

$$\phi_p(l_p = 1) = Z, \quad \phi_p(l_p = 0) = 0, \quad p \in (0, \theta), \quad \theta \in [0^\circ, 360^\circ) \quad (61.11)$$

Similarly, the nodes in the last column must be outside the object and are assigned label 1:

$$\phi_p(l_p = 0) = Z, \quad \phi_p(l_p = 1) = 0, \quad p \in (r_{max}, \theta), \quad \theta \in [0^\circ, 360^\circ) \quad (61.12)$$

Shown in Fig. 61.4.  $Z$  must be a high value in order to make sure that the initial labels to the first and the last columns could not change in the minimization step. To find the optimal path in the polar edge map, we apply the planar graph cut algorithm [6] to minimize the energy function  $G(M)$ , given in Eq. (61.6).

The global optimization process divides the polar edge map into two parts: left side (inside the object) and right side (outside the object), see Fig. 61.3f, g. Finally, we map the resulting binary segmentation back to the Cartesian space to obtain the desired result. The boundary between the left (label 0) and the right (label 1) parts in the polar space corresponds to the closed contour around the input point in the Cartesian space, see Fig. 61.3f, e.

## 61.6 Experimental Results

We evaluated the performance of the proposed algorithm on 100 image pairs along with their ground-truth segmentation, which is created by identifying and segmenting the most outstanding object of interest for each image pair manually.

**Table 61.1** The performance of the proposed method for the test images

For test images	F-measure
Without motion	$0.65 \pm 0.02$
With motion	$0.94 \pm 0.01$

**Fig. 61.5** Row 1: The original images with an input point. Row 2: The segmentation using the static visual cues only. Row 3: The segmentation for the same input point after combining static visual cues with motion cue

The final segmentation of our algorithm is compared with the ground-truth segmentation according to the empirical F-measure:

$$\Omega = \frac{2PR}{P + R} \quad (61.13)$$

$P$  symbolizes the precision which calculates the score of our segmentation overlapping with the ground-truth segmentation, and  $R$  symbolizes the recall which calculates the score of the ground-truth segmentation overlapping with our segmentation.

See Table 61.1; there has been considerable improvement in the proposed algorithm after combining static visual cues with motion cue. With the visual cues only, the actual boundary contour is not drawn precisely owing to the strong internal edges (see row 2 of Fig. 61.5). Nevertheless, the affiliation of optical flow cue makes the internal edges fade away, and the correct contour is found (see row 3 of Fig. 61.5).

On the other hand, we evaluate the performance of the proposed algorithm by comparing it with state-of-the-art methods [1, 2, 7, 8]. We use the public Alpert image database [8] for our comparative experiments. We perform better than [1, 2] and extraordinarily close to [7, 8], as clearly described in Table 61.2.

**Table 61.2** The performance of the proposed method compared with other methods

Algorithm	F-measure
Bagon [7]	$0.87 \pm 0.010$
Alpert [8]	$0.86 \pm 0.012$
Ours	$0.84 \pm 0.018$
NCut [2]	$0.72 \pm 0.012$
MeanShift [1]	$0.57 \pm 0.023$

## 61.7 Discussion and Conclusion

In this chapter, we have presented an efficient image segmentation algorithm. The framework cleverly uses all the visual cues to distinguish the internal contours from the actual boundary contours and then transforms the edge map from Cartesian space to polar space in order to find the optimal segmentation. As the future extension of this work, the algorithm can be used to segment large number of images, and the extracted regions can be studied for the high-level processes.

## References

1. Tu, Z. W., & Zhu, S. C. (2002). Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(5), 603–619.
2. Shi, J. B., & Malik, J. (2000). Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8), 888–905.
3. Mishra, A. K., Aloimonos, Y., Cheong, L. F., & Kassim, A. A. (2012). Active visual segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34(4), 639–653.
4. Luccheseysz, L., & Mitray, S. K. (2001). Color image segmentation: A state-of-the-art survey. *Proceedings of the Indian National Science Academy (INSA-A)*, 67(2), 207–221.
5. Martin, D., Fowlkes, C., & Malik, J. (2004). Learning to detect natural image boundaries using local brightness, color and texture cues. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(5), 530–549.
6. Boykov, Y., & Kolmogorov, V. (2004). An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 26(3), 359–374.
7. Bagon, S., Boiman, O., & Irani, M. (2008). What is a good image segment? A unified approach to segment extraction. In *ECCV* (Vol. 5305(3), pp. 30–44). Heidelberg: Springer.
8. Alpert, S., Galun, M., Basri, R., & Brandt, A. (2007). Image segmentation by probabilistic bottom-up aggregation and cue integration. In *Computer Vision and Pattern Recognition* (Vol. 0, No. 2, pp. 1–8).

# Chapter 62

## Image Restoration via Nonlocal $P$ -Laplace Regularization

Chen Yao, Lijuan Hong, and Yunfei Cheng

**Abstract** Image restoration technology can be applied in a lot of fields including image communication, image archive restoration, and image editing. In this chapter, we try to solve image restoration with a nonlocal regularization point of view. Similarity between different image pixels is measured by a nonlocal  $p$ -Laplace operator. We use minimum least square with a regularization term to formulate the whole procedure of image restoration. In the solvent of cost function, a linear Gauss–Jacobi iterative method is utilized for unknown pixel solvent. The complexity of iterative solvent is controlled by a step threshold. Finally, experimental results highlight our superior performance over previous methods from subject visual perception or object image quality assessment.

### 62.1 Introduction

In consumer electronic application, image is often degraded by cameral sensor noise or error-prone communication channel. Image restoration is a technology which uses non-corrupted image region to recover corrupted pixels. With the application of image and video, image restoration problems are bringing up more and more concerns [1, 2]. A lot of researches have been carried on image restoration. Interpolation-based methods are adopted for finding good model of natural images [3–5]. Anisotropic filtering, partial differential equation, total variation, and image decompositions on fixed bases such as wavelets are also used to implement image restoration [6–9]. Generally, a smoothness assumption is built for these technologies. More recently, the research on image self-similarity is carried on nonlocal filtering, learned sparse model, sparse dictionary learning, Gaussian scale mixture, fields of experts, and block matching with 3D filtering

---

C. Yao (✉) • L. Hong • Y. Cheng  
The Third Research Institute of Ministry of Public Security, Shanghai 200031, China  
e-mail: [yaochensing@126.com](mailto:yaochensing@126.com)

(BM3D) [10–15]. The image self-similarity technology brings more motivation for image restoration.

In this chapter, inspired by forerunners [16, 17], we present a new image restoration method based on nonlocal  $p$ -Laplace regularization scheme. In this chapter, image is viewed as discrete graph. Every pixel is regarded as graph vertices. The similarity between different vertices is represented by a nonlocal weight. A  $p$ -Laplace operator is defined as an operation of vertex Hilbert space. The diffusion of discrete vertices is formulated with an energy minimum function. For a given energy minimum cost function, a Gauss–Jacobi iterative method is used for unknown variable solvent. A constant threshold is selected as algorithm stop criterion.

In the rest part of this chapter, the energy minimum solving framework, nonlocal weight, and iterative solving method are reviewed and introduced in Sect. 62.2. Abound experimental results are provided in Sect. 62.3. And finally, Sect. 62.4 concludes the chapter.

## 62.2 Proposed Algorithm

In this section, the energy minimum framework, Gauss–Jacobi iterative method, nonlocal weight, and iterative solvent are introduced.

### 62.2.1 Energy Minimum Framework on Weighted Graphs

In this section, we use an energy minimum framework to describe pixel diffusion processes. Let  $H(V)$  be the Hilbert space of real-valued functions defined on the vertices of a graph. Function  $T$  is a mapping from discrete vertices to Euclidian space. Function  $S$  is a general function defined on graphs of the arbitrary topologies. The regularization of function  $S$  is formulated as following energy minimum framework:

$$\min_{T \in H(v)} \left\{ E(T, S, \lambda) = R(T) + \frac{\lambda}{2} \|T - S\|^2 \right\}, \quad (62.1)$$

where parameter  $\lambda$  is a Lagrange multiplier, which specifies the trade-off between two competing terms. And, regularization control factor  $R(T)$  is denoted as

$$R(T) = \frac{1}{p} \sum_{u \in v} |\nabla_w T(u)|^p, \quad (62.2)$$

where  $R(T)$  is a  $p$ -Laplace operator on weighted Hilbert space. In the formulation of  $R(T)$ ,  $\nabla_w T(u)$  is detailed as

$$\nabla_w T(u) = \frac{1}{2} \sum w(u, v)(T(u) - T(v)). \quad (62.3)$$

where  $u, v$  is denoted as different vertices. And,  $w(u, v)$  is the edge weight between  $u$  and  $v$ . The definition of  $w(u, v)$  is described in the following section. In this chapter, a nonlocal similarity metric is adopted as vertex weight in Hilbert space.

### 62.2.2 Nonlocal Weight Description

Inspired by Mahmoudi [16], we formulate edge weight in  $p$ -Laplace regularization factor using nonlocal-mean Gaussian description. The edge weight definitions are denoted as follows:

$$w(u, v) \propto \frac{1}{Z} \exp \left\{ -\frac{\|P(u) - P(v)\|_{G_\sigma}^2}{h^2} \right\}, \quad (62.4)$$

where the parameter  $h$  controls the decay of exponential function.  $G_\sigma$  is model Gaussian variance.  $P(u)$  is a  $5 \times 5$  block centered on vertex  $u$ . Analogously, the definition of  $P(v)$  is similar with  $P(u)$ .  $Z$  is normalizing factor. It is given by

$$Z = \sum_{u, v \in V} \exp \left\{ -\frac{\|P(u) - P(v)\|_{G_\sigma}^2}{h^2} \right\}. \quad (62.5)$$

We can use this nonlocal weight to measure the similarity between current block and searching block. Obviously, the searching block is more similar to current block if the computed weight is bigger. Here, nonlocal weight is used to model correlation between different blocks in  $p$ -Laplace regularization factor. The introduction of nonlocal weight model is served as a key factor in the problem solving of pixel similarity metric.

### 62.2.3 Iterative Solvent for Energy Minimum Framework

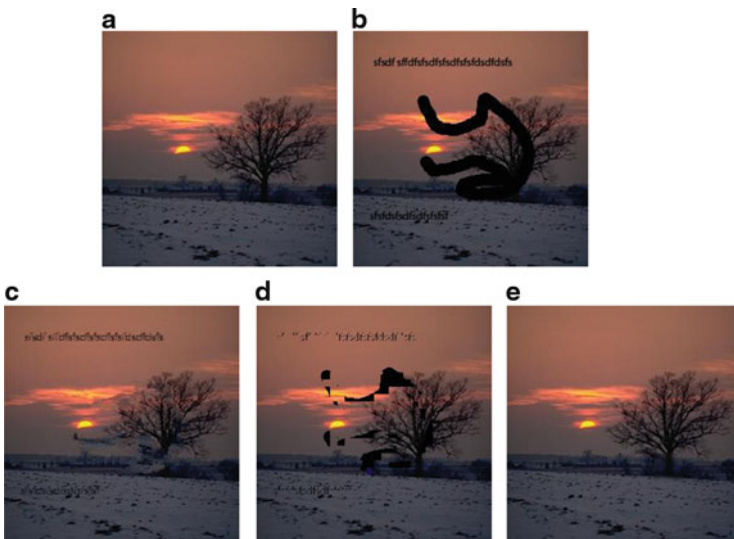
A linear Gauss–Jacobi iterative method is utilized for solving Eq. 62.1. We set  $n$  as iteration step. Then, the iterative solving method is given by the following:

$$\begin{cases} T^{(0)}(u) = S(u) \\ T^{(n+1)}(u) = \frac{\lambda T^{(0)}(u) + \sum w^t(u, v) T^{(n)}(u)}{\lambda + \sum w^t(u, v)}. \end{cases} \quad (62.6)$$

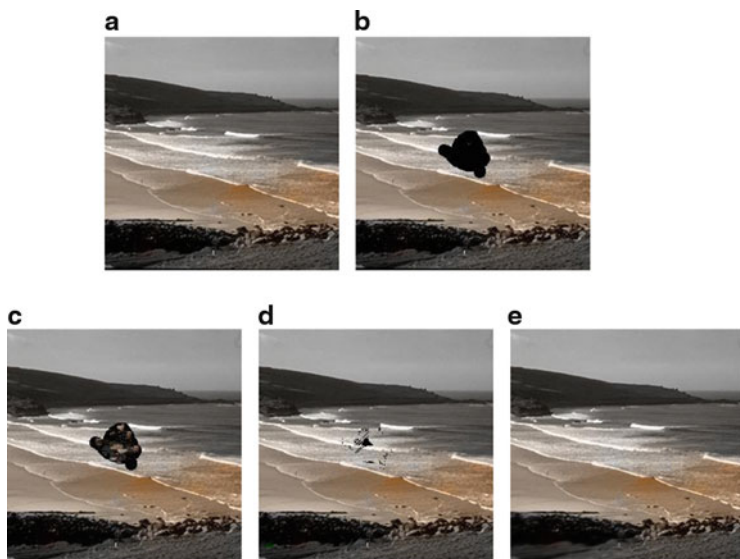
where to get the convergence of the process, a classical stopping criterion is  $\|T^{(n+1)} - T^{(n)}\| < \tau$ . In our implementation,  $\tau$  is a small fixed constant. At each iteration step, the updated value  $T^{(n+1)}$  depends on the initial value and weighted average of the filtered value in neighborhood.

### 62.3 Experiments

In this section, we demonstrate the image restoration results of the proposed algorithm while providing its comparisons with some other image restoration algorithm [19, 20]. Two test color images with size  $512 \times 512$  are used as input images in our implementation. Image reconstruction results under different image corruption masks are shown in Figs. 62.1 and 62.2. Objective image quality assessments are shown in Tables 62.1 and 62.2. PSNR [21] and structural similarity index (SSIM) [18] are used as evaluation tools. In Fig. 62.1, figure (a) is a original color image. Figure (b) is a color image with human-modified corruption. We use black region to represent image degeneracy. Figure (c) is the result of method [19]. Figure (d) is the result of method [20]. Figure (e) is our processed result. In order to highlight our performance, Fig. 62.2 gives results based on another corruption mask. Moreover, PSNR and SSIM are computed during the process of image restoration.



**Fig. 62.1** Experimental results: (a) Original image. (b) Corrupted image. (c) Processed result with method [19]. (d) Processed result with method [20]. (e) Processed result with our proposed algorithm



**Fig. 62.2** Experimental results: (a) Original image. (b) Corrupted image. (c) Processed result with method [19]. (d) Processed result with method [20]. (e) Processed result with our proposed algorithm

**Table 62.1** Performance comparison in PSNR (Db)

Test image	Method [19]	Method [20]	Our proposed method
Fig. 62.1	11.16	25.50	47.17
Fig. 62.2	20.10	27.44	39.99

**Table 62.2** Performance comparison in SSIM

Test image	Method [19]	Method [20]	Our proposed method
Fig. 62.1	0.28	0.88	0.99
Fig. 62.2	0.95	0.97	0.98

Table 62.1 illuminates the PSNR results as compared to Criminisi's and Li's [19, 20]. Table 62.2 illuminates the SSIM results as compared to Criminisi's and Li's [19, 20]. It can be seen that we achieve the highest PSNR and SSIM gain over previous algorithm. Subjective results shown in Figs. 62.1 and 62.2 also indicate that corrupted image processed by our method has good visual quality. We attribute the improvement of PSNR and SSIM to the use of nonlocal  $p$ -Laplace regularization scheme. The corrupted pixels can be well reconstructed by the most similar non-corrupted pixel. As compared to previous image restoration algorithm [19], our algorithm obtains both objective and subjective improvement. And our image restoration method does not depend on parameter adjustment either.



## 62.4 Conclusion

This chapter presented a novel image restoration algorithm built on nonlocal  $p$ -Laplace regularization strategy. We built an energy minimum framework by introducing nonlocal weight and Gauss–Jacobi iterative computation. Finally, the convergence speed was controlled by step threshold in iterative solvent. Experiments demonstrated the effectiveness and efficiency of nonlocal  $p$ -Laplace regularization image restoration algorithm.

**Acknowledgements** This work was supported by Science and Technology Innovation Foundation of Science and Technology Commission of Shanghai Municipality (12DZ0503300).

## References

1. Awate, S. P., & Whitaker, R. T. (2006). Unsupervised, information-theoretic, adaptive image filtering for image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(3), 364–376.
2. Bertalmio, M., et al. (2000). Image inpainting. In *Proceedings of the 27th Annual Conference on Computer Graphics and Interactive Techniques* (Vol. 20(1), pp. 417–424). New York: ACM Press/Addison-Wesley Publishing Co.
3. Gunturk, B. K., Altunbasak, Y., & Mersereau, R. M. (2002). Color plane interpolation using alternating projections. *IEEE Transactions on Image Processing*, 11(9), 997–1013.
4. Paliy, D., et al. (2007). Spatially adaptive color filter array interpolation for noiseless and noisy data. *International Journal of Imaging Systems and Technology*, 17(3), 105–122.
5. Zhang, L., & Wu, X. (2005). Color demosaicking via directional linear minimum mean square-error estimation. *IEEE Transactions on Image Processing*, 14(12), 2167–2178.
6. Perona, P., & Malik, J. (1990). Scale-space and edge detection using anisotropic diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 12(7), 629–639.
7. Rudin, L. I., & Osher, S. (1994). Total variation based image restoration with free local constraints. In *Proceedings of the IEEE International Conference on Image Processing, 1994 (ICIP-94), Austin, TX* (Vol. 1, pp. 31–35).
8. Mallat, S. (1999). *A wavelet tour of signal processing* (pp. 321–345). Chicago: Academic Press.
9. Kim, S. (2006). PDE-based image restoration: A hybrid model and color image denoising. *IEEE Transactions on Image Processing*, 15(5), 1163–1170.
10. Buades, A., Coll, B., & Morel, J.-M. (2005). A non-local algorithm for image denoising. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 2005* (Vol. 2, pp. 60–65).
11. Elad, M., & Aharon, M. (2006). Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image Processing*, 15(12), 3736–3745.
12. Mairal, J., Elad, M., & Sapiro, G. (2008). Sparse representation for color image restoration. *IEEE Transactions on Image Processing*, 17(1), 53–69.
13. Portilla, J., et al. (2003). Image denoising using scale mixtures of Gaussians in the wavelet domain. *IEEE Transactions on Image Processing*, 12(11), 1338–1351.
14. Roth, S., & Black, M. J. (2005). Fields of experts: A framework for learning image priors. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, CVPR, 2005* (Vol. 2, pp. 860–867).
15. Dabov, K., et al. (2007). Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8), 2080–2095.

16. Mahmoudi, M., & Sapiro, G. (2005). Fast image and video denoising via nonlocal means of similar neighborhoods. *IEEE Signal Processing Letters*, 12(12), 839–842.
17. Lezoray, O., Ta, V. T., & Elmoataz, A. (2008). Nonlocal graph regularization for image colorization. In *19th International Conference on Pattern Recognition (ICPR), 2008* (pp. 1–4).
18. Wang, Z., et al. (2004). Image quality assessment: From error visibility to structural similarity. *IEEE Transactions on Image Processing*, 13(4), 600–612.
19. Criminisi, A., Perez, P., & Toyama, K. (2003). Object removal by exemplar-based inpainting. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings, 2003* (Vol. 2, pp. 11–721).
20. Li, X. (2011). Image recovery via hybrid sparse representations: A deterministic annealing approach. *IEEE Journal of Selected Topics in Signal Processing*, 5(5), 953–962.
21. Hore, A., & Ziou, D. (2010). Image quality metrics: PSNR vs. SSIM. In *20th International Conference on Pattern Recognition (ICPR), 2010* (pp. 2366–2369).

# Chapter 63

## Analysis and Application of Computer Technology on Architectural Space Lighting Visual Design

Yiwen Cao

**Abstract** Based on “green building,” we use computer to make model analysis regarding building actual environment, adapting natural light, and using artificial lighting rightly to achieve green energy-saving building goal. According to living example making, we adopt 3d max and ECOTECT model analysis software to analyze and simulate building room natural light, artificial lighting and distribution. Analysis states: at the stage of the architectural sketch, we use 3d max modeling, uniting inter-room facility draft and analysis real environment data by ECORECT, getting the natural light effect, and uniting the two sides to get a more accurate room light environment data. Through this supporting data, we can get better daylight design, not only to satisfy the in-room light requirement but also to make it more beautiful and energy saving.

### 63.1 Introduction

Light is the origin of life, which is also an essential substance to humans. With the rapid development of artificial lighting technology, what people demand now is not only the simple illumination but the colorful and charming world which caters to people’s aesthetic taste. Artificial light applied to architectural space becomes even more important in contemporary architectural design. As the world’s energy problem and environmental pollution are coming up, how to save electric lighting and make full use of natural light in the building space attracts the attention of international architecture and lighting industry. With the development and the popularization of digital technology, computer technology plays an important role in the lighting technology and art design, especially in solving the problem of light

---

Y. Cao (✉)

Digital Media Department, Huaruan Software College of Guangzhou University,  
Guangzhou 510000, China  
e-mail: [caoyiwen116@yeah.net](mailto:caoyiwen116@yeah.net)

wasting, light pollution, and light modeling. Computer technology helps to evaluate and predict the effect so that the work can be easily checked. The application of computer technology has a great influence on architectural space lighting visual design.

The green building we refer to means saving energy, but we usually forget building itself is the base of the green building; if we only consider how to change lighting saving into green building, the building itself is not “green.” First, a good green building itself should be called a perfect energy-saving building and then it can make more details about the saving facility building. A good digital analysis technology can combine the fictitious room of the building, and good analysis of the actual environmental data makes the building achieve the “green” index.

There are many kinds of software combination of uniting digital technology building room daylight effect simulation and model analysis. According to the stylist’s own ability, we can use 2D software, such as PS, to draw up real and model picture of the building. Or we can make 3D model simulation referring to the architectural drawings and the actual environment, and getting the frames through model light. We can use digital technology to reflect the state of the inner or outside room, structure of the room and the texture and frames of the decorative material. At the same time, using ECOTECT arcology building design software, we can get the room natural lighting coefficient, which is got from the lighting coefficient computer. When using natural lighting, in-room illumination is changing all the time with outdoor illumination. So when we want to make sure the in-room natural light illumination level, we should consider the outside room illumination.

## **63.2 The Application of Computer Technology to Architectural Space Lighting Visual Art**

Computer technology is developing rapidly as well as graphics and image software. Computer technology combined with modern information and visual art forms a modern comprehensive art design which is based on computer technology but stresses on visual art; it is also becoming the dominant trend of modern art design. Nowadays, energy and environmental problems are extremely urgent to solve. People put forward a great many effective ecologic solutions to make use of natural light and save electric lighting. However, if a project is to be successful, it must be carried out through a serious course of survey, conceiving, confirming, modifying, design, and implementation. Some unreasonable designs and errors only appear in the course of implementation while others won’t come up until the construction has been completed. At this time, it is extremely difficult to solve various problems. Therefore, we need a computer to perform simulation analysis in order to obtain a better effect of visual design. In the course of project design, taking advantage of simulation method, the computer evaluates the plan and predicts the result so as to directly check the project. Apply the computer technology to the building space visual design and perform the simulation programming and data analysis during the

construction to form a set of complete digital simulation space, by modifying the design errors to reduce unnecessary waste and save the energy.

The simulation software for architectural space lighting effect helps the contemporary architectural designers to achieve the design goal. Digital simulation software, 3d max, as the commonly used software, makes the designers perform the professional 3D modeling, material designing color and lighting effects in the same environment. The software provides us a variety of lights to simulate natural ones, giving off real light and artificial light. The designers are able to directly carry out the 3D building plan so that computer technology simulating the real scene has come true.

### ***63.2.1 Natural Light Simulating Lighting in 3D Software***

In the real world, natural light is mainly composed of sunlight. The sunlight, which is a kind of parallel light, irradiates in a straight line and to a certain direction (Fig. 63.1). The light shadow in architectural space is changeable. The time and angle of the sun are changing, and so is the sunlight. The skylight is the result of sunlight and diffuse light, which goes well distributed and slowly changed without direction. The skylight reaching building space must be softer and weaker than the sunlight.

In 3d max, we can simulate parallel light by using the standard light; we can also use sunlight IES and skylight IES (in the photometry) to simulate real scene. The parallel light projects to the same direction [1]. Parallel light includes target light and free light.

Parallel light is mainly used to simulate sunlight. Adjust the light color and position by rotating light in three-dimensional space. The parallel light is in the form of a circular, rectangular, or prism. The figure above is the effect of standard parallel light simulating sunlight in 3d max, standard skylight tracing simulation, and IES sunlight and IES skylight simulating natural light (Fig. 63.2).

### ***63.2.2 The Simulation Analysis of Artificial Lighting in 3d max***

Besides the sunlight and skylight, people use more artificial illumination which can redesign the shadow change, modeling and redividing the building space in order to make it more colorful and charming.

In 3d max, standard light is being used to render space. The calculation method is direct, that is, simulation calculation without the data. So what we need is to arrange the light in the scene, through people's sense to evaluate the rendering effect. This type of light includes floodlights and spotlights [1] (Fig. 63.3).

Fig. 63.1 Parallel light in 3d max

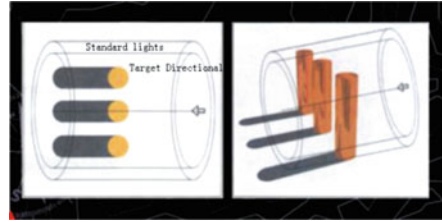


Fig. 63.2 Rendering effect of parallel light simulating sunlight

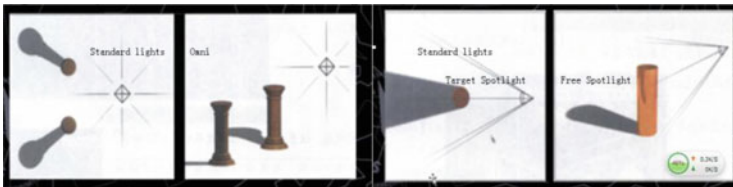
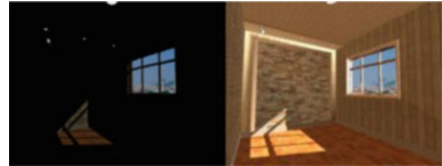
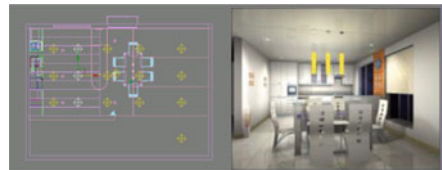


Fig. 63.3 Standard light in 3d max

Fig. 63.4 Standard lighting in 3d max



1. Floodlights: It is irradiated from one source to all directions. Floodlights add “auxiliary lighting” to the scene or simulate point light source. Floodlights can cast the shadow and projection. The single floodlight to cast shadows is equivalent to six spotlights to do so from the center to the outside.
2. Spotlights: It includes target lights and free lights. The former one is like a flash to cast focus beam which has its target point as the position, while the latter one hides this position (Fig. 63.4).

In addition to the standard light source, another commonly used simulation source is photometric light in 3d max. The light is based on simulation operational mode, by simulating real scene to get the rendering effect and, at the same time, to calculate the light distribution exactly. By shape, photometric lights include point source, line source, and surface source [1] (Fig. 63.5).

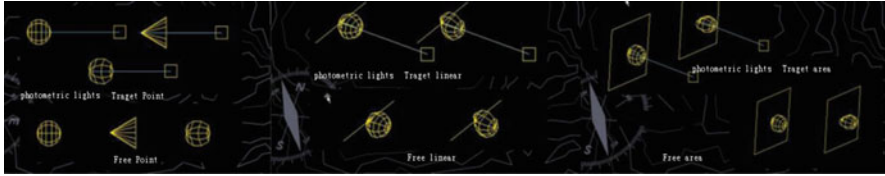


Fig. 63.5 Photometric light in 3d max

1. Point source: It is the simulation of incandescent lamp, for example, the ceiling lamps, wall lamps, and tube light. Electric light source radiates from itself.
2. Line source: It is used to simulate daylight lamps or strip-shaped light slots. Different from the point, the distribution of the line is based on the linear length, to emit in a line.
3. Surface source: It simulates the rectangular-shaped lights in real life such as ceiling lamps or grille lamps. The surface emits light in a rectangular area.

In addition to photometric shape simulation, light intensity and distribution are also important parameters. As the intensity and the irradiation shape to all directions are different, the lights illuminating to the buildings are different too. For example, the illumination from tube light and wall lamps to the surface of buildings produces different light sources and different shapes, so the light distribution of the object source is different.

### 63.3 Application of Computer Technology to Architectural Apace Lighting Design

An excellent green building is not only depending on its internal facilities to adjust light, but it must be an energy-saving one without any auxiliary equipment. If not so, such a building cannot achieve the goal of saving energy; the building is not the “green building.” The proper analysis of computer technology can combine the construction in the virtual space with the actual environment numerical value in order to simulate and analyze effectively. At present, there is a lot of computer software for simulation and analysis, such as thermal environment simulation and sunlight simulation. This paper analyzes the interior and exterior lighting illumination with ECOTECT software. Computer simulation technology of lighting and illumination analyzes thermal environment, light environment, sunshine environment, etc. The calculation results are accurate and intuitive. As for the architectures, it is essential to analyze the light effect. With the sunlight changing, the internal and external light intensity will alter. To make the architectures more green and energy saving, people need to modify the errors in the course of designing. It is necessary to use ECOTECT analysis so that the green goal can be perfect before the project is established.

ECOTECT provides a variety of functions for analysis, among which natural lighting is the most important, which influences the artificial lighting. This soft natural light is assumptive diffuse natural light in a closed space. As the sunlight is changing with time, occasionally the sun will also be prevented by objects, and the main building receives different light intensity, so the sun is not accepted as a reliable lighting source. Therefore, the best method to calculate lighting coefficient in ECOTECT software is Split Flux by Building Research Establishment. The calculation principle is based on the premise, without considering the influence of direct sunlight. Any point of natural light is composed of three parts in the space: the skylight—sky component (SC), the light through the internal glass objects directly irradiated the space; the external reflection light—externally reflected component (ERC), light reaches environmental objects, producing reflected light source; reflected component (IRC), direct light and reflected light reach a point in the space, together with other light spots around them to form a sort of internal reflection. CIE cloudy distribution in ECOTECT (Overcast Sky Distribution) model [2], through the analysis of the current grid or independent sensing, points to calculate the data of natural lighting. The lighting parameter is based on the coefficient of lighting and the sky illumination. The sky luminance values can be set according to the lighting analysis. When the calculation is being performed, semicircular lighting prediction figure will come up on the software interface. Through each light data in the same illumination, we can calculate the lighting vector on each point, which represents the direction of natural light. We can get the data through different colors and arrows on the interface. The analysis of natural light can be calculated by lighting coefficients, through which, it can calculate the illuminations on the certain days in a year and also can adjust its parameters.

## **63.4 Simulation Analysis of Architectural Space Lighting Design Combined with Computer Visual Technology**

### ***63.4.1 Simulation Analysis of Architectural Lighting Design in 3D Max***

In this case, analyze and simulate the photometric lights of indoor space [1].

Light energy transmission is a calculation way to simulate light, which is able to simulate and analyze refraction, reflection, and other indirect lighting to the objects' surface, through which the real scene can be beautifully and truly realized. It can also simulate the natural light which emits to the models' surface. The calculation of light emission can be completed before rendering. Calculation is different from rendering. If the refraction and reflection in the real scene can be realized, the calculation can't be skipped. In the following cases, we will calculate the light transmission through photometric lighting in indoor space. Such calculation is based on physical date, so the model establishing and lighting parameters are rigidly required. We must obey the rules to render.



**Fig. 63.6** Photometry  
simulating sunlight



**Fig. 63.7** Photometry  
simulation sunlight to  
render



In the case, we use IES to simulate the sun irradiation and adjust the ray tracing, and make the light complete the simulation through the calculative operation. The effect is after setting the light tracing (Fig. 63.6).

Set the lighting transmission parameter, set the IES sunlight as 2000lm, set three iterations and three filters, and use the general grid subdivision, rendering effect (Fig. 63.7).

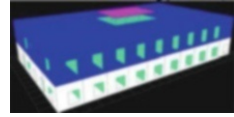
From the cases above, we can see photometric lights can analyze and simulate radiosity. The lighting method depends on the light source's direction and position in the real scene. Radiosity's calculating makes the light intensity distribution uniform and transferred conveniently, to simulate the good distribution of light in the scene. The effect of simulation light goes real and delicate, and it realizes a good atmosphere of real scene.

By the application of computer simulation software, we can simulate the effect in the design period, so that the designers can easily design the building space. Computer simulation can replace traditional hand-drawn drafts. It is precise and effective. What's more, it's convenient to modify and can greatly reduce the design time and improve work efficiency.

### **63.4.2 Computer Simulation Analysis on ECOTECH Light Environment**

As to energy-saving or illumination angle, natural lighting is very important for buildings, in large building space, except the room around the window, and other spaces such as the lobby, which is very dark. Due to the lack of sunlight conditions in the daytime, artificial lighting is needed. Thus, it is not an ideal space both for health and energy saving. Because the natural light is changed by time, sun altitude, atmospheric density and some other natural factors, therefore, we can use computers to simulate and analyze natural light, and get basic data of the internal natural

**Fig. 63.8** Creating rectangular skylights



light intensity. According to different light intensity in each region of interior space, the designer will use artificial light for light division and the secondary design in architectural space more effectively.

In this case, we use Learning Tutorial-Daylighting Analysis (proposed by Alexa. Ecotect), Ecological Simulation Analysis, and the transformation case for green building design company to state our opinion [3].

Analyze the construction space and set the contour as rectangular, single-storey building area as  $60 \times 30$  m. If two layers, floor height is respectively 6 and 7 m. Considering architectural space is wide and deep and the center is dark lighted, the skylights can be used for light selecting. The initial design stage will use the ECOTECT software, analyze the skylight lighting, and at the same time take rectangular roof as an example for analysis of fluorescent lighting.

The buildings' lighting window includes side windows and skylights. In accordance with the "Architectural Lighting Design Standards GB/T 50033-2001," the side window's window-to-floor proportion is 1/2.5, skylights' is 1/7.5, and rectangular skylights' is 1/3.5. With the partition between the two layers, the skylights' area covers the area of  $18 \times 60 = 1,080 \text{ m}^2$ .

The area of building window is estimated according to the computer lighting technology; as for the first floor, building lighting cannot calculate the specific area. So we need ECOTECT to simulate and analyze different shapes and areas of the skylights and to decide the form and area. Take rectangular skylight, for example, to simulate.

First, open ECOTECT to make sure the gridding and geographic position, and set up a 6 meters high, 3.5 meters wide rectangular wall. On the wall, we can set up a 3m high, 2.5m wide window and the height of windowsill is 0.9m. At the other side of the wall, we also create the window in the same way. At the top of the building, we set up a skylight and confirm its size and connect two layers. Then, we create a building interior wall, skylights, and door, and then we copy the door (Fig. 63.8).

Lighting Simulation of Rectangular Skylights. First open the Web management panel and then set up gridding. To avoid walling reflecting, we have to set up the gridding position to 1 M far from the wall. The First-floor of Rectangular Skylight Daylighting Simulation. Set the grid XY Axis offset of 600, which means analysis of grid is 600 mm from the ground. Select lighting coefficient and the intensity of illumination options, click calculation, and set simulation (Figs. 63.9, 63.10, and 63.11).

For the design of rectangular windows, first-floor lighting space is controlled between 2.5 and 15 %. About 80 % illumination space is around 6 %, with low intensity. Artificial lighting auxiliary should be added according to the actual analysis.

Fig. 63.9 Set up mesh parameter



Fig. 63.10 Simulation figure of first-floor rectangular window

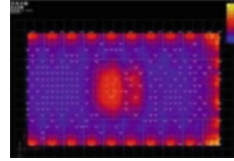
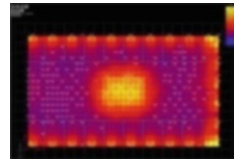


Fig. 63.11 Simulation figure of second-floor rectangular window



For the design of rectangular window, second-floor lighting space is controlled between 7.5 and 25 %. The light energy distribution is uniform in the rectangular window. If daylight is too strong, artificial shading measures should be taken to prevent the light.

By simulating daylighting coefficient, and we can know that there is no area that the daylighting coefficient is less than 2 %, that is, 100 % of the interior space daylighting coefficient is more than 2 %. It fully satisfies the rules of the green building evaluation technology in [4] (75 % of the interior space daylighting coefficient should be more than 2 %).

In the design of the sky illumination (5000 lux), we calculate indoor illumination, and we find the illuminance  $\geq 101$  lux, which is accordance with the rules (indoor natural light illumination of visual working class III (fine) is not less than 100 lux) in «architectural lighting design standards» (GB/T 50033-2001) [5]. Generally speaking, through the natural lighting simulation analysis of the office building standard layer, we find that the natural lighting design scheme satisfies the requirement of indoor visual, and it can achieve the ideal effect.

## 63.5 Conclusion

Humans live in the light environment and use light in all fields of life. However, the light works for humans while at the same time also consumes a lot of energy, causing serious pollution and destruction to the environment.

In the digital era, lighting design by computer technology has become an important part of people's life. In order to meet the people's aesthetic psychology, artificial light is becoming more and more important. As the energy problems appear, people pay more attention to lighting design in architectural space with computer technology. Computer shows the building space environment in the course of design so as to modify the improper part, evaluate and forecast the effect, and save unnecessary energy consumption. Computer technology's application to architectural space lighting design will have a further influence on human's life.

## References

1. Wang, X., Wang, K., & Shi, Y. (2004). *Learning by doing—3ds max effect drawing materials and lighting application techniques and examples* (p. 15). Beijing: People's Post & Telecom Press (in Chinese).
2. Yun, P. (2007). *Teaching series of computer technology on architecture—Ecotect building environment design tutorial* (p. 30). Beijing: China Building Industry Press (in Chinese).
3. Alexa, Ecotect (Ecological construction master). (2011). *Learning tutorial—Lighting analysis "Shenlvshu" ecological simulation analysis*. Green Architecture Design Company (p. 5). ShenZhen: China building industry Press (in Chinese).
4. China Academy of Building Research. (2007). *The technical details of the green building evaluation* (p. 65). Beijing: China building industry Press (in Chinese).
5. State Standard of PRC. (2001). *Architectural light-selection design standard GB/T 50033-2001* (p. 7). Beijing: China building industry Press (in Chinese).

# Chapter 64

## Improving Online Gesture Recognition with WarpingLCSS by Multi-Sensor Fusion

Chao Chen and Haibin Shen

**Abstract** In order to achieve the better online gesture recognition rate, a multi-sensor fusion method is proposed in this chapter. After the dimension reduction and quantization, we first measure the performance of every single sensor in training phase and use this prior knowledge to determine the weight vector; then we do the fusion of multiple sensors according to the weight vector which indicates each sensor's importance in recognition. The core algorithm we use for online gesture recognition is WarpingLCSS, which is demonstrated to be an efficient template matching method for gesture spotting. We do the experiments on the OPPORTUNITY Activity Recognition Datasets, and the results show that the recognition rate of multi-sensor fusion method achieves 61 %, which outperforms the single sensor's performance about 11 %. This demonstrates that our proposed multi-sensor fusion method is efficient in improving the performance of online gesture recognition.

### 64.1 Introduction

Multiple sensors can give a more comprehensive description about the system than single source data; thus multi-sensor systems play an important role in improving the accuracy of pattern recognition. In recent years, there are a large number of applications about multi-sensor fusion such as robot system [1], mechatronics [2], and human computer interaction [3].

In 2012, Long-Van Nguyen-Dinh proposed a new template matching method called WarpingLCSS [4]. He compared this new method with dynamic time warping (DTW) [5], a widespread method to do online gesture recognition, and

---

C. Chen • H. Shen (✉)  
Institute of Very Large Scale Integrated Circuit Design, Zhejiang University,  
Hangzhou 310027, China  
e-mail: [shb@vlsi.zju.edu.cn](mailto:shb@vlsi.zju.edu.cn)

his analysis shows that WarpingLCSS has higher accuracy than DTW. However, the recognition rate is still not high.

In order to improve the accuracy of online gesture recognition, we extend WarpingLCSS with the multi-sensor fusion method. After the preprocessing step such as dimension reduction and quantization, we test the performance of every single sensor and then do the fusion of multiple sensors based on this prior knowledge.

The results of our experiments show that this multi-sensor fusion method will highly improve the accuracy and the recognition rate is about 61 %, which outperforms the current WarpingLCSS about 11 %.

The rest of this chapter is organized as follows: In Sect. 64.2, we review some related works. In Sect. 3, we review the basic concepts about WarpingLCSS. In Sect. 64.4, we describe multi-sensor fusion method proposed by us. In Sect. 64.5, we give the experimental results of the proposed method. Finally, conclusions and future work are discussed in Sect. 64.6.

## 64.2 Related Works

Data acquired from sensors are basically a kind of time series; therefore we prefer to use time series analysis method to spot gestures. One of the most widely used methods is DTW [6]; it calculates the similarity distance between two time series, one is samples for test, and the other is the template. If the similarity distance is below a certain threshold, the two time series are considered to be the same class of gestures.

However, DTW shows a bad tolerance to outlier noise [7]. In order to recognize gestures with noise more accurately, the Wearable Computing Lab in ETH proposed a new method called WarpingLCSS based on the Longest Common Subsequence (LCSS) algorithm [4]. This method shows better tolerance to outlier noise and requires less hardware resources than DTW; in his work, one sensor attached to the upper arm is used and the accuracy achieves about 50 % which outperforms the DTW method about 12 %. But 50 % accuracy is still not that good.

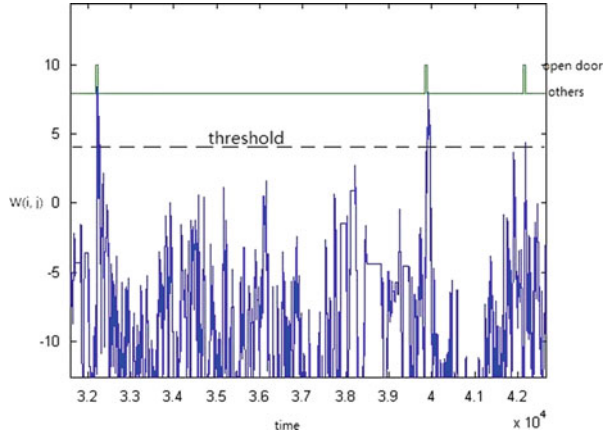
Multi-sensor fusion method can achieve better recognition rate since it combines information together. There are several methods we can choose to do the fusion [8]—the raw data level, the feature level, and the decision level. In this chapter, we do the fusion in the decision level since we want to diminish the wrong decisions of each sensor after applying the WarpingLCSS algorithm.

## 64.3 WarpingLCSS

The WarpingLCSS algorithm [4] is derived from the LCSS problem and can be computed efficiently by dynamic programming [9].

Let  $S_t$  be a gesture template generated in training phase and  $S_m$  be the streaming sensor data.  $W(i, j)$  in (64.1) gives the measurement of similarity between the  $i$ th

**Fig. 64.1**  $W(i, j)$  between a template and a string of preprocessed sensor data



symbol in template  $S_t$  and  $j$ th symbol in  $S_m$ .  $p$  is used to penalize the mismatch of two symbols, and  $d(l, m)$  is the Euclidean distance between the two corresponding centroids.  $W(i, j)$  will increase rapidly when lots of similarities exist between the new incoming sample and the template string and decrease rapidly due to the function of penalty parameter and thus highlights the endpoint of predicted gestures.

In Fig. 64.1, we give a case to show the function of WarpingLCSS. The line on the top indicates when the open door gestures happen, and  $W(i, j)$  varies rapidly during that period of time and will have some local maximum values. We just need to set a proper threshold like the figure shows, and the local maximum values above the threshold indicate the possible happening of open door gestures:

$$W(i, j) = \begin{cases} 0, & \text{if } i = 0 \text{ or } j = 0 \\ W(i - 1, j - 1) + 1, & \text{if } S_t(i) = S_m(j) \\ \max \begin{cases} W(i - 1, j) - p \cdot d(S_t(i), S_t(i - 1)) \\ W(i, j - 1) - p \cdot d(S_m(j), S_m(j - 1)) \end{cases}, & \text{otherwise} \end{cases} \quad (64.1)$$

### 64.4 Multi-Sensor Fusion Method

The whole process of recognition phase of our proposed multi-sensor fusion method is shown in Fig. 64.2. Raw sensor data series are first preprocessed and are calculated with WarpingLCSS algorithm separately. Then the processed data come to the multi-sensor fusion module, which combines the results of WarpingLCSS to diminish those unexpected noise and gives the outputs according to each sensor's importance in recognition. If the outputs suggest that an instance may belong to multiple classes, then the last module should decide which template has the highest similarity with the incoming stream and finally outputs the class of the gesture.

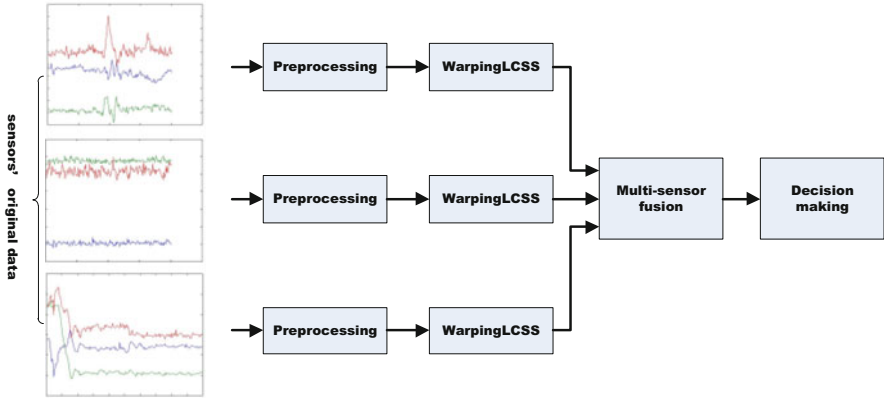


Fig. 64.2 The architecture of multi-sensor fusion system

### 64.4.1 Preprocessing

Raw data are generated by 3D accelerometers or other sensors. In order to speed up the processing rate, we first use a sliding window of size 6 to combine six sample vectors to one mean vector, overlap three samples each time, and then use k-means to cluster the mean vectors to their closest cluster centroids. Thus, we could get a string of symbols to represent the raw data. This process is the same in training phase and recognition phase.

### 64.4.2 Training Phase

For gestures belonging to the same activity class, we use LCSS to calculate the similarity between them and choose one to be the template if it has the highest similarity with others. And we repeat this process to get templates for all types of gestures and all types of sensors. When we get all these templates, for each gesture class, we check the accuracy of every sensor and choose the candidates for multi-sensor data fusion.

Equation (64.2) gives a method to do multi-sensor fusion.  $V$  represents the number of sensors to be fused.  $WV$  is a positive definite weight vector which gives certain  $W(i, j)$  more weight to improve the overall performance of online gesture recognition. We can test the accuracy of each sensor in training time, get the prior knowledge about the importance of sensors when spotting certain activities, and set the  $WV$  according to this prior knowledge:

$$MW = \sum_{v=1}^V WV(v) \cdot W(i, j)(v) \tag{64.2}$$



### 64.4.3 Recognition Phase

In recognition phase, we have almost the same process for incoming data as in training phase. When multi-sensor streams come, these data will first decrease their dimension in preprocessing stage; then each sensor data compare with the templates created in training phase and fuse according to (64.2). Gestures will be spot if the results of (64.2) are over the threshold.

### 64.4.4 Decision Making

If an instance is spotted as multiple gestures, we have to decide which one is best matched and output the decision. We first use the trace-back process [4] to get the start point of the gesture and then calculate the normalized similarity through dividing the similarity by the longer length between the template and the spotted string. The class with the highest normalized similarity is chosen as the final output.

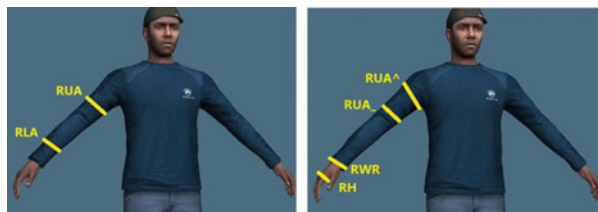
## 64.5 Experiment Setup and Simulation Results

### 64.5.1 The Datasets

All our experiments are doing on the OPPORTUNITY Activity Recognition Dataset [10]. We do not need the whole datasets, so we select part of the sensor data. The names and locations of chosen sensors are depicted in Fig. 64.3.

### 64.5.2 Evaluation Methods

To evaluate the performance of the method, we use accuracy, which is the proportion of correctly predicted samples over the number of total samples. However, the



**Fig. 64.3** The names and locations of inertial measurement units (*left*) and acceleration sensors (*right*)

**Table 64.1** The performance of the whole OPPORTUNITY Activity Recognition Datasets

Activity	Sample number	Accuracy	F1
Open door 1	10	0.90	0.72
Open door 2	9	0.88	0.40
Close door1	9	0.78	0.64
Close door2	9	0.67	0.75
Open fridge	31	0.61	0.60
Close fridge	33	0.61	0.52
Open dishwasher	12	0.83	0.60
Close dishwasher	14	0.64	0.41
Open drawer1	12	0.92	0.4
Close drawer1	12	0.67	0.53
Open drawer2	12	0.41	0.45
Close drawer2	12	0.67	0.67
Open drawer3	12	0.41	0.38
Close drawer3	12	0.75	0.4
Clean table	33	0.5	0.27
Drink cup	36	0.52	0.59
Toggle switch	21	0.33	0.22
The whole OPPORTUNITY datasets			
Accuracy		F1	
		0.61	0.48

accuracy may vary due to the selection of threshold, so we add another measurement called F1 to help to access the performance [4]:

$$F1 = \sum_i 2 \times W_i \frac{precision_i \cdot recall_i}{precision_i + recall_i} \quad (64.3)$$

$$precision_i = \frac{\text{correctly predicted samples}}{\text{total predicted samples}} \quad (64.4)$$

$$recall_i = \frac{\text{correctly predicted samples}}{\text{total samples of class } i} \quad (64.5)$$

### 64.5.3 Results

For each class of gesture, we first measure the accuracy and F1 of every sensor and choose the proper weight vector. Repeat the above process for every activity, and the results are shown in Table 64.1; we test the overall performance on the OPPORTUNITY Activity Recognition Datasets, which achieves 61 % in accuracy and 48 % in F1; the results are 11 % better than mentioned before [4], which demonstrate that multi-sensor fusion is an efficient way to improve the accuracy of activity recognition.

## 64.6 Conclusions and Future Work

In this chapter, we have proposed a multi-sensor fusion method based on the WarpingLCSS algorithm. The results show that our method can highly improve the accuracy of gesture recognition and outperforms the single sensor's performance mentioned before about 11 %.

In future work, we want to explore what features will best represent those daily activities and improve the accuracy of recognition significantly.

## References

1. Castellanos, J. A., Neira, J., & Tardós, J. D. (2001). Multisensor fusion for simultaneous localization and map building. *IEEE Transactions on Robotics and Automation*, 17(6), 908–914.
2. Luo, R. C., & Chang, C. C. (2012). Multisensor fusion and integration: A review on approaches and its applications in mechatronics. *IEEE Transactions on Industrial Informatics*, 8(1), 49–60.
3. Reddy, B. S., & Basir, O. A. (2010). Concept-based evidential reasoning for multimodal fusion in human–computer interaction. *Applied Soft Computing*, 10(2), 567–577.
4. Nguyen-Dinh, L. V., Roggen, D., Calatroni, A., Troster, G. (2012). Improving online gesture recognition with template matching methods in accelerometer data. In *2012 12th International Conference on Intelligent Systems Design and Applications (ISDA)* (pp. 831–836). IEEE.
5. Hartmann, B., & Link, N. (2010). Gesture recognition with inertial sensors and optimized DTW prototypes. In *2010 I.E. International Conference on Systems Man and Cybernetics (SMC)* (pp. 2102–2109). IEEE.
6. Stiefmeier, T., Roggen, D., Troster, G., Ogris, G., & Lukowicz, P. (2008). Wearable activity tracking in car manufacturing. *IEEE Pervasive Computing*, 7(2), 42.
7. Vlachos, M., Hadjieleftheriou, M., Gunopulos, D., & Keogh, E. (2003). Indexing multi-dimensional time-series with support for multiple distance measures. In *Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (pp. 216–225). ACM.
8. Jeon, B., & Landgrebe, D. A. (1999). Decision fusion approach for multitemporal classification. *IEEE Transactions on Geoscience and Remote Sensing*, 37(3), 1227–1233.
9. Sakoe, H., & Chiba, S. (1978). Dynamic programming algorithm optimization for spoken word recognition. *IEEE Transactions on Acoustics, Speech and Signal Processing*, 26(1), 43–49.
10. Roggen, D., Calatroni, A., Rossi, M., Holleccek, T., Forster, K., Troster, G., et al. (2010). Collecting complex activity datasets in highly rich networked sensor environments. In *2010 Seventh International Conference on Networked Sensing Systems (INSS)* (pp. 233–240). IEEE.

# Chapter 65

## The Lane Mark Identifying and Tracking in Intense Illumination

Yanyun Xing, Bo Yu, and Fangqun Yang

**Abstract** In order to enhance image contrast and ensure accurate identifying and tracking in intense illumination case, this chapter uses the algorithm of histogram cone-shaped, which can enhance image contrast effectively. With the algorithm of histogram cone-shaped, the scope of the gray value increases obviously. And the chapter introduces a first-order differential operator two-direction Prewitt operator to enhance the edge for image; the enhance effect is favorable, and the compute time is short. Then the algorithm of 2-D gray histogram is used to segment image. The chapter uses Hough transformation to identify the lane mark's two edges and account its intercept and slope and then draws the midline as the last identifying result. In order to reduce the count time, the chapter uses the algorithm of area of interesting to track the lane mark. The experiment results show that the lane mark can be tracked dependably in intense illumination and the algorithms are of real time; moreover, when the tracking algorithm is a failure, the system can also recover in time and lock the tracking target accurately again.

### 65.1 Introduction

Lane mark identifying and tracking is one of the research areas for vehicle safety driving assist system. It can be employed not only in lane departure warning systems but also in vehicle positioning. In the long run, the technology can also be used in the autonomous navigation of vehicle. Now the researches of lane mark

---

Y. Xing (✉)

College of Automotive and Transportation, Tianjin University of Technology  
and Education, Tianjin 300222, China  
e-mail: [jlucicheyb@hotmail.com](mailto:jlucicheyb@hotmail.com)

B. Yu • F. Yang

Automobile Engineering Research Institute, China Automotive Technology  
and Research Center, Tianjin 300000, China

identifying and tracking are mostly based on machine vision. The changes of environments have a great influence on the quality of machine detection method. Most of the working time of vehicle is daylight, especially in summer. It is very difficult to identify lane mark in the circumstance of long intense illumination. In order to solve this problem, the algorithm of histogram cone-shaped, which can enhance image contrast greatly, is adopted in this chapter. The results show that the extending algorithm is much better for this kind of images compared with other extending algorithm. The chapter introduces a first-order differential operator two-direction Prewitt operator to enhance the edge; the result is good as LOG operator whose convolution mask is  $5 \times 5$ , but the count time is much shorter. So the consequent image processing is same as the general image [1].

## 65.2 Preprocessing of Intense Illumination Image

The contrast of intense illumination image is very low due to the insufficient brightness range; as a result the useful information cannot be picked up accurately. In order to solve this problem, we need to enhance image contrast.

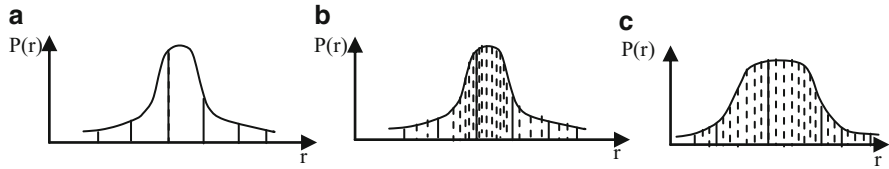
### 65.2.1 Extending Image Contrast by Histogram Cone-Shaped Algorithm

The histogram cone-shaped extending is an algorithm which can fill gray value in histogram cone-shaped unequally and transform the gray histogram axis into unequal field. Thus gray histogram is filled with more gray value in high-frequency region and less gray value in low-frequency region. Then a new gray histogram is set up according to the count of gray histogram axis, and a new image is reconstructed based on the new gray histogram. Since most of the histograms are shown as a single-peak or a multi-peak shape, the extending range is bigger in the histogram peak than in other area and the extending amplitude distribution is tapered; this histogram is called extending image contrast algorithm based on histogram cone-shaped. The principal sketch map of the algorithm is shown in Fig. 65.1 [2].

Histogram of nonuniform filling is a process that extends N layers of gray-axis histogram uniform to K layers unequally. If the gray range of image is  $m_0 \sim m_n$ , the corresponding value for gray distribution is  $H_{m_0} \sim H_{m_n}$ . Normally the gray-axis histogram is evenly divided into N equal segments; the total area of the histogram envelope curve can be obtained according to formula 1:

$$S = \sum_{i=0}^{n-1} [(H_{m_i} + H_{m_{i+1}})/2 \times 1] \quad (65.1)$$

If the histogram is extended into K ( $K > N$ ) layers and the gray range changes into  $g_0 \sim g_k$  as well, the algorithm needs to interpolate between  $m_i$  and  $m_{i+1}$ , so



**Fig. 65.1** The sketch map extending image contrast algorithm based on histogram cone-shaped. (a) Histogram of common image. (b) Histogram of cone-shaped. (c) Histogram of extending interpolation image

**Fig. 65.2** The original image



the interval  $m_i \sim m_{i+1}$  is segmented again. If the count of subsection is  $k_i$  for each interval, for the purpose of equal area of rezoned region and the total area of  $S$ , the area for every interval  $S'$  can be obtained according to formula 2:

$$S' = \frac{(H_{m_i} + H_{m_{i+1}})/2 \times 1}{k_i} = \frac{(H_{m_{i+1}} + H_{m_{i+2}})/2 \times 1}{k_{i+1}} \tag{65.2}$$

And also accord to the formula 3 at the same time.

$$n + k_0 + k_1 + \dots + k_{n-1} = k \tag{65.3}$$

Partition number is proportional to the gray value distribution in each interval, so the intervals of high gray value segmented are more than the ones of low gray value. The gray axis of the histogram with nonuniform filling is homogenized according to the points in the interval. Then the extending result of the histogram cone-shaped is accomplished. The histogram cone-shaped extending means that the area of the interpolation interval is cone distribution and the intervals of high gray value are extended higher than the one of the low. If  $m_i$  corresponding to a gray value changed into  $m_i'$  after extending,  $m_i'$  can be obtained according to formula 4:

$$m_i' = g_0 + \sum_{l=0}^{i-1} k_l \tag{65.4}$$

The original image and the extended image is shown in Figs. 65.2 and 65.3.

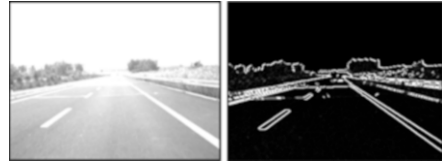
### 65.2.2 Enhancing the Edge by Prewitt Operator

The Prewitt operator is a first-order differential operator edge enhancement. For convolution mask of  $3 \times 3$ , it can estimate gradient in eight directions and can give

**Fig. 65.3** The extended image and its histogram



**Fig. 65.4** Comparison between the original image and the disposed image



gradient direction for the maximum amplitude convolution. Here are the  $3 \times 3$  mask templates for the Prewitt operator in the four directions [3]:

$$\begin{aligned}
 h_1 &= \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix} & h_2 &= \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix} \\
 h_3 &= \begin{bmatrix} 0 & 1 & 1 \\ -1 & 0 & 1 \\ -1 & -1 & 0 \end{bmatrix} & h_4 &= \begin{bmatrix} -1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 1 \end{bmatrix}
 \end{aligned} \tag{65.5}$$

Among them,  $h_1$  and  $h_2$  are the horizontal and vertical direction mask template,  $h_3$  is the  $45^\circ$  direction mask template, and  $h_4$  is the  $135^\circ$  direction mask template.

The lane marks of image are directional in this chapter. The lane mark of left side is slanting to the right, while the right side is slanting to the left, coinciding and matching with the template  $135^\circ$  and  $45^\circ$  Prewitt operator templates; as a result, we can use  $h_4$  template for the left half of the image and  $h_3$  template for the right half. This method is called two-direction Prewitt operator. The comparison between the original image and the disposed image can be obtained according to Fig. 65.4, in which the histogram cone-shaped extending algorithm and the enhancement edge by the two-direction Prewitt operator are shown.

### 65.3 Using the Algorithm of 2-D Gray Histogram to Segment Image

The pixels in the target areas or in the background areas have a great correlation in many images; the gray value of every pixel is approached with its adjacent area pixels, while the gray value between the edge of target and the background or between the noise and its adjacent area are obviously different. Every pixel can

**Fig. 65.5** The segmented image by the algorithm of 2-D gray histogram



have a gray group with two elements: the first is its gray value, and the other is the average gray value for its adjacent areas [3, 4].

If the gray value of the image is divided into  $L$  grade, the average gray value of the neighborhood area pixels is also divided into  $L$  grade. Hypothesis of the frequency for a gray group  $(m,n)$  is  $f_{mn}$  in the image; we can define the corresponding joint probability density as

$$p_{mn} = f_{mn}/N \quad m, n = 0, 1, 2, \dots, L - 1 \quad (65.6)$$

$N$  represents the count of the pixels of image,  $\sum_m \sum_n p_{mn} = 1$ . Based on  $m$  and  $n$  as independent variables and  $p_{mn}$  as dependent variable, a 2-D gray histogram could be formed.

The segmented image by the 2-D gray histogram algorithm is shown in Fig. 65.5.

## 65.4 Using Hough Transformation to Identify the Lane Mark

### 65.4.1 The Fundamental of Hough Transformation

Beeline equation  $y = mx + b$  can be represented with polar coordinates as

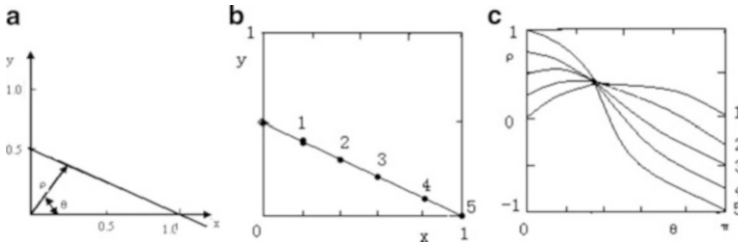
$$\rho = x \cos \theta + y \sin \theta \quad (65.7)$$

In the formula,  $(\rho, \theta)$  defines a vector from origin to the closest point on the beeline [5].

There is a two-dimensional space which is defined by parameters  $\rho$  and  $\theta$ . Every beeline in  $(x, y)$  plane corresponds to a point in the space. Therefore every beeline in  $(x, y)$  plane converts to a point in  $(\rho, \theta)$  space after Hough transformation (Fig. 65.6).

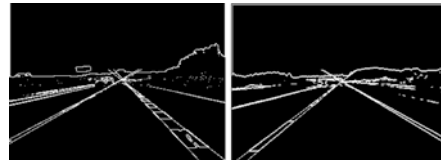
There is a given point  $(X, Y)$  in  $(x, y)$  plane. Countless beelines can cross the given point, and each beeline corresponds to a point in  $(\rho, \theta)$  space. Therefore the trajectory of the points where all the beelines are in  $(x, y)$  space is a sine curve in the parameter space, and each point in  $(x, y)$  plane is a sine curve in  $(\rho, \theta)$  space.





**Fig. 65.6** The fundamental of Hough transformation. (a) The polar coordinates for a beeline. (b)  $(x, y)$  plane. (c)  $(\rho, \theta)$  plane

**Fig. 65.7** Using Hough transformation to identify the lane mark edge



In order to find out the point of beeline, the  $(\rho, \theta)$  space is quantized into many small lattices. We can account every  $\rho$  according to the quantization of  $\theta$  for the point  $(x_0, y_0)$ . The value falls in one small lattice; the count accumulator of the small lattice will be plus 1. After all the point is accounted in  $(x, y)$  space, the small lattices will be checked up. If the value of one small lattice is large, it is a point that a beeline corresponds to, and we can employ it as a linear fitting parameter [5].

### 65.4.2 Using Hough Transformation to Identify the Lane Mark

The chapter uses Hough transformation to identify the two edges of the lane mark. The identifying result is shown in Fig. 65.7. The left lane mark is exemplified to illustrate the identifying process. Firstly, we use Hough transformation to find out the longest white beeline in the left half of the image and record it. Secondly, we assign 0 to the point in the white beeline as its new gray value. Then we account a longest white beeline again in the left-half image. So two white beelines are found in the left-half image at last that are the two edges of the left lane mark. And we can find the two edges of the right lane mark similarly.

The two beeline can be used to account its intercept and slope, which are found using Hough transformation. Then we can obtain the intercept and slope of its midline and the disappeared point of the lane mark. And by using the intercept and slope of the midline, combined with the disappeared point, we can draw the lane mark at last. The result is shown in Fig. 65.8.

**Fig. 65.8** Identifying the lane mark by Hough transformation



## 65.5 The Algorithm of Lane Mark Tracking

The algorithm of lane mark tracking is same with the algorithm of lane mark identifying essentially. Its characteristic is that when we identify the lane mark in the current image, we use the last identifying result to limit the identifying searching image area. That means we use the last identifying information to adjust the area of interesting (AOI) dynamic.

### 65.5.1 Algorithm of Setting Up AOI

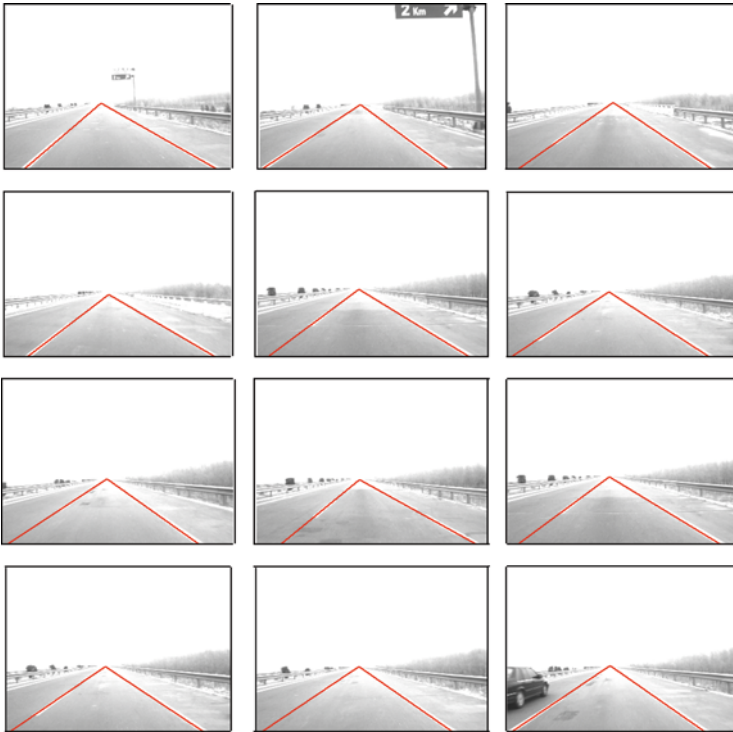
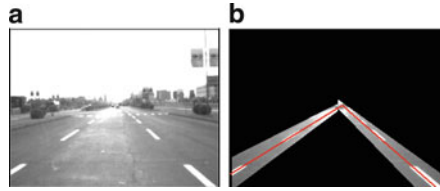
In a road image, the roadside information can interfere with the identification of lane mark, and most of the information on the road is useless for the identification. So the effective area for all algorithms is the area of interesting; we call it AOI [6]. In fact, the AOI only focuses on a small area nearby the road mark. In order to improve the real time the chapter uses AOI. That means the area, which is far away from the lane mark and is useless obviously for the lane mark identification, is removed directly without any treatment. Thus the system can save a lot of computation time. Figure 65.9b is the AOI for Fig. 65.9a. As we can see from the image, the AOI is very small compared with the full image, so it can improve the computer speed effectively and improve the real time for lane mark identification.

### 65.5.2 Recovery Algorithm After Failure

The AOI shows a good effect in the experiments, but when the lane mark is kept out seriously by something such as road vehicles or other objects, the AOI algorithm may cause great error or even failure. Moreover, when the vehicle changes lane or turns, the algorithm may be a failure. In practical application, in order to ensure that the road identification can recover the correct lane mark identification in the error or the failure case, the system needs to check the algorithm is a failure or not first and then determine whether to start the lane mark identification module or not. In this chapter, the method to determine the algorithm failure is as follows.

When the system checks the white points' number for one side or both sides of the lane mark is less than a fixed value in successive five-frame images or the information changes abruptly in the current frame relative to the last frame, we

**Fig. 65.9** Original image and its AOI. (a) Original image. (b) AOI



**Fig. 65.10** The video captures of real-time lane mark tracking

consider that the algorithm is a failure. Then the system needs to start the initial identification algorithm. That means the system needs to identify the lane mark in the whole image again and ascertain the AOI again, and then it recovers the correct path tracking.

### 65.5.3 Lane Mark Tracking

The experiment was made in a highway at noon, and the illumination is intense. The video captures of real-time lane mark tracking are shown in Fig. 65.10. The interval for contiguous two captures is 50 frames.

## 65.6 Conclusion

The chapter uses the algorithm of histogram cone-shaped extending to enhance image contrast first. Then it uses two-direction Prewitt operator to enhance the edge for image and uses the algorithm of 2-D gray histogram to segment image. The Hough transformation is used to identify the lane mark, and the algorithm of AOI is used to track the lane mark at last. Proved by the experiment of vehicle real driving on the highway, the algorithms are dependable and effective essentially in intense illumination case. Even in the case that the tracking target is error, the algorithm can also recover in time and lock the tracking target accurately again.

**Acknowledgements** Supported by Foundation of Tianjin University of Technology and Education Research Development Project Number KJ2008032.

## References

1. Massimo Bertozzi, Alberto Broggi, Massimo Cellario, Alessandra Fascioli, Paolo Lombardi, Marco Porta. *In Proceedings of the IEEE IV. 90(7)*. Jul 2002:1258–1271.
2. Yu-lin, S., & Zheng-xi, G. (2004). Algorithm of histogram cone-shaped extending applied to image contrast enhancement. *Computer Engineering*, 30(16), 153–154 (In Chinese).
3. Xiao Zhang, & Dong Yan-xue. (2005). *Digital image processing technology* (pp. 180–182, 188–190). Beijing: Metallurgical Industry Press (In Chinese).
4. De-kui, Y., Bao-min, Z., & Lian-fa, B. (1999). 2D Gray value transformation enhancement for infrared images. *Infrared Technology*, 21(3), 25–29 (In Chinese).
5. Sun, F.-r., & Liu, J.-r. (2003). Fast Hough transform algorithm. *Chinese Journal of Computer*, 24(10), 1102–1109 (In Chinese).
6. Rong-ben, W., Tian-hong, Y., Bai-yuan G., & Lie, G. (2006). Research on linear lane mark identification and track method based on edge. *Computer Engineering*, 32(16), 195–196 (In Chinese).

# Chapter 66

## Classification Modeling of Multi-Featured Remote Sensing Images Based on Sparse Representation

Xiaoting Hao, Chunmei Zhang, Jing Bai, Mo Dai, Wenxing Bao,  
and Wei Feng

**Abstract** A framework for multi-featured classification modeling of remote sensing images with sparse representations is proposed in this chapter. The problem for extracting features to build sparse learning optimal dictionaries is solved by using the remote sensing image spectral values combined with their transformation's characteristics such as Normalized Difference Vegetation Index and K–T transformation. This framework employs sparse representation on dimensional reduction and feature refinement for the remote sensing images. Experiments show that by using our approach the classification accuracy and Kappa coefficient are greater than the support vector machine and conventional sparse representation methods. This result is based on remote sensing images from the sand lake in Pingluo County, which is located in the Ningxia Hui Autonomous Region.

### 66.1 Introduction

The study of multispectral and hyperspectral data has become a major field of research in analyzing remote sensing images. Usually different materials such as sand, water, grass, roads, and buildings reflect different electromagnetic energies in a particular bandwidth. Based on this promise, we can differentiate the materials according to their spectral properties. But due to the biodiversity and terrain multiplicity, there is a dilemma in the remote sensing images. On occasions, same type of material reflects different spectral values and different materials

---

X. Hao • C. Zhang (✉) • J. Bai • W. Bao • W. Feng  
School of Computer Science and Engineering, Beifang University for Nationalities,  
Ningxia 750021, China  
e-mail: [chunmei66@hotmail.com](mailto:chunmei66@hotmail.com)

M. Dai  
Institut EGID, Université Michel de Montaigne—Bordeaux 3, 1 Allée Daguin,  
Pessac Cedex 33607, France

reflect same spectrums [1]. Therefore, we need extra features beyond spectral values for identification of the different types of materials.

In this work we introduce two other series features of Normalized Difference Vegetation Index (NDVI) and K–T transformation. Both of them are important indicators for eco-environment. Thus, with these key features, we are facing even higher dimensional space of data. In order to reduce dimensions of the multi-featured remote sensing images, we propose to learn the dictionaries to be used for sparsity-based classification. Experiments show that sparse representation method with learning dictionary can improve both classification accuracy and computational efficiency.

The contents of this chapter are arranged as follows: In the second section, we introduce the principle of NDVI and K–T transformation and describe the proposal of the multi-feature modeling of multispectral remote sensing images for wetland. In the third section, we introduce the sparse representations and dictionary learning classification method. In the fourth section, we present the specific steps of the algorithm of this chapter. The experimental results are analyzed in the fifth section.

## 66.2 Multi-Feature Modeling of Remote Sensing Images

In this section, by introducing the theory of NDVI and K–T transformation, we present the framework of multi-feature modeling.

### 66.2.1 *The Basic Principles of NDVI and K–T Transformation*

NDVI is an image-enhanced method for extracting vegetation information. NDVI is a global vegetation index, especially focusing on plant growth and related soil. NDVI is also helpful to interpret and analyze the characteristics of agriculture. NDVI is therefore always an important indicator of the characteristics of the wetland.

As an important component to present vegetation state, NDVI is well known as the main large-scale indicator for eco-environment. It is defined as the ratio of the subtraction and addition of the two bands which are near-infrared band and red band within visible light. The formula is as follows:

$$NDVI = (NIR - Red)/(NIR + Red) \quad (66.1)$$

where *NIR* is the reflectivity of near-infrared band, *Red* is the reflectivity of visible light band [2]. NDVI can eliminate most of the irradiant distortion including instrument calibration, sun angle, terrain, cloud shadow, and atmospheric

conditions to improve response speed of vegetation. It is quite sensitive to the bias of soil background and very efficient on monitoring vegetation and ecological environment. It also has a wide detection range of the vegetation coverage and a good spatial and temporal adaptability. So it has wide applications on vegetation and phenology studies for remote sensing images.

NDVI is a method to monitor the land cover by remote sensing images. The number of vegetation indexes over the past 20 years has increased to more than 40. The NDVI has a linear correlation with vegetation distribution density. Therefore it is considered currently the best indicator of land cover.

K–T transformation is also called Tasseled Cap transformation. It is a kind of empirical linear transformation of multiband images, which can be used as an interpretation for different vegetations and crops in different regions. It was put forward by Kauth and Thomas [3] on MSS image research in which they investigated the processing of the growth of the crops and vegetation. The mathematical expression is shown below:

$$u = R^T x + r \quad (66.2)$$

where  $R$  is the coefficients of the K–T transformation,  $x$  represents the gray values of different bands,  $r$  represents constant offset, and  $u$  is the gray values of different bands after the K–T transformation [3].

The K–T transformation provides several components. The top three components are closely related to the ground scenery. The first one is brightness index, which is the comprehensive reflective effect of the ground materials. The second component is the green index, which focuses on the coverage of the vegetation, leaf area index, and biomass. The third component is the humidity index, which refers to the water condition of the ground, especially soil humidity saturation. The rest of the components are the yellow index and noises.

## 66.2.2 The Model of Multi-Features

According to the multidimensional spectrum distribution structures of the multi-spectral remote sensing information such as the soil and vegetation by using NDVI and K–T transformation, the vegetation discrimination can be enhanced and the background features and texture features can be highlighted. Even more, the vegetation monitoring will be improved, and especially, the information extraction for scattered green patches can be more efficient.

Based on the above discussion, we use spectral feature TM 1–5 bands and TM 7 band of remote sensing images combined with the first component of the brightness index, the second component of the green index, and the third component of humidity index of the K–T transformation and NDVI, such as ten feature vectors, as the basic characteristics for classification.

## 66.3 Dictionary Learning and Sparse Decomposition

Sparse representation has a great flexibility for adaptive pattern recognition with the redundancy of dictionaries to capture the inherent structures of the natural signals. On one hand, sparse representation provides sufficient choices to design new classification models for dimensional reduction; on the other hand, the similarity matrixes can be built with the sparsity of sparse decomposition, which has a big potential on statistical pattern recognition.

There are two ideas about sparse representation: one's dictionaries have fixed bases (such as curvelets, contourlets), so their structures are fixed and the adaptability is not good. Another is the learning dictionary; this kind of dictionaries need not know the prior information of the signal; they can be set up directly with a small amount of training samples, and then through the machine learning method, a new self-adaptive dictionary can be developed by iterative update.

### 66.3.1 Dictionary Learning

Dictionary learning is a process which looks for the best atomic bases under the sparsest representation. It not only can satisfy the uniqueness constraints of sparse representation but at the same time can get the sparser and more accurate representation as well. The learning process is as follows.

We can get the training sample from TM remote sensing image data set  $X = \{x_i | x_i \in R^m, 1 \leq i \leq n\}$  to form the initial dictionary  $D \in R^{m \times p}$  (each column  $D_j \in R^m$  is an atom). This dictionary  $D$  can capture spectral characteristics of the pixels. And each pixel  $x_i$  can be linear represented by the atoms in the dictionary  $D$ . Then the dictionary learning problem can be transformed into the optimization problem as follows:

$$\begin{aligned} \min_{D, \alpha} \frac{1}{2} \|X - D\alpha\|_2^2 + \lambda \|\alpha\|_{1,1} \\ \text{s.t. } \|D_j\|_2 \leq 1 \quad \forall j \in 1, \dots, p, \quad p_2 \leq 1, \forall j \in 1, \dots, p \end{aligned} \quad (66.3)$$

where  $\alpha \in R^{p \times n}$  is the sparse matrix (each column  $\alpha_i \in R^p$  is a coefficient vector), the first polynomial of the above formula is the reconstruction error, and the second is sparse penalty function. Parameter  $\lambda$  represents a compromise between sparsity and data reconstruction.

If the sparse coefficient  $\alpha$  is known, formula (66.3) is a convex optimization about dictionary  $D$ ; if the dictionary  $D$  is known, the formula is a convex optimization about the coefficients; but it is not feasible to solve coefficient  $\alpha$  and dictionary  $D$  at the same time. To solve such optimization problem, we can assume that one of them is known; then  $x$  and  $D$  can be updated alternatively, and we can get the learning dictionary  $D$  by iteration [4].



### 66.3.2 Sparse Decomposition of the Characteristics

We can get the sparse decomposition coefficients  $\hat{\alpha}(x)$  of the training samples by the learning dictionary  $D$  according to formula (66.4). Then we can represent each characteristic of pixel  $x \in X$  sparsely and get their linear decompositions:

$$\hat{\alpha}(x) = \arg \min_{\alpha} \|x - D\alpha\|_2^2 + \lambda \|\alpha\|_1 \quad (66.4)$$

The above formula is a  $l_1$ -regularized least squares problem; the solution is a sparse vector in which only a few elements are not equal to zeros. Since the pixels that belong to the same materials have very similar structure, the values of sparse coefficients, which are obtained from dictionary  $D$  to represent the pixels' characteristics, are similar. According to the relevant reference [5], we can put  $\hat{\alpha}(x)$  as another new expression of pixel  $x$  which represents the sparse characteristics of pixel  $x$  in the dictionary  $D$ .

### 66.3.3 Classification of Image Sparse Representation

Different classificatory dictionaries can be built with different types of training samples according to the sample category. The projective coefficients of the test sample  $\hat{x}$  will not be zeros when the sample is decomposed with its own classificatory dictionary; otherwise, the coefficients should be zeros with other dictionaries so that the category of the samples can be classified.

Actually, due to the noise or the modeling error, the projective coefficients of the test sample  $\hat{x}$  are not zeros both with its own dictionary and with other dictionaries. Therefore, the category of the test sample  $y$  needs to be determined. If  $\delta_i(\hat{x})$  is a new vector and its nonzero coefficients only include these components in which  $\hat{x}$  has the relationship with the category  $i$ , then we can use  $\hat{y}_i = D\delta_i(\hat{x})$  to approximate  $y$  when  $y$  belongs to the category  $i$ . In another words, the smaller the distance between  $\hat{y}_i$  and  $y$ , the greater the possibility of  $\hat{y}_i$  belonging to the category of  $i$ th. So, the method of identifying the category of  $y$  is as follows:

$$\min_i r_i(y) = \min_i (\|y - D\delta_i(\hat{x})\|_2) \quad (66.5)$$

where  $r_i(y)$  is the residual to reconstruct  $y$  by training sample of the  $i$ th category.

## 66.4 The Algorithm of This Chapter

This chapter uses American landsat TM multispectral remote sensing images as the main data sources in which we can get the remote sensing image of Yinchuan wetland in October 2010. At first, the image is processed by some pretreatments, e.g., adjusting, deionizing, and cropping, and then we can get a  $400 \times 400$  pixel

wetland remote sensing image. We use our approach as described before to deal with this image. The algorithm steps are as follows:

#### **Algorithm Multi-feature image classification algorithm**

Input:

The image data set  $X = \{x_i | x_i \in R^m, 1 \leq i \leq n\}$  and use the data set to form the initial dictionary  $D \in R^{m \times p}$  (each column  $D_j \in R^m$  is an atom).

Output:

The decomposition coefficients  $\hat{\alpha}(x)$  and the residual errors  $r_i$ .

Procedure

1. Choose water, desert, vegetation, and Gobi wasteland as four kinds of samples in remote sensing image; the amount of each kind of samples  $x = 484$  ( $x \in X$ ), so the total number is  $484 \times 4$ .
2. The samples  $X$  are divided randomly into two parts as training set and testing set.
3. Get learning dictionary  $D$  with training set by using SPAMS toolbox [6].
4. Use lasso algorithm to decompose the test samples sparsely to get the sparse representation coefficients  $\hat{\alpha}(x)$  of pixel  $x \in X$  [7].
5. Compare the residual errors  $r_i$  of the test samples and their sparse representations; the minimum  $r_i$  is the category of the samples.

## 66.5 Experimental Results and Analysis

In order to evaluate effectively the experiments, the quantities of each type of samples are divided randomly into two equal parts as training set and test set. The experiments will be repeated ten times, and the final results are the average of the ten experimental results.

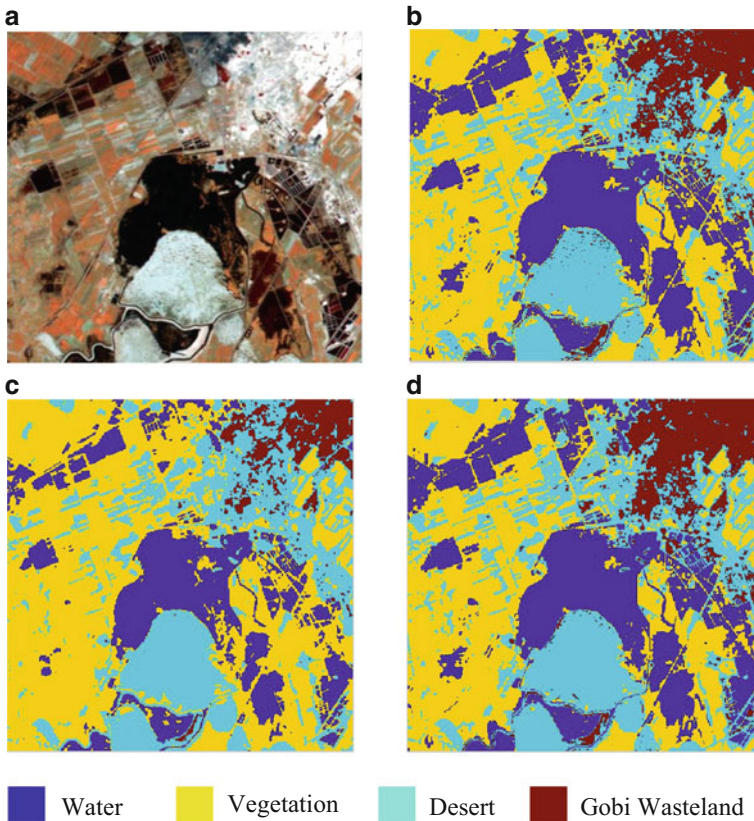
There are three different experiments in contrast to the support vector machine (SVM) and sparse representation.

1. The first group of the experiments takes six spectral values as its features, which are TM1 to TM5 bands and TM7 band. In this first group, we use the method of sparse representation classification.
2. The second group of the experiments takes the same features as the first group, but uses the SVM method.
3. The third group uses the sparse representation classification method. The total features of this group are ten. They include six spectral values, one NDVI value, and three K-T transformation components.

Table 66.1 shows the classification results: (1) The accuracy of the first group is much higher than the second. The overall accuracy of the first group is increased up to 92.56 % and Kappa coefficient is 0.9008, while the overall accuracy of SVM is only 88.74 % and Kappa coefficient is 0.8499. So the sparse representation is more

**Table 66.1** The classification accuracies are shown below with different characteristics and different methods

Characteristics	Classification method	Overall accuracy (%)	Kappa coefficient
TM1–5, TM7	Sparse representation	92.56	0.9008
	SVM	88.74	0.8499
TM1–5, TM7, NDVI, and K–T transformation	Sparse representation	95.05	0.9339

**Fig. 66.1** Classification effect is shown in the above figures, and we can compare the visual effects directly. (a) Sand lake original picture. (b) Only spectral information of sparse representation. (c) Classification of SVM with spectral information. (d) The results of our approach

robust. (2) The classification accuracy of the third group is higher than the first and the second. Its classification accuracy is up to 95.05 % and Kappa coefficient is 0.9339. So the results show that the features of K–T transformation and NDVI improve the classification accuracy and verify the above discussion (Fig. 66.1).

## 66.6 Conclusion

In this chapter, we have proposed a new framework for multispectral remote sensing image classification based on sparse representation which uses remote sensing image spectral values and its transformed features together. The simulation results demonstrate that our algorithm is better in both classification accuracy and visual effect than the conventional sparse algorithm and SVM methods.

**Acknowledgements** This work is supported by National Natural Science Foundation of China (61162013) and (61163016), Natural Science Foundation of Ningxia (NZ11143), and Projects of Beifang University of Nationalities (2012 XYC053). It is also supported by National Expert Introduction Projects.

## References

1. Fang, Y., Wang, X. L., Feng, Z. K., Wang, H. P., Nie, M. N., & Wang, B. (2010). One method for model calculation of Quickbird image extraction of urban green space information. *Forest Inventory and Planning*, 35(4), 19–23 (In Chinese).
2. Wang, Q., Wang, N., & Cao, X. F. (2013). Region vegetation cover change analysis of remote sensing monitoring based on TM image. *Yellow River*, 35(2), 70–71 (In Chinese).
3. Kauth, R. J., & Thomas, G. S. (1976). The Tasseled Cap—a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. In *Machine Processing of Remotely Sensed Data* (pp. 41–51). West Lafayette, IN: LARS, Purdue University.
4. Song, X. F., & Jiao, L. C. (2012). Classification of hyperspectral remote sensing image based on sparse representation and spectral information. *Journal of Electronics & Information Technology*, 34(2), 268–272 (In Chinese).
5. Raina, R., Battle, A., Lee, H., Packer, B., & Y. Ng, A. (2007). Self-taught learning: Transfer learning from unlabeled data. In *International Conference on Machine Learning* (pp. 759–766). Corvallis, OR.
6. Mairal, J., Bach, F., Ponce, J., & Sapiro, (2010). Online learning for matrix factorization and sparse coding. *Journal of Machine Learning Research.*, 11(1), 19–60.
7. Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society.*, 58(1), 267–288.

# Chapter 67

## A Parallel and Convergent Support Vector Machine Based on MapReduce

Yingying Ma, Liming Wang, and Longpu Li

**Abstract** In order to improve the performance of the traditional support vector machine (SVM), this chapter proposes one method referred as MR-SVM to parallelize SVM on MapReduce and mitigates the convergence problems brought by data partitioning and distributed computation. By splitting the large dataset and concurrently calculating the support vector set of each chunk across map units, MR-SVM improves the process capability and efficiency. Then the partial support vector sets are combined as the training set of the global training in reduce phase, and the current global optimum solved by reducing operations is fed back to each map units to determine whether MR-SVM should proceed with another pass. This process iterates until MR-SVM converges to the global optimum. In theory, it has been proved that MR-SVM converges to the global optimum within finite iteration size. Experimental results show that MR-SVM can improve the data processing capability and efficiency of the traditional counterpart and guarantee its high accuracy.

### 67.1 Introduction

Support vector machine (SVM) is a common method for constructing classification rules as well as one of the most robust and stable algorithms with outstanding accuracy in the field of machine learning. However, the standard SVM algorithm could only deal with some small sample datasets. So faced with the large-scale datasets produced by the practical problems nowadays, there are an increasing number of researches focusing on developing an efficient and high-speed SVM algorithm.

---

Y. Ma (✉) • L. Wang • L. Li  
Information Engineering Institute, Zhengzhou University, Zhengzhou 450052, China  
e-mail: [mayingyingch@163.com](mailto:mayingyingch@163.com)

In order to improve the performance of SVM, it is a widespread approach that splitting the dataset and training the individual chunks in parallel. But there is a variety of problems in these literatures. For example, Graf et al. introduced the Cascade SVM [1] which guarantees the global convergence, but it is poorly scalable and flexible as it requires predefining the network topology and size; Cao et al. proposed one parallel implementation of sequential minimal optimization (SMO) based on MPI [2], which focused on the efficiency but ignored the accuracy; Li et al. proposed a parallel incremental SVM algorithm on GPU [3], which is not preferable because of the lack of a theoretical convergence bound; Caruana et al. proposed one parallel SVM algorithm based on MapReduce for scalable spam filtering [4], but the operation in reduce phase is designed so simply that it needs human expertise to minimize the accuracy degradation, which leads to loss of automation in the approach. However, Chu et al. proved that SVM fits well into the MapReduce framework [5].

According to the deficiencies and inspirations in the previous researches, MR-SVM is focused on providing a scalable and flexible method for classifying large-scale database, improving the data processing capability and efficiency of the traditional SVM algorithm and solving the convergence problem in some similar parallel researches.

## 67.2 The Analysis and Design of MR-SVM

It is a general solution to improve the capability and efficiency of the standard SVM for data partitioning and distributed processing, but it always leads to convergence problems, resulting in a decline in classification accuracy. It is a problem worth studying to know how to make a better balance between the efficiency and the accuracy of the improved approaches.

### 67.2.1 Nonlinear SVM Classification Algorithm

The solving progress of the standard nonlinear SVM classification algorithm is concluded as below. Firstly, the Lagrange multipliers  $\lambda_i$  are obtained by evaluating the dual-problem formulation of the Lagrange function of the SVM optimization problem [6]:

$$\max L_D = \sum_{i=1}^N \lambda_i - \frac{1}{2} \sum_{i,j} \lambda_i \lambda_j y_i y_j K(x_i, x_j) \quad (67.1)$$

$$\text{Subject to } \sum_{i=1}^N \lambda_i y_i = 0, \quad \lambda_i \in [0, C], \quad i = 1, 2, \dots, N \quad (67.2)$$

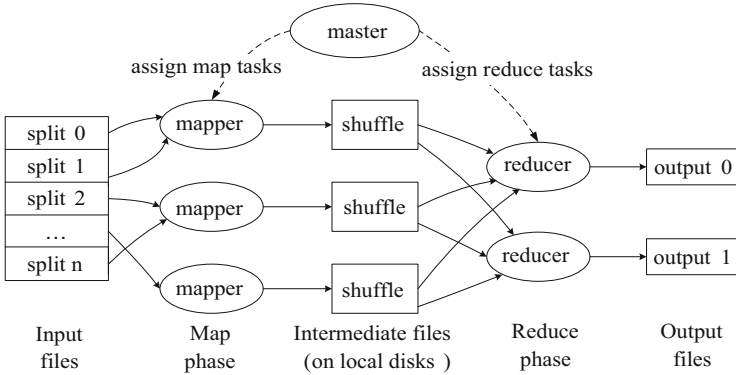


Fig. 67.1 Overview of Mapreduce programming framework

Most of the Lagrange multipliers are to be zero, and samples corresponding to nonzero multipliers are called the support vector set (SVs). Then the feasible solutions of the decision boundary parameters ( $w, b$ ) are obtained by the following formulations [6]:

$$w = \sum_{i=1}^N \lambda_i y_i \Phi(x_i) \quad (67.3)$$

$$y_i \left( \sum_{i=1}^N \lambda_i y_i K(x_i, x) + b \right) - 1 = 0, \quad \lambda_i \in (0, C) \quad (67.4)$$

These two formulations indicate that the decision boundary parameters ( $w, b$ ) depend on SVs.

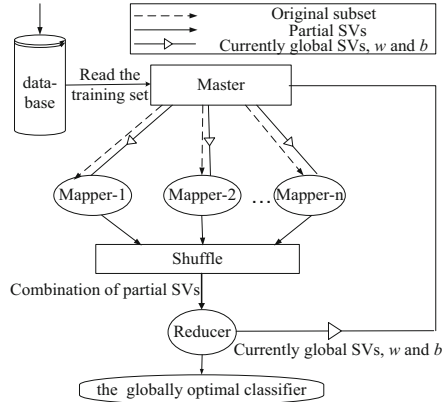
Finally, the classification function of nonlinear SVM is obtained as follows [6]:

$$f(x) = \text{sign} \left( \sum_{i=1}^N \lambda_i y_i K(x_i, x) + b \right) \quad (67.5)$$

## 67.2.2 MapReduce Programming Model

MapReduce is a parallel and distributed programming model in support of large dataset processing [7]. Figure 67.1 shows the overview of a MapReduce programming framework. The MapReduce library partitions the input data into  $n$  splits; then the *master* invokes *mapper* units distributed across multiple machines to process the input splits in parallel; then the intermediate key/value pairs produced by the *Map* function are written to local shuffle; and finally, the *reducer* units

**Fig. 67.2** The system model of MR-SVM



iterate over the intermediate data and append the output of the *Reduce* function to a final output file [7]. It is a highly scalable and flexible data processing tool which is increasingly popular.

### 67.2.3 System Model of MR-SVM

Based on the standard nonlinear SVM classification algorithm and integrating it with the MapReduce framework, the system model of MR-SVM is presented in Fig. 67.2.

#### 67.2.3.1 Parallel Strategy in Map Phase

Since the phase of evaluating the Lagrange multiplier dominates the time and space overhead of the SVM algorithm, it is a wise choice to execute this step in parallel. Accordingly, in MR-SVM, a large dataset is split into chunks horizontally, and then the individual chunks are distributed to the currently idle work units which will evaluate the Lagrange multipliers in parallel. As a result of the decision boundary decided by the SVs, every partial SVs in map phase is stored in local shuffle to represent the corresponding subset. In addition, to improve the efficiency further, the SMO algorithm [8] is adopted to computing the Lagrange multipliers, for its time complexity is  $O(n^2)$  while that of some other algorithms is  $O(n^3)$ .

Compared with the cascade structure [1], MR-SVM is highly scalable and flexible. Because it is built on the MapReduce framework, MR-SVM can call the idle work units one by one according to the current assignments. But in a binary cascade architecture, every time it increases one layer, there will be  $2^n$  nodes being added into the network, where  $n$  is the number of layers in the previous network. What is more, in contrast with the parallel SMO algorithm based on MPI [2], the parallel nodes in MR-SVM run separately, which contributes to a decline in communication overhead.



### 67.2.3.2 A Global Training in Reduce Phase

Because of the equivalence between the SVs and its corresponding training set, the partial SVs obtained from each map unit can represent each respective subset. Therefore, to advance the optimization progress automatically, the representations of every subset are combined and entered as the training set of the global training in reduce phase to evaluate the current global SVs, decision parameters ( $w$  and  $b$ ), and optimal objective function. In this case, MR-SVM is superior to the enhanced parallel SVM algorithm [4] which could not complete the whole optimization progress automatically.

Generally speaking, the support set is only a small part of the training set, so the parallelization of extracting the support vector of each subset in map process is an efficient strategy to reduce time and space overhead for the global training in reduce phase.

### 67.2.3.3 Feedback Loop Mechanism

Nevertheless, the union of the partial SVs is not equivalent to the whole training set, so the result of the reduce operation is not the global optimum. Therefore, the feedback loop mechanism is employed to guarantee that MR-SVM converges to the globally optimal solution. The results of the reduce phase are fed back to every map unit, and then each map unit employs the Karush–Kuhn–Tucker (KKT) conditions determined by the current optimum to test whether there are any samples violating the current KKT conditions [9]. If all samples in map phase satisfy the KKT conditions, it indicates that MR-SVM has converged to the global optimum; otherwise, the samples violating the current KKT conditions will be combined with the current global SVs and enter the next optimization pass to optimize the separating margin further.

The testing for KKT conditions not only plays a role in determining whether MR-SVM is converged to the global optimum but also screens the samples which should incorporate into the next training pass, contributing to the reduction in timing and space overheads of the following training phase.

## 67.3 Proof of Global Convergence

To illustrate the global training in reduce phase and the feedback loop mechanisms guarantee that MR-SVM converges to the global optimum in finite iteration size; give the below analysis and proof on the basis of the introduction of the MR-SVM system model and the executing flow in Sect. 2.3.

The identifications utilized during the proving process are listed as follows, where the superscripts  $j$  ( $j = 0, 1, 2, \dots, l$ ) indicates that the  $j$ -th iterative training pass.

$\Omega$ : The whole training dataset.

$m_i^j$ : The training subset of mapper -  $i$  unit during the  $j$ -th iterative training pass.

$M^j$ : The family of the training subset in map phase during the  $j$ -th iterative training pass,  $M^j = \{m_i^j | i = 1, \dots, n\}$ .

$r_i^j$ : The partial SVs which mapper -  $i$  outputs during the  $j$ -th iterative pass.

$R^j$ : The training set of reduce phase during the  $j$ -th iterative pass,  $R^j = \{r_i^j | i = 1, \dots, n\}$ .

$SV_{Global}^j$ : The current global optimal SVs of the  $j$ -th iterative pass.

$P(X)$ : The optimization function over dataset  $X$ .

According to the theorems proved by Graf [1], assume that the training set  $S$  is a subset of the training set  $F$  and  $Sv(S)$  is defined as the SVs of  $S$ ; in this case,  $\forall S \subseteq F$ ,  $P(S) = P(Sv(S)) \leq P(F)$ . Moreover, assume that the training set  $S$  is a proper subset of the training set  $F$ ; it is obvious that  $\forall S \subset F$ ,  $P(S) = P(Sv(S)) < P(F)$ .

In addition, assume that  $F$  is a set of training subsets; the subset  $f^*$  which achieves the greatest  $P(f)$  will be called the best subset in family  $F$ , defining  $P(F) = P(f^*)$  [1].

In this case, there will be a sequence  $(M^j)$  through the iterative training. So the test for the global convergence in map phase can be described as whether the formulation  $P(M^j) = P(\Omega)$  is established.

**Corollary:** The sequence  $(M^j)$  of MR-SVM converges to the global optimum within finite steps, namely,  $\exists j^*, \forall j > j^*, P(M^j) = P(\Omega)$ .

*Proof.* For  $\forall j, j = 0, \dots, l - 1$ , there is a set  $r_i^j \in R^j$  containing the support vector of the best set  $m_i^{j*}$  in  $M^j$ , that is,  $Sv(m_i^{j*}) \subseteq r_i^j$ , so  $P(m_i^{j*}) = P(Sv(m_i^{j*})) \leq P(r_i^j)$ . Therefore,  $P(M^j) = P(m_i^{j*}) \leq P(r_i^j) \leq P(R^j)$ . For  $\forall j, j = 0, \dots, l - 2$ , the  $m_i^j$  is the union of the violators of the current KKT conditions and the current  $SV_{Global}^j$ , so the current  $SV_{Global}^j$  is a proper subset of all  $m_i^{j+1} \in M^{j+1}$ , that is,  $Sv(R^j) \subset m_i^{j+1}$ , so  $P(R^j) < P(m_i^{j+1}) \leq P(M^{j+1})$ .

It is obtained through the above analysis that for  $\forall j, j = 0, \dots, l - 2$ ,  $P(M^j) \leq P(R^j) < P(M^j)$ , so  $P(M^j)$  strictly and monotonically increases during the iterative progress. For all  $S \subset \Omega$ , the sequence  $P(M^j)$  always takes its value within the finite set of the  $P(S)$ . For  $\forall i, j, i = 1, \dots, n, j = 0, \dots, l$ , the set  $m_i^j$  is always a subset of  $\Omega$ , plus  $P(M^j) = P(m_i^{j*})$ , so  $P(M^j) \leq P(\Omega)$ ; this implies that  $P(\Omega)$  is the upper limitation of  $P(M^j)$ .

Therefore, the iterative training continues until every  $m_i^l (i = 1, \dots, n)$  meets the current KKT conditions in the map testing operation when  $j = l$ . In this case, for  $\forall i, i = 1, \dots, n$ ,  $m_i^l = SV_{Global}^l = Sv(R^{l-1})$ , so  $P(M^l) = P(m_i^l) = P(Sv(R^{l-1})) = P(R^{l-1})$ . Finally, because the union of every training subset  $m_i^j (j = 0, \dots, l)$  entered into map phase is the whole training set  $\Omega$ , we get the conclusion that  $P(M^l) = P(R^{l-1}) = P(\Omega)$ .

To sum up, MR-SVM converges to the global optimal within finite iteration size.

### 67.4 Experiment Studying

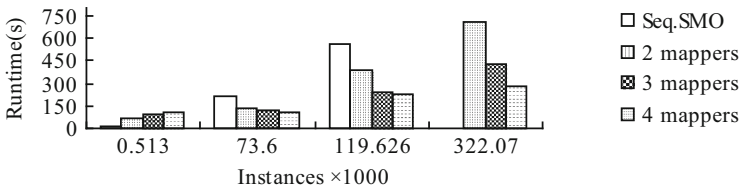
To assess the performance of MR-SVM, the experiments make a comparison with the sequential SMO. The configuration of the Hadoop cluster employed by MR-SVM is shown in Table 67.1. The nodes in the cluster were connected by Gigabit Ethernet. The SpamBase dataset with 4,601 instances was adopted, and the training sets of varying size were obtained by employing the Weka random re-sampling filter feature.

*A Comparison on Data Processing Capability and Efficiency.* MR-SVM made a comparison with the sequential SMO algorithm and executed among a varying number of nodes benefiting from the scalability of MapReduce framework. As listed in Fig. 67.3, the time required to train the SMO sequentially on a single node was  $\approx 554$  s while the time MR-SVM required was  $\approx 221$  s using the Hadoop cluster with four nodes when using 120K samples. Furthermore, the sequential SMO failed when training 322K samples. It indicates that MR-SVM exceeds sequential SMO on the performance of data processing capability and computing efficiency on large datasets.

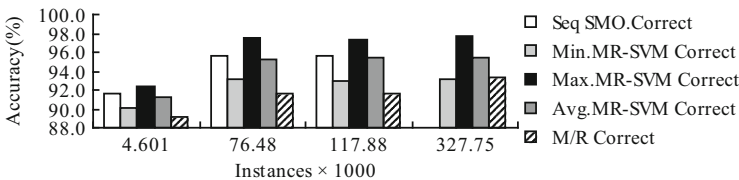
*A Comparison on Classification Accuracy.* We compared MR-SVM with the sequential SMO algorithm and the M/R algorithm [5]. The results in Fig. 67.4

**Table 67.1** The Hadoop cluster configuration of MR-SVM

Hardware environment			Software environment	
Nodes	Processors	Ram	O/S	Hadoop
Six nodes	Intel Core Duo	2GB	Ubuntu 12.04	Hadoop 0.20.2



**Fig. 67.3** A comparison on capability and runtime with sequential SMO



**Fig. 67.4** A comparison on accuracy with sequential SMO and M/R algorithm

indicate that MR-SVM keeps the same high accuracy level as the SMO and exceeds that of the M/R algorithm among different sizes of datasets. It is illustrated that the global training in reduce phase and the feedback loop mechanisms in MR-SVM are greatly helpful for improving the accuracy.

## 67.5 Conclusions

In summary, this chapter described a parallel and convergent SVM classifier based on MapReduce framework. It was proved that MR-SVM always achieves the global convergence within finite iteration size. The experimental results illustrated that MR-SVM can improve the data process capability and computing efficiency significantly on the basis of keeping the same excellent accuracy as the traditional sequential SMO.

However, the parameters, such as  $C$  and epsilon, were specified as default values directly in this chapter, so the parameter optimization selection problem in MR-SVM should be included in the future research work.

## References

1. Graf, H. P., Cosatto, E., Bottou, L., Durdanovic, I., & Vapnik, V. (2004). Parallel support vector machines: The Cascade SVM. *Advances in Neural Information Processing Systems (NIPS)*, 17, 521–528.
2. Salleh, N. S. M., Suliman, A., & Ahmad, A. R. (2011). Parallel execution of distributed SVM using MPI (CoDLib). In *International Conference on Information Technology and Multimedia (ICIM)* (pp. 1–4). IEEE.
3. Li, Q., Salman, R., Test, E., Strack, R., & Kecman, V. (2013). Parallel multitask cross validation for support vector machine using GPU. *Journal of Parallel and Distributed Computing*, 73(3), 293–302.
4. Caruana, G., Li, M. Z., & Liu, Y. (2012) *An ontology enhanced parallel SVM for Scalable spam filter training [EB/OL]*. <http://dx.doi.org/10.1016/j.neucom.2012.12.001>. doi:10.1016/j.neucom.2012.12.001#\_parent
5. Chu, C. T., Kim, S. K., Lin, Y. A., Yu, Y. Y., Bradski, G., Olukotun, K., et al. (2007). Map-reduce for machine learning on multicore. *Advances in Neural Information Processing Systems (NIPS)*, 19, 281–288.
6. Vapnik, V. (1995). *The nature of statistical learning theory* (pp. 131–162). New York, NY: Springer.
7. Dean, J., & Ghemawat, S. (2008). MapReduce: Simplified data processing on large clusters. *Communications of the ACM*, 51(1), 107–113.
8. Platt, J. C. (1999). *Fast training of support vector machines using sequential minimal optimization* (pp. 185–208). Cambridge, MA: MIT Press.
9. Wu, C. M., Wang, X. D., Bai, D. Y., & Zhang, H. D. (2009). Fast incremental learning algorithm of SVM on KKT conditions. *The Sixth International Conference on Fuzzy Systems and Knowledge Discovery* (pp. 551–554). IEEE.

# Chapter 68

## Vehicle Classification Based on Hierarchical Support Vector Machine

Mengwan Jiang and Haoliang Li

**Abstract** In order to solve the problem that the mature vehicle classification cannot meet the requirements of accuracy and speed concurrently, this chapter chooses the contour features and speeded-up robust features (SURF) features of vehicles and then adopts a hierarchical support vector machine (SVM) classifier for vehicle classification. At first, the system uses the contour features which are simple and fast in the first layer of classifier so that it will filter out the easy samples. Second, the system utilizes the rich information and stable SURF feature in the next layer of classifier. We conducted extensive experiments against a number of baseline methods; the accuracy of proposed method was increased by about 20 %, and the time was shortened by 2/3, significantly outperforming the baselines. The method of double features and hierarchical SVM has a good trade-off between speed and precision.

### 68.1 Introduction

Vehicle classification technology is an important component of intelligent transportation systems (ITS). It has been widely applied in many real-world applications, such as vehicle flow statistics, highway automatic toll station, management of all kinds of parking lot, and public security system. The traditional methods of vehicle classification are mainly based on the magnetic induction or ultrasonic technology. These methods encounter very high computational complexity. This may lead to a series of drawbacks, including time-consuming construction, expensive for creation and maintenance [1]. As a solution, we propose to apply image-processing technologies to support the classification system.

---

M. Jiang (✉) • H. Li

School of Information Engineering, Zhengzhou University, Zhengzhou 450001, China  
e-mail: [springmw@126.com](mailto:springmw@126.com); [iehlili@zzu.edu.cn](mailto:iehlili@zzu.edu.cn)

In the area of image processing, existing works of vehicle classification use vehicle's color, length, width, height, and outline features [2–4] to classify all vehicles into three sizes: small, medium, and large. But it is difficult to meet the requirements of accuracy and speed in making a distinction between passenger vehicles and lorries. In addition, it also cannot distinguish various small cars. Then the SURF feature can save the rich and important characteristic information [5], thus greatly improving the efficiency of classification. However, it has a relatively large amount of calculation and hereby does not meet the requirements as well. In terms of pattern classification, support vector machine (SVM) technique is a kind of pattern classification method [6]. It is believed to be outstanding in solving small sample, nonlinear, and high-dimensional pattern classification problems. This chapter studies the vehicle classification system, which chooses two features and uses a hierarchical SVM to meet the requirements of accuracy and efficiency.

## 68.2 The Algorithm of SURF Feature Extraction

On the selection of feature descriptor, speeded-up robust feature (SURF) point feature is fast [5], because it uses the Hessian matrix to improve the efficiency. It uses less time and is accurate to calculate the Hessian matrix. The key is the selection of location and scale space. Define a point  $X$  in the image,  $X = I(x, y)$ , in the scale space too; the Hessian matrix of pixel  $X$  under the condition of scale space for  $\sigma$  is as follows:

$$H(x, \sigma) = \begin{pmatrix} L_{xx}(x, \sigma) & L_{xy}(x, \sigma) \\ L_{xy}(x, \sigma) & L_{yy}(x, \sigma) \end{pmatrix} \quad (68.1)$$

In the formula  $L_{xx}(x, \sigma) = \frac{\partial^2}{\partial x^2} g(\sigma)$ ,  $L_{xy}(x, \sigma)$ , and  $L_{yy}(x, \sigma)$  in the same way.

SURF uses integral image algorithms so that the computing speed is increased. The SURF feature descriptor detects the neighborhood of each extreme value points which is  $4 \times 4$  regions and then calculates each Herr Wavelet response of the subdomain, a 4 d vector as

$$V = \left( \sum d_x, \sum d_y, \sum |d_x|, \sum |d_y| \right) \quad (68.2)$$

Finally all the  $4 \times 4$  neighborhoods and the 4 d vectors constitute 64 ( $4 \times 4 \times 4$ )-dimensional feature vectors. SURF is better off because it uses the method of integral image, and compared with other feature descriptor, SURF's speed of producing feature points has improved significantly.

### 68.3 Hierarchical SVM

SVM is a new kind of general statistical method based on structure risk minimization principle [6], which can solve the practical problems such as small sample, nonlinearity, high dimension, and local minimum point. There are two measures of controlling SVM's generalization: empirical risk and confidence limit. It uses the training error as the constraints of optimizing problems, and the goal of optimization is to minimize the confidence limit. As a result, the SVM is to solve a convex quadratic programming under leading constraints, making the solution of SVM the only, also globally, optimal solution.

The SVM is derived and developed from the optimal classification plane under the linear separable cases. The classification functions under the linear separable cases are shown below:

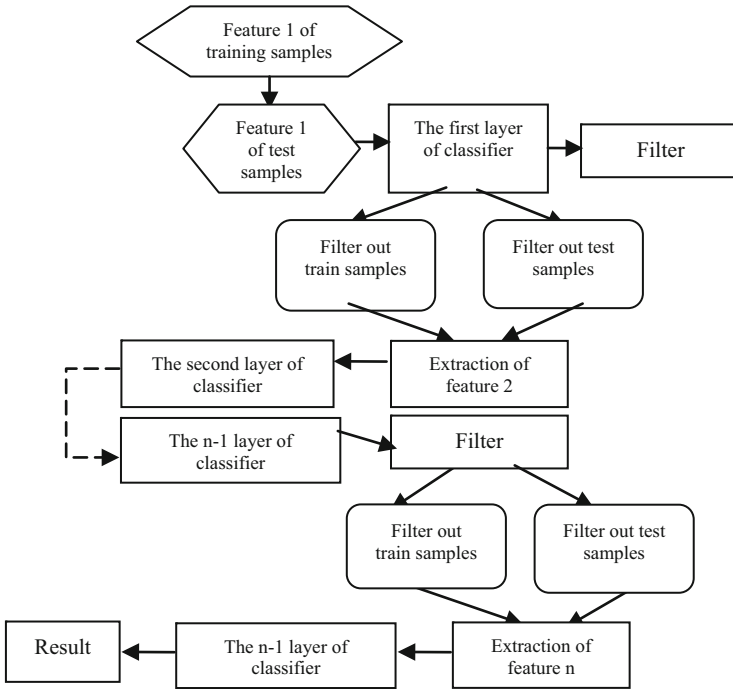
$$f(x) = \text{sign} \left( \sum_i y_i \alpha_i K(x, x_i) + b \right) \quad (68.3)$$

In this function,  $x_i$  is the  $i$  dimension of training space vector  $X$ ,  $y_i$  represents category  $x_i$  and its value is 1 or  $-1$ , and  $\alpha$  is the multiplier of Lagrange. For most  $i$ ,  $\alpha_i$  is 0.  $K$  is the dot product of the convolution kernel function. It can turn inner product operation of high-dimensional feature space into functional operation of low-dimensional input space. The curse of dimensionality caused by having nonlinearity of input space maps to high-dimensional feature space and can also be solved by kernel function. There are four basic kernel functions: linear, polynomial, RBF, and sigmoid kernel function. According to the features tested in this chapter, we finally choose the RBF kernel function. Its expression is  $K(x_i, y_i) = \exp(-\gamma \|x_i - y_i\|^2)$ ,  $\gamma > 0$ . The kernel parameter  $\gamma$  is determined by the results ( $\gamma = 0.15$  in this chapter).

At first SVM classification algorithm is only applicable to solve binary classification problems, lacking the ability to handle multiple classification problems [7]. In a binary classification problem, we are given a training set  $\{x_i, y_i\}$ ,  $i = 1, \dots, n$ ,  $x_i \in R^n$  is an  $n$  dimension real vector which represents sample properties.  $y_i \in \{1, -1\}^1$  is the category labels of data; then SVM can come down to an optimization problem as shown below:

$$\begin{aligned} \min_{w, b, \xi} \quad & \frac{1}{2} w^T w + C \sum_{i=1}^l \xi_i \\ \text{Subject to} \quad & y_i (w^T \phi(x_i) + b) \geq 1 - \xi_i \quad \xi_i \geq 0. \end{aligned} \quad (68.4)$$

$C > 0$  is a penalty parameter of error term.  $\phi$  is a mapping function, and  $K(x_i, x_j) = \phi(x_i)^T \phi(x_j)$  is the kernel function.



**Fig. 68.1** The flow chart of hierarchical support vector machine

In order to extend the SVM to the multi-classification, now there are two ways in general: one to one and one to many. In the SVM experiment given in this chapter a hierarchical SVM method with multi-features and one-to-many method in every layer of classification are used. At first, the hierarchical SVM method is to use the feature one of all the train samples to train the first layer of classifier [8] so as to filter out the samples which are not easy to identify according to the different threshold which is set at first. Then to extract feature two from the difficult samples after filtering, use the feature two to train the next layer of classifier. Test samples work in the same way, firstly to test the test samples in the first layer of classifier, filter out difficult samples according to the predicted value and threshold value, and then send them into the next layer of classifier. By such analogy, finally we get relatively the classification result through filtering step by step and centralized training. Specific flow chart is shown below (Fig. 68.1):

Compared with the traditional SVM, a hierarchical SVM has the following two advantages:

1. As the filter time increases, it is more difficult to classify the training sample which is filtered out, making the training more targeted. So we can filter out the easy samples in the previous layers of classifier and then use difficult samples, which are similar among samples to train SVM intensively, making it more targeted and more likely to get the optimal solution.



- Each layer can extract different features for classification, making the method more flexible. Due to a hierarchical SVM trains on each layer step by step, according to the need, it can try to extract different features from samples on each layer.

## 68.4 Test and Result Comparison

We collected 240 vehicle images, including 8 types of vehicles. There are 20 training samples and 10 test samples in every type, building the image dataset.

### 68.4.1 Feature Extraction

In this process, the system firstly extracts the outline features of all the samples and the SURF features of cars. The example's result of extracted features of outline feature and SURF feature shown as Figs. 68.2 and 68.3.

After extracting vehicles' features, the feature set will be later post-processed, through methods such as k-means algorithm and taking average by dividing area [9], in order to obtain a single feature vector, which is the characteristic classification to the corresponding vehicles.

### 68.4.2 Training and Testing the Classifier

In training, use the contour feature information of all training vehicles to train the first layer of classifier, filtering out large-type truck, medium-type truck, large buses, and medium bus, the four samples which are easy to be identified. Use the SURF feature of the cars (small vehicles) to train the second layer of classifier.

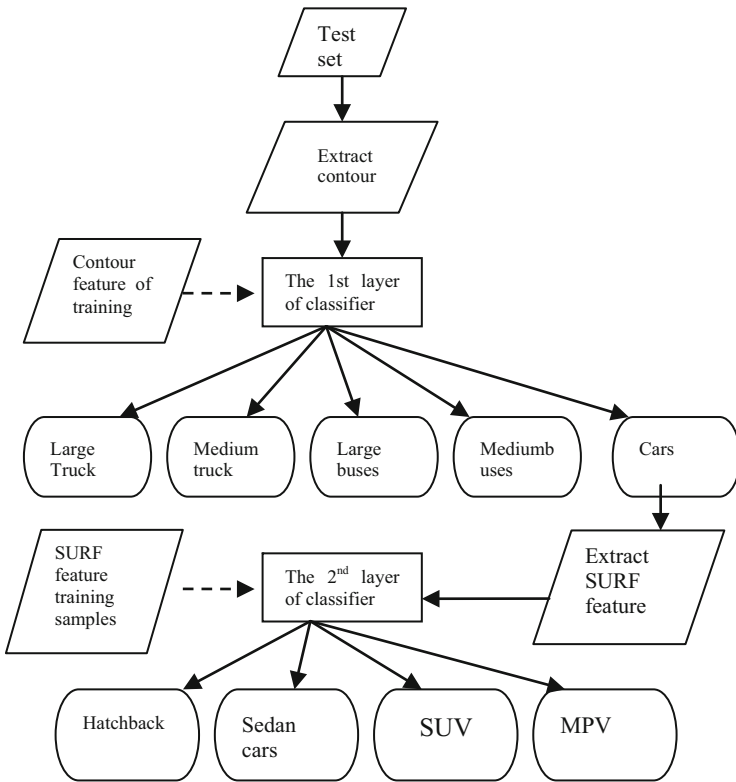
In the first layer of classification, send the contour features of the images after preprocessing and feature extraction into the first layer of classification. Use model

**Fig. 68.2** Outline feature of a medium-type truck



**Fig. 68.3** SURF feature of a multipurpose vehicle





**Fig. 68.4** Training and test flow of the double features and hierarchical SVM

code as output vector of the network to classify large-type truck, medium-type truck, large buses, and medium buses. Extract the SURF feature of the small vehicles, which are after filtering by the first layer of classifier, and then send them into the second precise layer of classifier to classify the vehicle by sedan car, hatchback SUV, and MPV.

The double features and hierarchical SVM training and testing flow are shown below. The dotted line is training process, and the solid line is testing process (Fig. 68.4).

### 68.4.3 Results and Comparisons

This chapter is implemented on MATLAB2010 with LIBSVM toolbox for the hierarchical SVM pattern classification. LIBSVM provides not only the source code but also some executable files under the Windows operating system, including svmtrain function, which is used to train SVM; svmpredict function which is to

**Table 68.1** Classification results of double features and hierarchical SVM

The number of training samples	The number of test samples	The number of support vector	Total training time (s)	Total test time (s)	Rate of classification
160	80	42	47.7	12.69	87 %

**Table 68.2** Comparison results of different methods

Methods of experiment	Total time of 160 training samples	Total time of 80 test samples	Rate of classification
Classification of single outline feature	31.8	9.17	69 %
Classification of single SURF feature	56.9	16.41	82 %
Traditional SVM classification of double feature	72.6	19.82	90 %
Hierarchical SVM classification of double feature	47.7	12.69	87 %

predict the test data according to the acquired training model; and `svmscale` function which is to normalize the training data and test data. In the experiments, we choose the default model parameters. The classification results of double features and a hierarchical SVM are shown in Table 68.1.

In the experimental results, total training time of 160 training samples is 47.7 s, less than a minute. And total test time of 80 test samples is just 12.69 s, one sample time about 0.16 s on average; therefore it can meet the speed requirement. Due to using the default model parameters of LIBSVM toolbox, it works well in terms of accuracy with classification rate of 87 %.

The contrast test of using different methods but all default values of classification with the same training samples and testing samples further verifies that the method used in this chapter is effective, including single contour feature classification, single SURF feature classification, and double features of the traditional SVM classification. Comparison results are shown in Table 68.2.

It can be seen from the result and comparisons of the experiment that double-feature hierarchical vehicle classification compared to single-feature recognition method, which is the fastest, is tens of seconds later, but the accuracy increased from 69 to 87 %. Compared to the double-feature traditional classification method which has the highest accuracy, the accuracy only reduced, three percentage points (from 90 to 87 %), but it is also reduced by 40 % time cost.

## 68.5 Conclusion

Compared with the single-feature or traditional classification method, this chapter used double-feature hierarchical vehicle classification. It can filter the easy samples with fast feature extraction by the first layer of classifier which is also trained by the fast feature, and then it sends the difficult samples into the second layer of classifier which is trained by rich information and high precision feature to classify by filtering step by step and centralized training. So the hierarchical classification can meet the requirements for high accuracy and fast speed.

## References

1. Ren, J. Q. (2011). Design of vehicle type recognition algorithm based on video sequence. *Computer Engineering*, 37(24), 245–250.
2. Luo, W. T., Hsieh, J. W., & Fan, K. C., (2007). Vehicle detection using normalized color and edge map. *IEEE Transactions on Image Processing*, 16(3), 850–864.
3. Zhang, Q. Y., Dai, G. M., & Chen, L. (2008). Vehicle recognition system design based on speed videos. *Microcomputer Information*, 24(31), 288–290.
4. Jia, Y. Q., & Zhang, C. S. (2009). Front-view vehicle detection by Markov chain Monte Carlo method. *Pattern Recognition*, 42(3), 313–321.
5. Bay, H., Tuytelaars, T., & Van Gool, L., (2006). SURF: Speeded up robust features. *Computer vision—ECCV 2006* (pp. 404–417). Graz, Austria.
6. Hsu, C. W., & Lin, C. J. (2002). A comparison of methods for multi-class support vector machines. *IEEE Transactions on Neural Networks*, 13, 415–425.
7. Castrillon Santana, M., & Vuong, Q. C. (2007) An analysis of automatic gender classification. In *Proceedings of the 12th Conference on Progress in Pattern Recognition, Image Analysis and applications* (pp. 271–280). Valparaiso, Chile: Springer-Verlag.
8. Li, K. L., & Liao, P. (2012). Face image gender identification based on cascade connection support vector machine. *Computer Engineering.*, 38(12), 152–154.
9. Zhang, S. S., & Zhan, Z. C. (2012). Research on vehicle classification system based on SIFT features and support vector machine. *Computer Knowledge and Technology*, 08(17), 4277–4280.

# Chapter 69

## Image Splicing Detection Based on Machine Learning Algorithm

Yan Xiao

**Abstract** Image splicing is a common method to construct forged image which decreases the authentication of the traditional image. Resizing operation is usually necessary to create a convinced forged image. Though the forged image leaves no visual clues, resizing operation using interpolation method destroys the relationship between neighboring pixels, thus leaving traces which can be captured by statistical feature. We first convert the traces left by resizing to feature and then feed features from enough sample images to support vector machines to train for detector. Finally, we use detector to determine whether the image is tampered and point out which parts of the image are tampered by block-wise method. Experimental results verify the effectiveness of our proposed method.

### 69.1 Introduction

Today's digital technologies make the image edition easily, so there are many forged images appearing in our daily life. Though elaborate tampering operations almost leave no visual clues, the correlations between adjacent pixels are disrupted, and such changes can be captured by statistical method just as forensics. The digital forensic method is used to check the validity of the image based on statistical fingerprint without inserting any evidence into the image.

In recent years, many forensic methods are designed for image tampering detection. Farid divided the forensic tools into five groups: (1) pixel-based technology which detects the fingerprint at the pixel level; (2) format-based technology mainly for JPEG compressed images; (3) camera-based and physically based technology; and (4) geometric-based technology [1]. Image splicing is a common

---

Y. Xiao (✉)

Southwest University of Science and Technology, Mianyang 621000, China

Mianyang Vocational and Technical College, Mianyang 621000, China

e-mail: [348105095@qq.com](mailto:348105095@qq.com)

forged manipulation which spliced two or more images into a forged image. Some authors found that the statistics on high-order Fourier can be used as fingerprint [2, 3].

Resizing operation is usually employed to create an applauded forged image. This process employs re-sampling operation which introduces period correlations between neighboring pixels caused by interpolation algorithm. Farid et al. utilized EM algorithm to reveal the above period relation and using this correlation to test whether the tested image is resized or not [4]. Babak Mahdian et al. demonstrated that the interpolated signal and its derivatives contained specific periodic properties and proposed a blind and efficient method to detect re-sampling and interpolation [5]. Andrew C. Gallagher also exploited a periodicity in the second derivative signal of interpolated images and employed discrete Fourier transform (DFT) to capture such periodicity in frequency domain [6]. Previous works [4–6] utilized periodicity as fingerprint left by interpolation; they did not deal with the situation that the periodicity contained in the rescaled image is faint such as when the rescaled factor is near to 1 (1.01, 0.99). In this chapter, the correlations between neighboring pixels are usually used as fingerprint for forensic tools. Chen et al. [7] proposed a blind and effective splicing approach based on a natural image model. They finally used feature sets and support vector machines to classify whether images are forged or not. The transition probabilities of Markov chains between neighboring pixels are modeled to construct feature set to detect median filtering [8].

As resizing operation will destroy the correlations between neighboring pixels, we use the method as [8] to construct the feature set and use the support vector machines (SVM in short) to learn the difference between original images and resized images. In order to locate spliced part, SVM is trained for a classifier on small-size images. The tested image is first divided into blocks and then using the classifier to test each block whether tampered or not.

The remainder of this chapter is organized as follows. In Sect. 69.2, we first construct the feature set and train for a classifier and then test the image using block-wise method. In Sect. 69.3, the experimental results show the effectiveness of our proposed method. Finally, we draw a conclusion at the end of the chapter.

## 69.2 Procedure of Our Proposed Method

In this section, we first introduce how to construct the feature to model the different correlations between rescaled images and non-rescaled images. Then the SVM are briefly introduced. Finally we give a detailed description of our method to detect image splicing.

When resizing an image, re-sampling method is always used. The bilinear and bi-cubic interpolation algorithm is a popular method accompanied with resizing operation. Let us assume that  $x(i,j)$  is the pixel in original image and  $y(i,j)$  is the

pixel in rescaled image. When using bilinear interpolation to compute the value of  $y(i,j)$ , the following formula is held:

$$y(i+u, j+v) = (1-u) \times (1-v) \times x(i, j) + (1-u) \times v \times x(i, j+1) + u \times (1-v) \times x(i+1, j) + u \times v \times x(i+1, j+1), \quad (69.1)$$

where  $u$  and  $v$  are the displacement from  $i$  and  $j$ , respectively, and  $0 \leq u, v \leq 1$ . From Eq. (69.1), it is easy to infer the neighboring pixel value of  $y(i,j)$ . Taking  $y(i+1+u, j+v)$  for example:

$$y(i+1+u, j+v) = (1-u) \times (1-v) \times x(i+1, j) + (1-u) \times v \times x(i+1, j+1) + u \times (1-v) \times x(i+2, j) + u \times v \times x(i+2, j+1). \quad (69.2)$$

It is obvious that there are two same pixels ( $x(i+1, j)$ ,  $x(i+1, j)$ ) that are contained in Eqs. (69.1) and (69.2). Thus, the correlations of neighboring pixels between interpolated images and non-interpolated images are different. Kirchner et al. model the fingerprint left by median filtering using first-order difference for adjacent pixels and utilized subtractive pixel adjacency matrix to construct the feature set [8]. Inspired by [6], we model the neighboring pixels by second-order difference and utilize subtractive pixel adjacency matrix to construct feature set. We model the adjacent pixels by second-order difference along horizontal and vertical directions as Eq. (69.3):

$$d_{i,j}^{k,l} = I(i+2k, j+2l) + I(i+k, j+l) - 2I(i, j), \quad (69.3)$$

where the superscript  $(k,l) \in \{(0,1), (0,-1), (1,0), (-1,0)\}$ . As the bilinear interpolation method considers two neighboring pixels, we consider the correlations among three neighboring pixels. The correlations of neighboring pixels are expressed by transition probabilities of second-order Markov chains as Eq. (69.4):

$$P_{\alpha,\beta,\gamma}^{k,l} = Pr(d_{i+2k,j+2l} = \alpha | d_{i+k,j+l} = \beta, d_{i,j} = \gamma), \quad (69.4)$$

where  $Pr(\cdot)$  in Eq. (69.4) means the probability.

Finally, we compute the feature by averaging four directions as Eq. (69.5):

$$P_{\alpha,\beta,\gamma} = \left( P_{\alpha,\beta,\gamma}^{0,1} + P_{\alpha,\beta,\gamma}^{0,-1} + P_{\alpha,\beta,\gamma}^{-1,0} + P_{\alpha,\beta,\gamma}^{1,0} \right) / 4. \quad (69.5)$$

If the pixel value ranges from  $[0, 255]$ , the residual pixel in ranges  $[-255, 255]$ , there are many elements in feature sets. It is unfeasible to feed such enormous feature into SVM to train for a model, so the truncation must be adopted to decrease the dimensionality of the feature. In order to lower feature's dimensionality, truncation of is executed. When the absolute value of difference is larger than a

threshold  $T$ , we truncate it. In the experiment, we set  $T = 3$  which results in 343 dimensionality feature.

The SVM is employed to train for the detection classifier. SVM is a classical supervised learning algorithm which can be used for binary classification [9]. In order to be a better classifier, SVM find an optimal hyper-plane which makes the nearest distance between the training samples and hyper-plane as large as possible, which can be formulated as an optimization problem as Eq. (69.6):

$$\begin{aligned} \min_{w, b, \xi} \quad & \frac{1}{2} w^T w + C \sum_{i=1}^l \xi_i \\ \text{St. } & y_i (w^T \varnothing(x_i) + b) \geq 1 - \xi_i, \xi_i \geq 0, \end{aligned} \quad (69.6)$$

where  $x_i$  is the training feature,  $y_i$  is the training label, and  $i = 1, 2, \dots, l$ .  $\varnothing(x_i)$  is the mapping function which will be used in kernel function.  $C > 0$  is the penalty parameter of the error term. The solution of Eq. (69.6) is usually transferred to its dual-quadratic problem (69.7) to solve

$$\begin{aligned} \min_{\alpha} \quad & \frac{1}{2} \alpha^T Q \alpha - e^T \alpha \\ \text{St. } & y_i^T \alpha = 0, C \geq \alpha_i \geq 0, \end{aligned} \quad (69.7)$$

where  $e$  is a unit vector and  $Q$  is an  $l \times l$  positive semi-definite matrix computed by kernel functions. The other symbol is the same as in Eq. (69.6). After solving the problem (69.7), the decision function is used to predict the label of the testing samples.

LIBSVM is an integrated software for support vector classification; we use C-SVC with RBF kernel to learn the model [10]. The non-forged image is given a label 0, and the forged image is labeled 1. As the shape of the pasted object is not regular and the position of the pasted object is also unknown, the block-based method is employed to test whether the image is tampered or not. So the classifier is obtained on small-size image for highly accurate detection results. The size of trained images is designed to be the same as the size of blocks. In general, the larger size of trained image will get better results; however, it will cause losses of detail position at location.

Once getting the classifier, the block-wise method is employed to test the whether the image is spliced or not. First, the test image is divided into non-overlapping blocks of size  $S \times S$  with step size  $S$ . For an  $M \times N$  image, there are totally  $\text{floor}(M/S) \times \text{floor}(N/S)$  blocks ( $\text{floor}(x)$  is the function for the largest integer value which is smaller than  $x$ ). In general, the smaller block will get more detailed detection results; however, the accuracy of classifier obtained from small-size image is not high. So the proper size of the block is an important factor. Then, compute the feature for every block. Finally, the classifier is used to predict each block's label. If the block's label is 1, the block is considered as a forged block and mark it with red color box.



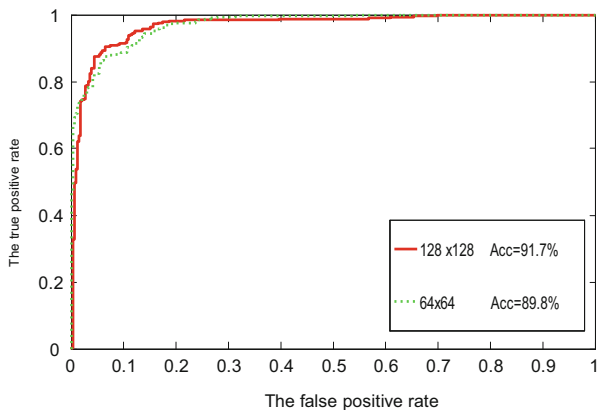
At last, we summarize our proposed method as follows:

- (1) Computing the feature sets from large number of small-size trained images. In our experiment, we set two kinds of images ( $128 \times 128$  and  $64 \times 64$ ) and then used SVM to train for a classifier on the above training set.
- (2) Dividing the tested image into blocks and computing each block's feature.
- (3) Feeding each block's feature into the trained classifier to predict whether each block is tampered or not.
- (4) Locating the tampered object by joining all tampered blocks.

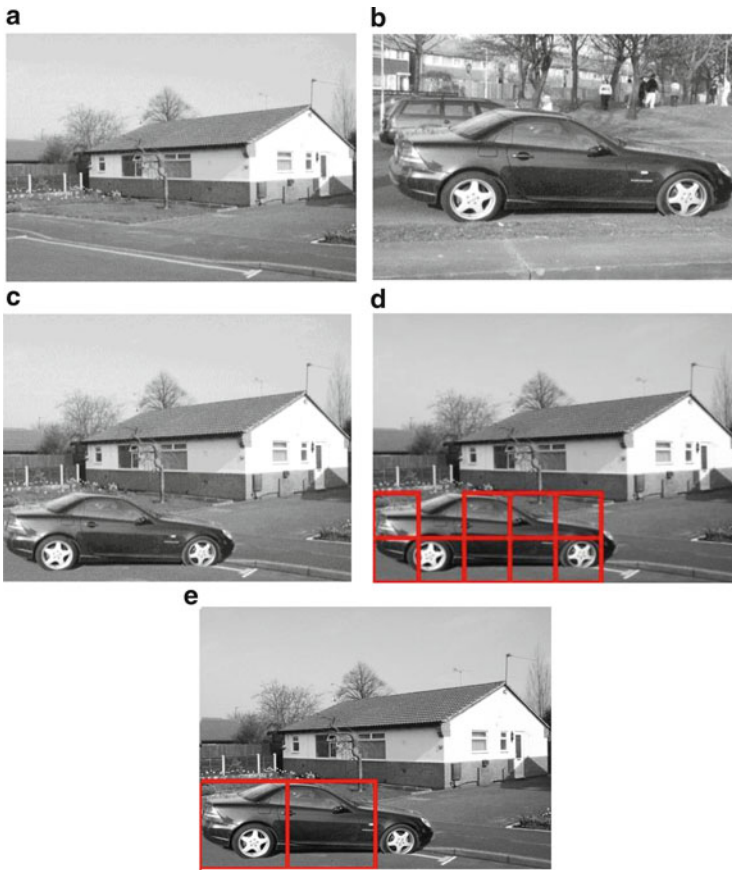
### 69.3 Experimental Results

In this section, we first test the efficacy of our proposed method on small-size images and demonstrate the performance by receiver operating characteristic curves (ROC in short) in Fig. 69.1. Then we use our proposed method to detect a spliced image, and the detection results are shown in Fig. 69.2.

In order to learn more knowledge from images, we prepare large number of diverse images to train. The UCID [9] image database which is popular in image forensics contains 1,338 images of size  $512 \times 384$ . All training samples in our experiment are shown by 8-bit gray image. We first convert all images into 8-bit gray images; then crop  $128 \times 128$  and  $64 \times 64$  blocks from each image, and then rescale blocks using bilinear interpolation method (scaling factor is randomly set to 0.95 or 1.05). Four groups of images are obtained: 1,338  $128 \times 128$  original images; 1,338  $128 \times 128$  rescaled images; 1,338  $64 \times 64$  original images, and 1,338  $64 \times 64$  rescaled images. First, we randomly select 338 original images and 338 rescaled images to test our proposed method's efficacy for each size of image. The training set is as follows: 1,000 original images + 1,000 corresponding rescaled images for each size. In general, more training samples and higher dimensionality of feature will obtain higher detection accuracy but will result in



**Fig. 69.1** The detection results shown by ROC curves for  $128 \times 128$  images and  $64 \times 64$  images. The testing samples are 676 images which consist of 338 original images and 338 rescaled images for each size



**Fig. 69.2** The detection results for first forged image. Panel (a) is the original image, and the car in (b) is cut and pasted into (a) to form a forged image shown in (c). Panel (d) shows the detection results of  $64 \times 64$  detection blocks. Panel (e) shows the detection results of  $128 \times 128$  detection blocks. The *square boxes* are to indicate forged part

more computing time. In order to balance the computing time and detection accuracy, the dimensionality is set to 343 and the amount of training sample is 1,000. We then compute the 343-D feature for each image in the above four groups. In order to get the best parameter  $C$  and  $r$ , fivefold cross-validations are executed on the grid:  $\{(C,r) \in (2^i, 2^j), i, j \in \mathbb{Z}\}$ . The ROC curves are used to assess our proposed method's performance. The accuracy as follows is used to show our results at scale:

$$\text{Acc} = \frac{|\text{correctly predicted data}|}{|\text{total testing data}|} \times 100\% \quad (69.8)$$

Figure 69.1 shows that our proposed method obtains high detection accuracy for both kinds of images. Even for small  $64 \times 64$  images, our proposed method

detection accuracy is as high as 89.8 %. When the false-positive rate is low ( $<10\%$ ), the positive rate of our method is also acceptable. In previous work [2, 10, 11], they also got high detection accuracy, but the size of tested image is larger than our tested image. If the scaling factor is larger than 1.05, better results will be obtained. The same results will be held for other two interpolation methods (nearest interpolation and bi-cubic interpolation). Figure 69.1 indicates that our proposed method will be useful when detecting parts of image are resized.

## 69.4 Image Splicing Detection

In this experiment, we also first prepare training samples to train for classifiers. In order to test the influence of the size of block, we get two classifiers on different size of image. One classifier is obtained on  $128 \times 128$  images; the other classifier is obtained on  $64 \times 64$  images. The training samples are also from UCID database. The training set is composed by 1,336 original images and 1,336 corresponding rescaled images for each size. After computing the 343-D feature of trained images, the training feature set is a  $2,672 \times 343$  matrix. We label 1,338 original block as 0 and 1,338 rescaled images as 1 for each size. The setup of SVM is the same as first experiment. The SVM training will implement on the above feature set and get the predicted model.

After getting the model, we will use the classifier to detect the spliced image. In order to get a convincing result, we first introduce how to create the forged image in our experiment. The two original pictures in Fig. 69.2 come from the UCID database. These two images are not contained in the training samples, so the number of training samples is  $2 \times 1,338 - 2 = 2672$ . The original images in Fig. 69.2a, b are  $512 \times 384$  in TIFF format. The Adobe Photoshop CS 2 is employed as editorial tool. When the car is pasted into the first picture, resizing operation must be employed to lessen the size of car for a convinced forged image. The forged image saved as TIFF format showed in Fig. 69.2c is the same size as original image. It is hard to distinguish whether the image in Fig. 69.2c is forged or not by human's eyes. When using detecting blocks with size  $64 \times 64$ , there are totally 48 blocks that need to be tested. We first compute the feature for each block. Then we use the classifier obtained on training samples of size  $64 \times 64$  to predict these 48 blocks. The same procedure is executed for detection block of size  $128 \times 128$ .

Figure 69.2d shows that our proposed method locates the forged part with high accuracy; just only a part of car is not covered by  $64 \times 64$  block. Though the marked box is square, the people can easily take the car as forged part by empirical experience. For the  $128 \times 128$  block, Fig. 69.2e shows that it also obtains good results, but the detection results are roughly caused by the larger detection block. If the tested image is of large size and the spliced object is also large, larger detection block is preferable. Though the accuracy of classifier obtained on small-size training samples may not be very high, it is also the best choice to detect image splicing.

## 69.5 Conclusion

In this chapter, machine learning method was employed to detect image splicing. Most of the image splicing operations will resize the pasted object for a better tampering, which will bring re-sampling artifacts in the forged image. The modified SPAM feature is used to capture the traces left by re-sampling. The feature of training blocks is fed into SVM for a classifier. Block-based method was used to test which parts of the image are forged by the classifier. Experiments showed that the proposed method can detect forged parts of image efficiently. Making a robust method (such as against JPEG compression) deserves more attention in the future.

## References

1. Farid, H. (2009). Image forgery detection. *IEEE Signal Processing Magazine*, 5(3), 16–25.
2. Farid, H. (1999). Detecting digital forgeries using bi-spectral analysis. AI Lab, Massachusetts Institute of Technology, Tech. Rep. AIM-1657
3. Ng, T. T., & Chang, S.-F. (2004). A model for image splicing. *Proceedings of IEEE International Conference on Image Processing. Singapore* (Vol. 2, pp. 1169–1172).
4. Popescu, A. C., & Farid, H. (2005). Exposing digital forgeries by detecting traces of re-sampling. *IEEE Transactions on Signal Processing*, 53(2), 758–767.
5. Mahdian, B., & Saic, S. (2008). Blind authentication using periodic properties of interpolation. *IEEE Transactions on Information Forensics and Security*, 3(3), 529–538.
6. Andrew C. Gallagher. (2005). Detection of linear and cubic interpolation in JPEG compressed images, *Proceedings of the Second Canadian Conference on Computer and Robot Vision*. Vol. 33, pp. 65–72.
7. Shi, Y. Q., Chen, C. H., & Chen, W. (2007). A natural image model approach to splicing detection, MM and Sec'07, *Proceedings of the Multimedia and Security Workshop, Dallas* (pp. 51–62).
8. Kirchner, M., & Fridrich, J. (2010). On detection of median filtering in digital images, *Proceedings of SPIE, San Jose USA* (Vol. 7541, pp. 300–308).
9. Cristianini, N., & Taylor, J. S. (2000). *An introduction to support vector machines and other kernel-based learning methods* (pp. 230–250). London: Cambridge University Press.
10. Chang, C.-C., & Lin, C.-J. (2011). LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, 5(7), 230–235.
11. Schaefer, G., & Stich, M. (2004). UCID-An uncompressed color image database, *Proceedings of SPIE, Storage and Retrieval Methods and Applications for Multimedia, San Jose, USA* (pp. 472–480).

# Chapter 70

## A Lane Detection Algorithm Based on Hyperbola Model

Chaobo Chen, Bofeng Zhang, and Song Gao

**Abstract** In order to improve the problem of recognition rate and inaccurate in the curve, this paper proposed a lane detection algorithm based on hyperbola model, which uses Canny operator to detect the edge of the lane and wields the Hough transform to extract lane boundary points, and utilizes extended Kalman filter to reduce road scanning range. By fitting points on pair road boundaries into the hyperbola model, and completes the lane boundary reconstruction. Some experimental studies are conducted, and the results show that the accuracy of the algorithm has reached 93.4 % and the processing speed of each image needs 77.4 ms. Our method is able to make full use of lane boundaries with existence partial occlusion, blur and low contrast. Meanwhile, it can quickly and accurately identify lane line, and it has high performance and robustness.

### 70.1 Introduction

Vision-based intelligent navigation system is an important application field of computer vision [1]. At the present stage, machine vision and autonomous navigation often is used in intelligent vehicle road recognition [2]. The method is simple and practical and can adapt to the complex road environment [3]. Currently, road line detection algorithm consists of feature based and model based [4]. Feature-based algorithm mainly identifies some characteristics of road, such as color [5], texture, and shape. However, illumination changes, water stains, shadows, damaged and discontinuous road markings will influence detection effect [6]. Model-based algorithm primarily establishes road model and then according to image analysis determines model parameters; and model contains all of the information of lane line. The approach mainly differs in models, such as straight model, parabolic

---

C. Chen • B. Zhang • S. Gao (✉)  
Department of Electronic and Information Engineering, Xi'an Technological University,  
Xi'an 710021, China  
e-mail: [gaosong@xatu.com.cn](mailto:gaosong@xatu.com.cn)

model, and spline curve model [7, 8]. There are a lot of lane detection algorithms that use straight model, parabolic model, spline curve model, etc. The analysis of the performance of the algorithms indicates that the straight model is fast, while recognition accuracy of the curve is weak; Parabolic model is weak that steady of straight and curve lines and markings is able to deviate; spline curve model is too complex, slow and takes large computational resources. To solve the problems mentioned, we have proposed a lane detection algorithm based on hyperbolic model. Firstly, the algorithm uses Canny operator to detect the edge of the lane and Hough transform to extract lane boundary points, and utilizes extended Kalman filter predictive tracking algorithm to reduce road scanning range. By the way, it can improve the accuracy of extraction. Eventually, by the left and right lane boundary parameters match with the hyperbola model, meanwhile uses the least squares method to solve the model parameters, and completes the lane boundary reconstruction. It can overcome parabolic model in curve and straight lanes of joint discontinuous problem, and experiment in many different conditions, includes various weather and road. It has high performance and accuracy.

### 70.2 Road Model

The paper uses lane line combination model to build model. Near field applies to the linear model is  $u = bv - a$ , and far area applies to hyperbola model is  $u - u_H = \frac{k}{v-l}$ . As shown in Fig. 70.1. We establish a hyperbolic model on both sides of the road:

$$u - u_H = \frac{k}{v - l} + b(v - l) \tag{70.1}$$

where  $l$  represents position of the vanishing line of the road on the image plane,  $k$ ,  $b$  and  $u_H$  of which are hyperbolic model parameters,  $k$  represents the curvature of lane,  $b$  represents relative direction of lane,  $u_H$  represents the distance which is between the lane line and the vertical axis. These parameters may be get that they have been calculated the road image data with internal and external parameters of

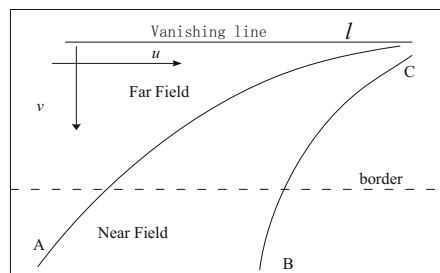


Fig. 70.1 Road model

the camera. One side of lane may be blocked by other vehicles, and leading to CCD camera can't be collected image information. Therefore, the paper proposes that use a pair of hyperbolic model, but we assume that the marking line is parallel and equal width. So the parameters can be solved with equation family on both sides of road line, the left and right road line equations are as follows:

$$u - u_H = \frac{k}{v - l} + b_l(v - l) \quad (70.2)$$

$$u - u_H = \frac{k}{v - l} + b_r(v - l) \quad (70.3)$$

As the boundary line on both sides of road is parallel, boundary parameters  $k$  and  $b$  value are the same, but  $u_H$  is different. The lane model can be formulated by an extended equation:

$$u - u_H = \frac{k}{v - l} + b_l(v - l) + b_r(v - l) \quad (70.4)$$

$b(l) = 0$ , it is on the right lane line and  $b(r) = 0$  if it is on the left one. This chapter combines both sides of the curve into a pair of hyperbola, and that curve parameters add only one. It is less increased calculation that it greatly improves the accuracy and robustness of the curve model, especially when the lane line is blocked, or partially damaged.

### 70.3 Lane Detection Algorithm Method

Vision-based lane line detection establishes the road model which is commonly used method for mostly scholars. The algorithm described in this paper is quite unique as it uses a combination of scan lane lines and improved Hough transform to match a hyperbola model. The combinatorial hyperbola model is rarely used, while most of the lane detection is conducted by template-based models. First, the algorithm converts the image to a grayscale. Due to the presence of noise in the image, so we apply the Bilateral Filtering algorithm. After this, the edge detector is used to produce an image edge by Canny operator, the image has detected which produce the left and right lane boundary. Next, the lane boundary scan uses the information in the edge image detected by Hough transform to perform the scan and obtain a series of points on the left and right lane lines, and using extended Kalman filter predictive tracking algorithm to reduce road scanning range. Finally, combinatorial hyperbola model is fitted to the date points and completed the lane boundary reconstruction (Fig. 70.2).

Fig. 70.2 Algorithm structure

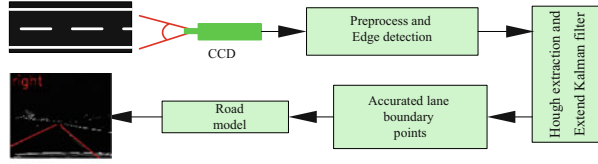
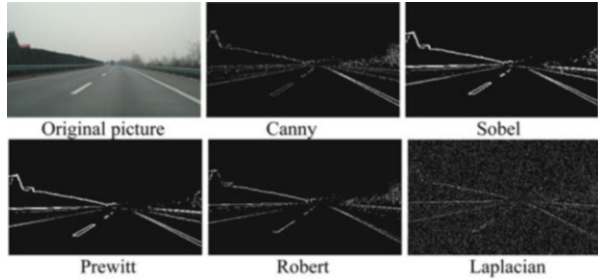


Fig. 70.3 Operator contrast figure



### 70.3.1 Lane Edge Extraction

Autonomous unmanned vehicle systems acquire pictures which are containing noise, but, presence of noise in our system will affect edge detection, so noise removal is very important, Bilateral Filtering is used to achieve edge-preserving de-noising. Now existing edge detection operator: Canny, Sobel, Robert, Prewitt, Laplacian, etc. Among Canny operator is a filter, enhancement and detection of multistage optimization operator. Detecting edges more complete and better positioning performance. Considering the performance and speed characteristics of the operator, the algorithm using Canny edge detection in this paper. Lane boundaries are defined by sharp contrast between the road surface and painted lines obviously form edges on the image. Thus, Canny edge detector was employed in determining the location of lane boundaries. Operator compare shown in Fig. 70.3.

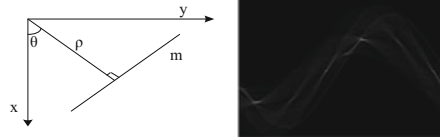
### 70.3.2 Lane Line Extraction

Lane extraction is used a standard Hough transform by a restricted search space. Its character is accurate extraction and high anti-interference ability. First, Hough transform obtain roadway centerline and vanishing line, then extracting road on both sides of the boundary line with real-time road of the same width. Hough transforms searches for lines using the equation shown in Fig. 70.4.

$$\rho = x \cos \theta + y \sin \theta \tag{70.5}$$



Fig. 70.4 Hough transform



Formula  $\rho$  represent from straight line  $m$  to the origin distance,  $\rho = \pm\sqrt{x^2 + y^2}$ . Normal line and  $x$ -axis angle  $\theta$ , it ranges  $-90^\circ \leq \theta \leq 90^\circ$ . In reality we reject any line that falls outside a certain region, for example a horizontal line is probably not lane boundary and can be rejected. The Hough transform was modified to limit the search space to  $30^\circ$  for each side. Also the input image is searched separately and returning the most dominant line in the half image that falls within the  $30^\circ$  window. The horizontal line at this intersection is referred to as the vanishing line.

### 70.3.3 Extended Kalman Filter

Lane line parameters  $\rho$  and  $\theta$  were obtained by the Hough transform, as we use hyperbolic model and extended Kalman filter. Due to the lane line is different, dimension of  $\rho$  and  $\theta$  is different. However, taking advantage of extended Kalman filter to track boundary point, the dimension of the parameter must be unified. These are the extended Kalman filter equations for a nonlinear measurement equation, the EKF will provide us with a feedback  $K$ .

$$x_{t+1} = Ax_t + Bu_t + w_t, \quad y_t = h(x_t) + e_t \tag{70.6}$$

$$[D_x h]_{ij} = \frac{\partial h_i}{\partial x_j}, \quad C_t = X_t D_x h \left( \hat{x}_t |_{t-1} \right) \tag{70.7}$$

by the observing data substitute into the formulas (70.6) and (70.7) to solve  $C_t$  matrix, then by means of the formulas (70.8) and (70.9) to get gain  $K_t$  and error  $P_t$ , finally, formula (70.10) output current estimates  $\hat{x}_t$ .

$$K_t = P_{t-1} C_t^T (C_t P_{t-1} C_t^T + Y_t R Y_t^T)^{-1} \tag{70.8}$$

$$P_t = A P_{t-1} A^T + Q - A K_t C_t P_{t-1} A^T \tag{70.9}$$

$$\hat{x}_{t+1|t} = A \left( \hat{x}_{t|t-1} + K_t [Y_t y_t - X_t h(\hat{x}_t |_{t-1})] \right) + B u_t \tag{70.10}$$

The predicted value of current state is the tracking results of previous state. A true value of the present state is the measured value. We can obtain track value of the state. The value is the predicted value of the next state, so the lane parameters

are circularly estimated. Usually, by the first three images road parameters average value is the initial value of forecast. We use lane tracking algorithm that in order to reduce the scope of lane line scan.

### 70.3.4 Fitting Hyperbola Model

We use the mid-to-side road scan method to obtain a point of the left and right sides of the two time series, the hyperbola fitting phase uses the vectors of points from the lane. We could determine the vanishing point of the lane marking in the image space and obtain the parameter  $l$  of the improved hyperbola model. Left and right lane line boundary points are  $L_l = \{(u_1, v_1), (u_2, v_2), \dots, (u_n, v_n)\}$  and  $L_r = \{(u_1, v_1), (u_2, v_2), \dots, (u_m, v_m)\}$ , respectively. Then a least squares technique is used to fit hyperbola model. The left and right parameters of lane model and working out the parameters of the road model:

$$N_1 = \begin{bmatrix} \frac{1}{v_1^r - l} & 1 & 0 & v_1^r - l \\ \vdots & \vdots & \vdots & \vdots \\ \frac{1}{v_n^r - l} & 1 & 0 & v_n^r - l \end{bmatrix}$$

$$N_2 = \begin{bmatrix} \frac{1}{v_1^l - l} & 1 & v_1^l - l & 0 \\ \vdots & \vdots & \vdots & \vdots \\ \frac{1}{v_m^l - l} & 1 & v_m^l - l & 0 \end{bmatrix}$$

where

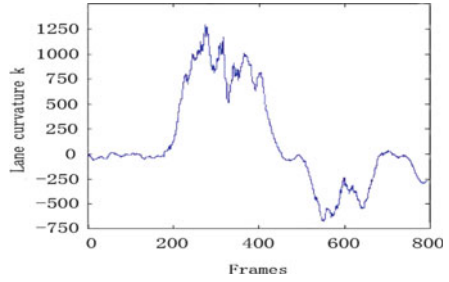
$$A = \begin{bmatrix} N_1 \\ N_2 \end{bmatrix}, \bar{X} = [k, u_H, b_l, b_r]^T, B = [u_1^{(r)}, u_2^{(r)}, \dots, u_n^{(r)}, u_1^{(l)}, u_2^{(l)}, \dots, u_m^{(l)}]^T$$

So solving matrix equation  $A\bar{X} = B$ , and calculating the model parameters  $k$  and  $b$ .

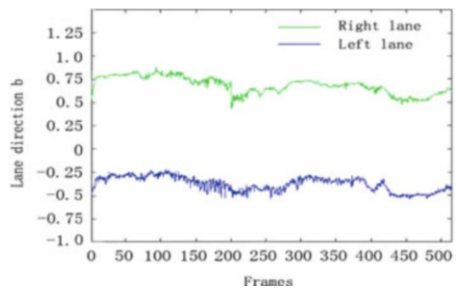
$$\begin{bmatrix} N_1 \\ N_2 \end{bmatrix} [k, u_H, b_l, b_r]^T = [u_1^{(r)}, u_2^{(r)}, L, u_n^{(r)}, u_1^{(l)}, u_2^{(l)}, L, u_m^{(l)}]^T$$

$k = 0$  means that straight road,  $k > 0$  means the road ahead turn left, and  $k < 0$  means the road ahead turn right. Amplitude indicates that the degree of bend of the left and right lane line;  $b$  represents the traveling direction of the lane. Simulation results demonstrate that hyperbolic model parameters provide curvature of the road ahead to unmanned vehicle systems and to judge bodywork turn left or right. It plays a vital role in the model matching and lane line fitting (Figs. 70.5 and 70.6).

**Fig. 70.5** Lane curvature simulation figure



**Fig. 70.6** Lane direction simulation figure

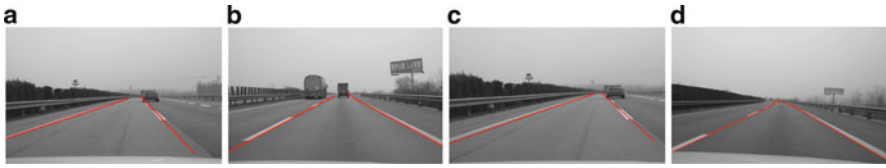


## 70.4 Experiment Result

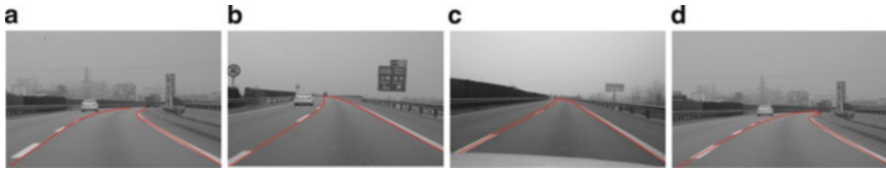
This lane detection algorithm has been simulated and tested by VC++ on real road images, in order to verify effectiveness and stability of the algorithm. These lane images include straight and curve road, with or without shadows and lane marks. During the experiments, the images are  $320 \times 240$  pixels with colors. The image processing speed needs 77.4 ms, accuracy reaches 93.4 %. We have tested at different locations and times. Partial test results are shown in Figs. 70.7 and 70.8. Partial lane is blocked by neighboring vehicles in Fig. 70.7, but the algorithm can extract the lane marking and it achieves a good result. Some curved road is tested in Fig. 70.8, but the algorithm can accurately orientate boundary line and curve position. The testing results indicate that the algorithm is good stability, high recognition rate, and robustness.

## 70.5 Conclusion

In this paper, we propose a lane detection algorithm based on hyperbola model. The algorithm extracts road markings, matches hyperbola model, and reconstructs the road markings by the least squares method. It can deal with the occlusion and imperfect road condition. We choose curve and straight road, with or without shadows lane condition. The experimental results indicate that it is able to



**Fig. 70.7** Straight-line test results



**Fig. 70.8** Curve test results

accurately and reliably recognize the boundary markings of the lanes in complex environment.

**Acknowledgements** This work was supported by the projection of National Natural Science Foundation of China (61271362) and Shaanxi Province Natural Science Foundation (12JK0502). We would like to thank them.

## References

1. Khalifa, O. O., Khan, I. M., & Assidiq, A. A. M. (2010). A hyperbola-pair based lane detection system for vehicle guidance. *Proceedings of the World Congress on Engineering Computer Science* (pp. 20–22). WCECS2010, USA
2. Yu, B., & Jain, A. K. (2010). Lane boundary detection using a multiresolution Hough transform. *Proceedings of International Conference on Image Processing* (pp. 748–751). IEEE Computer Society, USA
3. Cheng, H. Y., Jeng, B. S., et al. (2006). Lane detection with moving vehicles in the traffic scene. *IEEE Transactions on Intelligent Transportation Systems*, 7(4), 571–582.
4. Gao, S., Qin, L., & Chen, C. B. (2013). Structured road lane marking identification algorithm. *Journal of Xi'an Technological University*, 33(1), 14–19.
5. Li, Q., & Zheng, N. N. (2004). Spring robot: A prototype autonomous vehicle and its algorithms for lane detection. *IEEE Transactions on Intelligent Transportation Systems*, 5(4), 300–308.
6. Sharma, U. K., & Davis, L. S. (2009). Road boundary detection in range imagery for an autonomous robot. *IEEE Transactions on Robotics and Automation*, 4(5), 515–523.
7. Lutzeler, M., & Dickmanns, E. D. (2009). Recognition of intersections on unmarked road networks. *Proceedings of the IEEE Intelligent Vehicles Symposium* (pp. 302–307). Washington, DC: IEEE Computer Society
8. Wang, Y. (2004). Lane detection using spline model. *Pattern Recognition Letters*, 21(3), 677–689.

# Chapter 71

## Comparisons and Analyses of Image Softproofing Under Different Profile Rendering Intents

Qingxue Yu, Yunhui Luo, Maohai Lin, and Quantao Liu

**Abstract** This paper presents some results from an experiment of image softproofing under different International Color Consortium (ICC) rendering intents, which will be available for selecting an appropriate rendering intent in printing processes. In a screen softproofing procedure, image is converted via ICC profile embedded in image itself to output device profile, then to proofing device profile. With the rendering intent of absolute colorimetry in the second conversion, the effects of different rendering intents in the first conversion have been investigated through a softproofing software developed in Matlab 7.0. A variety of testing images, including light tone, shadow detail, etc., are used for image softproofing. The color differences between original images and proofed images are calculated under the S-CIE  $L^*a^*b^*$  color difference formulae. Comparisons and analyses on the obtained images under four rendering intents show the effects of color characteristics of original images on rendering intent selection.

### 71.1 Introduction

Faithful color reproduction is the aim of printing industry. Due to color information transferring among different devices, color consistency is the vital key to high-quality prints [1]. Therefore color management is necessary for printing reproduction. International Color Consortium (ICC) profile-based color management system (CMS) is the primary tool for color control in printing industry [2]. In an ICC

---

Q. Yu • Y. Luo (✉) • M. Lin

Key Lab of Pulp & Paper Science and Technology, Ministry of Education,  
Qilu University of Technology, Jinan 250353, China  
e-mail: [lyh@spu.edu.cn](mailto:lyh@spu.edu.cn)

Q. Liu

Shandong Dazhong-huatai Printing Ltd, Jinan 250000, China

profile, there are four modes of rendering intent, i.e., perceptual, saturation, relative colorimetry, and absolute colorimetry, for color conversion between different devices. The four rendering intents correspond to different color mapping algorithms of gamut compression or reduction, which are suited for different cases [3, 4]. Appropriate rendering intents can result in better printing performance. Theoretically, the selection of rendering intents is relevant to source image characteristics, output device properties, and printing purposes. But in practice, it is usually determined by fuzzy experiences of color management technicians. Therefore, automatic selection for rendering intents is an attractive research topic in printing production.

There are some attempts for automatically selecting rendering intents in recent years [5–8]. Based on the state information and a current task in an imaging workflow, from a group of rendering intent including perceptual, saturated, relative colorimetric, and absolute colorimetric, automatically converting an original image to a new color space based on the color profile and the selected rendering intent, and making the converted image available for processing and output. This is a general procedure of automatic selection for rendering intents [5, 6]. At present, the digital proofing technology has increasingly become mature and been of great development [8]. It utilizes the color of displays to simulate the real output color of prints, and with the advantage of fast proofing speed, high convenience, and low cost, especially real-time proofing in different locations. In order to meet demand of printing production, automatically selecting rendering intents for image softproofing will affect efficiency of the color reproduction.

This work investigated the selection of four ICC rendering intents through a softproofing experiment. An experimental methodology and a Matlab software are developed for converting image color under different rendering intents. By real-time softproofing, color gamut of original image, printing device, and softproofing device are all presented in a visualized fashion. Color difference with S-CIE  $L^*a^*b^*$  space [9], which combines with space attributes of human vision system, are introduced to evaluate color reproduction performance. The experimental results show the effects of color characteristics of original image and output device on rendering intent selection.

## 71.2 Preliminaries

*Color Conversion of Softproofing.* During the softproofing procedure, two color conversions are involved, connecting with twice selections of rendering intents, as shown in Fig. 71.1. The two types of profile used in the first step are the profile embedded in image itself (e.g., the sRGB profile) and the profile of printing devices (e.g., an OKI 9800 printer profile). In the second step the profiles for printing output device and softproofing device (e.g., a display profile) are used. Since softproofing is to employ colors of a display screen to simulate real printing colors, the color range of a proofing equipment, i.e., color gamut, should usually be larger than that



**Fig. 71.1** Color conversion of softproofing

of the printing device in order not to loss color information. It means that the color gamut described by the profile of softproofing device should be bigger than that of source image so as to realize the purpose of proofing. Therefore, the rendering intent of absolute colorimetry should be utilized for softproofing in the second conversion, and the selection of rendering intent is just needed and available in the first color conversion.

*Calibration and Characterization.* Calibration is the premise of characterization and the key step of color management [10]. For a display used for softproofing, a typical calibration process is as follows: (1) Switch on the power of display to adjust lightness, contrast and RGB which conform to the conditions of calibration, and keep open more than 2 h to ensure relatively stable color gamut. (2) Individually calibrate display contrast, lightness, white balance, tone reproduction, cooperating with an Eye-One Pro spectrophotometer and the software of ProfileMaker 5.0. (3) Implement the sequence of calibration according to the prompts of the ProfileMaker software. Establishing profile of a device (by the ProfileMaker software) is called as characterization. Since each color input device or output device, and even color material, such as printing ink, dyeing phosphorus of display screens, etc., can only exhibit a specific color range (color gamut), the purpose of characterization is to establish color gamut for these devices or materials, and record its mathematical characteristics so as to carry on proper color conversions [1, 10, 11]. Color management is achieved by a universal and device-independent Profile Connection Space (PCS) of color. The device profile bridges its device color space and the PCS.

*Color Difference.* The color difference between original and proofed images is a traditional evaluation method for color reproduction performance [8]. Color difference formula of 1976 CIE  $L^*a^*b^*$  color space has a wide range of applications in the printing and imaging science field. But the calculation of color difference does not fully match the actual feeling of human vision. For example, when watching a halftone image at a distance, the effect felt by human vision is very close to the actual effect of continuous tone image. But when we calculate the color difference between halftone image and continuous tone image with the CIE  $L^*a^*b^*$  color difference formula, the values for these two cases are very different. Aiming to measure errors of image color reproduction, Zhang and Wandell proposed S-CIE  $L^*a^*b^*$  (Spatial-CIE  $L^*a^*b^*$ ) color space [9]. S-CIE  $L^*a^*b^*$  color space adds a space preprocessing step on the basis of CIE  $L^*a^*b^*$  color space, and it is suited for image quality assessments. S-CIE  $L^*a^*b^*$  color space combines the traditional color difference formula with space attributes of human visual system.

### 71.3 Experimental Setup

#### 71.3.1 Testing Images Selection

Testing images cover common types encountered in practical applications. Eight testing images are selected from ISO (the international standard organization), GATF (Graphic arts technology foundation of the United States), and Kodak Company, as shown in Fig. 71.2. These images are embedded with the sRGB profile.

*Shadow tone image.* The dominant tone of such images is shadow tone, including the red, green, blue and a variety of perfect woodiness grain tone.

*Light tone image.* As Image 3 in Fig. 71.2 illustrated, its dominant tone is soft light and white.



**Fig. 71.2** Testing images (1) shadow tone (GATF) (2) shadow tone (ISO) (3) light tone (GATF) (4) Group portrait (GATF) (5) neutral gray (GATF) (6) full tone (ISO) (7) full tone (Kodak) (8) memory color (GATF)



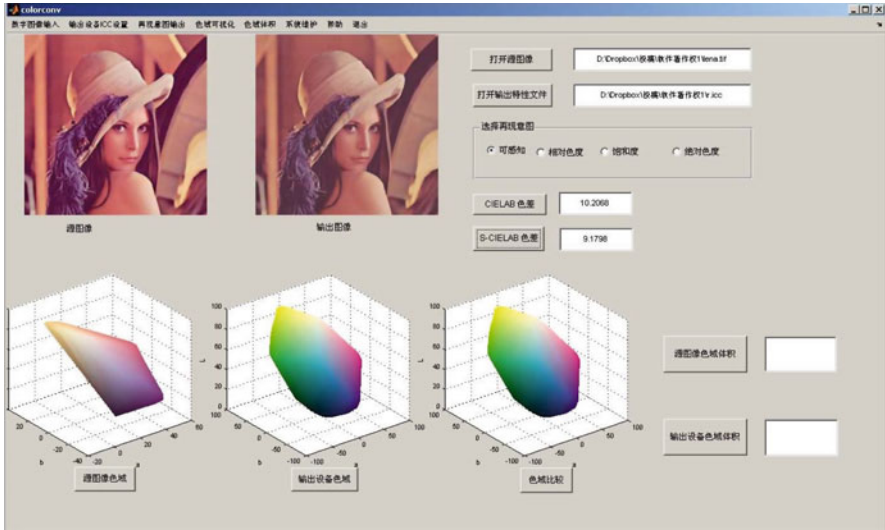


Fig. 71.3 GUI of the softproofing software

*Group portrait image.* In group portrait pictures, human skin color should be received much attention.

*Neutral gray image.* In neutral gray images, there still have some warm or cold gray tone besides neutral gray color. The saturated color is a key tone.

*Full tonal image.* An original full tonal image is usually taken by camera in a close range, as shown in Images 6 and 7 of Fig. 71.2. Both pictures have less object elements, providing a wide range of color and tone.

*Memory color image.* Image 8 of Fig. 71.2 shows an outstanding outdoor scenery of natural light color. It can arouse human reactions of memory color. Memory colors are processed by the human visual system and stored as memories. When readers see such images, they will first judge whether or not image colors are credible.

### 71.3.2 Softproofing

After device calibration and characterization, the color conversion will be implemented by a Matlab software for selected testing images. The specific output device profile is chosen to realize color conversions to simulate the actual output effects in a display, which is just so-called softproofing. The graphic user interface (GUI) of the proofing software developed in Matlab 7.0 is shown as in Fig. 71.3.

Three output profiles are used for testing experiments, which are an OKI 9800 digital printer profile, the USWebCoatedSWOP profile, and the JapanWebCoated

profile. The developed software can provide proofed images under the four rendering intents. And it also presents the functions of reading and showing images, selecting ICC profiles, showing color gamut of original image and output device, calculating color gamut volumes, selecting rendering intents, as well as computing color differences between original and proofed images.

## 71.4 Results and Discussion

Table 71.1 gives the calculation results of color difference, as illustrated in Fig. 71.4, where  $\Delta E$  is color difference under CIE  $L^*a^*b^*$  formula, and  $\Delta E_s$  under S-CIE  $L^*a^*b^*$  formula.

From results in Table 71.1, we can see that two kinds of color difference values ( $\Delta E$ ) of absolute colorimetry are very small in the four rendering intents for all of

**Table 71.1** Color difference under different rendering intents

I	P	Pc		RC		S		AC	
		$\Delta E$	$\Delta E_s$	$\Delta E$	$\Delta E_s$	$\Delta E$	$\Delta E_s$	$\Delta E$	$\Delta E_s$
1	O	9.1824	8.4582	6.1809	5.6173	9.1824	8.4582	5.8017	5.1207
	U	11.2083	10.0711	9.98	8.9668	10.8589	9.7939	6.404	5.6574
	J	11.0016	9.9907	8.8864	8.8864	10.5609	9.6557	6.2593	5.5461
2	O	7.7236	7.226	5.0603	4.7205	7.7236	7.226	3.8561	3.471
	U	10.3029	9.3176	9.5364	8.6129	10.1458	9.1847	4.7472	4.257
	J	9.7448	8.935	8.4715	7.7629	9.5454	8.7685	4.5876	4.1535
3	O	9.2044	8.3108	9.4993	8.7942	9.2044	8.3108	4.5012	4.2771
	U	9.7315	8.4912	19.8834	17.7913	19.6872	17.5998	12.6957	11.0366
	J	16.6087	14.9206	16.622	14.9852	16.5207	14.8616	9.7315	8.4912
4	O	8.022	7.1903	7.1535	6.623	8.022	7.1903	4.4031	3.8995
	U	12.6957	12.4205	13.5536	12.2659	13.5833	12.2181	6.1182	5.3619
	J	12.3009	11.0637	11.5951	10.5815	11.9453	10.7998	5.4399	4.7874
5	O	8.9893	8.7623	7.4581	7.2681	8.9893	8.7623	5.9545	5.8116
	U	12.0784	11.2882	11.5081	10.724	12.0716	11.2839	6.2993	6.1302
	J	11.7767	11.2005	11.077	10.5022	11.7513	11.1824	6.7447	6.6214
6	O	8.2202	7.4923	5.9906	5.4734	8.2202	7.4923	5.1105	4.3871
	U	11.7296	10.5297	10.635	9.5469	11.2639	10.1575	5.7357	4.9926
	J	11.0877	10.0266	9.0022	8.1818	10.463	9.5468	5.5072	4.7885
7	O	9.7536	8.2568	8.4941	7.3747	9.7536	8.2568	5.7435	4.7157
	U	14.9363	13.4799	14.614	13.143	14.7822	13.3763	10.3817	9.4473
	J	14.3781	13.0953	13.609	12.3419	14.1692	12.9617	10.5937	9.6664
8	O	10.0666	9.0852	9.0404	8.3656	10.0666	9.0852	6.1823	5.6012
	U	17.4486	15.8903	16.816	15.3668	16.8663	15.4044	13.8966	12.3266
	J	15.6582	14.2787	14.631	13.4386	15.017	13.7611	12.0472	10.7394

*I* Image, *P* Profile, *Pc* Perceptual, *RC* Relative Colorimetry, *S* Saturation, *AC* Absolute Colorimetry, *O* OKI 9800, *U* USWebCoatedSWOP, *J* JapanWebCoated

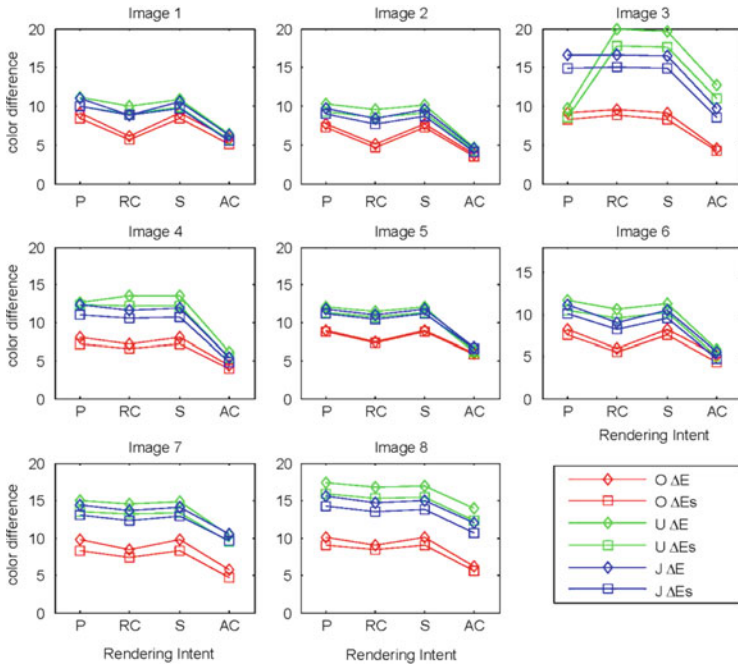


Fig. 71.4 Color difference under different rendering intents

the selected test images. Combined with human visual observation, the conversion effect of absolute colorimetry rendering intent is also the best. All above can prove that the absolute colorimetry rendering intent is suitable for image softproofing.

According to Table 71.1 and Fig. 71.4, the results with perceptual and saturation rendering intents are similar, and these two kinds of color difference values are all somewhat larger; Moreover, color difference values under relative colorimetry and absolute colorimetry rendering intents are somewhat smaller. These can illustrate that absolute colorimetry rendering intent is the best choice when color gamuts of source image and output device are nearly approached.

In addition, for a shadow tone image, the difference of  $\Delta E$  or  $\Delta E_s$  value between relative colorimetry and absolute colorimetry rendering intents is not very larger. For a light tone image, the  $\Delta E_s$  value among perceptual, relative colorimetry and saturation rendering intents have a little difference. The results of group portrait images, neutral gray images and memory color images are as the similar. For a full tonal images, the  $\Delta E_s$  value of four rendering intents have a large difference, which means the proper selection is the relative colorimetry or absolute colorimetry rendering intent.

## 71.5 Conclusion

This paper presented results from an experiment of image softproofing under different ICC rendering intents. Through a softproofing software developed in Matlab 7.0, a variety of testing images were used for color difference analysis under CIE L\*a\*b\* and S-CIE L\*a\*b\* formulae. Comparisons and analyses provided useful guidelines for selecting rendering intent of output device profiles in printing processes. Automatic selection for rendering intents is currently under study and progress will be reported in the future.

## References

1. Homann, J. P. (2008). *Digital color management: Principles and strategies for the standardized print production* (pp. 10–11). London: Springer. 35–36.
2. Fraser, B., Murphy, C., & Bunting, F. (2005). *Real world color management* (pp. 70–71). Berkeley: Peachpit Press.
3. Morovic, J. (2008). *Color Gamut mapping* (pp. 223–224). London: Wiley.
4. Bakke, A. M., Farup, I., & Hardeberg, J. Y. (2010). Evaluation of algorithms for the determination of color Gamut boundaries. *Journal of Imaging Science and Technology*, 54(5), 2–11.
5. Kulkarni, M. S., Borg, L. U. (2007, May 21). Automatic selection of color conversion method using image state information: US, 8014027.
6. Cai, S. Y., & Liu, R. F. (2004). An idea of automatically selecting rendering intents in color management. *Tian University of Science and Technology*, 19(3), 65–67.
7. Intwala, C., Clara, S. (2009, September 29). Color conversion preserving global and local image characteristics: US, 7965301.
8. Green, P. (2009). *Color management: Understanding and using ICC profiles* (pp. 55–57). London: Wiley. 87–88, 103–104.
9. Zhang, X. M., & Wandell, B. A. (1997). A spatial extension of CIELAB for digital color-image reproduction. *Journal of the Society for Information Display*, 5(1), 61–63.
10. Adams, R. M., & Weisberg, J. B. (2000). *The GATF practical guide to color management* (pp. 44–45). Pittsburgh: GATF Press. 123–125.
11. Sharma, A. (2006). Methodology for evaluating the quality of ICC profiles-scanner, monitor, and printer. *Journal of Imaging Science and Technology*, 50(5), 469–480.

# Chapter 72

## An Improved Dense Matching Algorithm for Face Based on Region Growing

Xin Xia and Shaoyan Gai

**Abstract** Traditional dense matching algorithms for face based on region growing have a lot of flaws. To generate a better disparity map, a novel improved method is proposed in this paper. Firstly, scale invariant feature (SIFT) algorithm is adopted to detect feature points for a pair of images, which are taken from two different angles. Secondly, this paper uses normalized cross correlation (NCC) to get match points and uses random sampling consensus (RANSAC) algorithm to eliminate mismatches. Several robust seeds are generated after this step. At last, by using an improved strategy of region growing, in which seeds are evaluated to help determine the locations and sizes of the search windows dynamically, the matching relations of seeds propagate to other parts of images. Experiments show that this method can obtain a good disparity map and has high computation speed.

### 72.1 Introduction

Three-dimensional information of face is widely used in the area of three-dimensional animations, face recognition, and so on [1]. Passive reconstruction systems for 3D information of face draw a wide attention for its flexible implementation and simple configuration [2]. This method reconstructs the 3D information of the face by using a pair of images taken from two different angles. This paper focuses on the stereo matching algorithm, which is the core of the reconstruction system.

---

X. Xia  
School of Automation, Southeast University, Nanjing 210096, China

S. Gai (✉)  
Key Laboratory of Measurement and Control for Complex System of Ministry of Education,  
School of Automation, Southeast University, Nanjing 210096, China  
e-mail: [cyoula@gmail.com](mailto:cyoula@gmail.com)

Scale invariant feature transform (SIFT) algorithm was proposed by Lowe in 1999 [3]. This method can detect feature points and keep the invariability for scaling, rotation, and several other transformations. Zhou used this algorithm for the research of face matching [4]. Because the SIFT feature points are based on the gradient magnitude calculated by the grayscale value of images, traditional matching algorithm based on SIFT has a poor performance when the difference of light distribution between two images is very large. In some condition, this algorithm cannot find any match points [5].

Image dense stereo matching algorithm based on region growing is a method to match the images by propagating the matching relations to other parts of the images [6]. This algorithm performs well for multi-texture images. Tang adopted this method for image matching [7]. But because this method relies on the amount and accuracy of the seeds, error-matched points would cost more time for the calculating of region growing and lead to the accumulation of mismatch. All these flaws may result in a poor disparity map and make it hard to determine the size of the search window.

Normalized cross correlation (NCC) algorithm is a matching method based on statistical principle [8]. This algorithm calculates the cross-correlation coefficients by using the grayscale values of pixels in the windows of the two images for matching. Compared to SIFT, this method is more stable for the difference of light between images. Random sampling consensus (RANSAC) is an algorithm using random uncertainty [9], which is capable of processing data with a large proportion of error points.

To solve all these problems above, this paper uses a method combining SIFT and NCC algorithms for the first match, then using RANSAC algorithm to eliminate the error-matched points to make the seeds more reliable. In addition, this paper proposes an evaluating system, which optimizes the choice of the size and location of the search window. This system helps to reduce mismatch and to increase the computational efficiency.

## **72.2 SIFT Feature Points Detecting and NCC First Matching**

Detecting of SIFT feature points of each image is necessary for the first match. This method of detecting contains four steps. First of all, Gaussian pyramid and DOG pyramid of each image should be built to detect extreme values and find out the approximate location of feature points by constructing scale space. Secondly, this method fits the local extreme values with 3D quadratic function to determine the location and scale precisely and eliminate points with low contrast on the unstable edge. Thirdly, this method assigns the major orientation of the feature points by using the distribution of the gradient direction in the neighboring windows of the feature points. Lastly, this method describes each feature point with 16 points, each

**Fig. 72.1** Initial images of Set 1



**Fig. 72.2** Initial images of Set 2



**Fig. 72.3** Feature points detected from Set 1



**Fig. 72.4** Feature points detected from Set 2

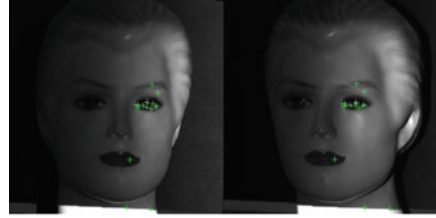


point has information in 8 directions and obtain 128 dimensions feature vectors. Figures 72.1 and 72.2 are two sets of initial images under different lighting conditions. Figures 72.3 and 72.4 show the SIFT feature points detected from the initial images.

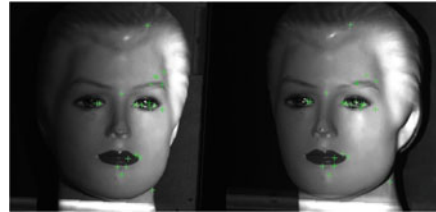
Traditional matching algorithms for the SIFT feature points are based on the Euclidean distance. Because the SIFT feature points are based on the gradient magnitude calculated by the grayscale value of images, traditional method has a poor performance when the difference of light distribution between two images is very large.

To solve this problem, this paper combines SIFT and NCC to process the first match. NCC algorithm is a matching method based on statistical principle, which

**Fig. 72.5** Match points in Set 1 from SIFT



**Fig. 72.6** Match points in Set 2 from SIFT



calculates the cross-correlation coefficients by using the grayscale values of pixels in the windows of the two images for matching. In this method, a neighboring window is opened around the feature point, which is to be matched in the left image. Then windows with same size are opened around every single feature point in the right image. Then the grayscale values of every pixel in those windows are taken into Eq. (72.2) to calculate the cross-correlation coefficients. The point with the largest coefficient and the feature point to be matched are a pair of match point.

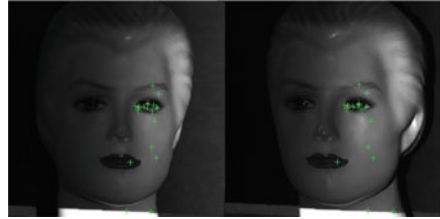
$$\rho(i_0, j_0) = \frac{\sum_{(i,j) \in W_{1,2}(i_0, j_0)} [G_1(i, j) - \bar{G}_1] [G_2(i, j) - \bar{G}_2]}{\sqrt{\sum_{(i,j) \in W_1(i_0, j_0)} [G_1(i, j) - \bar{G}_1]^2 \sum_{(i,j) \in W_2(i_0, j_0)} [G_2(i, j) - \bar{G}_2]^2}} \quad (72.1)$$

In this equation,  $W_1(i_0, j_0)$  and  $W_2(i_0, j_0)$  are the two neighboring windows around the feature point  $(i_0, j_0)$  in the two images.  $\rho(i_0, j_0)$  is the cross-correlation coefficient between those two windows.  $G_1(i, j)$  and  $G_2(i, j)$  are the grayscale values of point  $(i, j)$  in two images.  $\bar{G}_1$  and  $\bar{G}_2$  are the average values of the grayscale values in those two windows.

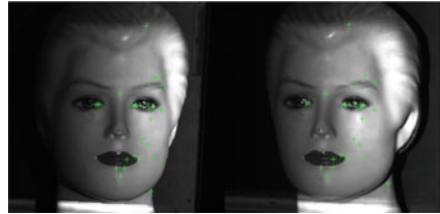
Compared to the common definitions of NCC coefficients, Eq. (72.1) subtracts the average value of grayscale in single window, which improves the performance of NCC and makes the process of matching not affected by the difference of distribution of light between images. Figure 72.5 shows the 13 pairs of match points in Set 1 computed from SIFT algorithm. Figure 72.6 shows the 22 pairs of match points in Set 2 computed from SIFT algorithm. Figure 72.7 shows the 16 pairs of match points in Set 1 computed from NCC method. Figure 72.8 shows the 28 pairs of match points in Set 2 computed from NCC method.



**Fig. 72.7** Match points in Set 1 from NCC



**Fig. 72.8** Match points in Set 1 from NCC



**Fig. 72.9** RANSAC experiment



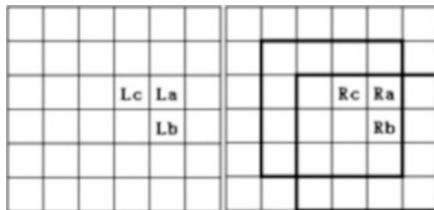
### 72.3 Eliminating Mismatch by RANSAC

After the first match, several match points are generated while some of them are error matched. In this paper, RANSAC algorithm is used to eliminate the mismatches. RANSAC is an algorithm using random uncertainty [9], which is capable of processing data with a large proportion of error points. Figure 72.9 shows an experiment using RANSAC to find the correct points in a line. There are 25 correct points and 50 error points in the figure, and the error rate is 67 %. After 500 iterations, 17 correct points were found out, and all error points are eliminated.

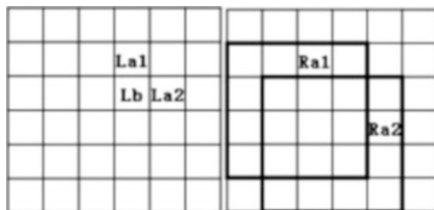
### 72.4 Region Growing

Region growing is a process that propagates the matching relations of seeds to other parts of images [10]. The basic strategy of region growing can be shown as Fig. 72.10. In Fig. 72.10, point  $L_a$  in left image and point  $R_a$  in right image are a pair of matched points. Point  $L_b$  and point  $L_c$  are around point  $L_a$  in the left image,

**Fig. 72.10** Basic strategy of region growing



**Fig. 72.11** Location of search window



which means point  $Rb$  and point  $Rc$ , the match points of point  $Lb$  and point  $Lc$ , are around point  $Ra$  too. So the search areas of those two match points can narrow down to neighboring windows of point  $Ra$ . With this algorithm, once a pair of robust match points is found, this matching relation can be quickly propagated to other parts of an image.

If two pairs of match points are known, the strategy can be shown as Fig. 72.11. In this figure, point  $La1$  with point  $Ra1$  and point  $La2$  with point  $Ra2$  are two pairs of match points. Point  $Lb$  in the left image is around both point  $La1$  and point  $Ra1$ , which means that point  $Rb$ , the match point of point  $Lb$  is around point  $Ra1$  and point  $Ra2$ . The search window is located by both point  $Ra1$  and point  $Ra2$ . Search window of a point can be determined by one pair of match points or several pairs.

But because the degree of accuracy and the computation speed of this method rely on the amount and accuracy of the seeds, to lower the negative effects of seeds with low robustness, this paper proposed an evaluating system for the reliability of the seeds.

$$R(i_0, j_0) = \frac{\sum_{(i,j) \in A_s} \rho(i, j)}{n} \tag{72.2}$$

In Eq. (72.2),  $R(i_0, j_0)$  is the evaluation for seed  $(i_0, j_0)$ .  $A_s$  is a neighboring area with the size of  $3 \times 3$ .  $\rho(i, j)$  is the NCC coefficient of the seed  $(i, j)$  in this area.  $n$  is the number of the seeds that in this area. During the process of region growing, seeds with large  $R$  will be taken into calculation first. The size of the search windows is inversely proportional to the  $R$ . In Eq. (72.3),  $N$  is the size of the search window for the next match point.  $R$  is the evaluation of the prior seed.  $\lambda$  is a coefficient which determined by the specific conditions.

$$N = \frac{\lambda}{R} \quad (72.3)$$

With this improved strategy of region growing, it can be ensured that growing happens around robust seeds first. Also the adaptive window can help to reduce the computation time and increase the veracity of the algorithm.

## 72.5 Algorithm Summary

Algorithm in this paper can be summarized as follow:

- (a) Detect SIFT feature points of images and use NCC to do the first match.
- (b) Eliminate the error-matched points by using RANSAC and obtain robust seeds.
- (c) Evaluate all seeds and put them into a sequence in order.
- (d) If the sequence is not empty, take out the first pair of seeds in the sequence and determine the size of the search window for the next match points. Calculate the match points of the 8 points around the seed.
- (e) Put the new match points into sequence in order.
- (f) Repeat step d and step e until the sequence is empty.

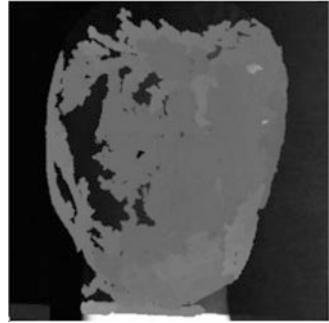
## 72.6 Experiments

Figures 72.12 and 72.14 are the disparity maps by using seeds from Figs. 72.5 and 72.6, which shows the result of the traditional algorithm. Figures 72.13 and 72.15 are the disparity maps by using this paper's method. Match points from Figs. 72.7 and 72.8 are processed with RANSAC and generate seeds for the region growing which uses the improved strategy with the evaluating system. It is obvious that the disparity maps generated by method in this paper are smoother and cover more area of the face. Table 72.1 shows the average computation time for two sets of experiments of both traditional algorithm and method of this paper. This paper's method is more efficient.

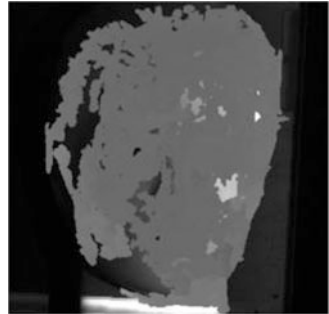


**Fig. 72.12** Traditional algorithm for Set 1

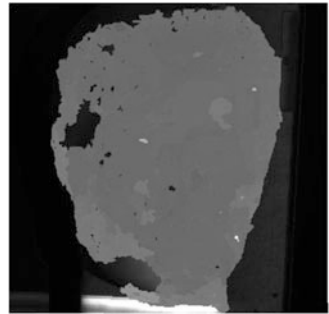
**Fig. 72.13** This paper's method for Set 1



**Fig. 72.14** Traditional algorithm for Set 2



**Fig. 72.15** This paper's method for Set 2



**Table 72.1** Computation time

Algorithm	Traditional	This paper
Computation time	5.8 s	3.9 s

### 72.7 Conclusion

This paper proposed an improved dense matching algorithm for face based on region growing. The algorithm combines SIFT and NCC for the initial match, and then uses RANSAC to eliminate the error-matched points to make the seeds more reliable. By using an improved strategy of region growing, in which seeds are

evaluated to help determine the locations and sizes of the search windows dynamically, the matching relations of seeds propagate to other parts of images.

From the experimental results, it can be concluded that compared to traditional algorithm, this paper's algorithm can generate better disparity maps, which are smoother and cover more area of the face. Also, this paper's algorithm has a higher computational efficiency.

## References

1. Liang, L. H., Ai, H. Z., Xu, G. Y., et al. (2002). A survey of human face detection. *Chinese Journal of Computers-Chinese Edition*, 25(5), 449–458.
2. Zhao, W., Chellappa, R., Phillips, P. J., et al. (2003). Face recognition: A literature survey. *ACM Computing Surveys*, 35(4), 399–458.
3. Lowe, D. G. (1999) Object recognition from local scale-invariant features. Computer vision, 1999. *The Proceedings of the Seventh IEEE International Conference on IEEE, England 2*, pp. 1150–1157
4. Zhou, Z. M., Yu, S. Y., Zhang, R., et al. (2008). An algorithm for face recognition based on SIFT descriptor. *Journal of Image and Graphics*, 13(10), 1882–1885.
5. Liu, X., & You, H. (2009). Matching feature points of multi-temporal space-borne SAR based on SIFT. *Science of Surveying and Mapping*, 1(015), 43–45.
6. Lhuillier, M., Quan, L. (2002). Quasi-dense reconstruction from image sequence. England, Computer Vision—ECCV 2002 (pp. 125–139). Berlin/Heidelberg: Springer
7. Tang, L., Wu, C. K., Liu, S. G., et al. (2004). Image dense stereo matching by technique of region growing. *Chinese Journal of Computers*, 27(7), 936–943.
8. Lewis, J. P. (1995). Fast normalized cross-correlation. *Vision Interface*, 10(1), 120–123.
9. Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, 24(6), 381–395.
10. Xing, Y. J., Meng, J. J., Sun, J., et al. (2006). An improved region-growth algorithm for dense matching. *Journal of Achievements in Materials and Manufacturing Engineering*, 18(1–2), 323–326.

## Chapter 73

# An Improved Feature Selection Method for Chinese Short Texts Clustering Based on HowNet

Xin Chen, Yuqing Zhang, Long Cao, and Donghui Li

**Abstract** Short texts have played an important role in the field of text data mining. Because of the problems arising from the complexity of Chinese semantics and data sparseness, which is an obvious characteristic of short texts, it is necessary to explore some new semantic-based methods to cluster Chinese short texts. An improved approach of feature selection based on HowNet is applied in this paper to address data sparseness of Chinese short texts. By redefining Vector Space Model in semantic level and merging generalized synonymy features, we present a new feature generation strategy. Experimental results show that by merging semantic similar feature, our method is effective in feature dimension reduction and gets better clustering performance. The proposed HowNet-based feature selection method is suitable for Chinese short texts clustering.

### 73.1 Introduction

With the popularization of Web2.0, instant message, image captions, twitter etc. which in the form of short texts have played an irreplaceable role in the area of public opinion analysis [1], Social Networking Services (SNS) [2], and topic tracking [3]. The clustering of short texts which is the basis of information analysis, however, presents great challenges. Because of the features of short texts, they cannot provide competent context information or word co-occurrence information for similarity measure [4], which is the basis of clustering methods [5]. Thus, the traditional methods of text clustering are not suitable for short text clustering when they are directly applied [6].

This paper is concerned with improving performance of the Chinese short texts clustering. We propose an effective feature selection method by word semantic

---

X. Chen • Y. Zhang (✉) • L. Cao • D. Li  
China University of Geosciences, Beijing 100083, China  
e-mail: [yqzhang@cugb.edu.cn](mailto:yqzhang@cugb.edu.cn)

similarity calculating based on HowNet [7]. We redefine Vector Space Model (VSM) [8] in semantic level and use a new feature generation strategy that merges generalized synonymy features. The experimental results demonstrate that the method obviously reduces the negative forces of sparse feature and improves the clustering accuracy.

## 73.2 Related Works

At the present, there are generally three types of approach in short texts clustering. The first proposed by Phan et al. [4] is mining the implicit topic of short texts based on the topic model and improving the co-occurrence probability of the feature terms. Quan et al. [9] also measured similarity between different feature terms in short texts with different topics based on the Latent Dirichlet Allocation (LDA) model. The second is achieving the short texts clustering based on similarity of the feature terms by using search engines by Sahami et al. [10]. The third presented by Gabrilovich et al. [11] is expanding feature terms by world knowledge to improving clustering performance in short texts. Hu et al. [12] using World knowledge such as WordNet or Wikipedia to reduce negative impact of data sparseness.

However, Chinese short texts are quite different from general short texts as a result of the complexity of Chinese semantic. Pu et al. [13] using Latent Semantic Indexing (LSI) for text preprocessing and then using Independent Component Analysis (ICA) to deal with Chinese short texts. The result expresses that the combination of LSI and ICA method is little better than ICA. Zuo et al. [14] improved the performance of Chinese short texts classification by selecting key words from training set. However, it is far from ideal effect. Ning et al. [15] proposed a method by using the domain word ontologies to deal with Chinese short texts. In order to solve the sparse feature keywords in Chinese short texts, Jin et al. [16] expanded semantic features to improve clustering effect. Our method is also committed to reduce the negative forces of sparse feature, but the core is improving feature selection method by word semantic similarity calculating.

## 73.3 Words Semantic Similarities Calculation

There are many approaches proposed to calculate Chinese words semantic similarities. HowNet as a widely adopted method is used in this paper to address words semantics.

### 73.3.1 HowNet Overviews

According to the official website of HowNet, the definition of HowNet is an online commonsense knowledge base unveiling inter-conceptual relations and inter-attribute relations of concepts as connoting in lexicons of the Chinese and their English equivalents [7]. The latest version covers over 65,000 concepts in Chinese and close to 75,000 English equivalents, and the relations include hyponymy, synonymy, antonymy, meronymy, attribute-host, material-product, converse, dynamic role, and concept co-occurrence [7]. The knowledge dictionary of HowNet is based on words and their concepts. Concept and sememe are the most important elements in the definition of HowNet. Concept, which is not we usually refer, describes one item of word semantics. Each word includes one or more items, and an item in the HowNet is called as sense corresponding. The concept of HowNet is described using a kind of knowledge representation language, each term used in the knowledge representation language is called as the sememe. A sememe refers to the smallest basic semantic unit. Every concept of a word or phrase and its description form one entry as follow:

NO = word/phrase ID

W\_C = Chinese word/phrase form

G\_C = Chinese word/phrase syntactic class

E\_C = example of Chinese word/phrase usage

### 73.3.2 Words Semantic Similarities Calculation Based on HowNet

One of the popular methods of Chinese word semantic similarity calculation based on HowNet was proposed by Liu [17]. For two words  $W_1$  and  $W_2$ , if  $W_1$  included  $n$  concepts  $C_{11}, C_{12}, \dots, C_{1n}$ , and  $W_2$  included  $m$  concept  $C_{21}, C_{22}, \dots, C_{2m}$ . The similarity between two words  $W_1$  and  $W_2$  defined as follows:

$$Sim(W_1, W_2) = \max_{i=1, 2, \dots, n, j=1, 2, \dots, m} Sim(C_{1i}, C_{2j}) \quad (73.1)$$

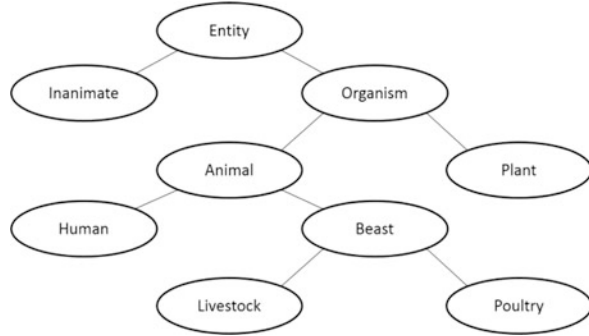
The similarity between two concepts defined as follows:

$$Sim(C_1, C_2) = \sum_{i=1}^4 \beta_i \prod_{j=1}^i Sim_j(C_{1i}, C_{2j}) \quad (73.2)$$

where  $\beta_i (1 \leq i \leq 4)$  means weight parameters of these four parts, respectively,  $\beta_i$  greater than 0.5 in generally,  $\beta_1 + \beta_2 + \beta_3 + \beta_4 = 1$  and  $\beta_1 \geq \beta_2 \geq \beta_3 \geq \beta_4$ .  $Sim_j$  represents the similarity of different section of concepts.



**Fig. 73.1** Hierarchy of sememe tree



The similarity between two sememe defined as follows:

$$Sim(P_1, P_2) = \frac{\alpha}{\alpha + dis(P_1, P_2)} \quad (73.3)$$

where  $dis(P_1, P_2)$  is the distance between  $P_1$  and  $P_2$  in sememe tree as shown in Fig. 73.1,  $\alpha$  is a smoothing factor.

In hierarchy of sememe tree, the depth of generalized sememe is smaller than the depth of specific sememe. Took into account depth information of the sememes and the semantic distance between two sememe nodes, we defined a father node of  $P_1$  and  $P_2$  in sememe tree called Least Common Node  $DN(P_1, P_2)$  and used the depth of  $DN(P_1, P_2)$  to replace the previous smoothing factor  $\alpha$ . Then similarity between two sememe redefined as follows:

$$Sim(P_1, P_2) = \frac{Depth(DN(P_1, P_2))}{Depth(DN(P_1, P_2)) + dis(P_1, P_2)} \quad (73.4)$$

### 73.4 General Framework for Chinese Short Texts Clustering

In this part, we introduce the presented general framework for Chinese short texts by an improved feature selection method based on HowNet. The framework consists of three phases, including data preprocessing, feature generation and selection, and clustering. As shown in Fig. 73.2, the first step of Chinese short texts clustering is Chinese word automatic segmentation. The matching process includes inserting the separator between Chinese terms according to specific standards, which is different from the space used in English. Moreover, some words are meaningless in Chinese just like “a” and “the” in English, which are called Stop Words. After data preprocessing, the bag of words should be segmented and exclude of Stop Words.

In phases 2, we use HowNet and the method presented in Sect. 73.3.2 to calculate the semantic similarity between each pair of words in bag. Then we get

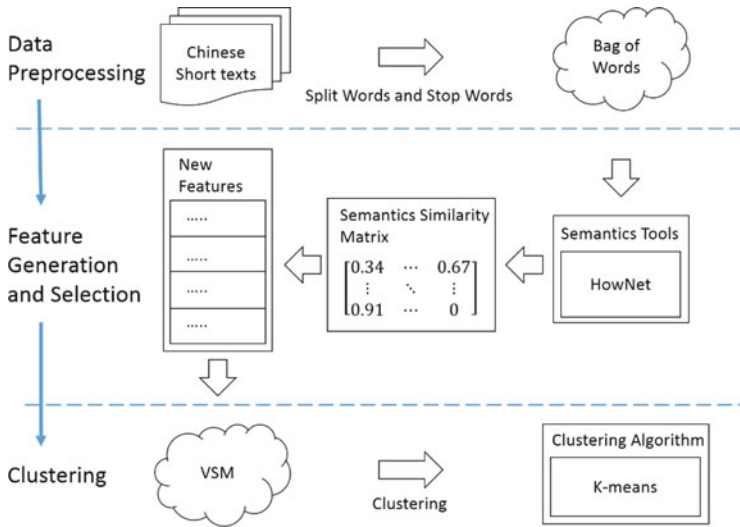


Fig. 73.2 Framework for Chinese short texts clustering

a Semantics Similarity Matrix, which dimension is equal to the number of words in bag. In order to reduce the matrix dimension, a threshold value  $K$  in our model is applied. If the semantic similarity between each pair of words is greater than  $K$ , we think that these two words are generalized synonymy in semantic. That is to say, their contribution can merge to computation as features. Therefore, the number of new feature will be reduced and the problem of data sparseness will be alleviated. The last step in this workflow is standard clustering algorithm and thus no more introduce.

### 73.5 Experimental Evaluations

In this section, we first present the experimental data used in our test bed and then introduce the evaluation criteria. Finally, we evaluate the experimental results.

#### 73.5.1 Date Sets and Evaluation Criteria

The microblog which is a typical representative of the Chinese short texts employed in our experiment as the benchmark datasets. We crawled twenty thousand microblogs from Tencent since August to September 2012, and manual tagged them to 15 cluster.

**Table 73.1** Average accuracy test condition

	Same class	Different class
Same cluster	<i>TP</i> (True positive)	<i>FN</i> (False negative)
Different cluster	<i>FP</i> (False positive)	<i>TN</i> (True negative)

**Table 73.2** Results with three feature selection methods

	5 Cluster		15 Cluster	
	F <sub>1</sub> -measure	Average accuracy	F <sub>1</sub> -measure	Average accuracy
Simple VSM	0.513	0.559	0.489	0.505
Synonymy dictionary	0.604	0.646	0.526	0.553
HowNet ( <i>K</i> = 0.6)	0.642	0.677	0.538	0.551

The evaluation standard F<sub>1</sub>-measure and Average Accuracy used in our experiments is based on the Confusion Matrix as illustrated in Table 73.1 for Clustering problem.

F<sub>1</sub>-measure is a measure of a test's accuracy which considers both precision and recall that measures a cluster contains only objects of a particular class and all objects of that class.

Average Accuracy is a statistical measure defined as follows:

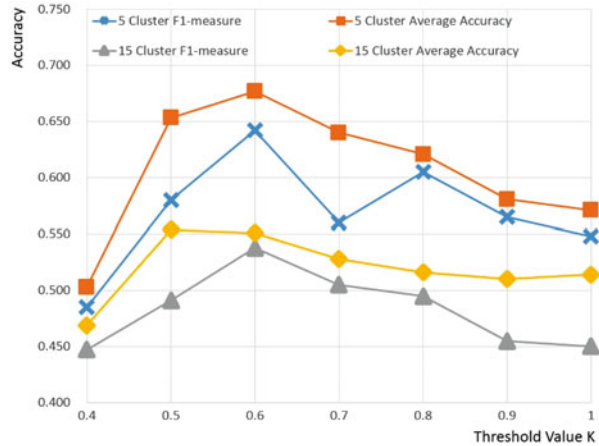
$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (73.5)$$

### 73.5.2 Performance Evaluation

Experimental results of the three feature selection methods on both 5 cluster dataset and 15 cluster dataset using the simple k-means clustering algorithms are reported in Table 73.2. Each result denotes an average of 20 test runs by randomly choosing the initial parameters for k-means algorithm. In this Tables, the feature selection based on HowNet with the threshold value *K* = 0.6 achieves better performance than other two methods except the Average Accuracy in 15 cluster dataset. Moreover, the performance of three methods in 5 cluster is obviously better as compared with them in 15 cluster.

In Fig. 73.3, we note that, in general terms, different threshold values achieve different performance. When *K* ≤ 0.4, threshold value is too low that many words which not similarity in semantic merge to one feature, and thus get a lower clustering accuracy. The most appropriate threshold value in our experiment is *K* = 0.6.

**Fig. 73.3** Results with different threshold values in HowNet method



## 73.6 Conclusion

This paper presented an improved feature selection method to address Chinese short texts. We redefined VSM in semantic level and used a new feature generation strategy that merges generalized synonymy features. Based on this new strategy, we used microblog which is a typical representative of the short text at present to validate the model. The experimental results indicated that the method can effectively reduce the negative forces of sparse feature and improve the clustering accuracy. Due to the various presentations of short texts, however, our experimental dataset was incomplete. Therefore, we should further expand the category of short texts and focus on the applicability of this method.

**Acknowledgements** Supported by “the Fundamental Research Funds for the Central Universities” of CUGB (2652013102, CUGB), Subject Supportive Project of CUGB (2013) and the Innovative Experiment Plan for College Students of China University of Geosciences, Beijing.

## References

1. Yang, Z., Duan, L., & Lai, Y. (2010). Online public opinion hotspot detection and analysis based on short text clustering using string distance. *Journal of Beijing University of Technology*, 36(5), 669–673.
2. Rosso, P., Errecalde, M., & Pinto, D. (2013). Analysis of short texts on the Web: introduction to special issue. *Lang Resources and Evaluation*, 47(1), 123–126.
3. Mo, Y., Liu, S., Liu, Y., & Cheng, X. (2012). An entropy based rule mining algorithm for filtering tweets by topics. *Journal of Chinese Information Processing*, 26(5), 1–6. 39.
4. Phan, X. H., Nguyen, L. M., Horiguchi, S. (2008). Learning to classify short and sparse text & web with hidden topics from large-scale data collections. In *Proceedings of the 17th International Conference on World Wide Web* (pp. 91–100). New York: ACM

5. Hu, J., Fang, L., Cao, Y., Zeng, H., Li, H., Yang, Q., Chen, Z. (2008). Enhancing text clustering by leveraging Wikipedia semantics. In *Proceedings of the 31st ACM SIGIR* (pp. 179–186). New York: ACM
6. Metzler, D., Dumais, S., & Meek, C. (2007). Similarity measures for short segments of text. *Lecture Notes in Computer Science*, 4425, 16–27.
7. Dong, Z., Dong, Q. HowNet knowledge database. <http://www.keenage.com/>
8. Salton, G., Wong, A., & Yang, C. S. (1975). A vector space model for automatic indexing. *Communications of the ACM*, 18(11), 613–620.
9. Quan, X., Liu, G., Lu, Z., Ni, X., & Liu, W. (2010). Short text similarity based on probabilistic topics. *Knowledge and Information Systems*, 25(3), 473–491.
10. Sahami, M., Heilman, T. D. (2006). A web-based kernel function for measuring the similarity of short text snippets. In *Proceedings of the 15th International Conference on World Wide Web* (pp. 377–386). New York: ACM
11. Gabrilovich, E., Markovitch, S. (2005). Feature generation for text categorization using world knowledge. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence* (pp. 1048–1053). San Francisco CA: Morgan Kaufmann
12. Hu, X., Sun, N., Zhang, C., Chua, T. S. (2009). Exploiting internal and external semantics for the clustering of short texts using world knowledge. In *Proceedings of the 18th ACM Conference on Information and Knowledge Management* (pp. 919–928). New York: ACM
13. Pu, Q., Yang, G. (2006). Short-text classification based on ICA and LSA. In *Proceedings of International Symposium on Neural Networks*. (pp. 265–270). German: Springer
14. Zuo, S., Wu, C., Zhou, Y., & He, H. (2006). Chinese short-text categorization based on the key classification dictionary words. *The Journal of China Universities of Posts and Telecommunications*, 13, 47–49.
15. Ning, Y., Fan, X., Wu, Y., & Text, S. (2009). Classification based on domain word ontology. *Computer Science*, 36(3), 142–145.
16. Jin, C., Zhou, H., & Bai, Q. (2012). Short text clustering algorithm with feature keyword expansion. *Advanced Materials Research*, 532–533, 1716–1720.
17. Liu, Q., & Li, S. (2002). Word similarity computing based on how-net. *Computational Linguistics and Chinese Language Processing*, 7(2), 59–76.

# Chapter 74

## Internet Worm Detection and Classification Based on Support Vector Machine

Huihui Liang, Min Li, and Jiwen Chai

**Abstract** This paper proposes a novel Internet worm detection and classification method. The behaviors of worms are different from each other's, and they are also different in terms of the normal Internet activities. So we can detect and classify worms by the extracted features of the network packets. At first, we sniff raw network packets from the local area network (LAN), and extract 13 features from the packet header, and then select 10 important features using the information gain algorithm. With the labeled features, we train Support Vector Machine (SVM) classifiers. The classifiers can classify the behaviors of the worm apart from the normal Internet activities. And this approach can also classify network attacks and Internet worms, although the network attacks and the Internet worms have similar behaviors. In the experiments, this approach performs well in worm detection and classification.

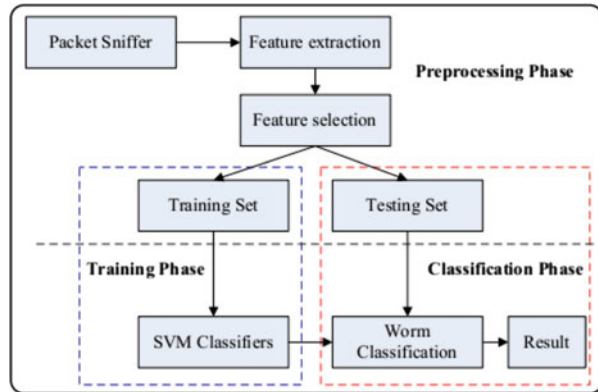
### 74.1 Introduction

With the increasing popularity of the Internet and the network attack technology continues to change, the network worm has become a critical threat in computer network [1, 2]. By self-propagating and fast spreading, the network worm attacks computer operating system or applications that have specific security vulnerabilities, so as to obtain control of some or all of the computers. Currently, many research works for Internet worm detection have been done [3, 4]. And most of the Internet worm detection methods are based on the intrusion detection system (IDS) [5]. The Internet worm-based IDS mainly has two categories, the network-based Internet worm detection and the host-based Internet worm detection. The host-based Internet worm detection approaches analysis the network packets that already reached the end-host. While the network-based Internet worm detection methods

---

H. Liang (✉) • M. Li • J. Chai  
Sichuan Electric Power Research Institute, Chengdu 610000, China  
e-mail: [liang\\_huihui@sohu.com](mailto:liang_huihui@sohu.com)

**Fig. 74.1** Overview of the framework



consider network packets before they reach an end-host. In this chapter, we detect Internet worms and classify the Internet worms apart from the Internet attacks by analyzing the network packets before they reach an end-host.

In this chapter, we propose a novel network packets based Internet worm detection method. In Fig. 74.1, we give an overview of the framework of our worm detection and classification approach. For worm detection, during the preprocessing phase, we sniff network packet data from our local area network, and we extract 13 features from each traffic data. And we give each extracted feature vector a label  $L$ . If the feature vectors are extracted from a normal traffic data, we set  $L = 1$ , else, we set  $L = c (c = 2, 3, \dots, C)$  for different Internet worms or Internet attacks,  $C$  is the number of the worm and the Internet attacks categories. In order to use some useful features to detect and classify the worms and the Internet attacks. We select ten features from the extracted features using an information gain method. Using this method, we collect our training data and testing data. We train a Support Vector Machine (SVM) classifier for each Internet worm category, each Internet attack category and the normal Internet activity class. During the classification phase, we can detect and classify the Internet worms, the Internet attacks, and the normal Internet activities.

The rest of this chapter is organized as follows. After describe related works of Internet worm detection methods in Sect. 74.2, we give an introduction of the abnormal behaviors of the Internet worms and the Internet attacks of the traffic packets in Sect. 74.3. In Sect. 74.4 we present the technique information of SVM. After introduce the experiment in Sect. 74.5, we conclude this paper in Sect. 74.6.

## 74.2 Related Works

Through normal network behavior to associate and compare with anomaly behavior of the network traffic, anomaly behavior could be detected. To detect unknown Internet worm, Data mining approach is a good choice. This section provides brief description of these existing detectors.

Wang et al. [6] built a model by training with these features extracted from malware behavior. They presented the results showed that the average detection rate is over 90 %, and the false alarm rate is around 6.67 % with n gram extraction. Siddiqui et al. [7] built a data mining model by train the features extracted from cleaned program and infected program. They presented good results with 95.6 % of detection rate for every end point and with false alarm rate close to 3.8 %.

In this chapter, we propose a novel network packets based Internet worm detection method. Different with most of the previous works, this work uses the one-against-all SVM classification method and trains an SVM classifier for each Internet worm category, each Internet attack category, and the normal Internet activity class. During the classification phase, we can detect and classify the Internet worms, the Internet attacks, and the normal Internet activities. And we get the better performance in our experiment.

## 74.3 Extract Features of Worms and Attacks

A lot of Internet worms have similar behaviors and payload as the Port Scan and Denial of Service (DoS) attacks. We take Blaster worm, Sobig.E worm, and Forbot-FU as the representation of the common worms. And we take UDP flood and HTTP flood as attack representation of DoS attack. In this chapter, we propose an approach to detect and classify Blaster worm, Sobig.E worm, Forbot-FU worm, Port Scan, and DoS attacks behaviors.

Port Scan is a technique to scan for available port or service from any users, it can be defined as an attack that sends client requests to a range of server port addresses on a host, with the goal of finding an active port and exploiting a known vulnerability of that service.

### 74.3.1 Feature Selection

In data classification, training set contain a large number of features can produce great computational complexity, and redundant feature vector can reduce the judgment probabilistic. So before we train the SVM classifiers, we select the extracted features. In this section, at first, we introduce the extracted feature records, and then, we select ten features as the classification feature vector using the information gain ratio method.

In order to detect and classify the Internet worm and the Internet attacks, we sniff packets from the local area network. Then we extract a lot of features from the collected packets. We define the 13 features that extracted from all of the packets collected in 1 s as a feature record. For a feature record  $X_i$ , we represent each feature component as  $x_{ij}$ ,  $j = (1, 2, \dots, 13)$ , and the detail information of each feature component is shown as follows:



- $x_{i1}$  Represent the number of the source IP address.  
 $x_{i2}$  Represent the number of the destination IP address.  
 $x_{i3}$  Represent the number of ICMP packet.  
 $x_{i4}$  Represent the number of TCP packet.  
 $x_{i5}$  Represent the number of UDP packet.  
 $x_{i6}$  Represent the number of the SYN flag.  
 $x_{i7}$  Represent the number of ACK flag.  
 $x_{i8}$  Represent the number of RST flag.  
 $x_{i9}$  Represent the number of source port.  
 $x_{i10}$  Represent the number of destination port.  
 $x_{i11}$  Represent the number of different packet size.  
 $x_{i12}$  Represent  $\frac{\text{the number of source port}}{\text{the number of destination port}}$ .  
 $x_{i13}$  Represent  $\frac{\text{the number of SYN flag}}{\text{the number of the destination IP}}$ .

The purpose of feature selection is in maintaining the accuracy of the premise, as far as possible reduce the number of features. In this paper, we use information gain algorithm to select features. Information gain algorithm is a feature selection methods that use the information gain of feature attributes to decision attributes as the importance measure of feature attribute.

At first gives the training feature record set  $X$ , the information entropy of training set  $X$  is defined as follows,

$$E(X) = - \sum_{c \in C} \frac{|X_c|}{|X|} \cdot \log_2 \frac{|X_c|}{|X|}, \quad (74.1)$$

where  $C$  is a subset of  $X$ .

Then we reduce the feature  $j$ , through measure the degree of entropy reduction (Information Gain)  $IG(X,j)$  to evaluate the information content of feature  $j$ , the formulation of information gain is shown as follows:

$$IG(X,j) = E(X) - \sum_{v \in V(j)} \frac{|X_v|}{|X|} \cdot E(X_v) \quad (74.2)$$

where set  $V$  contains all of the possible values of feature  $j$ .

Using this information gain formulation, we evaluate the network features. If the feature has a higher value, it represents this feature is more important than other features that have smaller information gain values. Based on the rank of the information gain values of those features, then we choose the top ten features as a feature vector. At last, based on the ten dimensional feature vectors and its corresponding labels, we train the SVM classifiers, detect and classify the Internet worms and Internet attacks.

## 74.4 Support Vector Machines

In this paper, we use the support vector machines (SVM) classifiers to classify the Internet worms and Internet attacks. SVM [8, 9] is a pattern recognition method developed from the stoical learning theory, and some works of worm detection that base on SVM have been done [10]. SVM algorithm was originally designed for binary classification problems, when dealing with many problems; it needs to construct a suitable multi-class classifier. Construct the multi-class SVM classifier method mainly has two kinds: the direct method and the indirect method. The direct method directly modifies the objective function, merges the parameter solution problems of multiple classification surface into one optimization problem, by solving the optimization problem to implement “one time” multi-class classification. This method looks like simple, but its computational complexity is very high, and it is difficult to implement. The indirect method achieves multi-class classification by combining multiple classifiers. Common methods include one-against-one and one-against-all two kinds.

The one-against-one SVM method designs an SVM classifier between any two classes of samples. For a training set of  $C$  categories, the one-against-one method needs to train  $C(C - 1)/2$  classifiers. When classifying an unknown sample, the category has the highest votes is the category of the unknown samples. The one-against-all SVM method needs design  $C$  classifiers when solve a  $C$  categories classification problem. During training, this method takes the samples of a certain category as one category and takes the other of the rest samples as another category, So that  $C$  kinds of samples are constructed  $C$  SVM classifiers. During classification, classify the unknown samples as the category that has the maximum value of the classification function.

In this paper, we use the one-against-all method to train SVM classifier. For each classifier, we use the Radial Basis Function (RBF) as a kernel to train the SVM classifier. The formulation of the RBF is shown as follows:

$$K(x, x_i) = \exp\left(-\frac{\|x - x_i\|^2}{2\sigma^2}\right), \quad (74.3)$$

where  $\sigma$  is the standard deviation, and  $\sigma > 0$ .  $x_i$  and  $x$  are feature vectors.

Based on the radial basis kernel function, using the Lagrange method, the SVM classification decision becomes the problem of finding the solution of the dual optimization problem as follows:

$$\max L(\alpha) = \sum_{i=1}^n \alpha_i - \frac{1}{2} \sum_{i,j=1}^n \alpha_i \alpha_j y_i y_j K(x_i, y_j) \quad (74.4)$$

where  $y_i = \pm 1$ ,  $i = 1, \dots, n$  are class indicator values and  $\alpha_i$ ,  $i = 1, \dots, n$  are Lagrange multipliers satisfying where  $C$  is the regularization parameter.

$$0 \leq \alpha_i \leq C, \quad \sum_{i=1}^n \alpha_i y_i = 0$$

The final discriminant function of the SVM classifier of  $C$  categories is shown as follows:

$$G_j(x) = \sum_{i \in SV} \alpha_i y_i K(x_i, x) + \omega_0 \quad (74.5)$$

where  $\alpha$  and  $\omega_0$  are the trained parameters,  $SV$  is the support vector set. The value of  $\omega$  is computed by the following function,

$$\omega_0 = \frac{1}{N_{SV'}} \left( \sum_{i \in SV'} y_i - \sum_{i \in SV, j \in SV'} \alpha_i y_i K(x_i, x_j) \right) \quad (74.6)$$

in which  $SV$  is the set of support vectors with associated values of  $\alpha_i$ , satisfying  $0 < \alpha_i \leq C$  and  $SV'$  is the set of  $N_{SV'}$  support vectors satisfying  $0 < \alpha_i < C$ .

At last, with the trained SVM classifiers  $G_j(x)$ ,  $j = 1, \dots, C$ , we can classify the Internet worms and attacks by the following function.

$$c = \operatorname{argmax}_j G_j(x). \quad (74.7)$$

Where  $c$  is the classify category of the unknown sample  $x$ .

## 74.5 Experiment

In this chapter, we set up a local area network and analog the normal working state of the computers. During collect data set, we insert actual blaster worm, Forbot-FU worm, and Sobig.E worm using a reliable on-line source. And generate Port Scan, HTTP flood, and UDP flood into this local area network using the Net Tool at the same time.

### 74.5.1 Generate Data Sets

In order to collect the experiment data set and the testing data set, we sniffer network packets from the local area network and extract the features from the packet header as describe as in Sect. 74.3. Then we split the collected data set into two parts: the original training set and the original testing set.

**Table 74.1** Comparison detection rate and true positive rate

Model	Bayesian network	Decision tree C4.5	Random forest	SVM
Detection rate (%)	96.9	98.6	99.4	99.6
Normal (%)	97.8	98.9	99.6	99.6
Blaster worm (%)	91.2	97.5	99.1	99.4
SoBig.E (%)	94.8	97.5	96.1	97.8
Forbot-FU (%)	98.2	98.2	98.2	98.2
UDP flood (%)	100	100	100	100
HTTP flood (%)	98.0	98.3	98.9	98.8
Port Scan (%)	99.8	99.8	99.8	99.8

The original data set contains redundant feature information. We use information gain method to select ten features from the training set as the final extracted feature vector. And then, we present the training set and the testing set as the ten dimensional feature vectors and its corresponding labels.

At last, we get a training data set of 6,000 feature vectors for normal class, and 1,000 feature vectors for each of the six class Internet worms and the Internet attacks. Totally, the training data set has 12,000 feature vectors. The same size with training data set, the testing data set also have 6,000 feature vectors for normal class, and 6,000 feature vectors for all class of the Internet worms and the Internet attacks. Totally, the testing data set has 12,000 feature vectors too.

## 74.5.2 Experiment Results

In the experiment, we compute the classification results of the worms and the Internet attacks. In order to evaluate the performance of our detection and classification model, we compare our detection rate, true positive rate, and false alarm rate with other three data mining models.

In Table 74.1, we show the detection rates of our algorithm and make a comparison between our model and other models. Obviously, our algorithm has a better performance. Our detection rate is 99.6 % which is the top score. When the situation is normal or the worm is Blaster worm or SoBig.E worm, we could get a higher rate than other algorithms. When the testing data set is Forbot-Fu worm, UDP Flood, HTTP Flood, or Port Scan, the positive rate is the same with other algorithms. In this case, the positive rate is high even 100 %, so to improve these rate is very difficult. Our algorithm presents a better result in most of the situations, and the same with other algorithms in many cases.

In Table 74.2, we show the false alarm rates of different algorithms. When the testing data is UDP Flood, HTTP Flood, or Port Scan, the false alarm rate is 0. It is the same with other algorithms. The rate of our algorithm is a little higher than Random Forest algorithm and lower than others when the worm is SoBig.E. However, we get the top score when the worm is Blaster Worm or Forbot-FU. In

**Table 74.2** Comparisons about false alarm rate

Model	Bayesian network	Decision tree C4.5	Random forest	SVM
Blaster worm (%)	2.1	0.3	0.5	0.3
SoBig.E (%)	4.2	2.1	1.5	1.6
Forbot-FU (%)	0.7	0.6	0.7	0.6
UDP flood (%)	0.0	0.0	0.0	0.0
HTTP flood (%)	0.0	0.0	0.0	0.0
Port Scan (%)	0.0	0.0	0.0	0.0

this case, the Decision Tree C4.5 has the same good performance with ours, and the Random Forest is a little bad. So, the result here is that our algorithm as a whole has a better performance than other algorithms.

## 74.6 Conclusion

In this paper, a novel method for worm detection and classification was proposed. At first, we sniffed data packets from the local area network, and we extracted a lot of features from the collected traffic data. Then we used information gain algorithm to select useful features. And then by using the selected useful feature vectors, we trained our SVM classifier. At last, we detected and classified the worms and Internet attacks. This algorithm performed well in worm detection and classification. From the experiment results, it can be concluded that our algorithm can correctly detect and classify Blaster worm, SoBig.E worm, Forbot-Fu worm, UDP Flood, HTTP Flood, and Port Scan.

## References

1. Magkos, E., Avlonitis, M., Kotzanikolaou, P., et al. (2013). Toward early warning against Internet worms based on critical sized networks. *Security and Communication Networks*, 6(1), 78–88.
2. Julisch, K. (2013). Understanding and overcoming cyber security anti-patterns. *Computer Networks*, 57(10), 2206–2211.
3. Rabinovitch, P., Chow, S. T. H., Abdel-Aziz, B. (2012). Worm detection by trending fan out U.S. Patent 8,095,981[P]. 2012-1-10.
4. Zheng, H., Lifa, W., Huabo, L., et al. (2012). Worm detection and containment in local networks. *Computer Science and Information Processing. Xi'an: IEEE*, 595–598.
5. Smith, C., Matrawy, A., Chow, S., & Abdelaziz, B. (2009). Computer worms: Architecture, evasion strategies, and detection mechanisms. *Journal of Information Assurance and Security*, 4, 69–83.
6. Wang, X., Yu, W., Champion, A., Fu, X., & Xuan, D. (2008). *Detecting worms via mining dynamic program execution. Security and privacy in communications networks and the workshops* (pp. 412–421). Nice: IEEE.

7. Siddiqui, M., Wang, M. C., & Lee, J. (2008). Detecting Internet worms using data mining techniques. *Journal of Systemics, Cybernetics and Informatics*, 6(6), 48–53.
8. Bishop, M. (2006). *Pattern recognition and machine learning* (pp. 325–359). New York: Springer-Verlag.
9. Webb, A. (2002). *Statistical pattern recognition* (2nd ed., pp. 123–200). Paris: John Wiley & Sons.
10. Nissim, N., Moskovitch, R., & Rokach, L. (2012). Detecting unknown computer worm activity via support vector machines and active learning. *Pattern Analysis and Applications*, 15(4), 459–475.

# Chapter 75

## Real-Time Fall Detection Based on Global Orientation and Human Shape

Shuangcheng Wang, Yepeng Guan, and Ruiyue Xu

**Abstract** Fall detection is an important problem in the research of abnormal behavior recognition. In this chapter, a novel real-time method is proposed to detect human fall with a single uncalibrated camera by the changes of global orientation and human shape. Our algorithm has three basic parts: moving object extraction, fall pre-detection, and fall confirmation. The overall orientations are derived from the combination of Gaussian mixture model and motion history image. The shape deformation is quantified from the silhouettes by an approximated bounding box. Standard deviations of the overall orientation and aspect ratio of bounding box are checked to pre-detect a fall, and a fall is confirmed by unmoving shape of the bounding box with the defined fall angle. Experimental results show that our method can detect all possible types of human fall accurately and successfully.

### 75.1 Introduction

Intelligent visual surveillance has got more research attention and funding due to increased global security concerns, whose key technology is abnormal behavior recognition. Fall detection is an important problem in the research of abnormal behavior recognition [1], because fall in elderly greatly threatens their health, especially who are living alone.

---

S. Wang (✉)

School of Communication, Information Engineering, Shanghai University,  
Shanghai 201305, China  
e-mail: [wsc36305@foxmail.com](mailto:wsc36305@foxmail.com)

Y. Guan

Key Laboratory of Advanced Displays and System Application, Shanghai 201305, China  
e-mail: [ypguan@shu.edu.cn](mailto:ypguan@shu.edu.cn)

R. Xu

School of Communication, Information Engineering, Shanghai University,  
Shanghai 201305, China

Here, a novel real-time fall detection algorithm is proposed, which detects fall by checking the variation of the global orientation and some changes in human shape with a single fixed uncalibrated camera. We notice that when people walk or run normally, the motion direction varies continuously at small scale, while during a fall, it changes greatly. Moreover, the rebound when the body hits on the ground and the movement of arms and legs can lead to great change in directions. The global orientation performs well on falls parallel or perpendicular to the camera optical axis in our work, which is derived from the combination of Gaussian mixture model (GMM) and motion history image (MHI). Meanwhile, the extracted moving object by GMM is approximated by a bounding box. Firstly, standard deviations of the overall orientation and aspect ratio of bounding box are checked to pre-detect a fall. And then a fall is confirmed by unmoving shape of the bounding box with the defined fall angle.

The organization of this chapter is as follows. Section 75.2 explains the related work on fall detection. Section 75.3 describes the details of the three basic parts of the method. Experimental results are represented in Sect. 75.4, and a conclusion is made at last.

## 75.2 Related Work

Traditional methods to detect falls use some wearable sensors [2]. However, the problem of such detectors is that people often forget to wear them and it's uncomfortable. Approach based on the computer vision is not only able to overcome these limitations but also capable of getting more information about fall, which may help to analyze the reason of a fall.

There are many related works in the field of fall detection based on the digital video processing techniques. For instance, a simple method was used only based on analyzing aspect ratio of the moving object's bounding box [1]. Some researchers have mounted the camera on the ceiling: Lee [3] detected a fall by analyzing the shape and 2D velocity of the person. Nait-Charif and McKenna [4] tracked the person using an ellipse and analyzed the resulting trajectory to detect inactivity outside the normal zones of inactivity like chairs or sofas. Rougier [5] and Miao Yu [6] both computed a motion coefficient based on the motion history in order to detect a fall. Rougier [5] combined the coefficient with shape descriptor ellipse fitting, while Miao Yu added head tracking using particle filter to find a fall with the speed of head. Other systems used the 3D information to infer events [7], but these mechanisms tend to be more complex, need more additional cost, and are difficult to be applied to real systems.

Unlike the above methods, in our system the global orientation and human shape are adopted to detect a fall. It has an acceptable detection rate, and it can work in real time using a webcam.



## 75.3 The Details of Our Method

Our algorithm has three basic parts: moving object extraction, fall pre-detection, and fall confirmation.

### 75.3.1 The Moving Object Extraction

Considering a relatively fixed background in our scene, we adapt an OpenCV algorithm BackgroundSubtractorMOG2 described by Z. Zivkovicin to segment the moving object [8]. In our work, four Gaussian distribution models are defined to describe the state of each pixel [8].

Then, morphological opening-and-closing operation is used to obtain a better result. And a denoising method is applied by means of the edge tracing method. The contours of updated binary foreground image are tracked, and the ones whose circumference is under a certain threshold (in practice, it is set to 80) are eliminated.

### 75.3.2 Fall Pre-detection

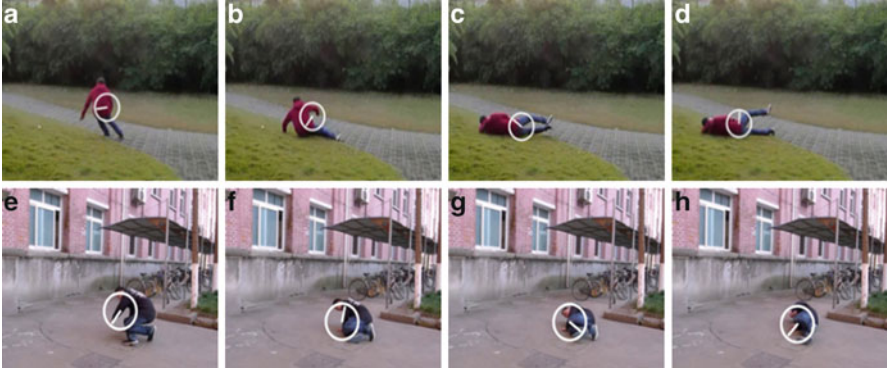
*Motion History and Moving Orientation* We notice that when people walk or run normally, the motion direction varies continuously at small scale, while during a fall, the direction changes obviously relative to walking or running. Moreover, the rebound when the body hits on the ground and the movement of arms and legs can lead to great changes in directions as it is shown in Fig. 75.1 (the radius indicates the direction). So the global motion orientation is needed.

MHI [9] is a simple and powerful method that records movement information of object by means of time stamp. The MHI is an image where the pixel intensity represents the recency of motion in an image sequence and therefore gives the most recent movement of a person during an action. Contrary to the optical flow [5], which is complex, time-consuming, and bad anti-noise, MHI is brief and real time.

The  $MHI_{\tau}(x,y,t)$  can be computed from an updated function  $D(x,y,t)$ :

$$H_{\tau}(x,y,t) = \begin{cases} \tau & , \text{ if } D(x,y,t) = 1 \\ \max(0, H_{\tau}(x,y,t-1) - \delta) & , \text{ otherwise} \end{cases}, \quad (75.1)$$

where  $(x,y)$  and  $t$  show the position and time and  $D(x,y,t)$  is a binary sequence of motion regions that is extracted from the original image sequence  $I(x,y,t)$ , which are already obtained by GMM. The duration  $\tau$  decides the temporal extent of the movement (normally, in terms of frames), and  $\delta$  is the decay parameter [9]. The result is a scalar valued image where more recently moving pixels are brighter and vice versa.



**Fig. 75.1** Changes in direction when falling. (1) Falling perpendicularly to the camera optical axis, large changes in direction. (a) 137th frame, (b) 143th frame, (c) 152th frame, (d) 161th frame. (2) Falling parallelly to the camera optical axis, also large changes in direction. (e) 120th frame, (f) 128th frame, (g) 135th frame. (h) 142th frame

Since the MHI records silhouette at different times, the overall motion orientation will be obtained by calculating the gradient of MHI. The gradient  $\text{Orient}(x,y)$  is computed as follows:

$$\text{Orient}(x,y) = \arctan(D_y(x,y)/D_x(x,y)), \quad (75.2)$$

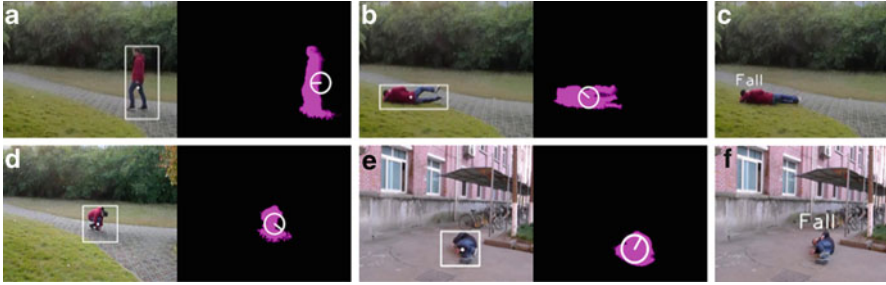
where  $D_x(x,y)$  is the difference in x direction and  $D_y(x,y)$  is the difference in y direction. The average direction is computed from the weighted orientation histogram, where a recent motion has a larger weight and the motion occurred in the past has a smaller weight, as recorded in MHI.

In the light of above work, we can easily obtain the  $\alpha$  standard deviation  $\sigma_\alpha$  from N frames (N is set to 20 in our work). In our work, we set the fixed threshold of  $\sigma_\alpha$  as 1.0 according to our datasets.

*Rectangular Approximation for Shape Analysis* It is obvious that the human shape changes greatly during a fall. Due to the well performance of global orientation, we simply approximate the human shape with a minimum bounding box [10]. When a person falls, the height and width of bounding box change drastically. We simply defined the aspect ratio  $\rho$  as

$$\rho = w/h \quad (75.3)$$

where  $w$  is the width of the bounding box and  $h$  is the height of the bounding box. We derived its standard deviation  $\sigma_\rho$  from fixed number of frames, which in our work is set to 20. We use  $\sigma_\rho$  to represent the changes in shape. It will stay at a high value when a fall happens, while it will stay at a low value when people walk normally. Its threshold for our datasets is set to 1.5.



**Fig. 75.2** Some examples on various situations (a) Walking normally:  $\sigma_\alpha = 0.0501$ , (b) Falling perpendicularly:  $\sigma_\alpha = 1.3297$  (c) Result  $\sigma_\rho = 0.5001$ ,  $\rho = 2.0894$ ;  $\sigma_\rho = 2.3211$ ,  $\sigma_\beta = 0.652$ ,  $\rho = 0.340$  (d) Crouch normally:  $\sigma_\alpha = 0.7805$ , (e) Falling parallelly:  $\sigma_\alpha = 1.2504$ , (f) Result  $\sigma_\rho = 0.8424$ ,  $\rho = 1.0216$ ,  $\sigma_\beta = 0.9532$ ;  $\sigma_\rho = 1.6033$ ,  $\rho = 0.8598$ ,  $\sigma_\beta = 0.7236$

Since  $\sigma_\alpha$  and  $\sigma_\rho$  need tens of frames to initialize, there may exist false alert at the beginning when a person first appears. Obviously, when normally walking,  $\rho$  stays at a low value, while during a fall,  $\rho$  stays at a high value. The parameter  $\rho$  is used to avoid these faults with a fixed threshold (in our work, we choose 1.0).

### 75.3.3 Fall Confirmation

The last step of the algorithm is to check whether the person remains motionless on the floor for a few seconds after falling. And the centroid of object [11] is needed.

The center of the bounding box is obtained by computing the coordinates of the center of mass with the first- and zero-order spatial moments [12]. We define fall angle  $\beta$  as the angle of centroid of object with respect to horizontal axis of the bounding box [11]. The unmoving shape of the bounding box is checked by the defined fall angle  $\beta$ . When a person is walking,  $\beta$  varies from  $45^\circ$  to  $90^\circ$  (depends on their style and speed of walking). But after a fall, the  $\beta$  changes very little, so  $\sigma_\beta$  (standard deviation of  $\tan \beta$ ) for ten frames will be low.

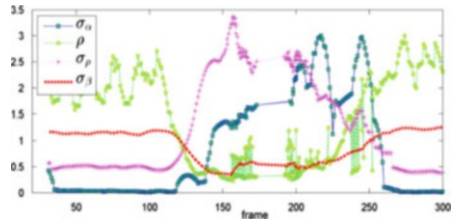
After a fall is pre-detected, if  $\sigma_\beta < 0.7$ , a fall is confirmed.

## 75.4 Experiment Results

Our algorithms implemented in the project are written in C++ using the OpenCV library. Our system is designed to work with a single uncalibrated camera and it can work in real time using a webcam. The camera can be placed with horizontal viewing angle or a wild angle of more than  $70^\circ$ .

Figure 75.2 shows some examples and detection results on various situations. In our testing project, when a fall is confirmed, we put out the text in the frames.

**Fig. 75.3** Variation of the coefficients during a fall



**Table 75.1** Detected results

	Detected	Not detected	Sensitivity/specificity (%)
Falls (sensitivity)	23	2	92.0
Daily activities (specificity)	4	19	81.8

**Table 75.2** Comparison of fall detection approaches in terms of sensitivity and specificity

	Our method (%)	The shape analysis [5] (%)	Posture analysis [13] (%)
Falls (sensitivity)	92.0	84.0	88.0
Daily activities (specificity)	81.8	82.6	78.2

**Table 75.3** Comparison of fall detection approaches in terms of time complexity and time consumption

	Time complexity <sup>a</sup>	Time consumption per frame (ms)
Our proposed method	$O(n)$	60
The shape analysis [5]	$O(n)$	80
The posture analysis [13]	$O(n \log n)$	–

<sup>a</sup> $n$  is the total number of pixels in one frame

In Fig. 75.2a when walking normally,  $\sigma_\alpha$  and  $\sigma_\rho$  are low; meanwhile,  $\rho$  is high, so no fall is detected. Figure 75.2b shows a person falling perpendicularly; it is obvious that  $\sigma_\alpha$  and  $\sigma_\rho$  are high and above the fixed thresholds; meanwhile,  $\rho$  is low, but  $\sigma_\beta$  is a little high, so a fall is pre-detected; after some frames  $\sigma_\beta$  becomes low enough, and a fall is confirmed and it is shown in

Figure 75.2c. A person is crouching normally in Fig. 75.2d;  $\sigma_\alpha$  and  $\sigma_\rho$  are higher than those walking normally, but they are below the fixed thresholds, so no fall is detected. A parallel falling happens in Fig. 75.2e; we notice that  $\sigma_\alpha$  and  $\sigma_\rho$  are above the fixed thresholds and  $\rho$  is low, so a fall is pre-detected; and then after some frames  $\sigma_\beta$  becomes low enough, and a fall is also confirmed, just as Fig. 75.2f shows.

The thresholds we choose are based on practice. Figure 75.3 shows an example of video sequence of a perpendicular fall, which is the same sequence as in Fig. 75.2b. It's obvious that when the person is falling, at about 140th frame,  $\sigma_\alpha$

and  $\sigma_\rho$  are high and  $\rho$  is low. The thresholds of  $\sigma_\alpha$  and  $\sigma_\rho$  are set to 1.0 and 1.5; meanwhile, the threshold of  $\rho$  is 1.5. After a fall,  $\sigma_\beta$  is still low, so a fall is confirmed. 0.7 is chosen as the threshold of  $\sigma_\beta$ . When the person starts to stand up about the 235th frame,  $\sigma_\alpha$  is still high, while  $\sigma_\rho$  is not as high as that when the person is falling due to its motion intensity, but  $\rho$  becomes higher; what's more  $\sigma_\beta$  is higher than 0.7, so no fall is detected.

To verify the feasibility of the proposed approach, we have used 45 video clips as our test targets, which include 23 daily normal activities (walking, running, crouch, sitting down) and 25 falling video clips (forward falls, backward falls, perpendicular falls, parallel falls). Sensitivity and specificity are obtained according to the method described by B. Töreyn [10].

Table 75.1 shows that our system gives very good results. We get a good detection rate with a sensitivity of 92.0 % and an acceptable rate of false detection with a specificity of 81.8 %. The specificity is not very high due to one brutal sits-down activity on a sofa with rebounding, one brutal crouches-down activity, and one brutal squat-down activity were detected as falls, which have large enough changes in motion detections and box ratio.

We compare our proposed method with the shape analysis approach using ellipse [5] and posture analysis technique [13]. Our proposed method can achieve better performance in terms of sensitivity and time consumption, despite the relatively low specificity, as Tables 75.2 and 75.3 show. The detection rate of the shape analysis [5] is lowest because of the instability of motion coefficient. Posture analysis [13] has high sensitivity, but low specificity, due to the combination of two different analyses used. The posture analysis [13] has the highest time complexity, which means more time consumption, while our method is less time-consuming than the shape analysis approach [5] due to less computation; the data were obtained with a PC of Intel i3 3.3 GHz CPU and 4G RAM.

## 75.5 Conclusion

In this chapter, we propose a new real-time fall detection approach based on global orientation and human shape. The global orientation performs well in our work. Our approach of fall detection has been proven to be robust and of real time through testing on realistic image sequence of simulated fall and daily activities.

**Acknowledgments** This work is supported in part by the National Natural Science Foundation of China (Grant no. 11176016, 60872117).

## References

1. Popoola, O. P., & Wang, K. (2012). Video-based abnormal human behavior recognition—A review. *Systems, Man, and Cybernetics-Part C: Applications and Reviews*, 42(6), 65–878.
2. Mubashir, M., Shao, L., & Seed, L. (2013). A survey on fall detection: Principles and approaches. *Neurocomputing*, 100(16), 144–152.
3. T. Lee, A. Mihailidis. An intelligent emergency response system: preliminary development and testing of automated fall detection. *Journal of telemedicine and telecare*. 11(4), 194-198 (2005)
4. Nait-Charif, H., McKenna, S. J. (2004). Activity summarization and fall detection in a supportive home environment. *Proceedings of the 17th International Conference on Pattern Recognition. IEEE Computer Society, Washington DC* (Vol. 4, pp. 323–326).
5. Rougier, C. (2007). Fall detection from human shape and motion history using video surveillance. *Proceedings of 21st International Conference on Advanced Information Networking and Applications Workshops. IEEE Computer Society, Washington DC* (Vol. 2, pp. 875–880)
6. Yu, M., Naqvi, S. M., Chambers, J. (2009). Fall detection in the elderly by head tracking. *Proceedings of IEEE/SP 15th Workshop on Statistical Signal Processing. IEEE, Cardiff Wales* (pp. 357–360).
7. Rougier, C. (2012). 3D head tracking for fall detection using a single calibrated camera. *Image and Vision Computing*, 31(3), 246–254.
8. Zivkovic, Z. (2004). Improved adaptive Gaussian mixture model for background subtraction. *Proceedings of the 17th International Conference on Pattern Recognition. IEEE Computer Society, Washington DC* (Vol. 2, pp. 28–31).
9. Ahad, M. A. R. (2012). Motion history image: Its variants and applications. *Machine Vision and Applications*, 23(2), 255–281.
10. Töreyn, B., Dedeoğlu, Y., & Cetin, A. (2005). HMM based falling person detection using both audio and video. *Computer Vision in Human-Computer Interaction*, 3766, 211–220.
11. Vishwakarma, V., Mandal, C., & Sural, S. (2007). Automatic detection of human fall in video. *Pattern Recognition and Machine Intelligence*, 4815, 616–623.
12. Flusser, J., & Kautsky, J. (2010). Implicit moment invariants. *International Journal of Computer Vision*, 86(1), 72–86.
13. Chen, Y. T., Lin, Y. C., Fang, W. H. (2010). A hybrid human fall detection scheme. *Proceedings of IEEE International Conference on Image Processing. IEEE, Hong Kong*. pp. 3485–348

# Chapter 76

## The Classification of Synthetic Aperture Radar Oil Spill Images Based on the Texture Features and Deep Belief Network

Xixi Huang and Xiaofeng Wang

**Abstract** This chapter introduces a new method to classify the SAR oil spill images. That is Deep Belief Network (DBN). Through the experimental certification, it is shown that the SAR images' information extracted by Gray-Level Co-occurrence Matrix (GLCM) can have a better effect in classification than that extracted by Gabor wavelet features. And using DBN to classify 240 samples including oil slick, looks-like oil slick and seawater, we can reach high total classification accuracy up to 91.25 %. Finally, we get a result that the method of DBN with GLCM features can better meet the needs of the SAR oil spill images classification.

### 76.1 Introduction

The Satellite SAR images of oil spill are rich in texture information that can be used to increase the image classification accuracy. Many researchers use neural network (NN) to classify the SAR images [1–3]. However, the complexity of the NN system and the high theoretical knowledge requirements restrict the widely application of network. It is mainly because of the dimension and structure of the network, the depth of the training set, and some other aspects of NN [4].

In this chapter, we use a new model named Deep Belief Network (DBN) for the classification of sea SAR image. DBN classifier is composed of a number of Restricted Boltzmann Machines (RBMs). After training each group of RBMs from the bottom up to form the hierarchy, DBN can solve the problem that the complex structure in traditional NN and its depth makes it not easy to be trained.

---

X. Huang (✉) • X. Wang

Information Engineering College, Shanghai Maritime University, Shanghai 201306, China  
e-mail: [dreamhk5001@gmail.com](mailto:dreamhk5001@gmail.com); [xfwang@cie.shmtu.edu.cn](mailto:xfwang@cie.shmtu.edu.cn)

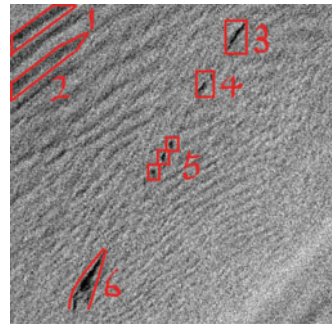
## 76.2 Experimental Data of SAR Oil Spill Images

In this chapter, the experimental data of SAR oil spill images comes from Infoterra global oil spills database. We select 95-King image of China Sea from ERS-1 and ERS-2 satellite with a resolution of 100 m. As shown in Fig. 76.1, zone 1 and zone 2 are looks-like slick areas; zone 3, zone 4, zone 5, and zone 6 are the real slick contaminated areas.

SAR images in this chapter are a series of preprocessed  $600 \times 600$  pix SAR oil spill images which have already been corrected, de-noised, and filtered. Because the SAR can cover very wide areas in high resolution, even a large area of sea still looks small in the SAR images. So the SAR images need appropriately and reasonably cut into  $40 \times 40$  pix images of oil slick, looks-like oil slick and sea water. There are 1,200 samples in all, and each kind of different category has 400 samples. Each image is saved as one-dimensional gray-level image. Then use histogram equalization to enhance the contrast of the gray-level of SAR images in the storing process of the three kind of classified images. In order to ensure that samples are randomly assigned, data are stored randomly, and will be disrupted before selecting the training samples and testing samples.

Before using DBN to classify the SAR images, we need to extract the texture feature information of the preprocessed images. In this chapter, we use Gabor wavelet [5–7] and Gray-Level Co-occurrence Matrix (GLCM) [8, 9] as the auxiliary texture feature information, and then use DBN to classify the feature information extracted from SAR image.

- (1) Gabor wavelet feature extraction: Create a Gabor filter with 5 scale by 8 orientation, convolving to 40 transformed images. Decimate the transformed images with the sampling rate at 8, and then we can get a feature matrix with 640-dimensional feature, denoted by  $F_{\text{gabor}}$ .
- (2) GLCM feature extraction: Extract energy, entropy, contrast, correlation, angular second moment, dissimilarity at four directions as  $0^\circ$ ,  $45^\circ$ ,  $90^\circ$ ,  $135^\circ$ . Use the value, mean, and variance of the six features to get a 36-dimensional feature matrix, denoted by  $F_{\text{glcm}}$ .



**Fig. 76.1** Oil spill image of SAR



Through the experiment we find that using GLCM for classification without setting learning rate can lead to a better result. And momentum has little effect on classification and the variation trend is more stable. When we use Gabor wavelet for classification, it has a bigger impact on learning rate and momentum, the result of classification swings great. So in this chapter, we choose GLCM as the extraction feature for DBN. The follow-up study is based on this.

### 76.3 Classification Process Based on Texture Features and DBN

#### 76.3.1 DBN

DBN is consisted of several RBMs. Its depth depends on the demand of experiment object on each layer. Actually three-layer DBN is enough, and Fig. 76.2a shows a DBN model. RBM is consisted of Markov Random Field with a visual layer and a hidden layer with no connection between the nodes in the same layer. Figure 76.2b shows a single-layer RBM structure.

The complication of RBM can be achieved by increasing the hidden layer nodes or superimposing a hidden layer [10]. The model can not only simulate original data but also process those representative data after pretreatment, such as the extracted features set.

Construct the training data as RBM network with visual layer and hidden layer. The pixel of image corresponds to the nodes of visual layer in the RBM network [11]. The features will be detected corresponds to the node of hidden layer [12]. A joint configuration, (v, h) of visible and hidden layer has an energy given by:

$$E(v, h) = - \sum_{i \in \text{visible}} a_i v_i - \sum_{j \in \text{hidden}} b_j h_j - \sum_{i,j} v_i h_j w_{ij} \tag{76.1}$$

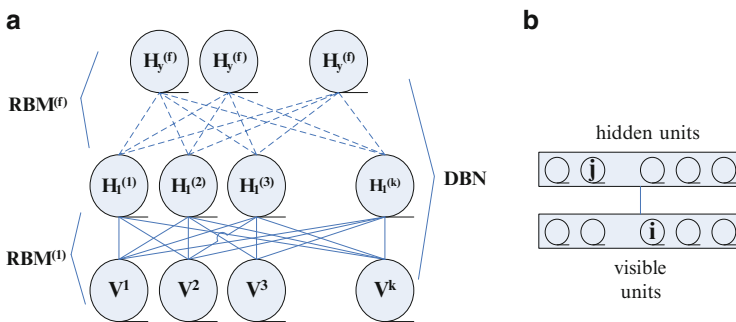


Fig. 76.2 The model of DBN and RBM

where  $v_i$  and  $h_j$  are the state of visible unit  $i$  and hidden unit  $j$ ,  $a_i$ , and  $b_j$  are their biases. And  $w_{ij}$  is the weight between them. According to the energy formula, the network assigns a probability to every possible pair of visible and hidden layer:

$$p(v, h) = \frac{1}{Z} e^{-E(v, h)} \quad (76.2)$$

where “partial function”  $Z$  is the sum of all possible pairs of visible and hidden layers:

$$Z = \sum_{v, h} e^{-E(v, h)} \quad (76.3)$$

Due to the nodes are independent from each other in one layer in RBM, the biases of  $\langle v_i h_j \rangle_{\text{data}}$  is easy to calculate. Randomly select a training sample  $v$ , the state  $h_j$  of each hidden layer is set to 1 with probability:

$$p(h_j = 1 | v) = \sigma \left( b_j + \sum_i v_i w_{ij} \right) \quad (76.4)$$

Given a hidden vector, then get an unbiased sample of the state of a visible layer:

$$p(v_i = 1 | h) = \sigma \left( a_i + \sum_j h_j w_{ij} \right) \quad (76.5)$$

Where the logistic sigmoid function is as follows:

$$\sigma(x) = \frac{1}{(1 + \exp(-x))} \quad (76.6)$$

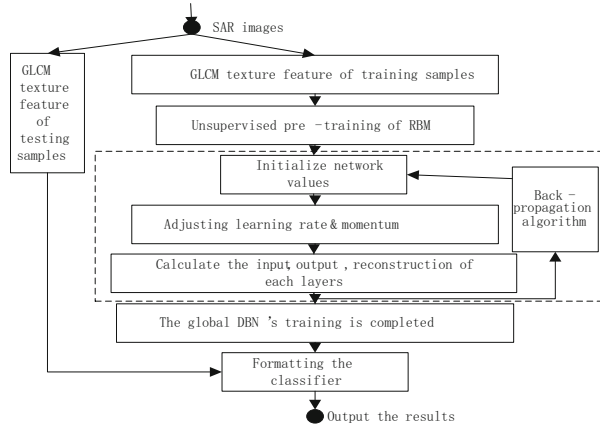
The change in a weight is then given by:

$$\Delta w_{ij} = \varepsilon \left( \langle v_i h_j \rangle_{\text{data}} - \langle v_i h_j \rangle_{\text{recon}} \right) \quad (76.7)$$

Where  $\langle v_i h_j \rangle_{\text{data}}$  is model value and  $\langle v_i h_j \rangle_{\text{recon}}$  is the “reconstruction” value.

DBN is a generative deep learning structure which is widely used [13]. It can solve the training problems in traditional multilayer neural network: (1) the requirement of huge amount of training set with label; (2) slower convergence rate; and (3) risk of falling into partial optimum due to an improper parameter selection [14].

**Fig. 76.3** Training and classification flow of DBN algorithm based on texture features



### 76.3.2 Learning and Classification Process of DBN Algorithm

In this chapter, we use 1,200 samples of SAR images. Choosing the quantity of training samples and test samples is also a part of what we will discuss in this chapter. So the choice of the quantity is gotten after the comparison through the experiment.

During the experiment, we firstly do the unsupervised pretraining on the bottom layer of RBM to get the initial value of the network. Then we use GLCM texture features to adjust the learning rate and momentum parameter of the DBN. Secondly, based on GLCM texture feature, we select an appropriate training number by comparing the results of different training number used in DBN. Finally, we get the most suitable quantity of the training samples by comparing the classification accuracy of DBN on the basis of the number of training samples.

Figure 76.3 describes the training process of DBN algorithm and the flow of how to use the trained DBN for classification.

- (1) Use the normalized features as the input data for unit  $V$  in the lowest layer and do unsupervised pretraining in RBM module to get the initialize parameter for network. Then adjust the learning rate and momentum.
- (2) The output  $h_i$  of Group  $K-1$  in RBM's hidden layer (Layer  $H$ ) becomes the  $k$  layer's input  $v_i$  in RBM's visible layer (Layer  $V$ ) and take the training in Layer  $K$ . Calculate the input, output, and reconstruction of each intermediate layer. Each hidden layer in RBM model is considered to be a feature detector and will be used to extract the representative features.
- (3) Repeat steps (2) to train each RBM group. When the last group of RBM in DBN is finished, do a back-propagation algorithm to fine-tune the overall DBN until getting a global optimal solution for the network. Then the training of DBN is complete and DBN becomes a classifier.
- (4) Put the testing GLCM texture features of training set into DBN, and finally calculate the classification results.

### 76.4 Result Analysis

From Sect. 76.2, we decide to use GLCM extracted from SAR images for DBN's classification. So now we go to analysis the experiment result coming from GLCM and DBN.

We can learn from Fig. 76.4 that with the increase of iteration times, the classification accuracy will be gradually improved and the trend goes more and more smooth. When the number comes up to 500 and 1,000, the average classification accuracy basically remains the same. So if the quantity of training samples is huge enough, we could choose a relatively small number of iterations, in order to reduce the amount of calculation as well as improve the calculation speed.

It can be seen from Fig. 76.5 that with the increase of training sample quantity, the average classification accuracy increases. In this chapter, we made a comparison of 40–80 % sample quantity at different training number. When the sample quantity is certain, and the training number is less than 1,000, the average accuracy

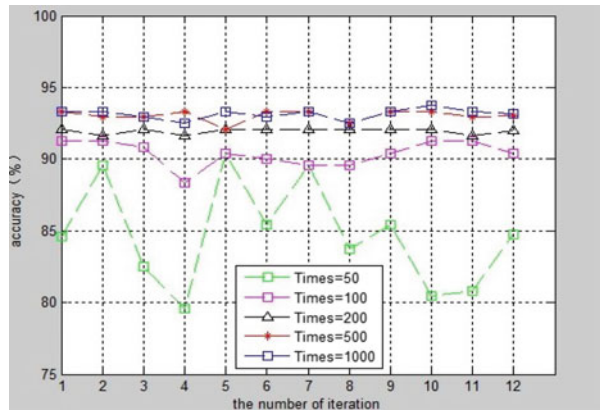


Fig. 76.4 Iterations impact on the result of classification

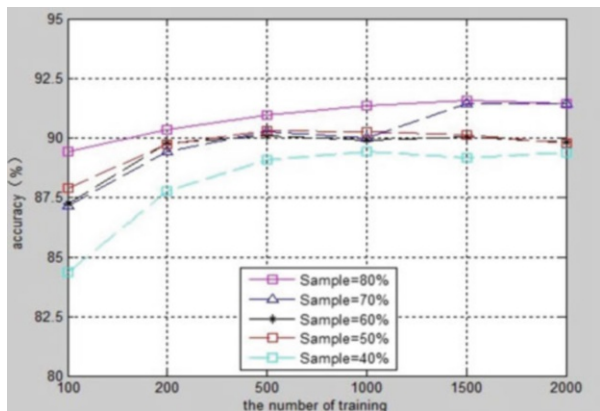


Fig. 76.5 Comparison of classification accuracy under different sample quantity by GLCM

**Table 76.1** Total comparison of the classification error statistics of 240 test samples

Category	Oil slick	Looks-like oil slick	Sea water	Total	Classification accuracy (%)
Oil slick	74	6	0	80	92.5
Looks-like oil slick	1	74	5	80	92.5
Sea water	0	9	71	80	88.75
Total	75	89	76	240	91.25

shows a process of surge. With the continued increasing of training time, the average accuracy remains unchanged, even shows a downward trend. When training number is certain, the increasing of training sample quantity will bring a relatively increase of average accuracy. So we come to the result that the effect of classification is most ideal when using 80 % sample quantity with 1,000 times of training for this experiment.

Take the data of oil slick, looks-like oil slick and seawater in Table 76.1 as Matrix A, calculate the total classification accuracy and Kappa coefficient. In the equation,  $a_{ij}$  means diagonal elements and correct classification number of A, N means the total amount of all types of samples.  $T_{i\bullet}$  and  $T_{\bullet j}$  equal to the sum of row i and column j. Total classification accuracy is different from Kappa coefficient in which total classification accuracy only takes correctly classified sample accuracy into account, but Kappa coefficient considers not only the samples which is correctly classified but also the incorrect and missing ones. The result of total classification accuracy (CA) and Kappa coefficient is shown below:

$$CA = \frac{\sum a_{ii}}{N} = 91.25\%, \text{Kappa} = \frac{N \cdot \sum_{i=1}^r a_{ii} - \sum_{j=1}^r T_{\bullet j} \cdot T_{i\bullet}}{N^2 - \sum_{i=1}^r T_{\bullet j} \cdot T_{i\bullet}} = 0.9205,$$

We can get the total classification accuracy is 91.25 %, respectively, is 92.5 % for oil slick, 92.5 % for looks-like oil slick, and 88.75 % for seawater which are correctly classified. The Kappa coefficient is 0.9205 with a high consistency. Although the low classification accuracy of sea water makes some effect on the final total classification accuracy. But being practical, sea water and looks-like oil slick belongs to nature phenomenon rather than man-made pollution caused by subjective factors in later stage. So we can use it only as a reference data. Thus, we can see that the DBN can meet the needs of the classification of SAR oil spill images in sea oil spill affairs, with a classification rate over 88 %.

## 76.5 Conclusion

In this chapter, we studied the classification method of SAR oil spill images based on the GLCM and DBN. According to the features of SAR images, we separately extract the GLCM and Gabor wavelet. Through experiment, we find that GLCM has a better effect in classification. When we use 80 % sample quantity and take the training 1,000 times, the classification of SAR images has a better effect by using GLCM and DBN. The classification accuracy is 4.5 % higher than choosing Gabor wavelet as features. Meanwhile we propose a new deep learning module named DBN. By cutting images into pieces, training batch by batch, each piece is a unit of RBM module, and the depth of learning can be reasonably added according to the actual needs. Eventually, we get total classification accuracy as 91.25 % when using 240 test samples to do the test. The classification accuracy of all category are higher than 88 %, and Kappa coefficient reaches 0.9205, which means DBN shows good classification of oil slick, looks-like oil slick and sea water.

To some extent, the associative level abstraction mode and fast learning method used in DBN are breakthrough of complexity and low computing efficiency of the traditional neural network. But in practices, image processing and recognition is a large-scale project, we still need a lot of further research and exploration to extract an efficient and executable computing ability based on the deep learning.

## References

1. Lijian, S., Chaofang, Z., & Peng, L. (2009). Oil spill identification in marine sar images based on texture feature and artificial neural network. *Periodical of Ocean University of China*, 39(6), 1269–1274. In Chinese.
2. Wencheng, X., Chuanqing, W., & Bin, W. (2008). Oil spill detection with SAR in South Korea's oil leak. *Remote Sensing Technology and Application*, 23(4), 410–413. In Chinese.
3. Topouzelis, K., Karathanassi, V., & Pavlakis, P. (2008). Dark formation detection using recurrent neural networks and SAR data. *International Journal of Remote Sensing*, 29(26), 4705–4720.
4. Marghany, M., Hashim, M. (2011). Comparative algorithms for oil spill detection from multi mode RADARSAT-1 SAR satellite data (pp 318–329). Computational Science and Its Applications-ICCSA, Santander, Spain: Springer Link.
5. Luo, X., & Xiaojun, W. (2011). Detection algorithm for infrared small and weak targets based on wavelet transform and Gabor filter. *Infrared and Laser Engineering*, 40(9), 1818–1822. In Chinese.
6. Jin, Q., Tong, X., & Bo, S. (2012). Study of Gabor feature selection for iris recognition. *Computer Engineering and Applications*, 48(19), 201–204. In Chinese.
7. Yi, N., Chen, D., & Yue, W. (2012). Improved Gabor wavelet transformation feature extraction method. *Computer Engineering*, 38(15), 145–147. In Chinese.
8. Zhou, Y., Lin, M., & Ma, T. (2010). SAR feature analysis of oil spill based on GLCM. *Marine Science Bulletin*, 29(4), 455–458. In Chinese.
9. Qing, X., Zheng, J., & Chen, Y. (2011). Application of texture analysis to identification of oil spills in SAR images. *Journal of Hehai University (Natural Sciences)*, 39(5), 569–574. In Chinese.

10. Nair, V., & Hinton, G. E. (2009). Implicit mixtures of restricted Boltzmann machines. *Neural Information Processing Systems*, 21(1), 1145–1152.
11. Hinton, G. E. (2010). Learning to represent visual input. *Philosophical Transactions of the Royal Society*, 365(1537), 177–184.
12. Hinton, G. E., Osindero, S., & Teh, Y. W. (2006). A fast learning algorithm for deep belief nets. *Neural Computation*, 18(7), 1527–1554.
13. Sun, Z., Xue, L., & Yangming, X. (2012). Overview of deep learning. *Application on Research of Computers*, 29(8), 2807–2810. In Chinese.
14. Mohamed, A., Dahl, G. E., & Hinton, G. E. (2012). Acoustic modeling using deep belief networks. *IEEE Transactions on Audio, Speech, and Language Processing*, 20(1), 14–22.

# Chapter 77

## The Ground Objects Identification for Digital Remote Sensing Image Based on the BP Neural Network

Shengkui Cao, Guangchao Cao, Kelong Chen, Chengyong Wu, Tao Zhang, and Jie Yuan

**Abstract** Spectral information of ground objects target in remote sensing image is complex, more noise, and highly nonlinear. It makes traditional data processing method no longer significant, effective, and efficient. The BP neural network classification-recognition method provides a more ideal solution. Using the TM remote sensing images as the example, this paper experimented the application of the BP neural network to the remote sensing image classification and recognition. Results showed that the classification precision of cultivated land was very low for both the BP neural network and traditional maximum likelihood methods because the spectrum difference between the new cultivated land and the bare land having low plant covered in this area was not significant. Maximum likelihood method wrongly regarded the bare land which had higher soil moisture content by lakeshore as water body. Except the grassland, the classification effect of the BP neural network was superior to maximum likelihood method. The overall classification accuracy by the BP neural network reached 81.79 %; however, the one by the maximum likelihood method was 79.08 %, indicating that the BP neural network classification and recognition was superior to the traditional maximum likelihood method.

### 77.1 Introduction

Recorded abundant surface spectral information by means of the electromagnetic waves characteristics is reflected or emitted by the ground objects. Remote sensing images provide the appropriate data for quick access to surface features relating to the space scale in human production and living. The remote sensing image represents the differences of the ground objects through pixel high or low value

---

S. Cao (✉) • G. Cao • K. Chen • C. Wu • T. Zhang • J. Yuan  
College of Life and Geography Science, Qinghai Normal University, Xining 810008, China  
e-mail: [243263340@qq.com](mailto:243263340@qq.com)



difference (reflecting spectral information of the ground objects) and spatial variation (reflecting the spatial information of the ground objects). Those are the physical basis to distinguish different images of the ground objects.

The recognition of the remote sensing images is the process of analyzing the spectral and spatial information of various ground objects in remote sensing image, selecting the features, and dividing each pixel in the image into different categories by means of the computer. Then to achieve the classification and recognition of the remote sensing images, we obtain the corresponding information of remote sensing image with the actual ground objects in order [1]. However, due to the complex of spectral information about the characteristics within the internal ground objects target, the effect of traditional data processing methods (such as parallel pipeline algorithm and the maximum likelihood method) is no longer significant [2]. With the diversification of network types, improvement of network structure, as well as perfect network learning algorithm, a parameterized artificial neural networks (ANNs) model family is gradually applied to remote sensing image recognition [3–5]. Moreover, the BP neural network and its variable forms are the most widely applied and successful [6, 7]. This article explored the effectiveness and advantages of the BP neural network classification method through the classification and identification of the land use of a certain local area of Hainan Prefecture in Qing province using the TM images as data, through comparison of the results classified and recognized by the traditional classification method and the BP neural network one.

## 77.2 Studying Methods

### 77.2.1 *Introduction of the BP Neural Network*

The BP (Back Propagation) network was proposed by Rumelhart and McClland in 1986. It is currently one of the most widely used neural network models. To conduct the error back propagation algorithm as its learning rules that processes the feedforward networks of supervised learning, it requires a considerable number of known samples to learning and training, to find out and remember the relationship between the input sample models and classification categories. It is usually need to be set that the features conditions of the ground objects will be classified, and it acts as a BP network input mode, then, we give the desired output mode, which plays the role in prediction type. Its structure includes input layer, hidden layer, and output one. The neurons in the same layer cannot interconnect, but the neurons in the different layers need full interconnection. The weights of the neural network are connected through a number of neurons (computing elements) by means of the feedforward or feedback. Those neurons lied in the hidden layer, and the input and output layers are connected by them [8].

### ***77.2.2 The BP Neural Network Algorithm***

The BP neural network algorithm is described as follows:

1. Initialized the network parameters: learning parameters, such as setting up a network initial weight matrix and learning factor
2. Provided training mode: not training the network until it meets the learning requirements
3. Forward propagation process: for the given training mode input, calculating output mode of the network and comparing with the expected pattern. If there was an error, step (4) would be executed; otherwise, and the program would be returned the step (2)
4. Backward propagation process: calculating the error in the same layer unit and correcting weights and thresholds; back to step (2)

### ***77.2.3 Selecting the Number of Hidden Layer Neurons***

When the network mapping is achieved using neural network, the number of hidden layer neurons directly affects the ability of learning and induction of the neural network. The smaller the number of hidden layer neurons, the shorter is each network's learning time. However, because of insufficient learning, the network could not remember all the learning content. When the number of hidden layer neurons is larger, the learning ability strengthened, but along with the longer times learning and the larger storage capacity of the network, resulting in the decline of inductive capacity of the network for the unknown input [9].

## **77.3 The Implementation of the Remote Sensing Image Classification and Recognition**

The BP neural network is used to recognize the feature category of remote sensing image. The main process is divided into sample training and pattern recognition.

### ***77.3.1 Training Sample***

Its purpose is to find out certain error conditions of weight matrices [10]. The essence is the input layer of neural network structure. Input items are values of the pixel ( $R_i$ ,  $G_i$ ,  $B_i$ ) that are known classification results. That is usually achieved through selection and drawing the training area of typical ground objects in the remote sensing image. Expected output is the classification-recognition

results in correspondence to each pixel, namely, the information categories of the ground objects in the study area.

### **77.3.2 *Pattern Recognition***

After the sample training is finished, weight matrix has met the requirements. This moment, put any pixel on the image (R, G, B) value as input, the output vector can be obtained by calculation. Vector component corresponds to probability value of classification types that the pixel appoints every ahead of time. The type corresponding to the biggest probability value is the type belonging to the pixel.

## **77.4 Application Examples**

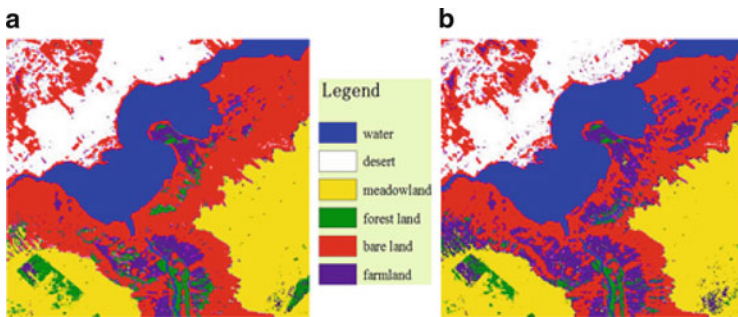
### **77.4.1 *The Experimental Data and Its Pretreatment***

Experimental data chose the TM remote sensing images at the Hainan Prefecture in the southeastern region of Qinghai province in July 2009 (Fig. 77.1). The ground objects intercepted were complex and varied terrain category. The study area where can more clearly reflect the land cover situation is selected. The TM images were selected six bands of TM1, TM2, TM3, TM4, TM5, and TM7, with the spatial resolution of all 30 m. The picture size is 157,994 pixels, and picture contrast is good, without a shadow of a cloud. Data preprocessing mainly included the atmospheric correction, geometric correction, and image enhancement. The atmospheric correction chose the histogram adjustment method; geometric correction used RTK measurement data as control points and used quadratic polynomial to do coordinate transformation and gray resample based on the nearest neighbor method and the precision control of the coordinate transformation within one pixel. Through the result of comparing the land use status, high-resolution remote sensing image, and the visual interpretation, this region can be divided into six types: the cultivated land, grassland, sandy land, bare land, forest land, and water body. To improve the spectral contrast of the land types in the remote sensing image, image enhancement was handled using a fast Fourier transform (FFT) and inverse FFT for weakening the noise in the image (Fig. 77.1).

### **77.4.2 *Sample Selection***

On remote sensing image, we set the neurons layer numbers of 6 input layer, 1 hidden layer, and 6 output layer using the means of human-computer interaction

**Fig. 77.1** The TM remote sensing image of experimental zone (5/4/3)



**Fig. 77.2** Classification results of different identification methods. (a) BP neural network classification. (b) Maximum likelihood classification

and the BP neural network algorithm. We selected multiple typical areas having the clear and determined types as the training samples (district), which each kind of sample cannot cross the boundaries of the different categories. According to the characteristics of the land types, eventually 152 cultivated lands, 744 grasslands, 589 sand lands, 600 bare lands, 130 woodlands and 1,052 water-body were selected as training samples. Because the sample quality had greater impact on the recognition accuracy, there required that all land types were homogeneous. Therefore, we need check the sample histogram. If the sample distribution closes to a normal one, it indicates that the sample quality is better. Otherwise, it needs to choose training sample.

The hidden layer number was determined by the constructor method, which validated the model prediction error one by one from the minimum until it reached the maximum value. At last, the hidden layer with the least model error was selected and set the number of 1. Set the training error  $E < 0.002$  in the classification. The result of the classification is shown in Fig. 77.2.

**Table 77.1** The precision comparing table of different classification methods

Class	Maximum likelihood method		BP neural network algorithm	
	Prod. acc.	Overall accuracy (%)	Prod. acc.	Overall accuracy (%)
Water	80.31	79.08	93.56	81.49
Sandy land	92.67		93.10	
Grassland	92.65		90.10	
Forest land	93.86		94.47	
Bare land	89.24		89.68	
Cultivated land	30.90		40.86	

## 77.5 Results and Analysis

The time of maximum likelihood classification is 3 s and the time of BP neural network classification is 56 s, which indicated that the time of BP neural network classification is obviously longer. Applying land use map at present as reference diagram, we evaluated the precision of maximum likelihood method and BP neural network classification through overall classification accuracy and mapping accuracy. The overall classification accuracy is equal to the value of correct classified pixel dividing the total pixel number. The mapping accuracy is equal to the value of pixel number of correct classification of each category dividing the number of the corresponding pixel on reference diagram.

Results showed that the classification precision of cultivated land was very low for both methods. The reason is that the spectrum difference between the new cultivated land and the bare land in this area is nonsignificant which leads to big error in classification. The reason for larger error is that the maximum likelihood method regarded the wetter bare land near the lakeshore as water body incorrectly. Except for the grassland, the classification effect of the BP neural network was superior to the maximum likelihood method based on the features of spectral information (Table 77.1). The overall classification accuracy by the BP neural network was higher than the one by the maximum likelihood method.

## 77.6 Conclusion

As a heuristic algorithm, the BP neural network used to improve the ability of nonlinear approximation. Under normal circumstances, the neural network speed is lower than the traditional remote sensing image classification method. Using neural network algorithm for remote sensing image classification and recognition, it could eliminate the ambiguity and uncertainty caused by the traditional image classification and improve the accuracy of classification. The handle of sample quality before the BP algorithm is very important. Gathering the sample, each ground object must be uniform in distribution in the remote sensing image. There needs to be a

gathering of 5–10 training area with at least greater than 100 pixels, and the sample histogram must be close to a normal distribution. Only in this situation, the image classification and identification become more reliable and accurate.

**Acknowledgements** This paper is cosponsored by National Social Science Foundation (10CJY015), Western Light Project of Chinese Academy of Sciences, Chunhui plan of the Ministry of Education (Z2012092), and Innovation Program of Qinghai Normal University (1294). The authors thank them and the anonymous reviewers.

## References

1. Yan, Y., Dong, X. L., & Li, Y. (2011). The Comparative study of remote sensing image supervised classification methods based on ENVI. *Beijing Surveying and Mapping*, 3, 14–16 (In Chinese).
2. Li, L. W., Ma, J. W., Ou, Y. B., & Wen, Q. (2008). High spatial resolution remote sensing image segmentation based on temporal independent PCNN. *Journal of Remote Sensing*, 12(1), 65–69 (In Chinese).
3. Simpson, J. J., & Mcintir, J. T. (2001). A recurrent neural network classifier for improved retrievals of areal extent of snow cover. *IEEE Transactions on Geoscience and Remote Sensing*, 39(10), 2135–2147.
4. Lin, J., Bao, G. S., Jing, R. Z., & Huang, J. X. (2002). A study of FasART neuro-fuzzy networks for supervised classification of remotely sensed images. *Journal of Image and Graphics*, 7(12), 1264–1268 (In Chinese).
5. Han, M., Cheng, L., & Tang, X. L. (2005). Application study of Fuzzy ARTMAP neural network in classification of land cover. *Journal of Image and Graphics*, 10(4), 416–419 (In Chinese).
6. Zhou, T. G., & Su, Y. C. (2004). Artificial neural networks-based study of vegetation's classification for aerial remote sensing image. *Journal of Southwest China Normal University (Natural Science)*, 29(6), 1037–1040 (In Chinese).
7. Yang, F. L., Liu, J. G., Zhao, J. H., & Du, Z. X. (2006). Seabed texture classification using BP neural network based on GA. *Science of Surveying and Mapping*, 31(2), 111–115 (In Chinese).
8. Xiu, L. N., & Liu, X. G. (2003). Current status and future direction of the study on artificial neural network classification processing in remote sensing. *Remote Sensing Technology and Application*, 18(5), 339–345 (In Chinese).
9. Lei, J. F., & Sun, J. Y. (2008). Research on image recognition based on artificial neural network. *Modern Electronic Technique*, 31(8), 127–130. 134 (In Chinese).
10. Chen, Y. M. (2002). Application of neural network to classification of remote sensing image. *Journal of Geomatics*, 27(3), 6–8 (In Chinese).

# Chapter 78

## Detection of Image Forgery Based on Improved PCA-SIFT

Kunlun Li, Hexin Li, Bo Yang, Qi Meng, and Shangzong Luo

**Abstract** In view of the problem existing in abusive using of image copy-move forgeries, this paper proposes an image forensics algorithm for detecting copy-move forgery based on improved PCA-SIFT. The present method works first by extracting features of an image and then reducing its dimensionality, and the method uses k-nearest neighbor to operate forgery detection. Owing to the similarity between pasted region and copied region, the descriptors are then matched between each other to seek for any possible forgery in images. Extensive experimental results are presented to confirm that the algorithm is able to precisely individuate the tampered image and quantify its robustness and sensitivity to image post-processing and offer a considerable improvement in time efficiency.

### 78.1 Introduction

With the availability and sophistication of digital imaging technology and image processing tools, it is possible to increasingly easily manipulate the image information. The authenticity and integrity of digital images are experiencing the serious threat and challenge. Therefore, image forensics is one of the key issues to be solved in maintaining information security.

Image forgery forensics techniques are generally divided into two categories: active methods and passive methods [1]. Active methods work when we have some prior information or its source about the image such as digital watermarking or signature. Passive (Blind) methods only make use of characteristics of images to identify image forgery which become the hot topics in the field of image forensics.

---

K. Li (✉) • H. Li • Q. Meng • S. Luo  
College of Electronic and Information Engineering, Hebei University, Baoding 071002, China  
e-mail: [likunlun@hbu.edu.cn](mailto:likunlun@hbu.edu.cn); [xinxinlanyi@163.com](mailto:xinxinlanyi@163.com)

B. Yang  
College of Mechanical Engineering, Yanshan University, Qinhuangdao 066004, China

Copy-move forgery is one specific type of image forgery where part of an image has been copied and pasted within the same image in an attempt to add or disguise feature in the scene [2]. It is very likely for the duplicated part to be subjected to a variety of post-processing operation. The detecting region duplication principle can be based on the similar characteristic presence of duplicated regions in the image.

There have been disappeared different techniques for detecting copy-move forgery. J. Fridrich [3] proposed an effective approach based on lexicographic sorting of quantized DCT coefficients of image blocks, which has the milestone significance, but the computational complexity is higher. B. Medan et al. [4] yield a fuzzy method to detect tampering based on blur moment invariants, which can resist fuzzy, noise, etc., but fail in the scale and rotated post-processing operation. M.A. Sekeh et al. [5] applied a coarse-to-fine approach to propose an efficient image duplicated region detection model using sequential block clustering, which improves time complexity. H.L. Huang et al. [6] detected image copy-move forgery using SIFT algorithm which suffers from the drawback that the search method is computationally complex, it not only reduces the image matching and retrieval speed but also affect the real-time performance.

In view of the problem existing in detection algorithms of the copy-move image forgeries, we proposed a novel image forensics algorithm for detecting copy-move forgery based on new PCA-SIFT, k-nearest neighbor to operate forgery detection.

The rest of the paper is organized as follows. In Sect. 78.2, we give a description of SIFT and PCA-SIFT algorithms and discuss the proposed method in detail. In Sect. 78.3, we apply our scheme to image forgery detection. Some experimental results and discussions are debated in Sect. 78.4, and conclusions are drawn in Sect. 78.5.

## 78.2 The Related Work

### 78.2.1 Review of the SIFT Algorithm

Scale Invariant Feature Transform (SIFT) algorithm [7] is essential to problems including image matching, object recognition. SIFT extracts distinctive features from images which are invariant to image scale and rotation and are robust to distortion, noise, illumination, and viewpoint. The major stages of computation are following.

Scale-space extrema detection: The first step searches over all scales and image locations. Given an input image  $I(x,y)$ , establish scale space  $L(x,y,\sigma)$  through Gaussian function  $G(x,y,\sigma)$  and implement the Difference-of-Gaussian (DoG) scale space  $D(x,y,\sigma)$  by using DoG function to identify potential interest points that are invariant to scale and orientation. Extrema are detected by comparing each sample point in DoG images to its 26 neighbors in  $3 \times 3$  regions at the current and adjacent scales in attempt to obtain candidate keypoints.



$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y) \quad (78.1)$$

$$\begin{aligned} D(x, y, \sigma) &= [G(x, y, k\sigma) - G(x, y, \sigma)] * I(x, y) \\ &= L(x, y, k\sigma) - L(x, y, \sigma) \end{aligned} \quad (78.2)$$

**Keypoint localization:** The next step performs a detailed model fit to determine location and scale. Candidate keypoints are localized to sub-pixel accuracy, reject keypoints with low contrast and eliminate edge responses to have stable ones.

**Orientation assignment:** The third step assigns the dominant orientations for each keypoint based on local image gradient directions to achieve invariance to image rotation. The keypoint orientation is calculated from an orientation histogram of local gradients from the closet smoothed image  $L(x, y, \sigma)$ . The gradient magnitude  $m(x, y)$  accumulate to the histogram according to orientation  $\theta(x, y)$ . An orientation histogram with 36 bins covering the  $360^\circ$  range of orientations is formed from the gradient orientations in the neighborhood of the keypoint. Peaks in the orientation histogram correspond to dominant directions of local gradients.

$$m(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))^2 + (L(x, y+1) - L(x, y-1))^2} \quad (78.3)$$

$$\theta(x, y) = \tan^{-1} \left[ \frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)} \right] \quad (78.4)$$

**Keypoint descriptor:** The final step builds a local image descriptor for each keypoint to guarantee descriptors being distinctive and robust to other variations, such as illuminations and viewpoint. Rotate the coordinate relative to the keypoint orientation to ensure rotation invariance, sample the magnitudes and orientations of the image gradient in a region around the corresponding keypoint, and accumulate orientation histograms with eight bins summarizing the contents over  $4 \times 4$  subpatches as the local statistics of gradient orientations to capture SIFT descriptor with 128 elements.

### 78.2.2 Brief of the PCA-SIFT

The algorithm for local descriptors (termed PCA-SIFT) [8] examines the local image descriptor used by SIFT. It can be summarized in the following steps: it consists of the highly restricted set of patches that passed through the first three stages of SIFT discussed above. So it accepts the same input as the standard SIFT: the sub-pixel location, scale, and dominant orientations of the keypoint; it explores alternatives to its local descriptor representation. Like SIFT, the descriptors encode the salient aspects of the image gradient in the feature point's neighborhood; however, instead of using SIFT's smoothed weighted histograms, the method extract the  $41 \times 41$  patch centered at the keypoint, and rotated to its dominant orientation, concatenate their horizontal and vertical gradient maps, then

the descriptor has  $2 \times 39 \times 39 = 3042$  elements, and PCA was applied to the normalized gradient patch to 20 elements. Details are given in the paper.

### 78.2.3 The Proposed Method

A novel method is presented to extract more challenging features. The method extract keypoints of the image based on SIFT to inherit stability of the SIFT, while refusing the traditional PCA-SIFT alternative to describing keypoints, it improves on SIFT by using PCA thoughts in a novel way in which different from traditional PCA-SIFT.

**Keypoint detection:** Given an image, we extract the keypoints  $X = \{x_1, x_2, \dots, x_n\}$  through three steps of SIFT algorithm. The first step identifies image scale-space extrema which robust to scale and rotation, preliminarily determine the position and the scale of the keypoints. The second step determines the position of the keypoints accurately and removes the low contrast keypoints and the unstable edge response to enhance the matching stability. The third step assigns orientations for each keypoint based on its local image patch. The assigned orientation(s), scale, and location to construct a canonical view for the keypoint that is invariant to similarity transforms.

**Keypoint representation:** Keypoint description using SIFT is created by sampling the magnitudes and orientations of the image gradient in the patch around the keypoint, and building smoothed orientation histograms to capture the important aspects of the patch. Then image descriptors  $D = \{d_1, d_2, \dots, d_n\}$  each with 128 elements are obtained.

The choices behind the SIFT descriptor is complicated, the descriptor matching will have heavy computation based on large eigenspace. Moreover, some characteristic of the eigenspace have very low contribution rate, they are susceptible to external interference, such as noise, blur. Principal Component Analysis (PCA) is used to reduce the eigenspace, simplify the dataset, and retain the identity-related variation while discarding the distortions induced by other effects. For well solve the problems above, it is not unreasonable to believe that the descriptor can be reasonably modeled by applying PCA to accurately represent them with a compact feature representation.

At first calculate the standardized data of input data  $d_i = d_i - \frac{1}{n} \sum_{i=1}^n d_i$ . Then the calculation of the covariance matrix should be done  $C = \frac{1}{n} \sum_{i=1}^n d_i d_i^T$ , and the eigenvalue and eigenvectors of the covariance matrix are extracted. Finally, sort the eigenvalues and select the first  $r$  eigenvalues as the main components, the corresponding eigenvectors to makeup matrix  $U$ . The transformation of input data that maps the data matrix  $D$  in a new matrix  $Y$  as  $Y = U^T D$ . Then the improved PCA-SIFT descriptors with  $r$  elements will be done.

### 78.3 Application to Image Forgery Detection

The proposed method is introduced into image forgery detection. Apply the simple and effective k-nearest neighbor using Euclidean distance as the similarity measure criteria to run image detection. Perform global traversal search to determine the matching decision by taking the ratio of the distance of the closest neighbor to that of the second closest one with respect to a global threshold. Given a keypoint  $x \in X$  with a similarity vector  $S = \{s_1, s_2, \dots, s_{n-1}\}$  that represents the distance with respect to the other descriptors. When  $s_1/s_2 < T$ , the keypoint is matched. Iterating on each keypoint in  $X$ , we can obtain the set of matched points which reveal the image forgery.

## 78.4 Experimental Results

### 78.4.1 Evaluation Metrics

Precision and accuracy are commonly used in evaluation standards, they are popular in evaluating various classification and detection methods, thereby we use them to evaluate detection performance of the new method, which are defined as follows:

$$\text{Precision} = \frac{\# \text{number of correct matches}}{\# \text{number of total matches}}$$

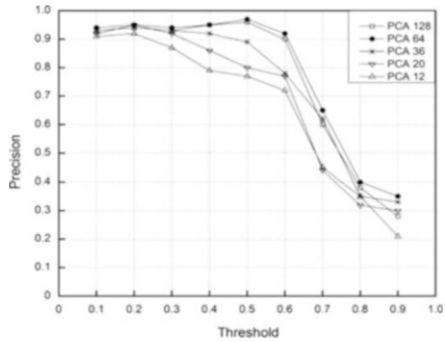
$$\text{Accuracy} = \frac{\# \text{images detected as forged being forged}}{\# \text{forged images}}$$

### 78.4.2 Test on a Large Dataset

In order to verify the effectiveness and stability of the proposed algorithm, the method has been tested on various images. Experiments are performed as follows.

Parameter determination: Figure 78.1 demonstrates that the performance of the improved PCA-SIFT as PCA dimension and threshold are various. Increasing the dimensionality of the space has a result in better accuracy, since the representation is able to capture the structure of the patch. Adding the threshold achieves better results too, but once it exceeds a certain size, the precision begins to decline. Note that the requirement for image description improves on image forgery detection. Using fewer components require less storage and faster matching. Thereby weighing the precision and detection time, this article chooses the threshold value  $T = 0.5$ , PCA space  $r = 64$ .

**Fig. 78.1** The performance of the improved PCA-SIFT as PCA dimension and threshold changes



**Fig. 78.2** The original image is shown in the first column; the tampered image is pictured in the second column; the detection results using SIFT and new algorithm are reported in the third and fourth column, respectively

Detection of image forgery with no post-processing: Comparison on the performance of SIFT [7] and new method in copy-move image detection are shown in Fig. 78.2. Threshold is set to 0.5. The new method shows the increased detection accuracy.

Detection of image forgery with post-processing: The dataset is composed by various post-processing in the copy-move forgery. Figure 78.3 shows that the new method can accurately detect and locate forgery region and has resistance ability for post-processing. An interesting phenomenon was found after repeated experiments, when the image is more distinct and the texture is stronger, it may bring about more keypoints and feature matching, the detection results will be better.

Detection accuracy: For statistical detection accuracy of SIFT and new method, each post-processing operations selected 200 tampering pictures including figures, landscapes, buildings, etc., and covering a variety of forgeries such as scale, rotation, and JPEG compression. Figure 78.4 presents that the detection accuracy of the new algorithm is superior to the SIFT algorithm. Furthermore, it greatly reduces the time complexity and improves the detection efficiency.

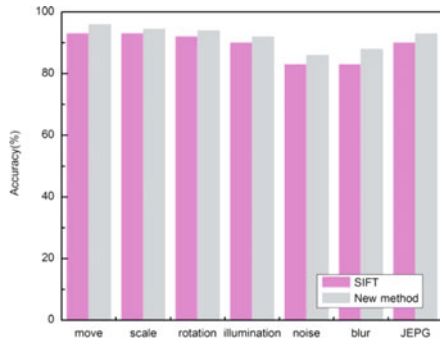
## 78.5 Conclusion

This paper proposed a novel algorithm for detecting copy-move forgery using the idea of improved PCA-SIFT. The presented method showed effectiveness with respect to detecting and locating forged images even if the images have undergone



**Fig. 78.3** Some examples of original images are shown in the first column; the tampered images (*rotation, scale, blur*) are pictured in the second column; the detection results are reported in the third column

**Fig. 78.4** Detection accuracy of SIFT and new method for each different attack (in percentage)



any post-processing. Furthermore, the algorithm also has lower computational complexity and it works well on all the basic image formats. It performed better than some of the previous methods. However, the algorithm still deserves further exploration to more accurately define the tampered region and enhance the robustness of the detection.

**Acknowledgements** Project supported by the National Science and Technology Support Plan Project (No. 2013BAK07B04) and Natural Science Foundation of Hebei Province, China (No. F2013201170).

## References

1. Gavin, L., Shih, F. Y., & Liao, H.-Y. M. (2013). An efficient expanding block algorithm for image copy-move forgery detection. *Information Sciences*, 239, 253–265.
2. Christlein, V., Riess, C., Jordan, J., Riess, C., & Angelopoulou, E. (2012). An evaluation of popular copy-move forgery detection approaches. *Information Forensics and Security*, 7(6), 1841–1854.
3. Jessica, Fridrich., David, Soukal., Jan, Lukas. (2003). Detection of copy–move forgery in digital images. *Proceedings of Digital Forensic Research Workshop (DFRWS'03)* (pp. 55–61). Cleveland, OH: IEEE Computer Society.
4. Babak, M., & Stanislav, S. (2007). Detection of copy–move forgery using a method based on blur moment invariants. *Forensic Science International*, 171(2–3), 180–189.
5. Mohammad Akbarpour, S., Mohd. Aizaini, M., Mohd. Foad, R., & Babak, M. (2013). Efficient image duplicated region detection model using sequential block clustering. *Digital Investigation*, 10(1), 73–84.
6. Hailing, Huang., Weiqiang, Guo., Yu, Zhang. (2008). Detection of copy-move forgery in digital images using SIFT algorithm. *Proceedings of IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application (PACIIA'08)* (vol. 2, pp. 272–276). Wuhan: IEEE Computer Society.
7. Lowe, D. G. (2004). Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, 60(2), 91–110.
8. Yan, Ke., Rahul, Sukthankar. (2004). PCA-SIFT: A more distinctive representation for local image descriptors. *Proceedings of 2004 I.E. Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04)* (vol. 2, pp. 506–513). Washington, DC: IEEE Computer Society.

# Chapter 79

## A Thinning Model for Handwriting-Like Image Skeleton

Shijiao Zhu, Jun Yang, and Xue-fang Zhu

**Abstract** In order to solve the letter skeleton problem with handwriting-like attributes, a thinning model is used in this paper. By introducing improved reservation and eliminating produces, additional pixels are constrained by thinning nearest pixels. In the experiment, the proposed method is compared with others in the literature English letters by using the one pass thinning algorithm (OPTA) and Hilditch methods. Empirical results show that the proposed model can thin handwriting-like skeleton in terms of reserving topology and eliminating extra pixels.

### 79.1 Introduction

Handwriting provides useful personal information that may be used to describe the property of an image [1]. While it is relative easy to describe print character, but difficult for personal handwriting character especially for widely use of handset devices. The first thing of reorganization is to preprocess images with thinning handwriting outlines.

Research in thinning methods has been interest in recent years due to the first step of retrieval features [2, 3]. Some algorithm has been used in some applications [4, 5]. An ideal thinning method has the properties of keeping connections, topology, and relative position. It is best to keep the thinning result outline in middle of handwriting skeleton. Some methods have been proposed, such as Hiditch [6], one pass thinning algorithm (OPTA) [7], and so on.

---

S. Zhu (✉) • J. Yang  
School of Computer and Information Engineering, Shanghai University of Electric Power,  
Shanghai 200090, China  
e-mail: [zhusj707@hotmail.com](mailto:zhusj707@hotmail.com)

X. Zhu  
Department of Information Management, Nanjing University, Nanjing 210093, China  
e-mail: [mediate@163.com](mailto:mediate@163.com)

But we all know that each method has its better performance based on its characteristic in special field. A widely used method is FAT (fast thinning algorithm) [8] which is based on eight nearest pixels. This method use local pixels properties. From the point compute iterate type, it is classified as serial algorithm and parallel algorithm. Serial algorithm is depended on preprocess order, while parallel algorithm is based on last produce. Therefore parallel algorithm is superior to serial one. For improved contour-based thinning method, it uses shape characteristics of text to get skeleton of nearly same as the true character shape [9]. Our proposed method is based on parallel algorithm.

In the paper that follows, we begin with the definition model in Sect. 79.2. Meanwhile, the produce and algorithm is presented. A framework of the thinning schema is discussed. In Sect. 79.3, we give experimental result and comparison of different algorithm. Section 79.4 briefly summarizes the paper.

## 79.2 An Improved Model

OPTA algorithm is a typical method for thinning binary image. It uses reservation and elimination templates to apply for an original image.

The whole produce banks of produce of thinning method are described in Fig. 79.1.

For original image, it is made by produce of binary, and then to removing superfluous pixels of skeleton by bank of keeping templates and removing templates.

The whole produce can be described as follow:

- (1) For current pixels, find nearest pixels in removing templates, if true then next, otherwise go to step (3);
- (2) Make judgment of keeping template, if pattern in keeping method then keep current pixel, otherwise make tag of removing and go to step (1);
- (3) No action for current pixel, go to step (1);
- (4) Process for next pixel;
- (5) Make judgment of the last time parallel produce, if has pixel removed, then continue next loop. Otherwise stop.

For handwriting skeleton, it has some of its own properties. Firstly, it has some subblocks of contiguous regions having similar characteristic such as color, intensity, or edges. Secondly, they are often vertical or horizontal alignment. Thirdly,

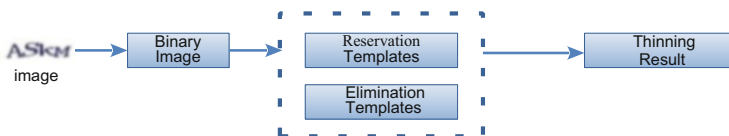
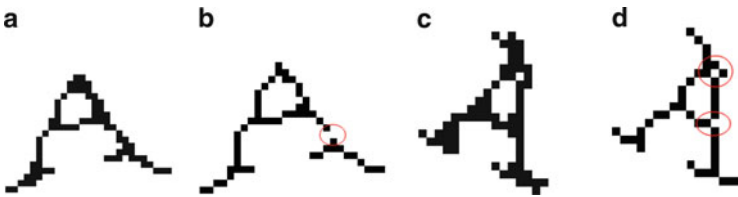


Fig. 79.1 Thinning produce



**Fig. 79.2** Elimination templates \* = 0/1

<b>a</b>	<b>b</b>	<b>e</b>	<b>f</b>
0 0 0	0 * 1	* 0 0	0 0 *
* ● *	0 ● 1	1 ● 0	0 ● 1
1 1 1	0 * 1	* 1 *	* 1 *
<b>c</b>	<b>d</b>	<b>g</b>	<b>h</b>
1 1 1	1 * 0	* 1 *	* 1 *
* ● *	1 ● 0	0 ● 1	1 ● 0
0 0 0	1 * 0	0 0 *	* 0 0



**Fig. 79.3** Two result of OPTA marked by circle. (a) Original image. (b) Broken line. (c) Another original image. (d) Close line

**Fig. 79.4** Pixel and its near pixels

p1	p2	p3	p4
p5	●	p6	p7
p8	p9	p10	p11
p12	p13	p14	p15

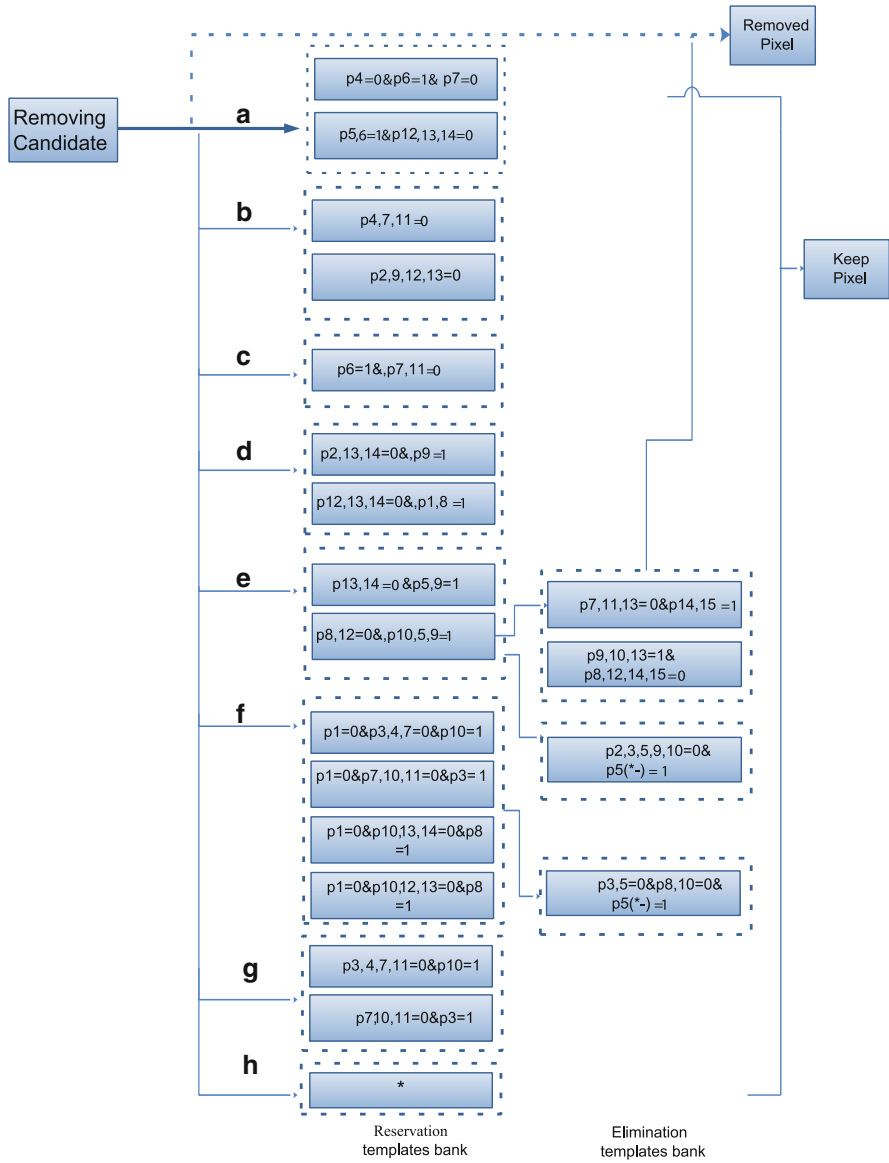
they are often been placed in center of an image. For above-stated reasons, we define some templates as follow (Fig. 79.2).

In paper of the thinning algorithms [2], it gives reservation templates. When pixel and its near pixels compared to the templates, if it meets condition, pixel is kept otherwise it is removed. From experiments, we find it has three negative sides. One is destroy topology structure for some outlines because it uses strict templates for constraint. Next one is invalid for closure structure. The last one is time consume for big size of reservation templates (Fig. 79.3).

In proposed method, we define keeping produce instead of keeping template. Firstly we define its position as symbolized letters as below:

The circle point in Fig. 79.4 is used to represent the location of the current point.

According to elimination templates, model gives candidate pixels with parallel computation. In the next step, reservation templates banks and with option of elimination templates bank was presented (Fig. 79.5).



**Fig. 79.5** Proposed bank of thinning. Where (\*-) symbolized as left pixel

The proposed method can be described as:

---

Binary image initial from image

---

Repeat until done:

- 1a. parallel processing current pixel and its nearest around pixels.
  - 1b. if judgment can meet removing templates and not meet keeping templates, then it is marked as erased pixel.
  - 2a. If judgment is empty, exit.
  - 2b. Make a new binary image based on pre-image and its judgment.
  - 2c. Use new binary image for next loop.
  - 3a. Update binary image.
- 

As mentioned, pixels of near candidate pixel are filtered by templates. In our implementation, these points are tagged by temporary array with the ability of parallel computation. The image pixels are, therefore, updated (pixel by pixel) at each iteration of the filtering algorithm.

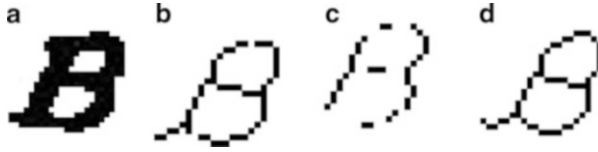
### 79.3 Experimental Results

In order to verify and analysis experiment, here we apply our algorithm to a variety of handwriting-like images. In each gray images, their size are  $160 \times 40$ . The computer used in the experiment:

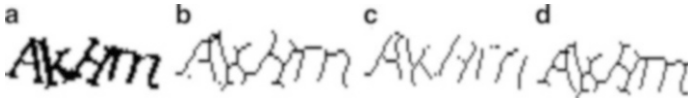
- CPU: Intel Core i5-2430M 2.40 GHz.
- Memory: DDR3 1066 MHz 2 GB.
- Operating System: Windows 7.
- Compiler: Visual C++ 2008.

Where possible, we make comparisons to previously proposed methods. We perform our experiment on the well-known OPTA and Hilditch methods. In experience, we use two category images: one is image with single letter and another one is multi-letters. A single character set is primarily to the validity of algorithm corresponding to the character topology. Multi-letters set is to verify validity of Connectivity between the characters. Figures 79.6 and 79.7 give the result.

From Fig. 79.6, we can notice that the topology of letter can be reserved while comparison of little disconnection by improved OPTA method and many disconnect region by Hilditch method. For multi-letters thinning in Fig. 79.7, Line with more than one pixel was generated in (b) by method of improved OPTA method and disconnection in sub-image with letter M. In (c), connection lines were broken into some regions. In (d), we can notice topology and connection of letters were reserved.



**Fig. 79.6** Single Letter result. (a) Original single image. (b) Improved OPTA method. (c) Hilditch method. (d) Proposed method



**Fig. 79.7** Multi-letters result. (a) Original single image. (b) Improved OPTA method. (c) Hilditch method. (d) Proposed method

## 79.4 Conclusion

In this paper, a model for thinning handwriting-like image was proposed. Based on the proposed model, thinning binary image was generated with better topology structure and connection. It is superior to the OPTA and Hilditch proved by experiments. What's more, the method is based on produce instead of only templates. It can accelerate the thinning speed.

However, this is not necessarily the best algorithm. Thinning binary image and reserving its topology may be able to further improve the post-process correctness based on keeping handwriting skeleton accuracy. It remains to be confirmed by further studies and experiments.

## References

1. Jawahar, C. V., Balasubramanian, A., Million, M., & Namboodiri, A. M. (2009). Retrieval of online handwriting by synthesis and matching. *Pattern Recognition*, 42(7), 1445–1457.
2. Cheng, J., Wang, J., Jiang, S., Zhou, Z.-H., & Hancock, E. (2011). Special edition on semi-supervised learning for visual content analysis and understanding". *Pattern Recognition*, 44(10–11): 2242–2243.
3. Shang, L., Yi, Z., & Ji, L. (2007). Binary image thinning using autowaves generated by PCNN. *Neural Processing Letters*, 25(1), 49–62.
4. Xing-kui, F., Lin-yan, L., & Zu-quan, Y. (1999). A new thinning algorithm for finger print image. *Journal of Image and Graphics*, 4(10), 835–838.
5. Peter I. Rockett, "An Improved Rotation-Invariant Thinning Algorithm," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 27, no. 10, pp. 1671–1674, Oct. 2005.
6. Jia, Yu., & Yaqin, Li. (2009). Improving Hilditch thinning algorithms for text image". 2009 International Conference on E-Learning, E-Business. Enterprise information Systems and E-Government. pp. 76–79.

7. Imiya, A., & Saito, M. (2006). Thinning by curvature flow. *Journal of Visual Communication and Image Representation*, 17(1), 27–41.
8. Ravi, J., Raja, K. B., & Venugopal, K. R. (2009). Fingerprint recognition using minutiae score matching. *International Journal of Engineering Science and Technology*, 1, 35–42.
9. Bag, S., & Harit, G. (2011). An improved contour-based thinning method for character images. *Pattern Recognition Letters*, 32(11), 1836–1842.
10. Jinhai, C. (2012). Robust filtering-based thinning algorithm for pattern recognition. *The Computer Journal*, 55(7), 887–896.

# Chapter 80

## Discrimination of the White Wine Based on Sparse Principal Component Analysis and Support Vector Machine

Rong Wang, Wu Zeng, and Jiao Ming

**Abstract** In allusion to the urgency of the white wine identification, and the key shortcoming of PCA (Principal Component Analysis), whose all loadings are non-zero, the sparse PCA (SPCA) is employed to enhance the explanatory and remove unnecessary variables. By using elastic net Zou Presented, SPCA can seek sparse factors and explain the maximum amount of variances in the data. For illustration, a comparison of strategy between PCA and SPCA, which combined with the specified categorizer—Support Vector Machine (SVM) and Infrared (IR) Spectroscopy, is utilized. The finally classified result of white wines based on SPCA is up to 93 %, but the PCA's 83 %, and directed categorizer's 80 %. Obviously, the SPCA can extract characteristics more effectively, which benefits the classification accuracy of SVM.

### 80.1 Introduction

Up to today, white wine, which blended delicately combined with storage perfectly in many years, has become a luxury slowly. The costliness price spilt over into confusion. In consequence, the fraud means emerge in endlessly. Thus, the quality of wine has attracted great importance. As a rapid and non-destructive methodology, NIR technique has been paid much attention recently. Studies have been reported on using NIRS technique to analyze quality of food oil, fruit, and honey [1–3]. IR spectroscopy is absorbed primarily by hydric groups, which is main composition of the wine. To qualitatively detect the category, IR can yet be regarded as a simple but effective method. However, IR acquisition area includes hundreds of wavelength points. The correlation among certain wavelength points surely increases the complexity of the problem, so we must take measures to select

---

R. Wang (✉) • W. Zeng • J. Ming  
Department of Electrical & Electronic Engineering, Wuhan Polytechnic University,  
Wuhan 430000, China  
e-mail: [340689241@qq.com](mailto:340689241@qq.com)

variable. PCA chances to transform original variables into a few new variables, thus realizes reducing the dimensionality of spectral data in order to extract main information.

PCA seeks linear combinations of the data variables (called principal components PCs) by capturing a maximum amount of variance. However, all PC coefficients (called loadings) are non-zero. This is to say, although PCA concentrates the information in a few key PCs for the convenience of the model interpretation and visualization, the PCs are still constructed by all original variables. To enhance the explanatory and remove unnecessary variables, Zou [4] presented SPCA. In SPCA, we seek a trade-off between the two goals of expressive power (explaining most of the variance or information in the data) and interpretability (making sure that the factors involve only a few coordinate axes or variables) [5].

The outline of the paper is as follows. After brief background introduction, data and methods needed are described in Sects. 80.2 and 80.3, separately. Then in Sect. 80.4, we demonstrate this mixed method and discuss its potential through several experiments. At the end, we conclude and give a general overview to future's work.

## 80.2 Materials

A total of five bottles of wine samples were bought from a hypermarket in Wuhan, including “guan-gong-fang” wine (ggf), “fen-jiu” aged 10 and 5 years (f10 and f5), “er-guo-tou” aged 8 years, and another “er-guo-tou” (egt8 and egt65d). Just because those wine are famous in china, thus it is more likely to be cheated. Moreover, there are different flavor, alcohol, and year so that selected samples are more typical.

The IR spectra were collected in the reflectance mode using a NEXUS 670 FT-IR spectrophotometer (Nicollet, USA) with an InGaAs detector in the range from 4,000 to 650  $\text{cm}^{-1}$ . The sample spectrum was acquired to subtract from a reference spectrum which scanned in an empty cell without sample to remove background noise. The spectrum of each sample was obtained by taking the average reading of 16 scans. In case of abnormal data, more than 50 samples were collected from each bottle. But 50 samples derived from each wine were selected as final experimental data. In multivariate analysis, the wine samples were divided into calibration set and validation set, which was used to train and evaluate the model separately. To testify the classified prediction, the concentrations of each category were arranged in randomly.

## 80.3 Methods

### 80.3.1 The Overall Idea of Algorithm

The paper adopts SVM based on the IR spectroscopy technique to realize perfect classification of white wine, whose overall idea is as follows:

1. To remove the abnormal samples;
2. To extract feature used PCA and SPCA;
3. To take the extracted PCs as the input of SVM and build a prediction model;
4. To predict the samples on the model.

### 80.3.2 PCA

PCA was used here mainly to extract information from the wine spectrum data, and meanwhile to reduce the dimensionality of the spectra. By calculating the eigenvectors of the covariance matrix of normalized calibration data, PCA transforms the original variables into new axes and calculates PCs as new variables to replace the original data [6].

In detail, the data  $X$  is an  $n \times p$  matrix containing  $p$  features for each observation. Without loss of generality,  $X$  is normalized. Then the singular value decomposition (SVD) of  $X$  is to be:

$$X = UDV' \quad (80.1)$$

So  $Z = UD$  is the PCs, and the columns of  $V$  are the corresponding loadings.

Notice that PCA has another function to realize clustering. Through observing the score plot, we can find the same kind of data will gather together while deviate far when they are different. Those dates deviated farther are thought to be abnormal. The paper is to use this method to eliminate abnormal samples.

### 80.3.3 SPCA

As an improve method, SPCA seeks sparse factors or linear combinations of the data variables to explain a maximum amount of variance in the data while having only a limited number of non-zero coefficients [5].

Elastic net, as a regression-type optimization problem, combines Ridge with Lasso, both of which are a penalization technique based on the  $l_n$  norm. Through the  $l_2$  norm, Ridge regression contracts coefficients of regression to achieve the purpose of reducing error. But Lasso increases  $l_1$  penalty function based on general linear least square method to make the sum of the absolute coefficient less than a constant.



Due to the constraint of natural property, some coefficients turn to zero. Elastic net integrates both such that removing the limitation of each other, and then simultaneously chooses variables and reduces the model prediction error. Meanwhile, it generally does not compress regression coefficients excessively.

To address these concerns, Elastic net regression attempts to mix the  $l_1$  and  $l_2$  penalties, as shown in (80.2).

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \left( \|Z_i - X\beta\|^2 + \lambda\|\beta\|^2 + \lambda_1\|\beta\|_1 \right) \quad (80.2)$$

The goal in elastic net is to produce a sparse feature space with the  $l_1$  penalty, but improve stability and retain correlated features with the  $l_2$  penalty [7]. Like the lasso this is not a computationally simple problem, but efficient methods for solving it which have been developed. There is also an additional free parameter in  $\alpha$ , which determines the relative strength of the penalties. Previous studies have shown this technique to be effective in classification of functional magnetic resonance imaging (fMRI) data [8].

### 80.3.4 SVM

SVM is a powerful classification technique first proposed by Vapnik, whose theory is based on Vapnik-Chervonenkis theory and structural risk minimization [9]. An optimum network structure and a better generalization performance could be achieved with SVM.

The main idea of SVM is to first map the data points into a high-dimensional feature space by using a kernel function, and then to construct an optimal separating hyperplane between the classes in that space. The primary advantage of SVM over the traditional learning algorithm is that the solution of SVM is always globally optimal and avoids local minima and over-fitting in the training process.

## 80.4 Results and Discussion

### 80.4.1 Spectra Investigation

During the process of collecting spectral data, there may be emerge abnormal values, which will affect the classification accuracy. In order to reduce the negative effect, we must take measures to remove these abnormal samples before feature extraction. According to PCA for all data, we find and remove three samples deviating farther, and then arrange new data. Figure 80.1 shows the mean spectral of each class for the original data mapping of five wines. As seen in Fig. 80.1, the

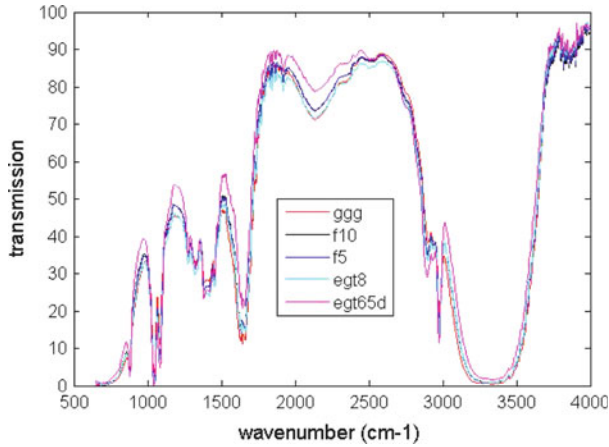


Fig. 80.1 The mean spectral of each class

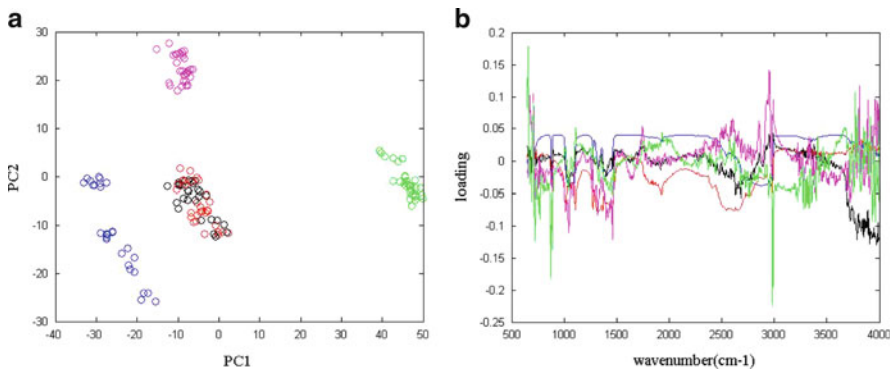
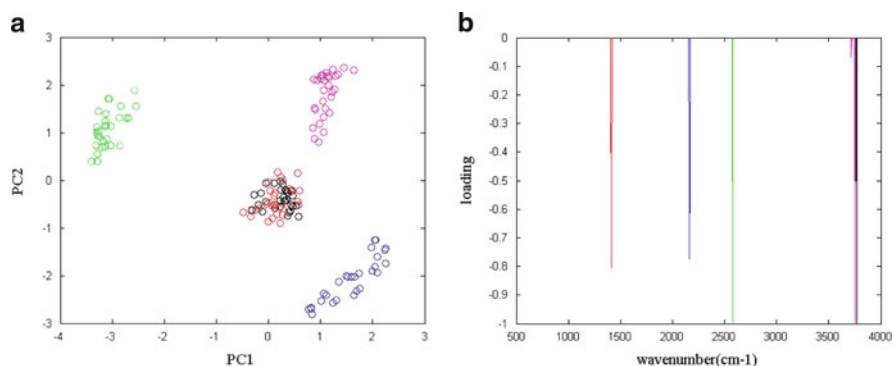


Fig. 80.2 The scores plot of PCA (left) and the loading of PCA (right)

features of wine spectral were similar generally because of the influence of water in the wine, which affects the result of classification seriously. So we must make measures to select features and improve the classified precision.

### 80.4.2 PCA and SPCA

Although both of them cannot be used as a classification tool, this behavior may indicate the data trend in visualizing dimension spaces. For visualizing the data trends and the discriminating efficiency, a scores plot using the first two PCs of training samples and the loading of top five PCs were obtained which are showed in Fig. 80.2. As it shows, the scores plot could reflect cluster relation at some degree. Then the sparse feature of SPCA was showed in Fig. 80.3.



**Fig. 80.3** The scores plot of SPCA (*left*) and the loading of SPCA (*right*)

Comparing Figs. 80.2 and 80.3, we can discover that although both scores plot getting into four categories, in which the two of “fenjiu” cannot be separated, the SPCA loadings are sparse enough so that it can clearly point out the strongest relevance between PCs and band. Thus, the property of sparse enhances the explanation and removes the redundance.

### 80.4.3 SVM

According variance contribution on request, the first extracted 17 PCs by PCA and SPCA, respectively, when training, contained more than 96 % of total variance. So they could almost express the total spectral information. Then the top 17 latent variables were put into the classifiers. As explained by Chih-Wei [10], RBF is the most reasonable choice because of its simplicity and ability to model data of arbitrary complexity. So we entirely focus on the RBF kernel in this paper. To improve SVM predictive performance, we adopt a “heuristic search” on parameters  $c$  and  $g$ .

After constructing the model, how to predict a new sample is the following work. According to the introduction of PCA and SPCA, we must take validation set to the same mapping and get PCs of it. The best way is to deal with Validation set based on projection matrix of calibration set. Then the obtained PCs of calibration set are taken into the model to predict.

Table 80.1 contains the prediction results for the SVM, PCA + SVM, and SPCA + SVM models. For pure SVM and PCA + SVM classified method, only 80 % and 83 % of true wine were correctly classified, respectively; while the SPCA + SVM models achieved 93 % correct classification. Obviously, the sparse PC as input of SVM could be produced better classification results than others.

**Table 80.1** Results for the calibration and validation set using hybrid SVM model

Methods	Data set	c	g	Classification accuracy
SVM	Calibration set	1	0.0313	100 %
	Validation set			80 %
PCA + SVM	Calibration set	4	0.0313	100 %
	Validation set			83 %
SPCA + SVM	Calibration set	32	0.0625	100 %
	Validation set			93 %

## 80.5 Conclusion

This work tried to improve enforcement of quality of Chinese white wine. The grade and commercial value of five breeds of wine is of great difference. So quickly, efficiently, and correctly identify wine has significant meaning. The study applied NIRS combined with SVM, and extracted optimal PCs using PCA and SPCA as the input of SVM. After comparing the abilities of two kinds of methods, we can conclude that the NIR spectroscopy technique based on SVM has high potential to identify the wine varieties in a rapid and nondestructive way and with a high degree of accuracy.

Although not achieving correct perfectly, the approach using SPCA still indicates that the sparse is effective and promising to classification in some degree. After all the penalty coefficient is selected by experience or man-made, which lack of comprehensive. If we can realize the adaptive selection, the classification accuracy will reach 100 %. So the adaptive selection turns to be the next step before us.

**Acknowledgements** The authors gratefully acknowledge the laboratory of electronic and information engineering, Huazhong University of Science and Technology (HUST), HuBei. This work has been financially supported by National Science & Technology Pillar Program of China (No. 2012BAK02B06).

## References

1. Du, C., & Zhou, J. (2011). Application of infrared photoacoustic spectroscopy in soil analysis. *Applied Spectroscopy Reviews*, 46(5), 405–422.
2. Ying, Y., & Liu, Y. (2008). Nondestructive measurement of internal quality in pear using genetic algorithms and FT-NIR spectroscopy. *Journal of Food Engineering*, 84(2), 206–213.
3. Cozzolino, D., Corbella, E., & Smyth, H. E. (2011). Quality control of honey using infrared spectroscopy: A review. *Applied Spectroscopy Reviews*, 46(7), 523–538.
4. Zou, H., Hastie, T., & Tibshirani, R. (2006). Sparse principal component analysis. *Journal of Computational and Graphical Statistics*, 15(2), 265–286.

5. Luss, R., & d'Aspremont, A. (2010). Clustering and feature selection using sparse principal component analysis. *Optimization and Engineering*, *11*(1), 145–157.
6. Yu, H., Lin, H., Xu, H., et al. (2008). Prediction of enological parameters and discrimination of rice wine age using least-squares support vector machines and near infrared spectroscopy. *Journal of Agricultural and Food Chemistry*, *56*(2), 307–313.
7. Kelly, J. W., Degenhart, A. D., Siewiorek, D. P., et al. (2012). Sparse linear regression with elastic net regularization for brain-computer interfaces. *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE* (pp. 4275–4278). IEEE.
8. Carroll, M. K., Cecchi, G. A., Rish, I., et al. (2009). Prediction and interpretation of distributed neural activity with sparse models. *NeuroImage*, *44*(1), 112–122.
9. Shao, Q., & Feng, C. J. (2012). Pattern recognition of chatter gestation based on hybrid PCA-SVM. *Applied Mechanics and Materials*, *120*, 190–194.
10. Hsu, C. W., Chang, C. C., Lin, C. J. (2003). A practical guide to support vector classification. Department of Computer Science, National Taiwan University.

**Part IV**  
**Cloud Computing**

# Chapter 81

## Design of Mobile Electronic Payment System

Ting Huang

**Abstract** Multi-bank mobile electronic payment system uses mobile terminals for electronic payments, which can circulate in multiple banks and cannot limit from the bank that issues the e-cash. The paper researches electronic payment on the withdrawal agreement, the pay agreement, the deposit agreement, the update protocol of the e-cash based on elliptic curve cryptography. The design of the system is more suitable for mobile payment terminals with limit of calculation capacity, storage, network bandwidth, and power supply, which meets the needs of the day-to-day transactions.

### 81.1 Introduction

Nowadays the mobile electronic payment is popular in the world that the user has done dynamic payment with his mobile terminal, such as Visa DPS. But a lot of e-cash cannot be in circulation in the number of financial institutions because the amount of e-cash issue is fixed and the transaction can only be done in the system [1]. Based on elliptic curve cryptosystem this paper proposes multi-bank mobile electronic payment system. The micro-payment transacts off line by the k-ary tree and the macro-payment transacts on line, which meet the needs of the actual transaction in the e-cash circulation of the banks. Divisibility of the mobile e-cash [2] is realized by the K-tree ( $K > 1$ ). The solutions of multi-bank electronic payment [3] are that multiple banks issue the e-cash.

---

T. Huang (✉)  
College of Computer and Information Technology, China Three Gorges University,  
Yichang, Hubei 443002, China  
e-mail: [14863403@qq.com](mailto:14863403@qq.com)

## 81.2 Design of a Project

The symbols of the whole text involve:  $x, y$  are the point of abscissa and vertical axis on the elliptic curve, respectively;  $S_U, S_C$  are the user and payee's private key;  $P_U, P_C$  are the user and payee's public key. Each bank is classified by the number, such as ICBC is the bank<sub>1</sub>, etc.; the other bank is the bank <sub>$n$</sub> , of which the corresponding private key and the signature private key are defined as  $S_{Bn}, S_{1n}, S_{2n}$  and the corresponding public key is  $P_{Bn}, P_{1n}, P_{2n}$ , ( $n \in (1, \infty)$ ). The model of the mobile electronic payment system is shown in Fig. 81.1.

The use cycle of the e-cash is defined as  $T$  ( $T$  is the time cycle  $T_1$  to  $T_3$ ). The time  $T_1$  generates when the user fetches the e-cash from the bank. The time cycle  $T_1$  to  $T_2$  is an effective use cycle of the e-cash  $T'$  ( $T'$  is the given value). When the time  $T_2$  reaches, the e-cash is unavailable. Only before the time  $T_3$  (The time cycle  $T_2$  to  $T_3$  is the updating cycle of the e-cash  $T''$  that is the given value) the user goes to the bank for updating the e-cash.

### 81.2.1 Design of Off-Line Payment(Micro-Payment)

#### 1. Withdrawal agreement

Step 1: The user fetches  $M$  yuan from the bank <sub>$n$</sub>  and gives the segmentation parameters  $K$  of the e-cash. Then he generates  $A \in_R Z_n$  and  $\rho = H(A + S_U)$ . At the same time, he gets the  $K$ -ary tree by  $M$  yuan, saves  $\alpha_{(n)}, \rho, K, T_1, B$  to the database, and sends to the bank <sub>$n$</sub>  after encrypting  $ID_{nU}, M, \alpha_{(n)}, \beta_{(n)}, \gamma_{(n)}, T_1$  by the shared key.

Step 2: The bank <sub>$n$</sub>  gets the root node of the divisible e-cash:  $\alpha_{(n)} = H[(MS_{1n} + \rho S_{2n}) || S_{Bn} || B || T_1]$ , and calculates  $\beta_{(n)} = H[\alpha_{(n)} || (CG)_x || M || T_1]$ ,  $\gamma_{(n)} = \beta_{(n)} S_{Bn} + C, B, C \in_R Z_n$ .

Then it saves  $\alpha_{(n)}, \rho, K, T_1, B$  to the database, encrypts  $ID_{nU}, M, \alpha_{(n)}, \beta_{(n)}, \gamma_{(n)}, T_1$ , and distributes to the user.

Step 3: The user decrypts  $ID_{nU}, M, \alpha_{(n)}, \beta_{(n)}, \gamma_{(n)}, T_1$  and stores.

#### 2. Payment agreement

Payments in two ways are described in Tables 81.1 and 81.2.

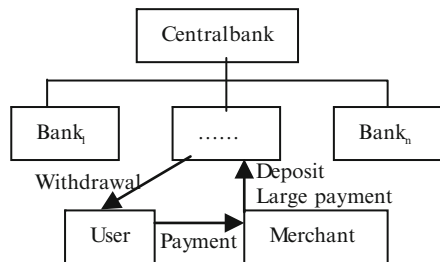


Fig. 81.1 The model of the mobile electronic payment system



**Table 81.1** Payment agreement ( $M' = M$  yuan)

User	Payee
$M, \alpha_{(n)}, \beta_{(n)}, \gamma_{(n)}, T_1$ is sent	The payee tests $\beta_{(n)} \stackrel{?}{=} H[\alpha_{(n)} \parallel (\gamma_{(n)} G - \beta_{(n)} P_{Bn})_x \parallel M \parallel T_1]$ , saves $M, \alpha_{(n)}, \beta_{(n)}, \gamma_{(n)}$ , computers $\eta = H[(S_C P_U)_x]$ , and sends $\eta$ , the user's name, $M, \alpha_{(n)}$
The user tests $\eta \stackrel{?}{=} H[(S_U P_C)_x]$ and deletes entire e-cash records of $\alpha_{(n)}, M$	

**Table 81.2** Payment agreement ( $M' < M$  yuan)

User	Payee
<p>The K-tree is generated by <math>M'</math>: For the branches <math>\alpha_{(n)1, \dots, 1S}</math>, the user takes <math>B_1, \dots, B_{k, \dots, k}, C_1, \dots, C_{k, \dots, k} \in_R Z_n, p = \{p, 11\}, \alpha_{(n)11} = H(B_{11} \parallel S_U \parallel \alpha_{(n)} \parallel T_1), \beta_{(n)11} = H[\alpha_{(n)11} \parallel (C_{11} G)_x \parallel p \parallel 11]</math>  <math>K \parallel M' \parallel T_1, \gamma_{(n)11} = \beta_{(n)11} S_U + C_{11}, \dots, \alpha_{(n)11, \dots, 1S} = H(B_{11, \dots, 1S} \parallel S_U \parallel \alpha_{(n)11, \dots, (S-1)S} \parallel T_1), p = \{p, 11, \dots, 1S\}, \beta_{(n)11, \dots, 1S} = H[\alpha_{(n)11, \dots, 1S} \parallel (C_{11, \dots, 1S} G)_x \parallel p \parallel 11] \dots \parallel p \parallel 1S \parallel K \parallel M' \parallel T_1, \gamma_{(n)11, \dots, 1S} = \beta_{(n)11, \dots, 1S} S_U + C_{11, \dots, 1S}</math>, saves <math>B_1, \dots, B_{k, \dots, k}, C_{1, \dots, k, \dots, k}</math> to the database. <math>M', p, \alpha_{(n)11, \dots, 1S}, \beta_{(n)11, \dots, 1S}, \gamma_{(n)11, \dots, 1S}, \alpha_{(n)}, T_1</math> are sent. The branches that are not used can be paid until all the K-tree branches are run out.</p>	<p>The payee tests  <math>\beta_{(n)11, \dots, 1S} \stackrel{?}{=} H[\alpha_{(n)11, \dots, 1S} \parallel (\gamma_{(n)11, \dots, 1S} G - \beta_{(n)11, \dots, 1S} P_U)_x \parallel p \parallel p \parallel 11 \dots \parallel p \parallel 1S \parallel K \parallel M' \parallel T_1]</math>, saves <math>M', p, \alpha_{(n)11, \dots, 1S}, \beta_{(n)11, \dots, 1S}, \gamma_{(n)11, \dots, 1S}, K</math>, computers <math>\eta = H[(S_C P_U)_x]</math>, and sends <math>\eta</math>, the user's name, <math>M', \alpha_{(n)11, \dots, 1S}</math></p>
After the user tests $\eta \stackrel{?}{=} H[(S_U P_C)_x]$ , the amount of e-cash will subtract $M'$ from the database and $\alpha_{(n)11, \dots, 1S}$ will be deleted.	

### 3. Deposit agreement

After the businesses send the relevant transaction information stored in the payment protocol to the bank, the bank verifies the correctness of this information to detect double-spending. When the merchant deposits to the bank<sub>n</sub>, the e-cash deposits to the bank<sub>n</sub>; merchant deposits to the bank<sub>m</sub> (not the bank<sub>n</sub>), the bank<sub>n</sub> sends depositing issued to the bank<sub>m</sub>. After the e-cash in the bank<sub>m</sub> is dealt the merchant finishes the depositing process. The central bank that has higher level than ordinary bank verifies the authenticity of the transfer information on a regular time, detects the untrue transfers, and increases the penalty.

In this section the payment node  $\alpha_{(n)1, \dots}$  is of an e-cash sent by the business as an example: The e-cash is checked whether it is out of date in deposit database, and whether it has the same path  $p$ . If it has the same path, you can conclude the reuse of the e-cash. Then the bank can find the dishonest user's identity who takes e-cash  $\rho$  through the withdrawal database. Otherwise, the bank accepts the deposits.

### 4. Update of the e-cash

When the usage time  $T'$  of the e-cash reaches, the user must update through the following protocol, while the e-cash will set aside more than the time  $T_3$ .

Step 1: The user regenerates  $A \in {}_R Z_n$ ,  $\rho = H(A + S_U)$ , regenerates the K-tree by the residual e-cash  $M$ , saves  $\rho$ ,  $K$  to the database. After they are encrypted by the bank and the user's shared key,  $\rho$ ,  $K$  are sent to the bank.

Step 2: The bank regenerates the root node of the divisible e-cash according to the user's update time  $T_1$ :  $\alpha_{(n)} = H[(MS_{1n} + \rho S_{2n}) \| S_{Bn} \| B \| T_1]$  and computes  $B, C \in {}_R Z_n$ ,  $\gamma_{(n)} = \beta_{(n)} S_{Bn} + C$ ,  $\beta_{(n)} = H[\alpha_{(n)} \| (CG)_x \| M \| T_1]$ , saves  $\alpha_{(n)}$ ,  $\rho$ ,  $K$ ,  $T_1$ ,  $B$  to the database. After they are encrypted,  $ID_{nU}$ ,  $M$ ,  $\alpha_{(n)}$ ,  $\beta_{(n)}$ ,  $\gamma_{(n)}$ ,  $T_1$  are sent to the user.

After it has been updated, the e-cash can be transacted by the previous protocol.

## 81.2.2 Design of On-Line Payment (Large Payment)

### 1. Withdrawal agreement

Step 1: The user withdraws  $M$  yuan from the bank<sub>n</sub>. After he encrypts  $M$  with the shared key of the user and the bank<sub>n</sub>, the user sends it to the bank<sub>n</sub>.

Step 2: The bank<sub>n</sub> gets the root node of the divisible e-cash:  $\alpha_{(n)} = H[(MS_{1n} + S_{2n}) \| S_{Bn} \| B \| T_1]$ , calculates  $\beta_{(n)} = H[\alpha_{(n)} \| (CG)_x \| M \| T_1]$ ,  $\gamma_{(n)} = \beta_{(n)} S_{Bn} + C$ ,  $B, C \in {}_R Z_n$ , and saves  $\alpha_{(n)}$ ,  $T_1$ ,  $B$  to the database. After he encrypts  $ID_{nU}$ ,  $M$ ,  $\alpha_{(n)}$ ,  $\beta_{(n)}$ ,  $\gamma_{(n)}$ ,  $T_1$ , the bank<sub>n</sub> will distribute them to the user.

Step 3: After he decrypts  $ID_{nU}$ ,  $M$ ,  $\alpha_{(n)}$ ,  $\beta_{(n)}$ ,  $\gamma_{(n)}$ ,  $T_1$ , the user will save them.

2. Payment agreement

Step 1: The user pays the e-cash to the payee,  $M, \alpha_{(n)}, \beta_{(n)}, \gamma_{(n)}, T_1$  are sent to the payee.

Step 2: The payee inspects the authenticity of the e-cash: The payee tests  $\beta_{(n)} \stackrel{?}{=} H\left[\alpha_{(n)} \parallel (\gamma_{(n)}G - \beta_{(n)}P_{Bn})_x \parallel M \parallel T_1\right]$ , and sends the user name,  $\alpha_{(n)}, M, T_1$  to the bank.

Step 3: The bank checks the authenticity of  $\alpha_{(n)}$ . After it gets the validation of  $\alpha_{(n)}$ , the bank will inform the payee.

Step 4: The payee saves  $M, \alpha_{(n)}, \beta_{(n)}, \gamma_{(n)}$ , calculates  $\eta = H[(S_C P_U)_x]$ , sends  $\eta$ , the user name,  $\alpha_{(n)}, M$  to the user.

Step 5: The user confirms that the payment is successful, tests  $\eta \stackrel{?}{=} H[(S_U P_C)_x]$ , deletes the entire e-cash records of  $\alpha_{(n)}, M$ .

3. Deposit agreement

The course is the same as the deposit agreement in off-line payment. For example, the business sends the e-cash  $\alpha_{(n)}$ , which is same as the payment node  $\alpha_{(n)11, \dots, 1S}$  of an e-cash in off-line payment.

4. Update of e-cash

The e-cash will set aside more than the time  $T_3$ . The user must update the e-cash through the following protocol.

Step 1: After he encrypts  $M$  with the shared key of the user and the bank, the user sends it to the bank.

Step 2: The bank regenerates the divisible e-cash according to the time  $T_1$  of the user updating the e-cash  $\alpha_{(n)} = H[(MS_{In} + S_{2n}) \parallel S_{Bn} \parallel B \parallel T_1]$ , calculates  $B, C \in_R Z_n, \beta_{(n)} = H[\alpha_{(n)} \parallel (CG)_x \parallel M \parallel T_1], \gamma_{(n)} = \beta_{(n)}S_{Bn} + C$ , saves  $\alpha_{(n)}, T_1, B$  to the database. After he encrypts  $ID_{nU}, M, \alpha_{(n)}, \beta_{(n)}, \gamma_{(n)}, T_1$ , the bank will distribute them to the user .

After it has been updated, the e-cash can be the same with the previous transaction protocol.

**81.3 Security and Efficiency Analysis**

1. Blind signature:  $\alpha_{(n)}, \beta_{(n)}, \gamma_{(n)}$  are the right signature of the e-cash  $M$  based on blind signature protocol.

*Proof:* If the validation of the e-cash  $\beta_{(n)} \stackrel{?}{=} H\left[\alpha_{(n)} \parallel (\gamma_{(n)}G - \beta_{(n)}P_{Bn})_x \parallel M \parallel T_1\right]$  is established, that proves  $\gamma_{(n)}G - \beta_{(n)}P_{Bn} = CG$ .  $\gamma_{(n)} = \beta_{(n)}S_{Bn} + C$ , so  $\gamma_{(n)}G - \beta_{(n)}P_{Bn} = (\beta_{(n)}S_{Bn} + C)G - \beta_{(n)}P_{Bn} = CG$ . Similarly we can prove that  $\alpha_{(n)11, \dots, 1S}, \alpha_{(n)11, \dots, 1S}, \beta_{(n)11, \dots, 1S}, \gamma_{(n)11, \dots, 1S}$  are correct signature of the e-cash  $M$ .

2. The signatures and the e-cash are unforgeable

The signatures and the e-cash embed the private key  $S$ . Any attacker must get the private key  $S$  so that he can forge the signatures and the e-cash, which must solve the ECDLP problem. This problem remains unsolvable, so the signatures and the e-cash are unforgeable.

3. When the e-cash is spent repeatedly, the bank will reveal the user's anonymity

When he consumes the e-cash, the user will send the e-cash and  $p$  to the merchant. If the user spends repeatedly, the bank will find the reuse of the e-cash or  $p$  when the business deposits. The bank inquires about the withdrawal database, thus it can identify the user's identity.

So it is not possible to forge a correct e-cash. When the e-cash is spent repeatedly, merchant can detect the user's identity through the bank. Therefore the protocol is safe and fair (The bank must be safe and reliable). If the user trades normally, the user's anonymity is guaranteed. The bank will reveal the user's identity only the illegal transaction. Thus the anonymity of the user will be ensured, the interests of the banks and the merchants will be ensured too.

The time of protocol implementation and storage consumption in multi-bank mobile electronic payment system are the key to efficiency. 160 bit length of the key in ECC has the same powerful functions as 1,024 bits length of the key in RSA [4]. The divisible e-cash protocol [5] bases on the zero-knowledge proof, strong RSA problem. The achieve agreement of this system bases on ECC. Compared to RSA, ECC need not exponentiation compute. The computation of ECC can be negligible. Therefore the efficiency of this protocol runs faster.

Paying a divisible e-cash [5] needs to save  $(S, \pi)$ ,  $S = g^{\frac{1}{s+j+1}}$ ,  $\pi = (R, \text{info}, \text{pk}_M, \text{time}, T, C_r, C_{\text{pk}_i}, C_S, C_t, C_d, C_j, \Phi)$ . The storage space needs at least  $1024 + 128 + 1024*7 + 1024*5 = 13440\text{bit} = 1680\text{byte}$ . While paying a divisible e-cash in this system simply saves  $M'$ ,  $\alpha_{(n)11,12,13}$ ,  $\beta_{(n)11,12,13}$ ,  $\gamma_{(n)11,12,13}$ ,  $p = \{\rho, 11, 12, 13\}$ ,  $K$  (Assuming the  $K$ -tree assigns to the third branch). The storage space is approximately  $32 + 128 + 128 + 192 + (128 + 32*3) + 32 \text{ bit} = 736 \text{ bit} = 92 \text{ byte}$ , which reduces the 94.5 % of the storage amount and the network bandwidth. The 2-tree is used [5, 6]. Because the divided branches in splitting the e-cash are too much, the amount of computation in generating the e-cash is increased.

The e-cash in this system contains the use cycle. When the end of the use cycle is coming, the bank can delete the e-cash. Efficient divisible e-cash scheme [7] needs storage  $(S, T, \prod s, I, \xi)$ ,  $S_{i,0}, \dots, S_{i,2^i-1}$ . Thus a lot of storage space is saved in this paper. The efficiency of the bank retrieving the database and the deposit agreement are improved. The reliability and maintainability of the data are improved too. When the end of the use cycle of the e-cash is coming, the user can update the e-cash in the bank through implementing update protocol of the e-cash, which guarantees the legitimate rights and interests of the user.

The implementation of the multi-bank e-cash [8] requires the central bank to issue the e-cash. A financial institution is added, which makes the protocol more complicatedly.

## 81.4 Conclusion

To sum up, the protocol of multi-bank mobile electronic payment system is safe, simple, efficient, and suitable for the mobile payment terminals of which calculation capacity, storage, network bandwidth, and power supply are very limited.

## References

1. Ziba Eslami, Mehdi Talebi. (2011). A new untraceable off-line electronic cash system. *Electronic Commerce Research and Applications*, 10(1), 59–66.
2. Ting Huang, Shou-zhi Xu. (2010). Study on mobile divisible e-cash based on elliptic curve. *Journal of Wuhan University of Technology*, 32(23), 150–153 (In Chinese).
3. Ting Huang. (2012). Study on multi-bank mobile electronic payment. *Theoretical and Mathematical Foundations of Computer Science* (Vol. 38, pp. 507–511). USA: Information Engineering Research Institute.
4. Menezes, A., Okamoto, T., & Vanstone, S. (1993). Reducing elliptic curve logarithms to logarithms in a finite field. *IEEE Transactions on Information Theory*, 39(5), 1639–1646.
5. Xinyu He. (2010). *Research on e-cash protocol based on multi-level proxy blind signature and each node can be paid*. Master dissertation for Yanshan University, Qin Huang Dao (In Chinese).
6. JiuHong Wang. (2010). *The research on efficient divisible e-cash based on ECC*. Master dissertation for Yanshan University, Qin Huang Dao (In Chinese).
7. Yong-bo Yu, Xiao-zhu Jia, Feng Qing-feng. (2010). Efficient divisible e-cash scheme based on one-way accumulator. *Computer Engineering and Applications*, 46(10), 206–208 (In Chinese).
8. Xiangwen Meng, Baohua Zhao. (2011). Fairness-based multi-bank e-cash. *Computer Applications and Software*.28, 163(10), 195–197 (In Chinese).

# Chapter 82

## Power Saving-Based Radio Resource Scheduling in Long-Term Evolution Advanced Network

Yen-Yin Chu, I-Hsuan Peng, Yen-Wen Chen, Chi-Fu Yi,  
and Addison Y.S. Su

**Abstract** It is well known that power saving is one of the most important issues for mobile device in accessing network services. The efficient conservation of energy for longer operation times of a mobile station is vital to the success of various mobile applications. This paper proposes a systematic approach to allocate radio resources in Long-Term Evolution Advanced (LTE-A) network by considering the channel condition and QoS requirements, while the power saving is a centric issue during the scheduling process. The proposed scheme includes the selection of component carriers (CC) and the allocation of radio resource to satisfy the QoS demands while minimize the power consumption of user equipment (UE). Additionally, exhaustive simulations were performed to examine the performance of the proposed scheme. Both http and video streaming traffic models were applied during the simulations. The experimental results show that the proposed scheme achieves better performance when compared to the other scheme.

### 82.1 Introduction

The mobile multimedia applications, such as video streaming and video telephony, enables the needs of broadband access in wireless mobile networks. Although Long-Term Evolution (LTE) network provides higher bandwidth than that of current 3G network, its channel bandwidth is limited to 20 MHz [1] without flexibility.

---

Y.-Y. Chu • Y.-W. Chen (✉) • C.-F. Yi

Department of Communication Engineering, National Central University, Taoyuan, Taiwan  
e-mail: [ywchen@ce.ncu.edu.tw](mailto:ywchen@ce.ncu.edu.tw)

I.-H. Peng

Department of Computer Science and Information Engineering, Minghsin University of Science and Technology, Hsinchu, Taiwan

A.Y.S. Su

Research Center for Advanced Science and Technology, National Central University, Taoyuan, Taiwan

The emerging LTE-Advanced (LTE-A) provides four additional technologies, which are carrier aggregation (CA), advance multi-input multi-output (MIMO), coordinated multiple points (CoMP) transmission and reception, and relaying, to improve transmission performance. Among them, carrier aggregation and advance MIMO are designed to overcome the issue of limited channel bandwidth [2]. The user equipment (UE) with LTE-A capability could aggregate more than one component carrier (CC) for more channel bandwidth. However, the LTE-A capable UE requires more antennas to receive the radio signal from different CC, which results in more power consumption. For the mobile devices, such as UE of LTE, it is critical to save power so as to lengthen the usage after charge.

The Orthogonal Frequency Division Multiple Access (OFDMA) technology is used for downlink transmission in LTE/LTE-A network. The eNode B (eNB) is responsible for the allocation of radio resource. The radio resource allocation is considerably more correlated and critical to the diversity channel conditions among UEs and the effective usage of the adaptive modulation and coding (AMC) scheme for transmission in OFDMA systems. A higher level of modulation and coding scheme (MCS) achieves higher spectrum efficiency and, therefore, higher system throughput if channel condition is acceptable. As the channel conditions of UEs may be different and changeable from time to time, if it is not well allocated, the service quality will be downgraded due to insufficient bandwidth. In LTE-A network, eNB can allocate either resource blocks (RB) of the same CC or the RBs that belong to different CC to LTE-A capable UE to satisfy its required bandwidth. However, the LTE UE can only be allocated with the RBs of the same CC. Then, for a specific LTE-A UE, if its channel conditions of all RBs are the same then it will save more power to be allocated with RBs of the same CC. Furthermore eNB may allocate RBs of the same CC to an UE even this allocation will sacrifice system throughput from the power saving point of view. Thus it is a tradeoff when the issues of system throughput and power saving are encountered. And the satisfaction of quality of services (QoS) is the compromise to balance these two issues. The objective of this paper is to propose a systematic scheme to properly allocate radio resource for LTE UE and LTE-A UE coexistence environment with different QoS requirement and under changeable channel condition.

The rest of this paper is organized as follows. Section 82.2 overviews some background and related works. The proposed radio resource allocation scheme is described in Sect. 82.3. Section 82.4 provides the comparison of simulation results with discussion. We conclude our study in the last section.

## 82.2 Related Works

The use of CA in LTE-A provides UE to aggregate the radio signals from different component carriers. And the component carriers can come from different spectrum bands. Thus UE can aggregate carriers from either the contiguous CCs or the

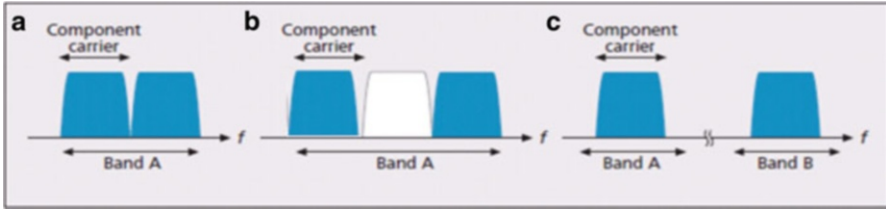


Fig. 82.1 Types of carrier aggregation

noncontiguous CCs of the same spectrum band. UE can also aggregate carriers from the CCs that belong to different spectrum bands as shown in Fig. 82.1 [3].

Each UE can use activation/deactivation mechanism to manage the CA function. And eNB can inform UEs about the allocation information, including CC and RB, through the radio resource control (RRC) message. UE can turn off the antenna without receiving the signals of physical uplink shared channel (PUSCH) and physical downlink control channel (PDCCH) that is temporally unused. As the carrier aggregation provides the flexibility in allocating radio resource, several schemes were proposed to study this issue [4–8]. The LTE-A UE and LTE UE are different in using CC, therefore, it is better to use different policies for both kinds of UE [4]. The authors also designed the measurement functions as the reference to be adopted in both independent packet scheduling per CC and cross-CC packet scheduling [4]. In addition to introducing the random selection and round robin (RR) selection for CC, the reference signal received power (RSRP) value was defined for the use of CC selection [5]. And the authors also presented the proportional fair (PF) and modified largest weighted delay first (M-LWDF) scheduling algorithms for fair resource allocation among UEs. The separated burst-level scheduling (SBLS) was proposed to improve the separated random user scheduling (SRUS), which may introduce unbalance load of CC [6]. The possible structures of LTE-A UE were proposed in with detail analysis of their power consumption [7]. The discontinuous reception (DRX) mechanism provides sleep mode for UE to save power. The authors proposed that LTE-A UE can arrange separated DRX parameters for each CC by considering the QoS requirement [8]. The authors gave their observations on the relationship between power saving efficiency and packet loss ratio by adjusting the inactivity timer and on duration in LTE-A environment. Thus the resource allocation can be designed from different view point such as QoS, fairness, throughput, power saving, etc. In this paper, we propose the radio resource scheduling scheme by trading off several criteria. Although the power saving is the centric in the proposed scheme, QoS satisfaction of each UE and the resource utilization shall also properly considered.



### 82.3 The Proposed Power Saving Centric Radio Resource Scheduling

The conceptual architecture of LTE UE and LTE-A UE coexistence environment is illustrated in Fig. 82.2. It shows that LTE-A UE can access more than CC, however, the LTE UE can only access one CC. The subcarriers of each CC can be grouped into physical resource block (PRB) and each CC consists of several PRBs. It is noted that LTE/LTE-A users have different channel condition, represented as signal to interference plus noise ratio (SINR), on each PRB. The objective of radio resource allocation is to properly arrange PRB(s) to UEs in accordance with their bandwidth requirement and channel condition. Each UE can either periodically or aperiodically report its channel condition, i.e. channel quality indication (CQI), to eNB through the Physical Uplink Control Channel (PUCCH) and PUSCH. The channel condition can be reported by either wideband CQI or subband CQI. The wideband CQI is the average channel condition over the entire channel and cannot precisely describe the quality difference among channels. On the contrary, channels are grouped in to resource block groups (RBG) in subband CQI and each RBG reports its CQI.

The proposed scheme can be subdivided into two parts. The part 1 deals with the selection of CC, and part 2 allocates radio resource to UE as shown in Fig. 82.3. In practical application, UE may have different applications. In this paper, we categorize the UEs into two types of services, which are guaranteed bit rate (GBR) and non-GBR. As the QoS requirements of GBR UE and non-GBR UE

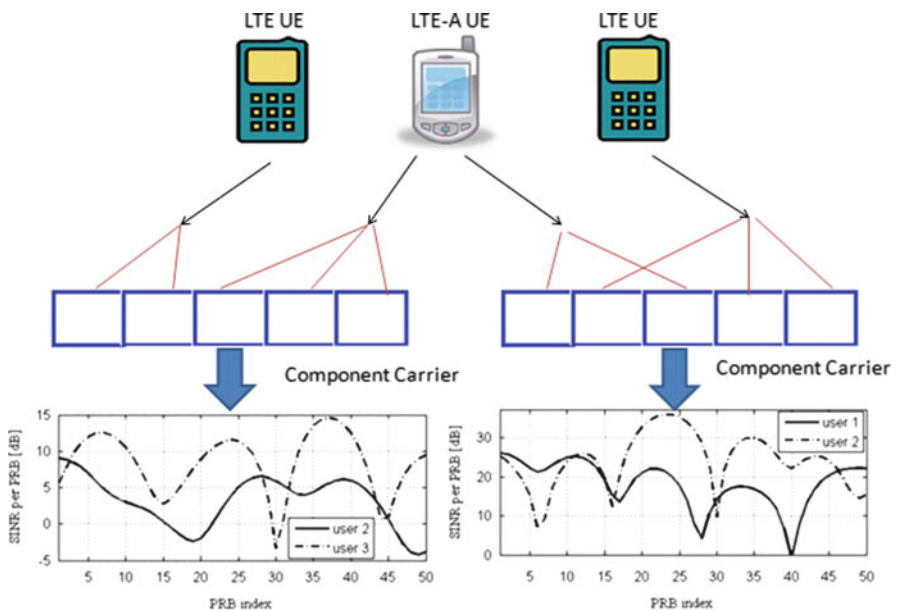
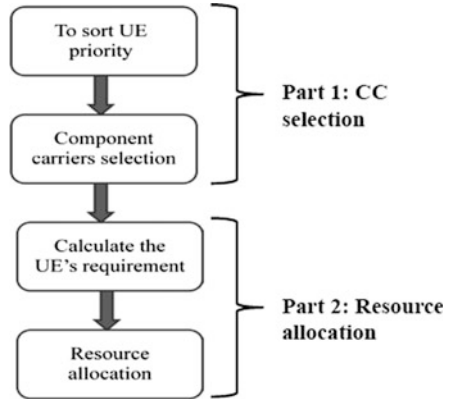


Fig. 82.2 Conceptual model of LTE and LTE-A resource allocation

**Fig. 82.3** The proposed radio resource allocation framework



are different, two weighting functions are designed to differentiate their scheduling priority. Generally, the delay budget is the major concern of GBR traffic, while the packet loss is critical for the non-GBR traffic. And in order to avoid wasting network resource, the GBR packet tends to be dropped if its queuing time exceeds the delay budget. The weighting functions of GBR UE and non-GBR UE are illustrated in (82.1) and (82.2), respectively. It is noted that, although the weighting functions of GBR and non-GBR are obtained in different way, their weightings are compared to each other to decide the priority.

$$W_{G(i)} = \left( \frac{\gamma^* M_{ave(i)}}{M_{max}} \right) \left( \frac{d_i}{D_i} \right) \tag{82.1}$$

$$W_{N(i)} = \left( \frac{M_{ave(i)}}{M_{max}} \right) \left( \frac{q_i}{Q_i} \right) \tag{82.2}$$

The  $M_{ave(i)}$  is the average MCS level of the UE  $i$  received for all un-allocated resource block groups (RBG) and  $M_{max}$  is the maximum MCS level. In (82.1),  $d_i$  and  $D_i$  denote the longest waiting time of the packets that are in the queue and the maximum delay budget of UE  $i$ , respectively. The values of  $q_i$  and  $Q_i$  indicate current queuing length and the queue size of UE  $i$ , respectively. The parameter  $\gamma$  is applied to adjust the balance of both weightings because, as the scales of delay time and queue length may be incompatible, it may introduce bias between the ratios of the second terms in both equations. The UE with larger weighting value has higher priority to choose CC. Each UE calculates its preferences of all CCs and chooses the CC with the highest preferences for radio resource allocations. The preference of UE  $i$  on CC  $n$  is calculated by the summation of the available bandwidth of all  $k$  RBGs of CC  $n$  by referring to the MCS level of UE  $i$  as (82.3).

$$P_{i,n} = \sum_{j=1}^k (R_{n,j} | m_{i,n,j}) \tag{82.3}$$

```

Algorithm: Power-saving Resource Allocation
1. Select a high-priority UE according to the priority queue
2. IF high-priority UE is LTE
3. use high priority CC for UE on the system
4. WHILE any RBG is not allocated or the UE doesn't get enough resource
5. IF the bandwidth of MCS level < UE_Requirement && MCS level != 1 THEN
6. Reduce the MCS level
7. ELSE
8. allocate resource to the high-priority UE
9. END IF
10. ENDWHILE
11. ELSE IF high-priority UE is LTE-Advanced
12. FOR priority CC on the system < LTE-A UE the number of antenna
13. WHILE any RBG is not allocated or the UE doesn't get enough resource
14. IF the bandwidth of MCS level < UE_Requirement && MCS level != 1 THEN
15. Reduce the MCS level
16. ELSE IF MCS level == 1 && use CC < LTE-A UE the number of antenna
17. Change CC
18. ELSE allocate resource to the high-priority UE
19. END IF
20. ENDWHILE
21. END FOR
22. END IF

```

Fig. 82.4 The proposed power saving-based resource allocation algorithm

where  $m_{i,n,j}$  denotes the MCS level of UE  $i$  in RBG  $j$  of CC  $n$  and  $(R_{n,j}|m_{i,n,j})$  represents the residual bandwidth of RBG  $j$  of CC  $n$  for UE  $i$  by referring to its MCS level. It is noted that each LTE UE can only choose one CC and the number of CC that the LTE-A UE can choose depends on the number of its antenna. The bandwidth to be allocated for each GBR UE follows its guaranteed bit rate. However, the bandwidth to be allocated for each non-GBR UE can be calculated in a statistic manner for fairness consideration. In this paper, we calculate the bandwidth to be allocated for the non-GBR UE  $i$  in each frame,  $B_i$  as the following equation

$$B_i = B_{ave(i)} * \left[ \beta + (1 - \beta) * \frac{q_i}{Q_{ave(i)}} \right] \quad (82.4)$$

where  $B_{ave(i)}$  and  $Q_{ave(i)}$  are the average allocated bandwidth and average queue length of UE  $i$  till now, respectively, and  $q_i$  is its current queue length. The parameter  $\beta$  is designed to adjust whether  $B_i$  tends to follow the previous allocated bandwidth  $B_{ave(i)}$  or to reflect current queuing condition. For example, if  $\beta$  is equal to 1, then the UE will always be allocated with the average allocated bandwidth.

The proposed power saving-based radio resource allocation algorithm is shown in Fig. 82.4. It is known that UE shall adopt the same MCS level for all received RBGs. If eNB allocates two RBGs to a specific UE and the channel conditions of these two RBGs are different, then the RBG with better channel quality shall yield to the RBG with worse channel quality and downgrade its MCS level. As LTE-A

UE has more than one antenna, it can choose RBG with the same MCS level from different CC without downgrading its MCS level to maximize its spectrum efficiency. However, from power saving point of view, eNB can allocate RBGs of the same CC to LTE-A UE and the UE only needs to open one antenna. It may introduce the downgrade of MCS level and then sacrifice some system throughput. It is noted that the more RBGs UE gets does not mean the more acquired bandwidth because of the downgrade of MCS level when getting more RBGs. Therefore eNB needs to check whether it is worthwhile or not to allocate additional RBG.

### 82.4 Experimental Simulations

In order to examine the efficiency of the proposed scheme, exhaustive simulations were conducted to compare the performance of the proposed scheme with the compared scheme [6]. The video streaming service and http service were applied as GBR and non-GBR traffic, respectively [9]. The channel bandwidth was assumed to be 10 MHz and the ITU Veh-A channel model was adopted during the simulations. The number of CC, each with 17 RBGs (3 RBs/RBG), and the number of antenna of each LTE-A UE were assumed to be 4 and 2, respectively. We fixed the number of non-GBR UE to be 20 and varied the number of GBR-UE in the simulations. The simulation results of mean throughput and the average numbers of antenna used per GBR UE and non-GBR UE are shown in Figs. 82.5 and 82.6, respectively.

It is noted that both of the proposed scheme and the compared scheme use the same priority decision and resource allocation scheme during the simulations. The only difference is the selection of CC. The proposed scheme selects CC according

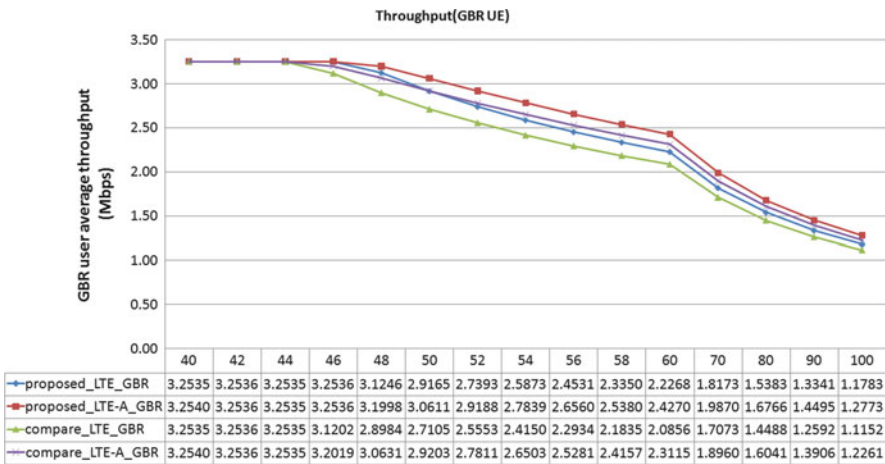


Fig. 82.5 Results of mean throughput

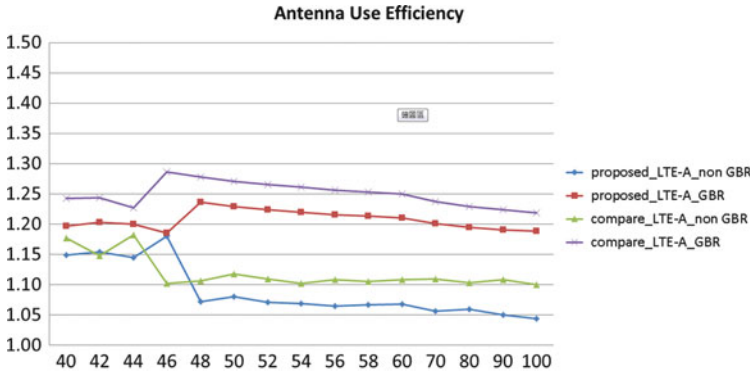


Fig. 82.6 Results of numbers of antenna used

to (82.3) while the least load first scheme is applied in the compared scheme. The simulation results show that the proposed scheme has higher mean throughput than the compared scheme. The proposed scheme uses less number of antennas than that of the compared scheme as shown in Fig. 82.6. It also shows that the number antenna decreases when the number of UE exceeds 48. This phenomenon is mainly owing to that eNB actively discards GBR packets, which exceeds the delay budget, and, as the load is a little decreased, LTE-A UE may use less number of antennas.

### 82.5 Conclusion

The power saving-based radio resource allocation scheme in LTE-A network is proposed in this paper. The proposed scheme considers the coexistence of LTE and LTE-A UE with GBR and non-GBR quality requirements. Exhaustive simulations were performed to examine the efficiency of the proposed scheme. The simulation results demonstrate that the proposed scheme can achieve the desired performance in throughput and number of antenna used when compared to the other scheme. As the power consumption model highly depends on the design of hardware architecture, the study of power saving shall be flexible enough to be adopted for new hardware design. And it may be a valuable research direction in the future.

**Acknowledgements** This research work was supported in part by the grants from the National Science Council (NSC) (grant numbers: NSC 98-2221-E-008-063, NSC 99-2218-E-159-001, NSC 100-2221-E-008-097, and NSC 101-2221-E-159-026) and Research Center for Advanced Science and Technology, National Central University, Taiwan, ROC.

## References

1. 3GPP TS 36.300 V10.3.0 Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); overall description; Stage 2.
2. Iwamura, M., Etemad, K., Fong, M. H., Nory, R., & Love, R. (2010). Carrier aggregation framework in 3GPP LTE-advanced. *IEEE Communications Magazine*, 48(8), 60–67.
3. Akyildiz, F. I., Gutierrez-Estevez, M. D., & Reyes, E. C. (2010). The evolution to 4G cellular systems: LTE-advanced. *Physical Communication*, 3(4), 217–244.
4. Wang, Y., et al. (2010). Carrier load balancing and packet scheduling for multi-carrier systems. *IEEE Transaction on Wireless Communication*, 9(5), 1780–1789.
5. Tian, H., Gao, S., Zhu, J., Chen, L. (2011). Improved component carrier selection method for non-continuous carrier aggregation in LTE-advanced systems. *IEEE Vehicular Technology Conference*, San Francisco, CA, 2011 (pp. 1–5).
6. Zhang, L., Zheng, K., Wang, W., & Huang, L. (2011). Performance analysis on carrier scheduling schemes in the long-term evolution-advanced system with carrier aggregation. *IET Communications*, 5, 612–619.
7. Wang, Y. J., Xiao, D. K., & Wang, W. J. (2010). A research on power consumption of receiver in CA scenarios. *ICIE*, 1, 247–250.
8. Yin, F. (2012). An application aware discontinuous reception mechanism in LTE-advanced with carrier aggregation consideration. *Annals of Telecommunications*, 67, 147–159.
9. 3GPP TR 25.892 V6.0.0, Technical Specification Group Radio Access Network, Feasibility Study for Orthogonal Frequency Division Multiplexing (OFDM) for UTRAN enhancement.

# Chapter 83

## Dispatching and Management Model Based on Safe Performance Interface for Improving Cloud Efficiency

Bin Chen, Zhijian Wang, and Yu Wang

**Abstract** In order to solve the performance problem of the cloud computing environment, a dispatching and management model (JDRMSP), which is based on safe performance interface is proposed in this chapter. By using the performance interface integrated into the safe DPI as the original basic data capture, agent-based job scheduling algorithm as the job dispatching method, and ant colony algorithm resource scheduling strategy as the resource management method, the integrated cloud performance is enhanced. For illustration, a simulation experimental example is utilized to show the effect of the model. From the experimental results we can get the conclusion that The JDRMSP model can analyze the cloud environment performance of various cloud components distribution pattern more accurately, and this is the basis of configuration control of the performance of the entire cloud environment, ultimately achieving the purpose of enhancement of the performance of the cloud. The JDRMSP model can effectively solve the performance data capture accuracy problem and take advantage of the dispatching and management algorithm to optimize the cloud environment.

### 83.1 Introduction

With the rapid development of cloud computing, more and more industries took cloud computing uses in business models into account [1]. As the cloud performance is deeply impacted by the system dispatching and resource management strategy, providing valid job dispatching and resource management model is essential, and the model must build on the premise of safety and efficiency. In this chapter, the performance data interface is mainly performance agent and server interface method (PASI), and it combines with the deep packets inspection (DPI).

---

B. Chen (✉) • Z. Wang • Y. Wang  
Computer & Information Engineering College of HoHai University, Nanjing 210098, China  
e-mail: [robininblue@hotmail.com](mailto:robininblue@hotmail.com)

Moreover, continuous monitoring and large system scale lead to the overwhelming volume of data collected by health monitoring tools. So, collecting exactly sufficient and valid data for performance analysis is necessary. It means , appropriate selection method and collection frequency of working data are very important for the effect and exactitude of the performance analysis. The collected information is used to take part in the analysis of enhancement and optimization of the dispatching and management model (JDRMSP).

## 83.2 Related Work

Job dispatching has become an important issue in cloud computing [2]. The job combination and dispatching strategy algorithm with dynamic programming (JCDS-D) was proposed to focus on the job allocation and the communication overheads minimizing in cloud system [3]. In contrast to the traditional job managers based on scheduler with push, light-weight job coordinator is designed to process data request by pull mode in the proposed job management system [4].

Meanwhile, the resource management is also a key factor for the performance of the cloud environment. There are various significant issues in resource allocation, such as maximum computing performance and green computing [5]. Green computing has become more and more important in recent times[6]. Energy saving is achieved by continuous consolidation of VMs according to the current utilization of resource, virtual network established between VMs, and thermal state of computing nodes [7].

The job dispatching and resource management strategy needs essential performance data as proof. The performance agent and server interface method (PASI) which consists of PMC, PMA, and PMS effectively evaluates the performance of the cloud center [8]. The PASI model can collect performance information automatically and give the analysis report according to the different granularities. However, the cloud architectures also increase dynamic data communications which inherently increase security risks. Deep Packer Inspections (DPI) is essential in protecting the cloud against malicious threats such as Web exploits, zero-day attacks, and mal-ware. [9].

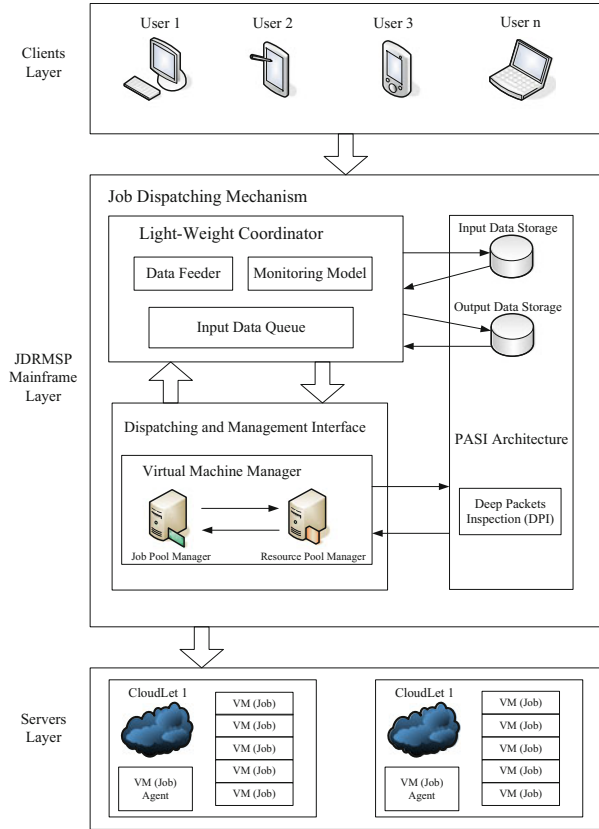
In this chapter, we focus on job dispatching combined with resource management (JDRMSP) based on the VMs which improve cloud efficiency.

## 83.3 JDRMSP Model for Cloud Computing

Job dispatching and resource management based on safe PASI architecture is divided into three layers as shown in Fig. 83.1. On the top is the clients layer; it is the portal of the cloud environment. Request messages from different terminals are gathered here, synchronized, and sent to the job dispatching mechanism.

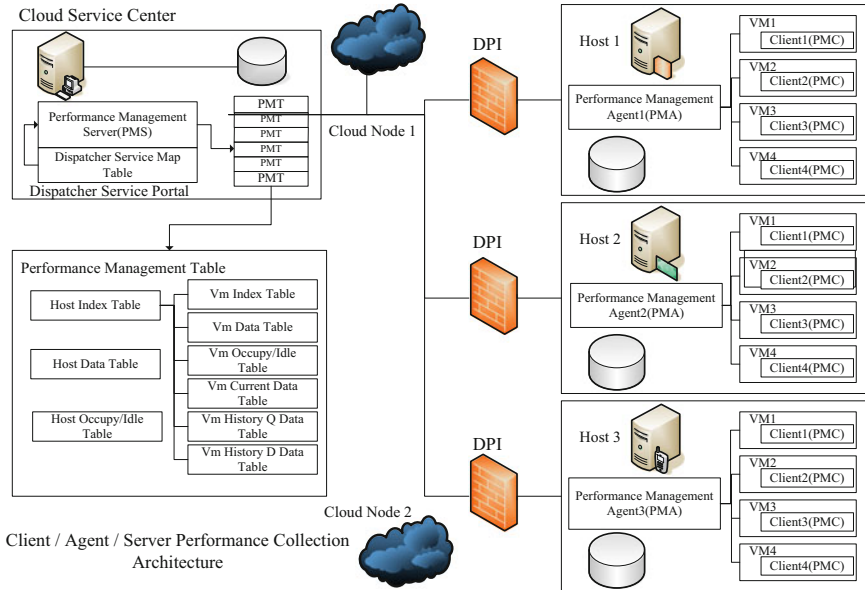


**Fig. 83.1** Job dispatching and resource management based on safe PASI architecture



The JDRMSP model uses the agent-based job scheduling algorithm as the job dispatching method, the ant colony algorithm resource scheduling strategy as the resource management method, and performance interface integrated into the safe DPI as the original basic data capture.

The job dispatching mechanism, which plays the central role in the JDRMSP Mainframe Layer, is consisted by Light-Weight coordinator, Data Storage, Dispatching, and Management Interface. The proposed job dispatching system is composed of a data feeder, queue of input data, and monitoring module. The coordinator communicates with the data storage system, which belongs to the PASI architecture, for necessary performance information exchange. The light-weight coordinator performs the essential resource analysis based on the performance statistic data and response to the dispatching and management interface for the core data source request. The essential performance analysis data is supplied by the PASI which is embedded in DPI [10]. The architecture is given in Fig. 83.2.



**Fig. 83.2** Architecture of performance client, agent, and server collection model embedded with DPI

### 83.4 Modeling and Analysis

In reality, the JDRMSP architecture plays as a powerful coordinate system for the cloud computing environment. We assumed that the jobs that come to the queue abide by the Poisson distribution process with an arrival rate of  $\lambda$ , while the process time to each job by the queue has a general distribution. So, we can build an M/G/1 queuing model with a non-preemptive system. It means that M jobs dispatching is done by the coordinator by 1 server at G time. To measure the characteristics of jobs and computing resources, we assume that jobs are in the same priority and are submitted to the cloud according to Poisson distribution with rate  $\lambda_i$ , and job dispatching system in the cloud will arrange some resources in the cloud to process each job with a general service time distribution of “T.” First, we can get the total rate  $\lambda_i$  of the arriving of the jobs:

$$\lambda = \sum_{i=1}^n \lambda_i$$

Then, we can get the relationship between the rate and time and the length of the queue in the system which is named  $L_S$ :

$$L_S = \lambda E\{T\} + \frac{\lambda^2 (E^2\{T\} + VAR\{T\})}{2(1 - \lambda E\{T\})}$$

The average number of jobs in each cloudlet waiting queue is  $K_q^i$ ;  $E[T]$  means service time probability,  $\mu$  means the service rate, and  $VAR[T]$  means the traffic intensity:

$$R_i = \lim_{\alpha \rightarrow \infty} \left[ \frac{M(T)E[X^2] + L(T)E[V^2]}{2T} \right] \quad W_i = R_i + \sum_{i=N_i}^{i-1} X_j$$

From the above equation, we can derive the  $K_q^i$  and we can get the result of the cost function of each cloudlet, and finally we can get the job dispatching cost function of the cloud environment  $\phi_i(\mu)$ :

$$K_q^i = \lambda_i * W_i \quad \phi_i(\mu) = \eta_i \mu_i + \psi_i K_q^i \quad \phi(\mu) = \sum_{i=1}^n \phi_i(\mu_i)$$

We assume that the maximum storage amount of resource is  $S$ . At the initial time  $T = 0$ , the resource pool is filled with  $S$  units of resource. In the resource management, only four kinds resource costs of cloud computing should be considered. The storage cost per resource unit in resource pool ( $E_1$ ), the maintenance cost per resource unit in resource pool ( $E_2$ ), the shortage cost of per resource unit within shortage time unit ( $E_3$ ), and the shortage cost of per resource unit ( $E_4$ ).  $F(S)$  means the total expectation cost in per unit time,  $P_n$  is the probability of the number of resources ordered but not delivered. To find the optimal resource pool capacity means to find  $S^*$  which makes a minimum worth of  $F(S)$ . We can derive the optimal resource in job dispatching function.

$$F(S) = E_1 \sum_{n=0}^S (S - n) P_n + E_2 S + E_3 \sum_{n=S}^{\infty} (S - n) P_n + \lambda E_4 \sum_{K=1}^N K q_K \sum_{n=S}^{\infty} P_n \quad JR = \phi(\mu) * F(S)$$

As the JDRMSP architecture is based on PASI which is embedded by DPI, the proportional allocation of cores for request  $G_i$  can be calculated by:

$$\theta_i(t) = size_i(t) \quad G_i = \frac{\theta_i(t)}{\sum_{j=1}^M \theta_j(t)} * R = \frac{size_i(t)}{\sum_{j=1}^M size_j(t)} * R$$

We let  $Q_i$  be the probability that core  $i$  has the largest multiplied hash value, and if we set  $\prod_{k=1}^N x_k = 1$ , Then we can derive:

$$Q_i = \sum_{j=1}^i \frac{\left(\prod_{k=1}^{j-1} x_k\right) \left(x_j^{N-j+1} - x_{j-1}^{N-j+1}\right)}{N-j+1}$$

$$Q_i = Q_{i-1} + \frac{\left(\prod_{j=1}^{i-1} x_j\right) \left(x_i^{N-i+1} - x_{i-1}^{N-i+1}\right)}{N-i+1}$$

The  $D_{cloud}$  is the most important index for the performance of the cloud [8].

$$D_{cloud_{Q/D}} = \frac{T_m * S_E(i)}{P_{Q/D}} \qquad S_E(i) = \sum_{j=i}^y S(j)$$

$$S(y) = Prob\{M_{y-1} \leq x \leq K\} \qquad S(r) = Prob\{M_{i-1} \leq x \leq M_i - 1\}$$

$$= \prod_{j=1}^{y-1} \eta^{M_j - M_{j-1}} \eta_y \left(\frac{1 - \eta_y^{K - M_{y-1} + 1}}{1 - \eta_y}\right) S_0 \qquad = \prod_{j=1}^{i-1} \eta^{M_j - M_{j-1}} \eta_i \left(\frac{1 - \eta_i^{M_i - M_{i-1}}}{1 - \eta_i}\right) S_0$$

The advantage of the model is the combined use of job dispatching and resource management strategy and making use of the PASI architecture as the performance data source and DPI as the safe guarantee [11]. We can reach the conclusion that the JDRMSP model can effectively solve the performance data capture accuracy problem.

### 83.5 Numerical Illustration

In this section, we will demonstrate the mathematical model of comprehensive effect of JDRMSP index  $Comp_{cloud}$ , which consists job dispatching factor  $\phi(\mu)$ , resource management factor  $F(S)$ , DPI factor  $Q_i$ , and PASI factor  $D_{cloud_{Q/D}}$ . From Tables 83.1 and 83.2, we can get the relationship between the change of JDRMSP factor and unidirectional and bidirectional compound cloud performance. Meanwhile, we can summarize the result from Fig. 83.3. It illustrates the influence brightly.

We can reach the following conclusion: The adjustment of PASI is the most effective method for the compound performance of the cloud environment. The job dispatching method influence is inferior to PASI. The control of resource management engenders less effect compared to dispatching of job and PASI. DPI adjustment produces a negative effect in the whole cloud environment, while the cloudlets' safe index is promoted.

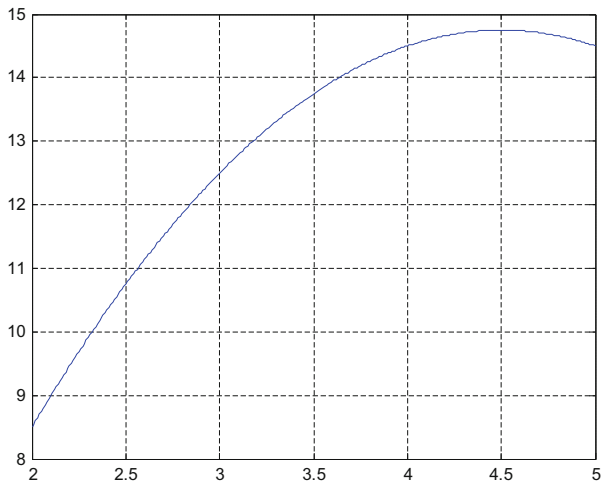
**Table 83.1** Unidirectional compound cloud performance coordinate result

$\phi(\mu)$	$F(S)$	$Q_i$	$D_{cloud_{O,D}}$	$Comp_{cloud}$
$\uparrow \Delta\phi$	$\uparrow \Delta F$	-	-	$\uparrow 2.08\Delta C$
$\uparrow \Delta\phi$	-	$\uparrow \Delta Q$	-	$\downarrow 0.11\Delta C$
$\uparrow \Delta\phi$	-	-	$\uparrow \Delta D$	$\uparrow 4.15\Delta C$
-	$\uparrow \Delta F$	$\uparrow \Delta Q$	-	$\downarrow 0.32\Delta C$
-	$\uparrow \Delta F$	-	$\uparrow \Delta D$	$\uparrow 3.36\Delta C$
-	-	$\uparrow \Delta Q$	$\uparrow \Delta D$	$\uparrow 1.73\Delta C$

**Table 83.2** Bidirectional compound cloud performance coordinate result

$\phi(\mu)$	$F(S)$	$Q_i$	$D_{cloud_{O,D}}$	$Comp_{cloud}$
$\uparrow \Delta\phi$	$\uparrow \Delta F$	$\downarrow \Delta Q$	$\downarrow \Delta D$	$\uparrow 0.33\Delta C$
$\uparrow \Delta\phi$	$\downarrow \Delta F$	$\uparrow \Delta Q$	$\downarrow \Delta D$	$\downarrow 0.94\Delta C$
$\uparrow \Delta\phi$	$\downarrow \Delta F$	$\downarrow \Delta Q$	$\uparrow \Delta D$	$\uparrow 1.47\Delta C$
$\downarrow \Delta\phi$	$\uparrow \Delta F$	$\uparrow \Delta Q$	$\downarrow \Delta D$	$\downarrow 1.28\Delta C$
$\downarrow \Delta\phi$	$\uparrow \Delta F$	$\downarrow \Delta Q$	$\uparrow \Delta D$	$\uparrow 1.12\Delta C$
$\downarrow \Delta\phi$	$\downarrow \Delta F$	$\uparrow \Delta Q$	$\uparrow \Delta D$	$\uparrow 0.28\Delta C$

**Fig. 83.3** Compound cloud performance chart



### 83.6 Conclusion

In this chapter, we have described a job dispatching and resource management model based on safe performance interface-PASI for improving cloudlets' efficiency. In this model, we can adjust the performance of the cloud environment by relevant dispatching and management method dynamically. The research result is useful for the analysis of the average performance of the current running cloud environment, affords some relevant revision, and adjusts the strategy for the cloud architecture.

In our future works, we will concentrate on the performance suppression strategy in the cloud environment, as the performance statistic methods cause some problem.

**Acknowledgements** [Foundation] The natural science foundation of Jiangsu Province in 2012 “The data in the cloud environment safe recovery of applied research.”

## References

1. Qinlong Jiang, Weibing Feng, Junjie Peng, Fangfang Han, Qing Li, Wu Zhang, et al. (2011). Inventory-based resource management in cloud computing, 2011. In *Tenth International Symposium on Distributed Computing and Applications to Business, Engineering and Science* (pp. 242–243).
2. Tai-Lung Chen, Ching-Hsien Hsu, & Shih-Chang Chen. (2010). Scheduling of job combination and dispatching strategy for grid and cloud system. *5th International Conference on Advances in Grid and Pervasive Computing, GPC 2010*, 23(2), 109.
3. Kushal Dutta. (2012). A smart job scheduling system for cloud computing service providers and users modeling and simulation. In *First International Conference on Advances in Information Technology* (pp. 192).
4. Haehyun Kim. (2011). Light-weight cloud job management system for data intensive science, 2011. In *Fourth IEEE International Conference on Utility and Cloud Computing* (pp. 625–630).
5. Huang, C.-J., Guan, C.-T., Heng-Ming Chen, Y.-W., Wang, S.-C. C., Li, C.-Y., & Weng, C.-H. (2013). An adaptive resource management scheme in cloud computing. *Engineering Applications of Artificial Intelligence*, 26(2), 141.
6. Younge, A. J., von Laszewski, G., Wang, L., Lopez-Alarcon, S., & Carithers, W. (2010). *Efficient resource management for cloud computing environments* (pp. 534). IEEE.
7. Beloglazov, A. (2010). Energy efficient resource management in virtualized cloud data centers, 2010. In *Tenth IEEE/ACM International Conference on Cluster, Cloud and Grid Computing* (pp. 303–305).
8. Chen Bin, Wang Zhijian, & Wang Yu. (2012). Performance collection model with agent and server interface for cloud computing, ICCCT2012. In *Seventh International Conference on Computing and Convergence Technology (ICCT, ICEI and ICACT)* (pp. 118–119).
9. Smallwood, D., & Vance, A. (2011) Intrusion analysis with deep packet inspection, 2011. In *International Conference on Cloud and Service Computing* (pp. 281–284).
10. Huang, C.-J., Guan, C.-T., Heng-Ming Chen, Y.-W., Wang, S.-C. C., Li, C.-Y., et al. (2013). An adaptive resource management scheme in cloud computing. *Engineering Applications of Artificial Intelligence*, 26(3), 45–48.
11. Matsumoto, H., & Ezaki, Y. (2012) Dynamic resource management in cloud environment. *Computer Programming—723 Computer Software, Data Handling and Applications* (pp. 367–371).

# Chapter 84

## A Proposed Methodology for an E-Health Monitoring System Based on a Fault-Tolerant Smart Mobile

Ahmed Alahmadi and Ben Soh

**Abstract** In the development of general system design approaches, the main concern is whether the approach meets the proposed system's specifications and the ability of the system to operate for a specified period within those specifications. However, with the expansion in the field of sensitive and complex systems such as e-health monitoring systems, a greater emphasis is placed on the behaviour of the system with the presence of fault (i.e., fault-tolerance). Consequently, when the system is being built, tasks such as fault-tolerance requirements are essential to ensure the quality of the resulting reliable e-health monitoring system. By considering the fault-tolerance requirements as functional requirements in the requirement phase, the completeness of reliability requirements for an e-health system can be developed. This paper proposes a methodology that conceptually studies fault-tolerance in relation to a smart mobile e-health monitoring system. The methodology aims to contribute towards standardising the fault-tolerant requirements of a reliable e-health monitoring system.

### 84.1 Introduction

The distractions of the operation of different major complex systems have underlined the need to develop and propose novel mechanisms that reduce the effect of disruptions and enhance the reliability level of these systems. An e-health monitoring system is one such complex system that needs to be studied and analysed in order to improve and develop its reliability. Fault-tolerance plays a significant role in achieving this aim. It may be defined as the operation of avoiding system failures in the presence of faults. However, fault-tolerant design is one of the most significant dependability attributes that are required to achieve reliability. As a

---

A. Alahmadi (✉) • B. Soh  
Department of Computer Science and Computer Engineering, La Trobe University,  
Melbourne, VIC 3083, Australia  
e-mail: [ahalahmadi@students.latrobe.edu.au](mailto:ahalahmadi@students.latrobe.edu.au)

result, the importance of fault-tolerance introduces the need to study and analyse the concepts of fault-tolerance and its mechanisms in such complex systems.

The clear expansion in the field of e-health systems, and the aim to initiate a highly reliable application relating to human life, has created the importance of such a performance requirement for both stakeholders and developers. To date, very little research has addressed fault-tolerance requirements for e-health monitoring systems. Accordingly, it appears that no clear conclusions can be drawn from previous work to clarify the mechanism of integrating the fault-tolerance concept and requirements into a smart mobile e-health monitoring system.

Thus, this paper proposes a methodology for an e-health monitoring system based on a fault-tolerant smart mobile. Our methodology aims to contribute towards standardising the fault-tolerant requirements for a reliable e-health monitoring system.

### 84.2 Proposed Methodology

With a view to studying and analysing the correlations between the attributes indicating the fault-tolerant design and the complex system QoS level required by users, we propose a methodology that is categorised into three phases, as shown in Fig. 84.1. First, phase 1 defines the concept of fault-tolerance in terms of definitions, attributes and requirements that need to be clearly understood and then developed to identify the QoS level expected by the patient, which is the second phase. The third phase describes a particular complex system—in our case, an e-health monitoring system—and the performance attributes that need to be recognised. One way to develop the system design methodology is to consider the characteristics of e-health monitoring systems and the theory associated with the concept of fault-tolerance. For instance, a design approach based on an understanding of the concept of fault-tolerance leads to the advantage of the capability of developing the design utilising the main essence of that concept. Qualitative studies and approaches are significant and required to afford a better understanding and design of a fault-tolerant system such as an e-health monitoring system [1].

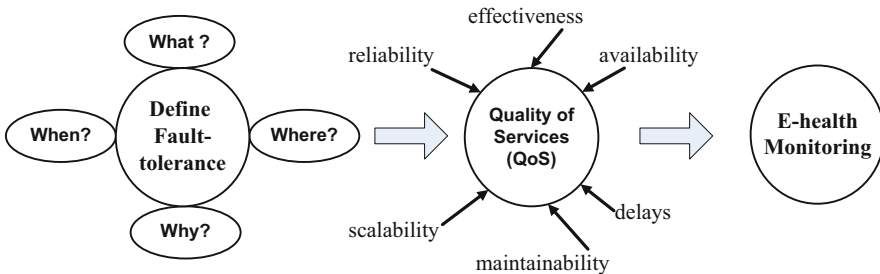


Fig. 84.1 Proposed methodology



### 84.3 Context of the Proposed Methodology

In this section, we discuss the context of our proposed methodology, which includes smart mobile e-health monitoring systems' perspectives, fault-tolerance concept clarification, and investigation and analysis of the fault-tolerance concept in an e-health monitoring system.

#### 84.3.1 Smart Mobile E-Health Monitoring System Perspectives

Towards the aim of investigating and analysing the correlation between the concept of fault-tolerance and the proposed e-health monitoring system, it is necessary to first identify and clarify the perspectives of the system. The network architecture of the system is proposed based on a high level of scalability. The design of three levels of interconnected networks is flexible enough to guarantee the availability of the system in all expected situations. Figure 84.2 shows the overall configuration of the system's network, which consists of body sensors connected via Bluetooth to a smart phone. The exact location of the patient in the home environment is recognised using radio-frequency identification (RFID) tags distributed in each part of the house. These tags also help to determine the current activity of the patient. Through the Internet, the smart phone can contact healthcare providers or paramedics directly if needed.

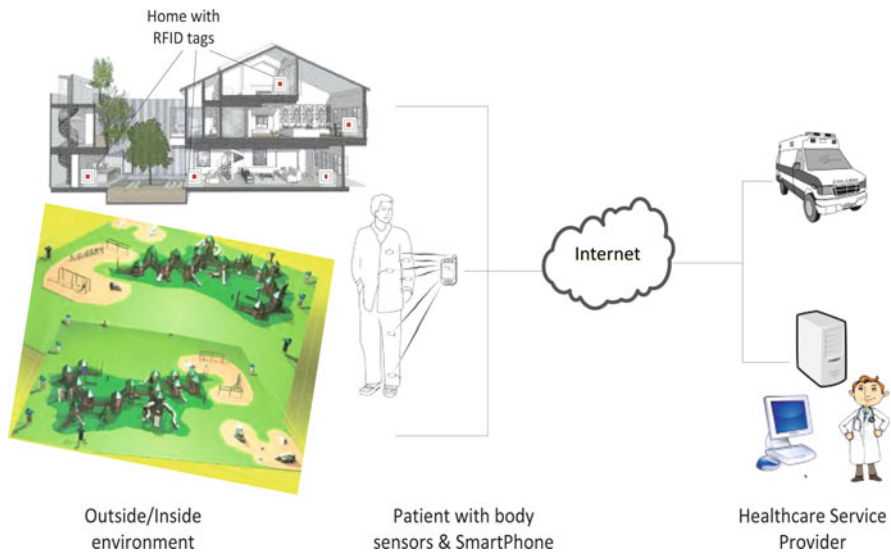
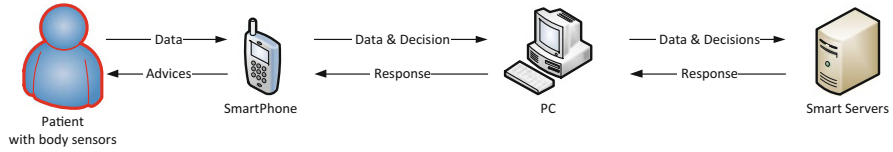


Fig. 84.2 Overall configuration of the system's network



**Fig. 84.3** The physical components and the system framework [2]

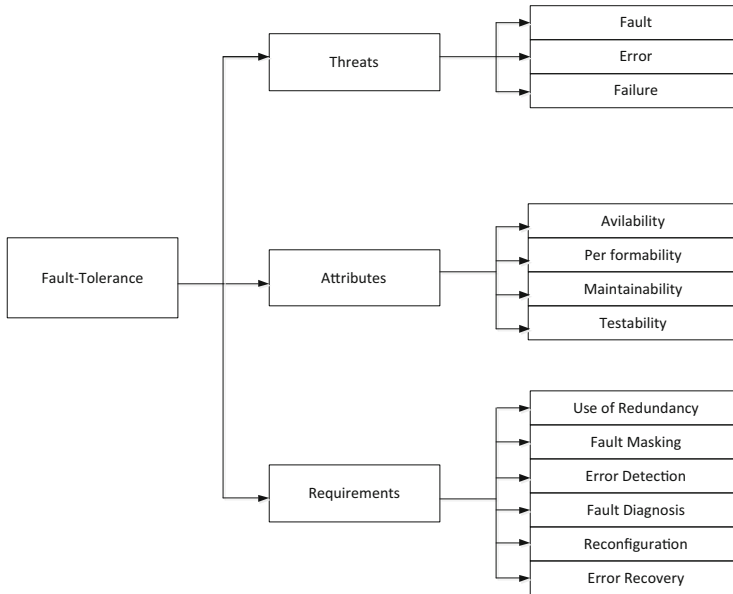
To study and identify the appropriate way of integrating fault-tolerance concept perspectives into a complex e-health monitoring system, the three main components of our system—physical, network and software—should be identified. The network component was introduced and discussed earlier. Both the physical components and the roles of each component (i.e., software component) must be clarified based on the specifications of the system and users’ requirements. In our system, the physical component and the system framework are shown in Fig. 84.3. More details about the proposed system can be found in our previous paper [2].

## 84.4 Investigation and Analysis of the Fault-Tolerance Concept in an E-Health Monitoring System

In this section, we contribute to identifying a clear understanding and pre-analytical study of the concept of fault-tolerance in a specific system—smart mobile e-health monitoring. Based on the proposed system objectives and requirements mentioned in Sect. 84.3 and the fault-tolerance conceptual analysis (Fig. 84.4), this section aims to clarify and analyze our system’s fault-tolerance definitions, threats, attributes and requirements in order to clarify the efficient methods of integrating the results into the design and implementation phases of our smart mobile e-health monitoring system.

### 84.4.1 Definition and Measures Analysis

Fault-tolerance is the ability to avoid the service failure; indeed, it is the capability of the system to keep executing the correct implementation of its programme and all other functions despite the occurrence of a fault [3, 4]. In some works, it is defined as the aim of delivering a specified service in the existence of an active fault [5, 6]. In addition, a fault-tolerant system has the ability to continue performing its operations with the appearance of faults. As mentioned in many works fault-tolerance is considered one of the significant requirements in order to obtain a dependable and reliable system [6–8].



**Fig. 84.4** The fault-tolerance concept taxonomy

The system fault-tolerance can be analysed and measured quantitatively and qualitatively [9]. Some design characteristics that qualitatively evaluate a fault-tolerant design include a safe-shutdown mode, no single point of failure and consistency specifications. In contrast, due to fault-tolerance being one of the main requirements of reliability, the quantitative measures are typically derived from the quantitative reliability parameters, including Mean Time Between Failure (MTBF), Mean Time To Failure (MTTF), Mean Time To Repair (MTTR), Failure Rate ( $\lambda$ ) and Repair Rate or Maintainability Parameter ( $\mu$ ). All of these parameters can be calculated and measured for the physical, network and software components of our system in order to demonstrate the level of fault-tolerance.

In our e-health monitoring system, the fault-tolerance concept is the ability to achieve the requirements of the system, patient and medical staff with the appearance of faults in any of the system’s three components (physical, network or software). As mentioned above, our system is considered a sensitive system due to the relations between the system objectives and patient’s health, consequently, life. As a result, denying its services completely (i.e., system failure) is not an acceptable scenario, especially for essential/critical functions such as emergency situations when a patient needs the paramedics’ attendance. The above-mentioned facts show the clear need for a fault-tolerant design in our e-health monitoring system to ensure the delivery of its services.

### **84.4.2 Threats (Faults, Errors and Failure Concept) Analysis**

The various definitions of the fault-tolerance clearly show that the main threat to fault-tolerant design is fault, followed by error and failure. These three terms, as well as the cause and effect correlations between them, are the basic and key terms in fault-tolerant design [7]. In our observation, accepting that a fault will occur in the system at some point is the secret behind the need for fault-tolerant design.

The concept of faults, errors and failures has been understood differently by different people. A system failure occurs when the achieved service deviates from the specified service, where the service specification is an agreed description of the expected service. A failure occurs because the system is erroneous: the error is the part of the system state that is liable to lead to failure (i.e., the delivery of a service that does not comply with the specified service). In its phenomenological sense, the cause of an error is a fault. Another observation is that since the relationship between faults, errors and failures is an inheritance (i.e., cause–effect) relationship, both of the concepts—fault-tolerance and error-tolerance—are usually used interchangeably.

### **84.4.3 Attributes Analysis**

Although many authors demonstrate different features of fault-tolerant design, some of the most common attributes are availability, performability, maintainability and testability [7, 9]. In some works, “graceful degradation” is defined as another attribute of fault-tolerance; however, the majority mention it as the system performability due to the very close relationship between them. Each attribute has its own definition, threats, features and means that introduce the need to study and analyse each one. Some of these attributes are discussed below.

**Availability:** Unlike reliability, availability is defined as the capability of the system to perform its requirements correctly at a given time  $t$  [3]. To remove ambiguity between the reliability and availability concepts, it is important to understand that availability measures the correctness of the required performance at a given instant of time, whereas reliability measures the correctness of the required performance during a specific period (interval) [7].

**Performability (Graceful Degradation):** The performability of a system is defined as the probability that it will either perform correctly or it will be allowed to continue operating at a diminished performance level.

While reliability is a measure of likelihood that all of the system functions are performed correctly, performability is the likelihood that a subset of the critical functions is performed correctly. However, the assumption in the analysis of availability and reliability concepts in relation to our system is that the system’s

possible states are binary—either up or down. This unsophisticated view holds true for our system if tolerating faults is not an option, but in our proposed fault-tolerant smart mobile e-health monitoring system design, many additional system states become significant—one for each possible fault prototype.

#### **84.4.4 Requirements Analysis**

Many studies show a variety of requirements for obtaining a fault-tolerant design [6, 7, 9]. In general, the use of redundancy is one of the logical and common techniques that play a significant role in achieving the goal of a fault-tolerant concept. In addition, fault-masking, error detection, fault-diagnosing and reconfiguration followed by error recovery are other techniques used in fault-tolerant design. Some of these attributes are discussed below.

**Use of Redundancy:** In fault-tolerance, redundancy is the use of ancillary elements in a system to achieve the same or similar functions as the main elements for the purpose of tolerating possible faults [3]. In a more detailed definition, it is the use of extra resources beyond what is required for a normal system operation. However, this involves some inherent costs, including hardware, software, effects on system performance and the penalties of space and power. This observation forces designers of fault-tolerant systems to cautiously and comprehensively answer the questions of “where, why and how” they use redundancy techniques in relation to their system objectives. Usually, redundancy techniques can be classified into hardware, software, information and time redundancy.

**Fault Masking:** Fault masking is the process of preventing faults from constructing an error [3]. From the above-mentioned concepts of faults, errors and failures, faults are initially the cause of errors in a system, and errors are the first effect of a fault. The aim of fault-masking techniques is to cancel the effects of faults (i.e., prevent the system from being erroneous). For example, for a negative value that will cause an error, the system should prevent its input rather than handle it.

**Error Detection:** This is the identification of errors in a system. The detection of errors ensures the system’s continual correct operation. Therefore, from a fault-tolerance design standpoint, the failure to detect an error will lead to a system failure. In general, error-detection mechanisms may be classified into two main categories: a Self-Checking (SC) mechanism, where a module can detect internal errors concurrently with normal operation before propagating results to the upper levels of the system; and a Non-Self-Checking (NSC) mechanism, where a module does not have internal error-detection capability and the error-detection mechanism is performed by the upper levels of the system.

**Reconfiguration:** Once an error has been detected and/or fault diagnosis has been carried out, system reconfiguration must be initiated, in that the faulty unit must be isolated and replaced (if spares are available). If spares are not available the mentioned performability/graceful-degradation attribute is needed. This graceful-degradation permits the system to continue operating with a reduced performance level.

## 84.5 Conclusion

The important factor that influences the quality of an e-health monitoring system is the level of dependability. Towards this aim, and respecting that faults will occur in the system at some point, we first need to obtain a system with an ability to tolerate faults. The significance of fault-tolerant design introduces the need to study and analyse the concepts of fault-tolerance and its mechanisms in a complex smart mobile e-health monitoring system. This paper presents a pre-analytical framework of the concept of fault-tolerance in relation to an e-health monitoring system. One of the best ways to develop our system design methodology is to consider the characteristics of e-health monitoring systems and the theory associated with the concept of fault-tolerance. This paper first clarifies our proposed e-health monitoring system by showing its three components with the system specifications and user requirements (physical, network and software). We then identify the concept of fault-tolerance, including definitions, threats, attributes and means. Based on these two steps, we propose a pre-analytical conceptual study of a fault-tolerant smart mobile e-health monitoring system as a third step towards implementing the results of our study in the execution phase of our e-health monitoring system in future work.

## References

1. Orlandi, E. (1990). Computer security: A consequence of information technology quality, *Proc. 1990 I.E. Int. Carnahan Conference on Crime Countermeasures, Security Technology*, 1990 (pp. 109–112).
2. Alahmadi, A., Soh, B. (2011). A smart approach towards a mobile e-health monitoring system architecture. *2011 International Conference on Research and Innovation in Information Systems (ICRIIS)*, 2011 (pp. 1–5).
3. IEEE Std 610. (1990). *IEEE Standard Computer Dictionary. A Compilation of IEEE Standard Computer Glossaries*.
4. von Neumann, J. (1956). Probabilistic logics and the synthesis of reliable organisms from unreliable components. In C. E. Shannon & J. McCarthy (Eds.), *Automata studies, annals of math studies* (Vol. 34, pp. 43–98). Princeton, NJ: Princeton University Press.
5. Edwards, N. (1994). *Building dependable distributed systems*. ANSA. Cambridge, UK: APM Ltd.
6. Avizienis, A., Laprie, J. C., Randell, B. (2001). *Fundamental concepts of dependability* (Research Report No. 1145). Toulouse, France: LAAS-CNRS.
7. Pradhan, D. (1996). *Fault-tolerant computer system design* (1st ed., pp. 5–14). Upper Saddle River, NJ: Prentice Hall Inc.
8. Oliveto, F. E. (1997). The four steps to achieve a reliable design, *Proc. 1997 National Aerospace and Electronics Conference, (NAECON)*, 1997 (Vol. 1, pp. 446–453).
9. Heimerdinger, W., Weinstock, C. (1992). A conceptual framework for system fault tolerance (Technical Report CMU/SEI-92-TR33. ESC-TR-92-033). SEI.

# Chapter 85

## Design and Application of Indoor Geographical Information System

Yongfeng Suo, Tianhe Chi, and Tianyue Liu

**Abstract** For the present situation of shortage in GIS indoor theories and insufficiency in indoor GIS applications, a set of indoor GIS research theories, indoor map cartography specifications, and related technologies closely integrated with fire-fighting industry were proposed in this paper. The indoor map cartography specifications included technological processes of the map cartography and matched data updating mechanism, convenient for fast, accurately, timely producing professional indoor map. The key technologies, such as symbol dynamic drawing, indoor outdoor seamless integration, map updating and path analysis, were preliminary applied in fire-fighting emergency rescue platform, so as to realize functionalities such as indoor and outdoor seamless expression, POI updating periodically, and the best rescue path analysis, and improve the transparent command level of the fire rescue site, and it also may have certain reference value to other emergency rescues.

### 85.1 Introduction

As indoor activities are increasingly frequent, the requirements of interior space service have become more and more urgent. When people get into the large complicated building like marketplace or airport, it is very easy to get lost or can't find the right destination because of distinguishing ambiguous direction. Therefore, the market needs indoor geographic information "indoor map" as the carrier of public space information personalized. At abroad, the Google takes the lead in

---

Y. Suo (✉)

Institute of Remote Sensing Applications, Beijing 100101, China

College of Navigation Jimei University, Xiamen 361021, China

e-mail: [yfsuo@qq.com](mailto:yfsuo@qq.com)

T. Chi • T. Liu

Institute of Remote Sensing Applications, Beijing 100101, China

releasing indoor maps of Android version based on Google Map, when the user get into the interior building, the detailed building floor plans will appear automatically. Indoor map production mainly adopted the traditional data processing methods at present, relying on data collection network accumulated by outdoor map. But as a result of field data acquisition channel and low accuracy, slow update cycle, it is difficult to ensure professional application requirements.

From the application-driven angle, this paper put forward indoor geographic information processing chain combined with the public security fire control industry with depth, and find out the indoor map data production specifications which are from data preparation to final work output. With the purpose of solving the difficult indoor data acquisition, updating large-scale batch production, long update cycle, and other key issues efficiency, the indoor GIS platform was designed, and also indoor map-related technologies were applied in fire emergency rescue, enhanced the complex building fire rescue efficiency and relief effect from the overall ascension.

## **85.2 Indoor GIS Theory Analysis**

Just like exploration in outdoor geographical space, people in research of interior space also need to develop suitable methods and technologies for interior space to ensure that the Geographic information system into the interior space with the scientific theoretical basis and the effective implementation method [1].

### **85.2.1 Conception**

Indoor GIS (short for “IDGIS”), relative for Outdoor Geographic Information System, it is a kind of computer system to collect, storage, manage, analysis, show the internal space information of the huge constructions, such as hotel, university, school, museum, trains, subway station, shopping center parking, and any other type of building. It is the general technical expression in a microcosmic and large-scale angle to express high precision space and attribute information of interior space entities. Indoor map is the direct expression of Indoor Geographic Information System. Indoor map included different floor POI (Points of Interest) information, the mapping of doors, windows and walls, different attributes of the polygon information, extended data itself, such as the height of the 3d properties and captured additional semantic information [2].



### 85.2.2 Indoor GIS Coding

Indoor GIS coding is composed of five levels, namely, “city area code,” “building,” “floor,” “room,” “entity,” the “entity” of room is minimum management unit, and each level corresponding to the only coding of indoor map, namely the division of indoor GIS coding code is from the “building” to “floor,” to “room” (or “corridor”), and finally to entity according to the hierarchical coding [3].

## 85.3 Design of Fire-Fighting Indoor GIS

### 85.3.1 Indoor Map Cartography

To meet the large-scale indoor GIS needs of firemen’s rescue work, the internal fire electronic map of building must embody POI point layer, partition wall line layer and room corridor surface layer. On that basis, setting the fire rescue plan of Key Unit of Fire Safety, provide data support for fire-fighting indoor GIS platform, and effectively carry out fire rescue work. Compared with the traditional field collection, indoor map field collection used the first acquisition with a combination of fire key position and fire control facilities Enterprise fill tool (short for “Enterprise fill tool”) which updated regularly. First acquisition information included important fire control facilities, fire unit-related information, etc.; “Enterprise fill tool” is the updated source of indoor map. The fire key units provided the fire regularly indoor change information, and Modified part of enterprise terminal will be real-time updated to fire terminal database, assuring indoor map data acquisition real-time and accuracy.

- Overall drawing efficiency  
In order to ensure overall working efficiency in drawing indoor map, the overall production process should include data preparation, field acquisition, office mapping, and quality inspection to achievements management, as shown in Fig. 85.1.
- Fire indoor two-dimensional map  
Fire indoor two-dimensional map element types concluded POI layer, partition layer, and surface layer. POI layer showed important fire-fighting facilities point information in indoor map, which was aimed to show different types of important fire control facilities with different point symbols, including fire-fighting facilities, alarm facilities, smoke facilities, evacuation indicator facilities, dangerous goods. Most of the fire POI symbol referred to the fire equipment graphic symbol national standard [4]; Partition layer showed the linear information such as partition way and wall material, which distinguished area effectively; surface layer distinguished reflect regional function division, through the filling different colors to show [5, 6].

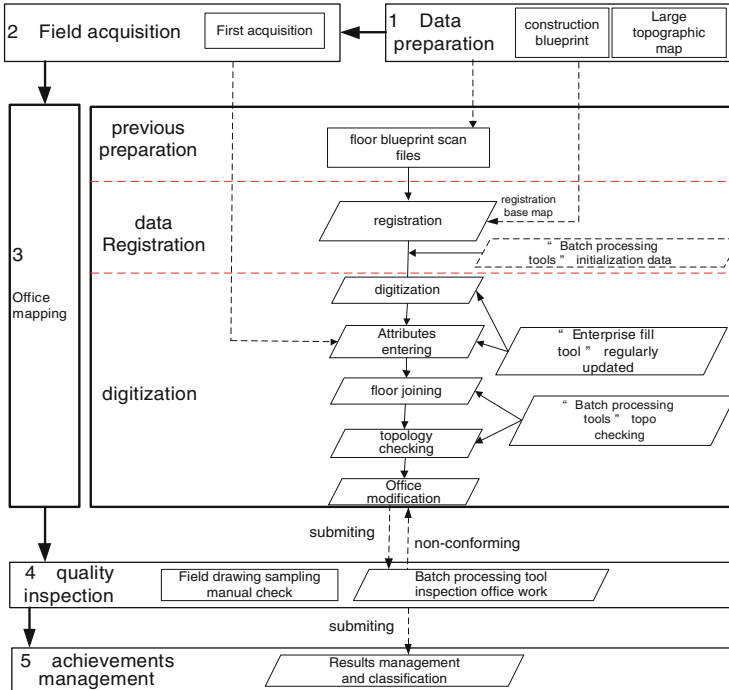


Fig. 85.1 The overall indoor map cartography

### 85.3.2 Overall Architecture

- Overall architecture design  
 To meet the demands in fire fighting and rescue work of the high-rise buildings, we need to build “fire-fighter indoor GIS platform” on the basis of indoor GIS, which contained five-layered architecture, fire-fighter GIS platform Running deck layer, and provided indoor and outdoor map-based data services and function services. The platform was based on WebService-based SOA architecture, and service was provided for all kinds of application model all the way. By using the unified description language, this platform can support application model integrated in cross-platform, professional field, heterogeneous environment [7–9].
- “Enterprise fill tool”  
 As the simply main source of indoor map data, “Enterprise fill tool” included regularly updating maintenance records such as the important fire control facilities and the latest attribute information from the fire safety key units, and realized online graphics editing function. Enterprise users created the task form, filled in or modify the returning data, and submitted it to fire department,

whose user submitted data and audited if agreed, they directly filed; if disagreed, then back to the submitter. The data modified by enterprise was finally real-time updated to fire database.

## **85.4 Key Technologies**

### ***85.4.1 Symbol Dynamic Drawing***

The map symbols were important part of the indoor map, which mainly were drawn in the following two ways. Styled Layer Descriptor (referred to as “SLD”) sign extension can be divided into three forms like point, line, and area. Different Layer corresponded to different SLD files, which were associated with the service of releasing map by GeoServer. Symbol dynamic render referred to those elements of point, line, and surface type can be redraw on the basis of extending Openlayers original symbols. For example, drawing the factor of every fault line was a dynamic way by transforming line to area, then mapping those areas, so as to produce different line symbols.

### ***85.4.2 Data Exchanging and Indoor and Outdoor Map Seamless Displaying***

Multi-source heterogeneous data exchanging was achieved through the specific trigger. The moment the trigger was met the requirements that it began to exchange, and the relevant settings in switching strategy description file documents used standard XML format, describing the data with the corresponding relations, the trigger strategy, rules script exchanging, processing plug-ins exchanging, and other information. At the same time, the platform also uses open-source tools for a variety of format conversion, such as the usage of the Ogr2Ogr exchanging ESRI Shapefile format geographic data into GeoJson format and usage of the shp2pgsql in PostgreSQL exchanging shapefile data into PostGIS database and so on. Indoor and outdoor map seamless displayed through the setting zoom level value of the indoor map, namely amplification to a certain proportion, ensuring in the premise of the same outdoor and indoor map projection, indoor map automatically overlay displaying outside the outdoor map.

### **85.4.3 Indoor Map Updating**

Indoor map data updating was proceeded by “enterprise fill tool,” the regular feedback mechanism, and the regular updated data was managed through the release mechanism. It can be divided into two kinds that one type was the “default version,” namely local area network indoor map version, also known as the father version; the other type was enterprise networks updated version the enterprise regularly updated Version, also known as subversion. Enterprise fill tool update adopted the Web Feature Service Transactions (short for “WFS-T”), this way of which also followed OGC standards, the client adding, modifying, deleting, saving operation through the WFS map Service elements.

### **85.4.4 Path Optimization Algorithm**

Based on the traditional outdoor network data structure and path analysis algorithm, path optimization algorithm comprehensively considered the open-source database software (PostGIS) and commercial GIS software (ESRI ArcGIS) team work abilities, increased stair, fire channel, fire door, fire shutter, and the fire special dynamic topology characteristic, formed the indoor and outdoor unified network data structure, and used node segment network to model more floor space between floors and buildings path analysis. Across the floors path analysis generated indoor channel network by the key nodes like stair, elevator, and then according to the network automatically generate starting points and end points and added the weight value to channel, which adopted indoor and outdoor integration variable weight path analysis algorithm, considering the fire shutter extreme situation that Fire shutter down cause weight changing. It was suitable for special emergency use. Across the buildings path analysis automatic generated indoor and outdoor unified network based on building entrances and outdoor road nodes.

In a building, optimal path analysis was firstly to call corresponding network data services; then we can call field analysis algorithm, each get the nearest point number in the network; according to entering the position of starting point and end point; at last, we can calculate the optimal path between start point to end point and any critical distance between nodes by querying statement and calling optimal path analysis algorithm. At the same time, we can also set end point and obstacle points through the network data, then recreate a new optimal path. The generating path results can control the generation of language description by establishing rich lexicon and rationalization of the rule base, specifically including generating direction and distance and so on along the route.

## 85.5 Implementation and Application Example

“Indoor GIS platform” mainly was realized by the Openlayers client technology, following OGC standard completely. This platform was simple, lightweight, and easy for coding and transmission which was as practical application in for fire fighting in Tianjin Binhai Hi-tech Zone.

The platform provided the fire commanders with the standardization and normalization of spatial information, meeting the demands of fire rescue in large buildings. Among them, “Enterprise fill tool” provided a method for rapid fire-fighting indoor rapid mapping, and increased overall efficiency of the mapping after verifications; on the other hand, “enterprises updating mechanism” was put forward for the first time, which was important to indoor facilities maintenance, guaranteeing to form a “live” fire indoor geographic information database every three months.

## 85.6 Conclusion

In view of that current indoor GIS theories were not very mature, cartography specification and application system have not formed unified standards. Combined with public security fire control industry, this paper formed a set of indoor map production standards, extended fire indoor GIS application to traditional map processing production. It will play a reference role in large-scale constructing other industry indoor map. The technology of indoor map visualization, map update, and optimal path analysis in fire indoor GIS platform can meet the preliminary application requirements in interior space. The indoor map production standards and key technologies can be copied and expanded from public security fire department to other fields, such as business, emergency, flood, quakeproof, disaster prevention, and earthquake terror. In addition, indoor geographic information system is an important technology to promote the construction of digital city or wisdom city. Geographic information system from outdoor to indoor is the inevitable trend in the future.

## References

1. Sen Xiao, Xiang Li. (2010). The research and application of geographic information system in indoor space. *Geomatics & Spatial information technology*, 33(5), 38–40 (in Chinese).
2. Summary of Indoor OSM [EB/OL]. (2012-07-15) <http://wiki.openstreetmap.org/wiki/IndoorOSM>
3. The 2000 technical guidance of Existing surveying and mapping results into the national geodetic coordinate system. (2010-04-25) <http://wenku.baidu.com/view/39d3067102768e9951e738e5.html>
4. GB/T 4327-2008. (2008). *Fire protection technical documents graphic symbol with fire-fighting equipment*. Beijing: China Standards Press (in Chinese).

5. Schafer, M. (2011). Automatic generation of topological indoor maps for real-time map-based localization and tracking, *Indoor Positioning and Indoor Navigation*. *Indoor Positioning and Indoor Navigation (IPIN), 2011 International Conference on, IEEE*, 2011, Guimaraes (pp. 1–8).
6. Bernhard Hohmann. (2010). A GML shape grammar for semantically enriched 3D building models. *Computers & Graphics*, *34(4)*, 322–334.
7. Jianjie Chen, et al., (2006). Implementation of spatial information web services based on ontology. *Journal of Zhejiang University (Engineering Science)*, *40(3)*, 376–380 (in Chinese).
8. Hailong Yu, et al., (2006). A study of integration between GIS and GIS-based model based on web services. *Science of Surveying and Mapping*, *35(2)*, 153–161 (in Chinese).
9. Lina Yang, et al., (2011). Design and implementation of digital city share-and-exchange platform based on SOA. *Science of Surveying and Mapping*, *36(6)*, 230–232 (in Chinese).

# Chapter 86

## Constructing Cloud Computing Infrastructure Platform of the Digital Library Base on Virtualization Technology

Tingbo Fu, Jinsheng Yang, Yu Gao, and Guang Yu

**Abstract** In order to improve hardware resource utilization, reduce maintenance and management costs, to build a new IT infrastructure platform for the user to provide a stable, efficient access to services. Taking the library of Harbin Institute of Technology, using VMware cloud computing solutions to build private clouds as an example, through the introduction of VMware vSphere to build a virtual architecture data center, integrate various application services, the introduction of VMware View software system provides “cloud + terminal” desktop cloud of office desktop solution. It is illustrated application of virtual technology by example. It can realize the unified management and deployment of hardware resources and the application of the data center and provide applications with high reliability, high availability, and service of mobile office environment. The IT platform can effectively gather or carrying spare computing capacity, corresponding the IT resource and service priority, improve IT management level.

### 86.1 Introduction

Modern digital library, we strengthen the construction of electronic resources in order to realize the resources to perfect, and use all kinds of high and new technology, through the powerful data statistical analysis to provide all kinds of personalized service to the readers [1]. With informationization of library constantly deepening, the traditional IT architecture can not gradually adapt to the rapid development of business needs. Virtualization technology is the most basic and core technology of cloud computing, construction of cloud computing infrastructure platform becomes the new development direction of Library Infrastructure Based on virtualization technology.

---

T. Fu (✉) • J. Yang • Y. Gao • G. Yu  
Library of Harbin Institute of Technology, Harbin 150001, China  
e-mail: [futb@hit.edu.cn](mailto:futb@hit.edu.cn)

## **86.2 Current Situation of Platform Infrastructure of Digital Library and Analysis of Virtualization Requirements**

Network, server, storage, and PC together constitute the infrastructure platform of Digital Library. This platform can run stably and reliably and meet the growing needs of applications and services. This is a guarantee to the normal operation of digital Library. Servers of Harbin Institute of Technology library mainly are based in X86, they are running more than 30 kinds of database and application system. 100 TB: local storage in IBM DS5020. PC clients mainly include more than 120 sets of office computer, nearly 500 sets for public are distributed in electronic reading room, multimedia classroom, the zone of terminal retrieval. With the development of digital resources increase and expand the application service demand of readers, equipments need to buy have gradually increased, this will give more investment and maintenance costs.

How to build the sharing system in the existing platform? Realize the unified management and deployment of hardware resource and the application of the data center are problems which Harbin Institute of Technology library faced. Based on the sufficient research, for economic and management security considerations, the VMware virtualization solutions are choice to optimization of digital library services of Harbin Institute of technology.

## **86.3 The Construction of Cloud Computing Infrastructure of Digital Library**

Cloud computing is representative of the IT technology. However, the deployment of cloud computing technology should be a gradual process; the way of overturning is not desirable. At present, according to the system application in our library and the physical equipment rate of existing and future business development needs to build a private cloud is the most realistic choice. In the construction of library virtual private cloud, virtualization is played a key role, but also can be said to be a part of cloud computing.

The overall design scheme is taken server, storage device, PC, and network seamless aggregated to “computing resources platform of distribution according to need” cloud by the VMware virtualization technology, as shown in Fig. 86.1. The first stage: it uses VMware vSphere to build a virtual architecture data center, integrates existing 30 kinds of application service, and through resource scheduling technology dynamically allocates and balances calculation resources, provides readers with the efficient 7 × 24 h access service. The two stages: library office desktop will be provided “cloud + terminal” desktop cloud solutions through VMware View software system it will realized all desktop unified management and deployment. Users install the client software in the PC, IPAD, MAC, and



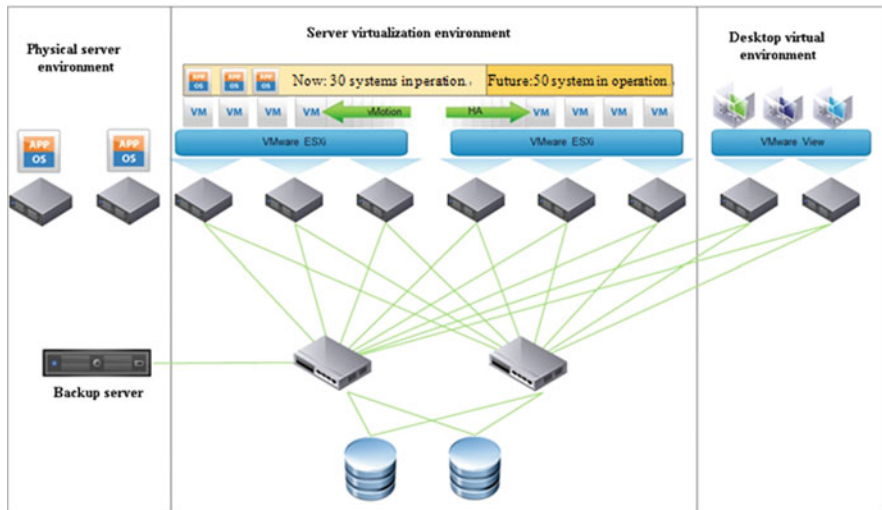


Fig. 86.1 Cloud computing infrastructure platform in library

ANDROID system. It is not restricted by time and place, user login to the server to obtain its own desktop environment, and realize application virtualization by the ThinApp component of VMware View. The application will be isolated and encapsulated to the executable file, it is separated from the underlying operating system, running in the data center server, in order to improve the compatibility and simplify application management. It is accessed through the shortcut virtual desktop.

This platform will be innovated and adjusted from the Desktop Deployment model of the original based on physical server and PC server to structure of physical and virtual combination, including physical server, server virtualization, and desktop virtualization environment. The new platform with unified computing resource pool, unified storage network, and uniform desktop environment can meet the library service on the increasing hardware infrastructure needs.

### 86.3.1 Server Virtualization Solutions

Server virtualization is taking the virtual hardware, operating system and application program “packaging” into a file, it is called a virtual machine (VM) [2].

The main part of the scheme is the six sets of installed the VMware vSphere software dual blade server. The internal connection of Virtualization architecture is as shown in Fig. 86.2 (Two sets as an example).

VMware vSphere is used for virtualization technology to construct the cloud computing infrastructure, and provides virtualization infrastructure, high

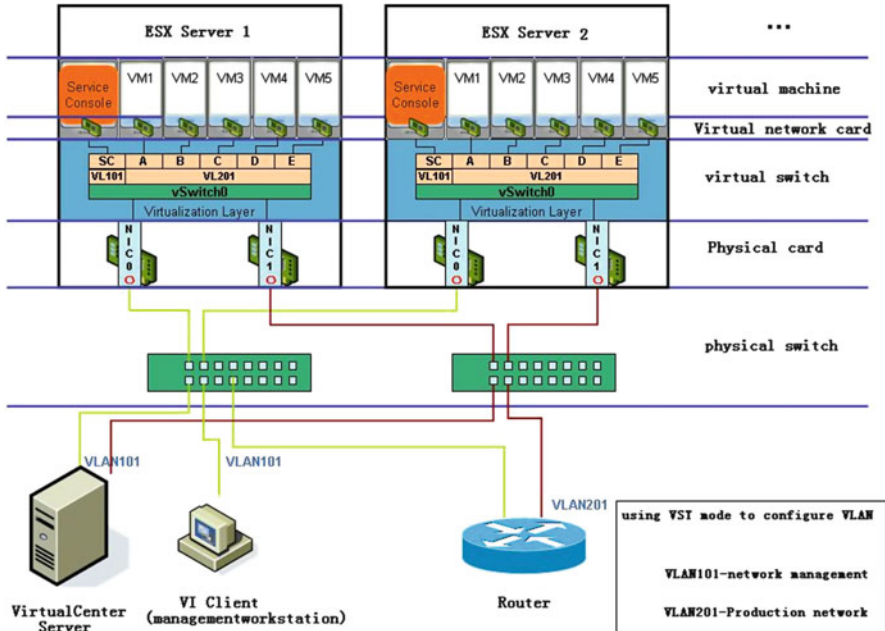


Fig. 86.2 Virtualization architecture internal connection diagram

availability, centralized management, monitoring, layer and a set of solutions. Including the key components of VMware vCenter server and ESX/ESXi, DRS, VMotion, HA, FT, Data Recovery, etc. ESXi Server is based on VMware virtual architecture suite vSphere components, which are mounted directly on the bare machine of the physical server. The processors, memory, storage, and cyber source in the physical server are abstracted to multiple virtual machines and across a large number of virtual machines sharing hardware resources [3]. Operating system windows, Linux, and other kinds of required application and database is installed on each virtual machine on the server, and then install a variety of application software, so as to make full use of the existing resources in the server, realizing the integration of business system in server level.

### 86.3.1.1 Hardware Platform

According to the analysis of performance, prior to the existing server application system and storage usage, identified by six IBM HS22 blade servers and IBM DS5020 storage array to build virtual platform, expansion after the six servers, each capable of generating 5–10 virtual server, configuration is shown in Table 86.1.

**Table 86.1** The table of the server virtualization platform configuration and extension

Cycle	Name	Unit	New number	The basic configuration or processing capacity requirements
Stage 1	IBM HS22 Blade Server	Set	Memory/CPU 6	Dual quad-core Xeon CPU, 96 GB Memory
	IBM DS5020 Storage Array	Block	10	Increase of 600 GB/SAS hard disk

### 86.3.1.2 Software Platform

The main configuration is in method of “12 + 1,” it is 12 sets of virtual enterprise architecture enhanced version of VMware vSphere (calculation of license according to the CPU number) and used authorization for one sets of VMware vCenter Management Server standard edition. VMware vCenter server is the core of virtual cluster management, for the IT environment provides centralized management, automation, resource optimization, and high availability. We should try to improve the stability and reliability of operation, so approaching the virtual machine (VM), there is a unified management by introducing the VMware vCenter virtual device.

### 86.3.1.3 The Server of Application Integration

The system integration included the library website, the papers submitted, integrated retrieval, collections search, mobile library, social networks, CALIS interlibrary loan, electronic reading room billing system, reference system, etc. In some not suitable for transfer to the virtual server platform, such as two IBM minicomputer which is loading of library automation system, for the larger demand of storage and system resource, retain the original physical server load mode.

### 86.3.2 Desktop Virtualization Solutions

Desktop virtualization can provide an independent virtual machine for each user to desktop computing, all desktop are running in the server platform, realize the unified management, when the user uses any terminal at any time and any place to login to server will call your own desktop environment, we can obtain the full PC experience at the same time.

**Table 86.2** Desktop virtualization platform configuration expansion table

Cycle	Name	Unit	Number	The basic configuration
Stage 2	Four IBM HS22 and two		Memory	Each set Up to 96 GB
	IBM 3650M3 server expansion		CPU	memory Dual CPU
	IBM DS5020 hard disk storage expansion	Block	10	2 TB/SATA

### 86.3.2.1 Hardware Platform

Desktop virtualization platform is used the four blade servers available and two IBM X3650M3 server, configuration is shown in Table 86.2. In desktop virtual environment, just take the terminal as access equipment, not for data processing, because of cost considerations, the original PC as terminal.

### 86.3.2.2 Software Platform

Software configuration is the 120 set of View 5 platinum edition software package, including vSphere for desktop user, vCenter x1, View Manager x1, Composer, ThinAPP, View Agent, View Client. VMware View is based on VMware vSphere virtualization platform built in the virtual machine, to build a complete desktop environment—operating system, application program, and configuration. The desktop is changed from the delivery of cloud service, key applications, and data on terminal equipment are safely packaged in cloud computing infrastructure with virtual environment of high number encryption [4]. Staff enjoys the hitherto unknown desktop access free, and improves the work efficiency greatly. Later it will consider providing cloud desktop service for the public and teachers who have special needs.

## 86.4 Application Effect of VMware Virtualization Technology in Digital Library

The introduction of virtual technology do not subvert the essence of the existing network architecture, it is based on the existing physical infrastructure to build virtual cluster and unified management. The original server device can be operated normally and with the virtual server together, building VLAN from network, data sharing, service isolation, etc. In the user access patterns, for the virtual server exchange data operation, it is equivalent to the traditional physical server access pattern will not cause any adverse impact on the business system.

### ***86.4.1 Reduce the Cost, Improve the Overall Resource Utilization Rate***

The number of physical devices was reduced through server consolidation. In the new IT system can satisfy 50 systems. The migration of more than 30 of the original application can be virtualized to virtualization platform. The average utilization server rate is increased from 5–15 % to 60–80 %.

### ***86.4.2 Improve Operational Efficiency***

All physical host, virtual machine, and virtual desktop were centralized management and maintenance, unified planning applications, and security configuration through the VMware vCenter Server, it can shorten the new application of on-line time, improve flexibility. The desktop is run on the server platform, which can ensure the data security, large maintenance work and can reduce the desktop, realize mobile office.

### ***86.4.3 Improve the Security of the System***

Platform management module provides VMotion function, it can immediately in the operation of virtual machine migration to another server, perform without disruption of the IT environment maintenance. With the use of HA and FT, maintain and upgrade zero downtime hardware, guarantee business system efficient and stable and uninterrupted operation [5]. The DataRecovery function can be backed up the management of all server and desktop by backup and management, further providing server and desktop data security.

### ***86.4.4 Improve the Service Level, so That the Resources and the Priority of Business Correspondence***

The VMware DRS function can balance the resource automatically and intelligently between the virtual machine, the deployment in a virtual machine running process play a role online, so that any one application can ensure the full resources to stable operation, at the same time, the application is not to use the resources at this time can also be other more resources application of temporary borrowing, it is a very good solution to some business system resource occupied a large quantity of problems in application of the peak.

## 86.5 Conclusion

Through the implementation of server virtualization and desktop virtualization solution, realizes the IT infrastructure platform in the management efficiency and resource efficiency increase, provides mobile office environment safe and convenient for the staff, while achieving system high availability and business continuity. Application of virtual technology in the library can effectively gather or carrying spare computing capacity, with the constant expansion of old, or buy high configuration server, expand the scale of virtual cluster gradually, it formed IT framework like a “cloud computing” architecture finally, the purpose was realized for raising the level of IT management.

## References

1. Ping Liang. (2012). Application of server virtualization technology in the library in Colleges and Universities. *China Computer & Communication*, 24(6), 116–117.
2. Qiusheng Dong, Wen Huang. (2009). Application of server virtualization technology in the digital library server consolidation. *Information Studies: Theory & Application*, 32(1), 119–121.
3. VMware. (2011). Getting started with ESX. <http://www.vmware.com>
4. Hua Chen. (2011). VMware View desktop virtualization solutions. <http://wenku.baidu.com/view/6226bd34ee06eff9aef807e6.html>
5. Haitao Zhu. (2012). Construction and application of virtual VMware system in the library in Colleges and Universities. *New Technology of Library and Information Service*, 16(1), 68–72.

# Chapter 87

## A New Single Sign-on Solution in Cloud

Guangxuan Chen, Yanhui Du, Panke Qin, Lei Zhang, and Jin Du

**Abstract** In order to deliver centralized visibility for login activity, reduce identity proliferation and confusion, increase user adoption and security, reduce administrative costs, and support for entire identity management lifecycle in cloud, a new single sign-on solution is proposed in this paper. By introducing OAuth protocol combined with identity federation mechanism and identity mapping, the new single sign-on model can give the cloud user that has succeed through an identity authentication the permission to access other cloud services in a reasonable time period without entering the username and password repeatedly. Empirical results show that the solution will be used as an impactful measure in scenarios where frequent interactions among different cloud services and clouds that result significant impact across multiple security domains. The OAuth-based single sign-on solution can effectively solve the problems of complexity of identity management and cross-domain authentication in cloud environment and thus increased the security and improved the user's efficiency.

### 87.1 Introduction

Since its emergence, the cloud has become one of most vigorous forces in the industry. Providing greater reliability, improved flexibility, simpler deployment, and lower costs, the cloud can bring benefits to both users and businesses [1].

---

G. Chen (✉) • Y. Du • J. Du  
People's Public Security University of China, Beijing 100876, China  
e-mail: [ericcgx@163.com](mailto:ericcgx@163.com)

P. Qin  
State Key Lab of Information Photonics and Optical Communications, Beijing 100876, China  
Beijing University of Posts and Telecommunications, Beijing 100876, China

L. Zhang  
The Logistics Academy, Beijing 100036, China

Apparently, cloud has undeniable prospect and potential. However, security remains a thorny issue and also the most convincing reason for users not moving their business to cloud.

In the cloud era, customers may use various kinds of cloud services at the same time or their data may distribute on different clouds. Thus, the customers may possess different identities in different clouds that may have different security mechanism. It is rather inconvenient to repeat the process of inputting the user name and password for each identity authentication when they need to call multiple resources that distributed on different clouds to jointly accomplish certain task. Meanwhile, it is really a nightmare for the user to manage numerous accounts and passwords. They may write down the passwords on the notepad or in the document. When the notepad is lost or the document is been accidentally deleted, then the passwords will be lost too. Furthermore, the passwords in the notepad and document are easy to be peeped by these with ulterior motives, resulting password disclosure.

Single sign-on (SSO) is just a good way to solve these problems. Single sign-on is a transparent user authentication mechanism whereby a single action of user authentication and authorization can permit a user to access multiple protected services and network resources without the need to enter multiple passwords. Considering the identity problems that obsessing the cloud users and cloud providers, single sign-on is of great significance to the promotion and development of cloud.

This paper proposed a new sign-on solution based on OAuth for cloud by combining identity federation mechanism with identity mapping. The proposed solution is shown to deliver a centralized visibility for login activity and reduce identity proliferation and confusion, while increasing user adoption and security in practice.

The remainder text is organized as follows: Section 2 introduces the principle of OAuth protocol. In Sect. 3, we provide a detailed analysis of single sign-on flow in cloud and propose a new single sign-on model for cloud. And Sect. 4 gives the experimental evaluation of the proposed solution. Finally, Sect. 5 concludes with a summary of the new solution and suggests future work.

## 87.2 Single Sign-on Based on OAuth

There are two standards making the implement of single sign-on in cloud available. One is SAML (Security Assertion Markup Language), an open standard based on XML evolved from Security Services Technical Committee of OASIS. The other one is OpenID, an open user-centric digital identity authentication frame that allows users to be authenticated by certain cooperating sites using a third party service [2]. Due to the deficiency of these two tentative approaches, we proposed the SSO solution in cloud based on OAuth in this paper.



OAuth is an open protocol to allow secure authorization in a simple and standard method from web, mobile, and desktop. It provides a method for the users that allow the third party applications to access users' protected resources stored on certain site without the need to provide the username and password to the third party application.

OAuth permits user providing a token rather than username and password to access their data stored on special sites. Each token authorize a special third party application to access certain resources in a certain period of time (e.g., allowing the application to browse the tutorials in a photo editing site in the next 3 h). Thus, OAuth allows users authorize the third party application access their data stored on other sites without sharing their accessing permission or the whole content of the data.

The authentication and authorization involve three roles:

- Service provider: providing software and hardware resource, platform and service for the users.
- User: the owner of the protected resources stored on the site provided by the service provider.
- Client: the third party application that want to access the resource of the service provider.

Figure 87.1 shows the flow of OAuth authentication.

The basic process of the OAuth authentication can be summed up as follow:

1. Client requests Request Token  
Client (third party application) asks OAuth service provider for unauthorized Request Token and makes a request to Request Token URL. The request includes `oauth_client_key`, `oauth_signature_method`, `oauth_signature`, `oauth_timestamp`, `oauth_nonce`, and `oauth_version` (optional).
2. Service provider grants Request Token  
OAuth service provider grants client's request and issues the `oauth_token` and corresponding `oauth_token_secret` that without the user's authorization to client.
3. Client directs user to service provider  
Client requests OAuth service provider for user-authorized Request Token and makes a request to User Authorization URL. The request includes the unauthorized token and secret key obtained from the previous step.
4. Service provider directs user to client  
OAuth service provider directs the user to authorize. The user may be prompted which protected resources are authorized for the client to access. The authorized Request Token may be return or not in this step.
5. Client requests Access Token  
When Request Token is authorized, client will request Access Token URL to replace the Request Token with Access Token. The request includes `oauth_client_key`, `oauth_token`, `oauth_signature_method`, `oauth_signature`, `oauth_timestamp`, `oauth_nonce`, and `oauth_version` (optional).

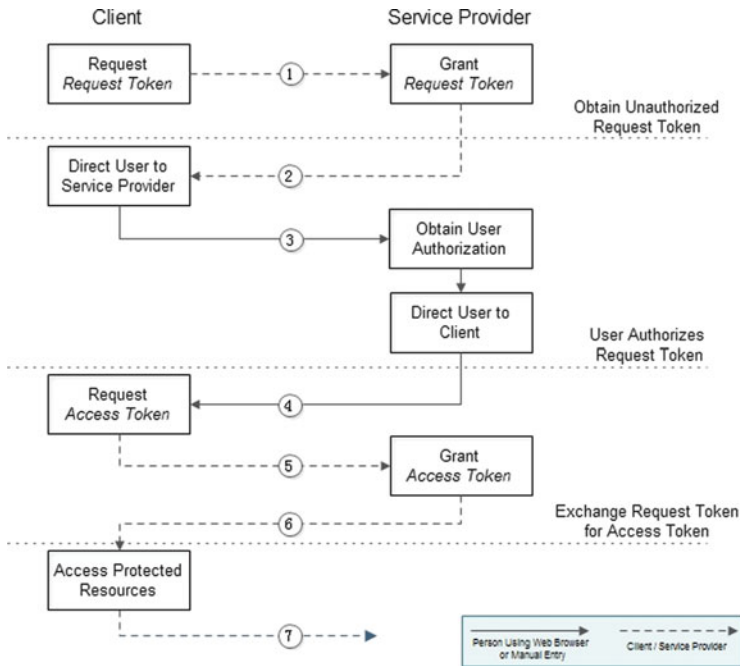


Fig. 87.1 The flow of OAuth authentication

6. Service provider grants Access Token

OAuth service provider grants the client’s request and issues the Access Token and corresponding secret key to client. The response includes `oauth_token` and `oauth_token_secret`.

7. Client accesses protected resources

Client now can use the returned Access Token to access the resources authorized by the user.

The entire OAuth authentication and authorization can be summarized as: Obtaining unauthorized Request Token; Obtains user-authorized Request Token; Replaces the authorized Request Token with Access Token.

### 87.3 Single Sign-on Model for Cloud

An effective identity authentication mechanism should give the entity that has succeed through a identity authentication the permission to access other resources in a reasonable time period without entering the username and password repeatedly. Such security mechanism must take the problems like federation of authentication domains and identity mapping that brought by the request of across multiple

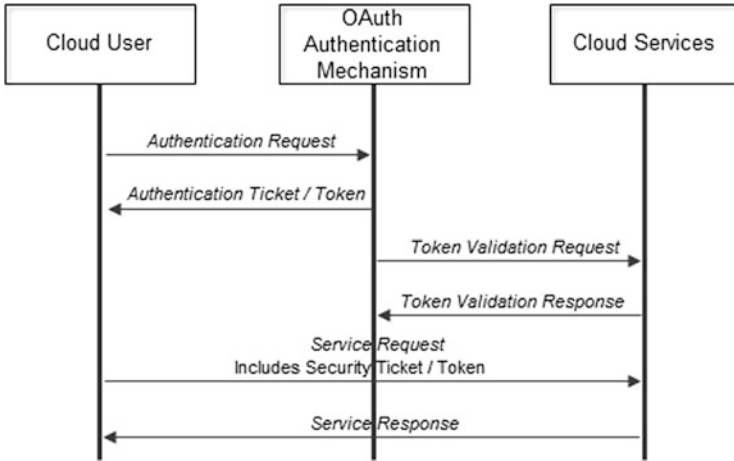


Fig. 87.2 OAuth-based single sign-on flow in cloud

security domains into consideration. So, we proposed the single sign-on solution in cloud [3]. The solution gives the user in a special logical security domain the permission to access the authorized resources through one identity authentication.

The principle of the solution of SSO in cloud is to share an SSO Token\_ID object. When the cloud user request for accessing an application system on the cloud, he will be directed to authentication system for identity authentication [4]. If the user is certified, he will receive a ticket, i.e., Token\_ID. The Token\_ID will be labeled as a user session mark and will be sent to the authentication system for checkout when user wants to access other cloud resources. As for Web application in the cloud, the Token\_ID will exist as cookie. According to this principle, we designed the SSO flow in cloud, showed in Fig. 87.2.

As can be seen from Fig. 87.2, the SSO flow in cloud can be concluded as: Cloud user requests and obtains unauthorized Request\_Token\_ID; Cloud user obtains authorized Request\_Token\_ID; Request\_Token\_ID validation.

In order to realize mutual authentication and multi-party authentication, this authentication mechanism has token full consideration of cross domains, agent, and identity federation. Through SSO, cloud user can access cloud services more conveniently.

According to the principle of SSO flow and OAuth protocol, we designed the SSO model of cloud, shown in Fig. 87.3. When a cloud user accesses cloud service A, cloud service A will direct him to a special OAuth-based authentication system. After authorized, he will receive a ticket that generated by the authentication system. In a reasonable time period which can be customized according to the security level, when the cloud user want to access cloud service B, he will send the ticket to cloud service B for checking up. Then the cloud service B will check the ticket with the authentication system. As the ticket can be recognized by the authentication system through identity federation mechanism, the cloud user now

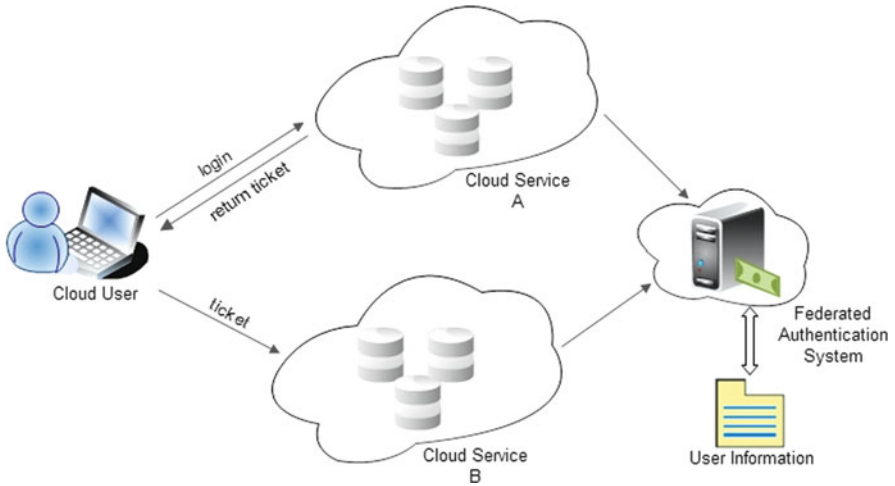


Fig. 87.3 Single sign-on model in cloud

```
GetFederationTokenRequest request = new GetFederationTokenRequest  
{  
    Duration = 3600 * 2;  
    Name = cloudUsername;  
    Policy = cloudPolicy;  
}
```

Fig. 87.4 GetFederationTokenRequest

can access cloud service B freely. In this process, the federated identity authentication will interact with the user information database so as to accomplish effective authentication.

Due to the characteristic of cloud, trust is essential to the realization of SSO among different cloud services which distribute in different corners [5]. So, we intruded federated authentication system that providing federated identity authentication service. The federated authentication system which contains one or more federated servers that sharing public trust policy can meet the requests of cloud users in the same domain or other clouds [6]. Here, we show the two important functions in identity federation example: GetFederationTokenRequest and GetFederationTokenResponse.

The mechanism retrieves the Access Policy for the authenticated user and requests temporary security credentials by calling GetFederationTokenRequest with a valid name, an expiration time set to 2 h and the policy, shows in Fig. 87.4.

The temporary security credential is returned as a response that contains: AccessKeyID (the access key identifier for the temporary credentials), SecretAccessKey (the key used to sign requests), and SessionToken (the security token).

## 87.4 Experimental Study

Here, we designed two applications (named App1 and App2) to simulate two separated cloud services. Actually, the number of application can be extended on demand. The single sign-on process that covers these two applications can be realized as follows:

When we want to access App1, we can enter the URL <http://localhost:8080/App1> and input the username and password for authentication. The OAuth authentication mechanism then will verify the correctness of the account. If verified, the Token\_ID for the App1 will be generated and the user is redirected to the URL of APP1. And thus, user can now access the resources of App1 freely. When the user wants to access App2 at the same time, he just needs to click the link of App2 (providing the URL of App2 is <http://localhost:8080/App2>) without reenter username and password. In the valid time period (we set is as 2 h), client can access these two service distributed in different “cloud” freely through a single sign-on mechanism.

In this process, the login activity presenting a centralized visibility and identity proliferation and confusion are weakened.

## 87.5 Conclusion

Summary, this paper analyzes the plight of the cloud user when they managing numerous accounts and password in their work. Then a new single sign-on solution combined with identity federation and identity mapping in cloud is proposed. This OAuth-based SSO solution delivers a centralized visibility for login activity and can reduce identity proliferation and confusion, increase user adoption and security in practice. Future work will focus on cross-cloud identity federation.

## References

1. Ravich, Y. I. (1995). “Selective carrier scattering in thermoelectric materials”, Chap 7. In D. M. Rowe (Ed.), *CRC handbook of thermoelectrics* (pp. 67–81). Boca Raton, FL: CRC Press.
2. Ravich, Y. I. (1995). “Selective carrier scattering in thermoelectric materials”, Chap 7. In D. M. Rowe (Ed.), *CRC handbook of thermoelectrics* (pp. 67–81). Boca Raton, FL: CRC Press.
3. Ravich, Y. I. (1995). “Selective carrier scattering in thermoelectric materials”, Chap 7. In D. M. Rowe (Ed.), *CRC handbook of thermoelectrics* (pp. 67–81). Boca Raton, FL: CRC Press.
4. Ravich, Y. I. (1995). “Selective carrier scattering in thermoelectric materials”, Chap 7. In D. M. Rowe (Ed.), *CRC handbook of thermoelectrics* (pp. 67–81). Boca Raton, FL: CRC Press.
5. Ravich, Y. I. (1995). “Selective carrier scattering in thermoelectric materials”, Chap 7. In D. M. Rowe (Ed.), *CRC handbook of thermoelectrics* (pp. 67–81). Boca Raton, FL: CRC Press.
6. Celesti, A., Tusa, F., Villari, M., Puliafito A. (2012). Evaluating a distributed identity provider trusted network with delegated authentications for cloud federation[C]. *Proceedings of the 2nd International Conference on Cloud Computing, GRIDs, and Virtualization* (pp. 80–85). UK: Curran Associates Inc.

# Chapter 88

## A Collaborative Load Control Scheme for Hierarchical Mobile IPv6 Network

Yi Yang, QingShan Man, and PingLiang Rui

**Abstract** With consideration of the invalid registration flows and load balance problems in hierarchical mobile IPv6 (HMIPv6) networks, a collaborative load control scheme (COLC) for HMIPv6 networks is proposed to reduce registration flows and balance load. In COLC, mobile anchor point (MAP) is allowed to transfer part of its packet delivery load to its neighboring MAPs with lower load, by which the invalid registration flows decrease, and more mobile nodes (MNs) register with their favorite MAPs without capacity expansion. The validities of the scheme in reducing registration flows of HMIPv6 and performing better load balance are examined in the simulations.

### 88.1 Introduction

The registration traffics supporting mobility lead to magnitude pressure in wireless/mobile networks. In order to reduce the registration traffics, the work group of Internet engineering task force proposed an improved mobile IPv6 protocol named hierarchical mobile IPv6 (HMIPv6) [1], which introduced the mobile anchor point (MAP) to separate micro-mobility from macro-mobility with the objective of reducing the registration traffics.

As important intermediate nodes in HMIPv6 networks, MAPs are often deployed hierarchically to avoid overload problem. And each mobile node (MN) is allowed to register with one of those MAPs whose domains cover MN's current location. Most MNs would like to register with the MAP covering larger domains in order to stay at a relatively longer time and avoid home registration

---

Y. Yang (✉) • Q. Man • P. Rui  
The 28th Research Institute of China Electronics Technology  
Group Corporation,  
Nanjing 210007, China  
e-mail: [yiyang0803@yahoo.cn](mailto:yiyang0803@yahoo.cn)

frequently, but it's hard to keep the load balance between MAPs. Moreover, the popular MAPs can't afford all the MNs' registration especially when the number of MNs is large. In this case, most of the registration will be refused, leading a large amount of invalid registration flows. Many threshold-based load control schemes [2–4] are proposed to decrease the invalid registrations and force MN to register with other MAPs when the popular MAP is busy. But in this way, most MNs can't register with its favorite MAP, which leads to frequent home registrations and brings heavy registration flows into the network. Some studies [5, 6] have proposed the point forwarding scheme to decrease the home registration frequency, but as the point forwarding scheme lengthens the routing path, the schemes are proper for the network with small flow of data packets only.

In this paper, we introduce a collaborative load control (COLC) scheme to solve the problem described above. The proposed load control scheme allows MAP to transfer part of its packet delivery load to its neighboring MAPs with lower load and helps MAP to accept more registration from MNs without capacity expansion. This scheme helps to decrease the invalid registration flows, and allow more MNs to register with their favorite MAP, and plays good effect on the registration flow. We evaluate the performance of the proposed scheme through simulations, which shows that the collaborative load control scheme has a significant improvement on decreasing registration flows especially for scale HMIPv6 network and has a better load balance performance.

## 88.2 Collaborative Load Control Scheme

In this section, the COLC scheme running by MAP is introduced, which helps to decrease the invalid registration flows and allow more MNs to register with their favorite MAP.

### 88.2.1 Load Control Scheme

In our scheme, each MAP saves the topology map locally, which records the deployment of local network and is updated with the periodic RA from the surrounding MAPs including their coverage and load. Using topology map, MAP can search for a capable MAP list (CML) according to MN's location dynamically, which contains information (address, distance, etc.) about MAPs whose domain covers the AR.

Figure 88.1 shows the key mechanism of COLC and describes the decision-making procedure of MAP when receiving a registration message from MN.

Receiving a regional registration message, MAP searches the MN in its binding cache (BC) with the RCoA encapsulated in the registration message. Binding cache holds all the MNs primarily served by the MAP with their LCoA and RCoA recorded.

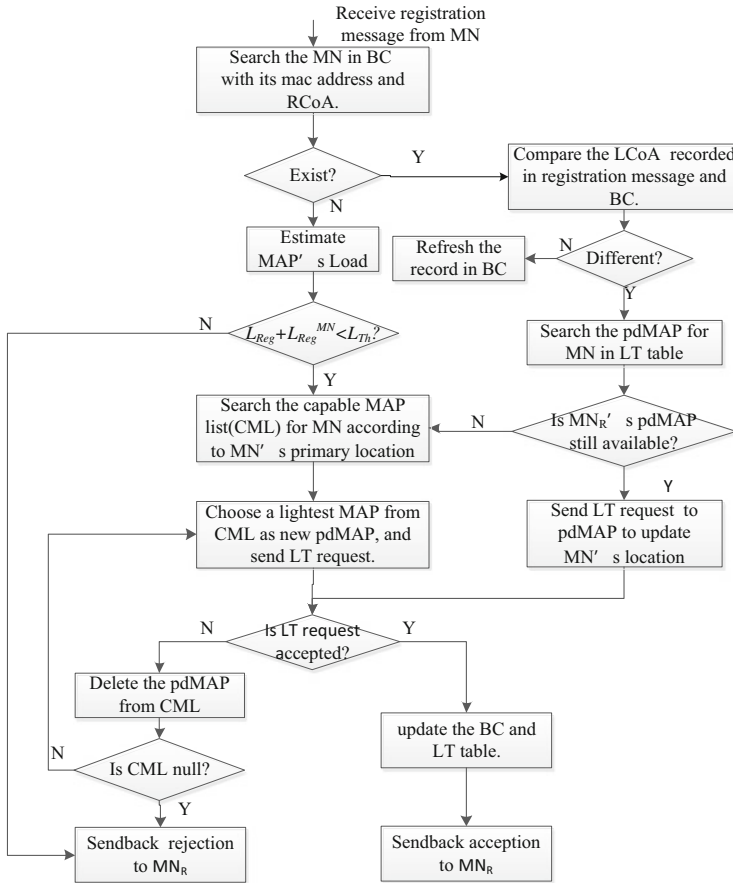


Fig. 88.1 Key mechanism of COLC

When finding an entry for MN in BC, MAP compares the LCoA recorded in entry and encapsulated in the registration message. If the two LCoA are the same, MAP just refreshes the entry. Otherwise, MAP will search the LT table to find the previous packet delivery MAP (pdMAP) for MN, which will direct packets to MN's new location. If the MN still stays in its previous pdMAP's coverage, MAP will start the load transition procedure to update the new LCoA to the previous pdMAP. If the MN has moved out of its previous pdMAP's coverage, MAP needs to choose a new pdMAP with the lightest load from MN's CML and then sends load transition message to the new pdMAP.

If there is no entry for MN in BC, the MN must be a newly accessed node. MAP should check out its load condition to determine whether to accept or reject the newly coming registration. In our scheme, MAP's load consists of packet delivery load (LPd) and registration processing load (LReg). Packet delivery load describes the occupied system capability for packet delivery, which can be computed



according to the packet delivery ratio of MAP. Registration processing load describes the occupied system capability for processing registration message, which can be computed according to the registration processing ratio of MAP. In order to ensure MAP's load does not exceed its capability, the total load threshold (LTh) of MAP is set in our scheme. The registration from newly accessed MN will be accepted with the following conditions met:

$$L_{Reg} + L_{Pd} + L_{Reg}^{MN} < L_{Th} \quad (88.1)$$

$$L_{Reg}^{MN} = \mathbf{Ratio}_{moving} \times \varnothing \quad (88.2)$$

In the above condition,  $L_{Reg}^{MN}$  is an estimation of the increasing registration processing load when MAP accepts MN's registration, which can be computed with the packet receiving ratio of MN as (88.2).  $\varnothing$  is the load estimation parameter. Then MAP needs to validate the RCoA with DAD detection and continue the process only when the detection is passed. After the DAD detection, MAP needs to choose a new pdMAP with the lightest load from MN's CML and then sends load transition (LT) request to the new pdMAP. Otherwise, the registration will be rejected.

When a load transition message is received, which is extended from ICMPv6 [7] with MN's LCoA and RCoA encapsulated, MAP will check whether its load meets the conditions as follows:

$$L_{Reg} + L_{Pd} + L_{Pd}^{MN} < L_{Th} \quad (88.3)$$

$$L_{Pd}^{MN} = \mathbf{Ratio}_{pktRec} \times \delta \quad (88.4)$$

$L_{Pd}^{MN}$  is the estimation of increased load when MAP accepts the load transition request, which can be computed with the packet receiving ratio of MN as (88.4).  $\delta$  is the load estimation parameter. If MAP's load meets the condition show in (88.3), the MAP will accept the load transition request and record the MN in LT table with the source address of load transition request as RMAP. Otherwise, the load transition request will be denied.

When the load transition response is received, MAP will continue its registration processing procedure. If the LT request is accepted, MAP updates binding cache with MN's new location and pdMAP. Otherwise, the pdMAP should be deleted from CML, and MAP continues the load transition procedure with the new pdMAP chosen from CML until CML is null.

### 88.2.2 Registration Procedure

The regional registration procedure in the COLC scheme is shown in Fig. 88.2.

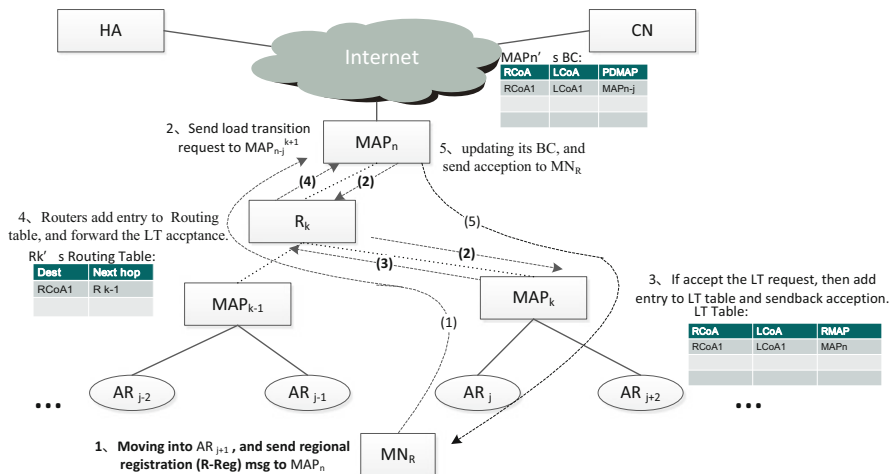


Fig. 88.2 Regional registration procedure

When an MN moves into the area covered by a new access router,  $AR_j$ , it firstly sends a regional registration message (binding update message) to the MAP covering  $AR_j$  (Step 1 in Fig. 88.2), which might be its previous MAP or a new one depending specially on MN's MAP selection scheme. The MAP-received registration message,  $MAP_n$ , decides whether to accept this registration following the decision-making procedure shown in Fig. 88.1 and sends load transition request to the selected pdMAP (Step 2 in Fig. 88.2),  $MAP_k$ . When the load transition request is accepted,  $MAP_k$  adds an entry to its LT table which records MN's RCoA-LCoA address pair and  $MAP_n$  and then sends the load transition response to  $MAP_n$  (Step 3 in Fig. 88.2). As load transition message is extended from ICMPv6 [7], each router on the way from  $MAP_k$  to  $MAP_n$  will extract the message and add a routing entry which will direct packets destined to MN's RCoA to  $MAP_k$  (Step 4 in Fig. 88.2). When  $MAP_n$  received the load transition response from  $MAP_k$ ,  $MAP_n$  updates its binding cache with MN's RCoA-LCoA address pair and pdMAP's address and then sends back the registration response to MN (Step 5 in Fig. 88.2). If  $MAP_n$  is not the previous MAP of MN, then MN needs to register with its HA to announce its new RCoA after the regional registration.

In our scheme, binding cache is extended with pdMAP field, which records the address of the MAP sharing the responsibility of packet delivery. If the pdMAP field is set null, then the MAP has to take both the responsibility of registration processing and packet delivery.

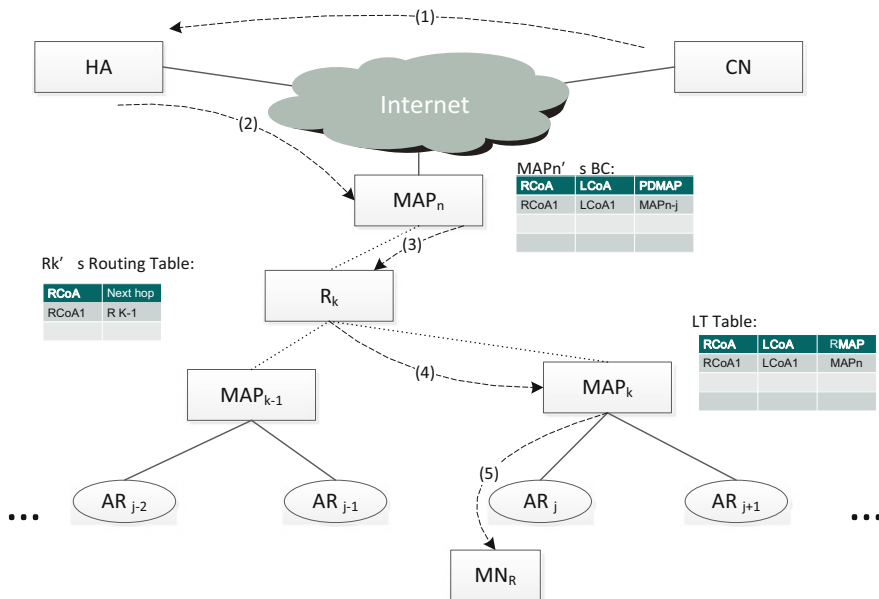


Fig. 88.3 Packet delivery procedure

### 88.2.3 Packet Delivery Procedure

After the regional registration procedure, MN continues its previous communication. The flow chart of the packet delivery procedure is shown in Fig. 88.3.

When a CN has packets to be sent to MN, the CN firstly sends the packets to the MN's home address (Step 1 in Fig. 88.3), which will be directed to MN's home agent (HA). Then, the HA intercepts the packets (Step 2 in Fig. 88.3) and tunnels them to the registered MAP of the MN (Step 3 in Fig. 88.3), MAP<sub>k</sub>. Since MAP<sub>n</sub> and all the routers on the way to MN's pdMAP maintain the routing information about RCoA, the received packets will be transferred directly to MN's pdMAP, MAP<sub>k</sub> (Step 4 in Fig. 88.3). MAP<sub>k</sub> maintains the mapping information between the RCoA and the LCoA in LT table; the MAP will re-tunnel the received packets to MN's primary location (Step 5 in Fig. 88.4).

In the above procedure, MN's pdMAP is responsible for packets re-tunneling, which will largely reduce the load of its registered MAP. In some cases, MN's pdMAP can also be its registered MAP.

### 88.3 Simulation

In this section, the performance of the proposed scheme is evaluated through the network simulator of OMNET++ [8] with extension of xMIPv6 [9]. By rewriting xMIPv6 with the load control scheme proposed in our paper, we simulate the HMIPv6 and the threshold-based load control scheme proposed in RFC3775, respectively, as to compare the performances of each scheme.

#### 88.3.1 Simulation Environment

Figure 88.4 shows the simulation topology consisting of 64 ARs and 21 MAPs deployed hierarchically.

We deploy 64 ARs uniformly with wireless radius of 50 m in the rectangular area of 720 m × 720 m, and three layers of MAPs are deployed upon the ARs, i.e., 16 first-layer MAPs, 4 second-layer MAPs, and 1 third-layer MAP. The first-layer MAPs connect directly to the ARs. Each first-layer MAP covers four neighboring ARs, respectively. The second-layer MAPs are two hops away from the nearest AR, which directly covers four neighboring first-layer MAPs, respectively. The third-layer MAPs are three hops away from the nearest AR, which directly covers four neighboring second-layer MAPs, respectively. Not all the MAPs from the same category do crossover.

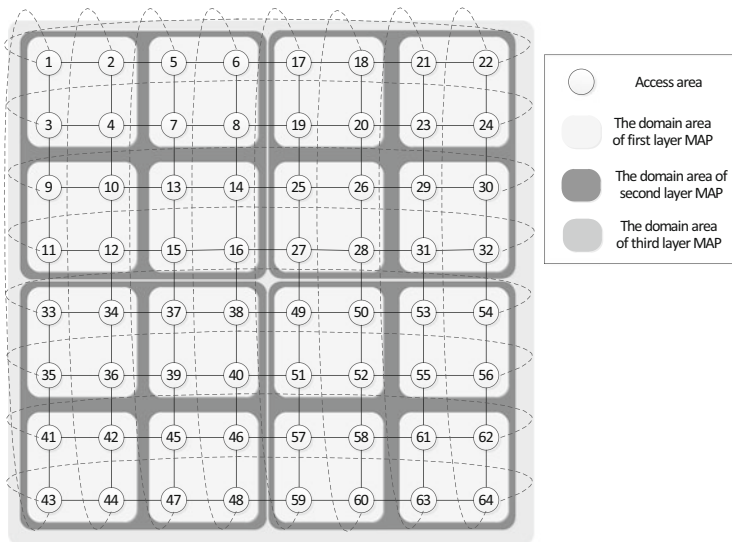


Fig. 88.4 Simulation topology

**Table 88.1** Parameter values for simulation

$\Phi$	$\delta$	$L_{Th}$	$Ratio_{pktRec}$	Init value of $Ratio_{moving}$
1	1	100	1	0.1 0.2
Number of MNs deployed				Speed range
50	200	600	1,000	5–15 m/s 15–25 m/s

The MN' mobility follows the random-walk mobility model [10], in which the routing probability for each direction is identical. In addition, the wrap around model [11] is adopted to eliminate the boundary effects. That is, the possible directions from cell 1 are 2, 3, 22, and 43. The velocity of the MN follows the uniform distribution. In order to examine the performance under deferent load condition, we simulate the mobile environment with 50 MNs, 200 MNs, 600 MNs, and 1,000 MNs with speed ranges of [5, 15] or [15, 25], respectively. Each MN follows the distance-based MAP selection scheme proposed in RFC 3775 [1], in which MNs will choose the farthest MAP from MN.

Table 88.1 shows the parameter values used in the simulation.

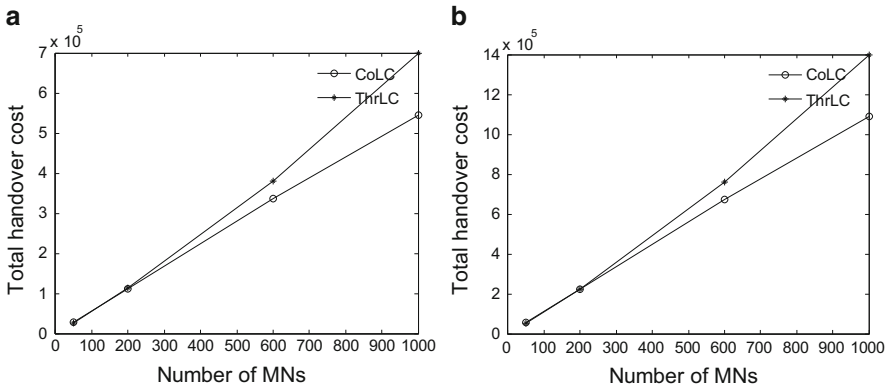
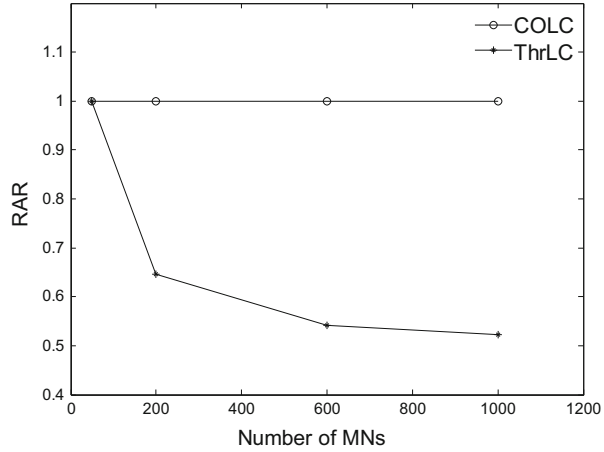
### 88.3.2 Simulation Result

Simulation evaluated the load control scheme with three performance metrics: registration admission ratio, registration cost, and load condition of each level MAP.

Figure 88.5 illustrates the average registration admission ratio (RAR) under different load conditions with speed ranges of [5, 15]. RAR is the rate of successful regional registration. The larger RAR there is, the less regional registrations are rejected by the MAP, which helps to reduce the registration flows in the network and decrease the delay spent for registrations. Comparing the performance of the two load control schemes, we can find that the performance of the two schemes is closely related with the load conditions. When there is 50 MNs simulated, the average RAR of two schemes is practically the same. With the increasing of MNs, COLC performs better and better than the threshold-based load control scheme (ThrLC) in RAR. When the number of MNs is increased to 1,000, the average RAR of COLC is significantly lower than ThrLC. It is because MAPs can transfer part of their packet delivery load to neighboring MAPs to accept more registrations in collaborative load control scheme. Comparing with the threshold-based load control scheme in RAR, we can conclude that our scheme can significantly improve the RAR of HMIPv6 network especially for the high-density networks.

Figure 88.6 illustrates the total registration costs under different load conditions with speed ranges of [5, 15] and [15, 25], respectively. Registration cost is a metric for the registration traffic in network, which can be computed based on the number and transmission range of registration messages as follows:

**Fig. 88.5** Average registration admission ratio (RAR)

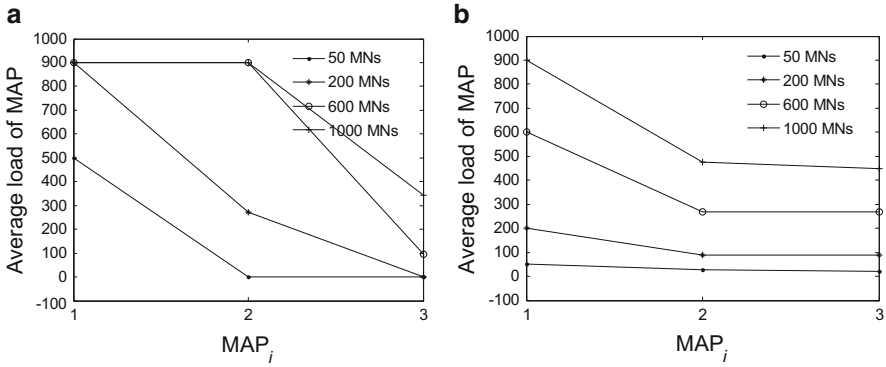


**Fig. 88.6** Registration cost. (a) Speed range [5, 15]. (b) Speed range [15, 25]

$$C_{Reg} = \sum Hops_{ri} \times Weight_{ri} \tag{88.5}$$

Where  $Hops_{ri}$  is the transmission hops of registration message  $ri$  and  $Weight_{ri}$  is the one-hop transmission cost of  $ri$ . In our simulation,  $Weight_{ri}$  is set to 1.

From Fig. 88.6a, we can see that there exist two conditions in registration cost comparison. When the number of mobile nodes is small (50 or 200 in our simulation), the total registration cost of COLC scheme is a little higher than threshold-based load control scheme. It is because that the registration procedure applied with COLC scheme is a little more complex than the threshold-based load control scheme, which leads to higher registration cost. However, the total registration cost of COLC scheme becomes much smaller than threshold-based load control scheme when the number of mobile nodes increases. It is because COLC scheme performs better and better in registration admission ratio with mobile nodes



**Fig. 88.7** Load balance performance. (a) Threshold-based load control scheme. (b) Collaborative load control scheme

increasing and leads to less invalid registrations and lower registration traffic. Comparing Fig. 88.6a, b, we can find that the velocity of mobile nodes has a significant effect on the registration cost. It is obvious that COLC scheme can perform much better than the threshold-based load control scheme in decreasing registration traffics for HMIPv6 network, which affords larger number of mobile nodes with high velocity.

Figure 88.7 illustrates the load balance performance of the two load control schemes under different load conditions. In our simulation, MAPs are sorted into three classes, which are identified as MAP<sub>1</sub>, MAP<sub>2</sub>, and MAP<sub>3</sub>, according to the distance from the MAP to its closest AR. MAP<sub>1</sub> is the farthest MAP from AR (three hops away) and manages the largest domain area. MAP<sub>3</sub> is the closest MAP from AR (one hop away) and manages the smallest domain area. As MAP's domain has an important effect on mobile node's regional registration, mobile nodes will firstly register with the MAP<sub>1</sub> to acquire lowest regional registration ratio, which always leads load imbalances. Comparing Fig. 88.7a, b, we can see that COLC scheme performs better on load balance of MAPs than threshold-based load control scheme.

From the simulation results above, we can get a conclusion that the collaborative load control scheme has a significant improvement on the registration admission ratio for the large-scale HMIPv6 network, which helps to reduce the registration traffic in the network. In addition, our scheme performs better on load balance.

### 88.4 Conclusion

In this paper, we proposed a collaborative load control (COLC) scheme for hierarchical mobile IPv6 to improve the registration admission ratio for mobile nodes, which helps to reduce registration traffics in the network. With COLC scheme, MAPs are allowed to transfer part of their packet delivery load to neighboring

MAPs with lower load, and the MAPs accept more registration from MNs without capacity expansion. In this way, the registration traffic is reduced. The simulation results show that our scheme markedly reduced the registration flows of HMIPv6 and has a better load balance performance.

## References

1. Soliman, H., Castelluccia, C., & EL Malki, K. (2005). *Hierarchical mobile IPv6 mobility management*. IETF RFC 4140.
2. Kim, Y., Kim, M., & Mun, Y. (2006). Performance analysis of the mobility anchor point in hierarchical mobile IPv6. *IEICE Transactions on Communications*, E89-B(10), 2715–2721.
3. Wang, Y. H., Huang, K. F., & Kuo, C. S. (2008). Dynamic MAP selection mechanism for HMIPv6. In *International Conference on Advanced Information Networking and Applications* (pp. 691–696). Okinawa: Institute of Electrical and Electronics Engineers.
4. Zhou, W., & Hong, P. L. (2008). A MAP-controlled load balance scheme for hierarchical mobile IPv6. In *International Conference on Wireless Communications* (pp. 691–696). Dalian: Institute of Electrical and Electronics Engineers.
5. Yi, M., & Hwang, C. (2004). A pointer forwarding scheme for minimizing signing cost in hierarchical mobile IPv6 networks. *Lecture Notes in Computer Science*, 32(7), 333–345.
6. Yi, M., Choi, J., & Yang, Y. (2007). A pointer-forwarding scheme for minimizing signaling costs in nested mobile networks. In *International Conference on Networks* (pp. 230–234). Adelaide: Institute of Electrical and Electronics Engineers.
7. Conta, A., Deering, S., & Gupta, M. (2007). *Internet control message protocol for the internet protocol version 6 specification*. IETF RFC 4443.
8. <http://www.omnetpp.org/omnetpp>
9. <http://github.com/zarrar/xMIPv6>
10. Camp, T., Bolen, J., & Davis, V. (2002). A survey of mobility models for ad-hoc network research. *Wireless Communication and Mobile Computing*, 2(5), 483–502.
11. Zeng, H., Fang, Y., & Chlamtac, I. (2002). Call blocking performance study for PCS networks under more realistic mobility assumptions. *Telecommunication Systems*, 19(2), 125–146.



# Chapter 89

## A High Efficient Selective Content Encryption Method Suitable for Satellite Communication System

Yanyan Xu, Bo Yang, Zhengquan Xu, and Tengyue Mao

**Abstract** Data transmitted by satellite communication system should be encrypted in order to provide confidentiality. A selective content encryption method suitable for satellite communication system is presented in this chapter, the key content information in the compressed stream is extracted and encrypted, and the variable modulus encryption method is proposed to solve the problem of variable length code encryption; thereby, the encrypted stream can be format compliant. This method can improve the efficiency of encryption and achieve fast, secure, and high efficient encryption of satellite communication system. The experimental results prove the effectiveness of our method.

### 89.1 Introduction

With the rapid development of broadband satellite communication system, its security is becoming an important issue. The data stream transmitted by satellite links is often very important and cannot be accessed by unauthorized users. Therefore, the data confidentiality is the most important security requirement of satellite network [1, 2].

At present, the most common encryption method in satellite communication is to use secure transmission protocols to do channel encryption, such as IPSec and TLS/SSL [2, 3]. Although IPSec has been successfully used in the Internet, it

---

Y. Xu (✉) • Z. Xu

State Key Lab of Information Engineering in Surveying, Mapping and Remote Sensing,  
Wuhan University, Wuhan 430079, China

e-mail: [xuyy@lmars.whu.edu.cn](mailto:xuyy@lmars.whu.edu.cn)

B. Yang

The Academy of Satellite Application, Beijing 100086, China

T. Mao

College of Computer Science, South-Central University for Nationalities,  
Hubei 430074, China

will cause some special problems when it is used in satellite networks, such as incompatibility with TCP performance enhancement technologies [4, 5]. SSL and TLS are only suitable for TCP connections and cannot be used in UDP connections. Moreover, the cost of secure transmission method using channel encryption is relatively high, the delay caused by encryption is high [6], and it is unacceptable to the multimedia service which has high requirement for real-time performance.

An encryption method based on DVB\_RCS satellite communication network is presented in this chapter. A selective encryption method is used to extract key content information, which has most important effects on data reconstruction, and a variable modulus encryption method is proposed to solve the difficult problem of variable length coding codeword encryption; therefore, ciphertext can be format compliant. The method reduces the amount of data needed to be encrypted and improves the encryption efficiency. The experimental results prove the effectiveness of the method.

## 89.2 The Secure Requirement of DVB-RCS Satellite Communication System

The DVB-RCS standard can provide real-time multimedia services and it has been widely used in commercial broadband satellite systems. The structure of traditional DVB-RCS satellite communication system is presented in Fig. 89.1. The media streams are encapsulated in IP packet to transmit, while the forward link and the return link are asymmetric and the transmission delay is high. Data stream between center station and VSAT needs one-hop link to transmit and the delay is about 600 ms, while data stream between two VSATs needs two-hop links to transmit and the delay is about 1,200 ms. Thus, the DVB-RCS system has high requirements to encryption efficiency in order to ensure that data stream can be processed in high efficiency and satisfies the real-time requirement of multimedia service. At present, most satellite systems encrypt all data stream to get data confidentiality [7]; however, multimedia information, especially video information transmitted by satellite communication systems, is massive; if the data is encrypted without differential, then it will result in time and computing resource consumption and lead to high delay.

Image and video information are often compressed first before they are transmitted in order to save bandwidth. The compressed media stream is composed of several kinds of information, such as flag information, padding information, coding control information, channel encoding information, and source content information. There exist some problems if all information is encrypted: (1) Encrypting fixed format fields will cause the known-plaintext attacks. (2) Most format information does not consist of real source content information. The encryption of format information will result in computing resource consumption and lead to unnecessary overhead. (3) The encryption of some channel information such as synchronization

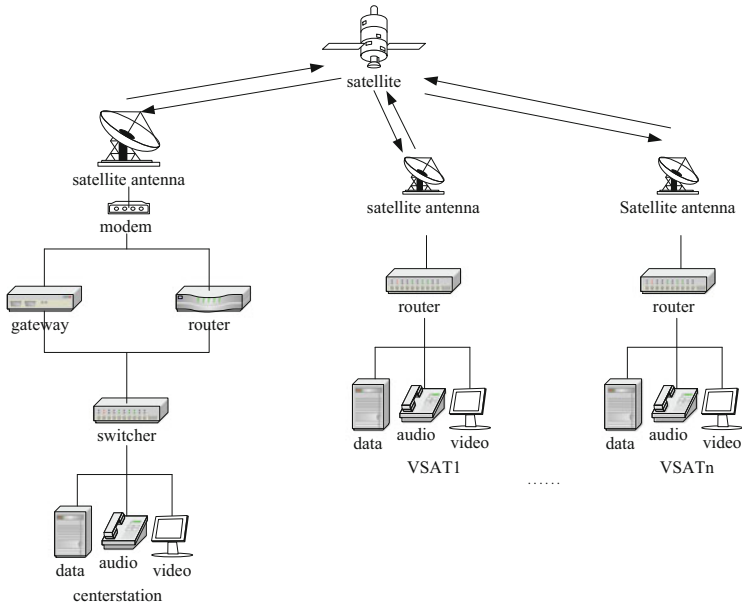


Fig. 89.1 Model of DVB-RCS satellite communication system

fields and fault-tolerant fields will cause the influence of the network adaptability and fault tolerance. According to these reasons, the encryption of DVB-RCS satellite communication systems should be format compliant, that is, the structure and syntax of ciphertext should be compliant with the standard.

Image and video information have the source feature, that is, the proportion of key data impacting on image reconstruction most is lower than 5% [8]. A selective content encryption method is proposed in this chapter, and the high-intensity encryption method is used to encrypt key data, which has most important influence on data reconstruction, while the lightweight encryption method is used to encrypt large amount of the remaining redundancy data. By this method, the data need to be encrypted is reduced and the computing complexity is low, and the delay caused by encryption is low, too. At the same time, the difficult problem of VLC codeword encryption is solved in the chapter. The syntax information in the compressed streams such as encoding format and channel coding information will not be changed by encryption so that the encryption will not affect data compress ration and the ciphertext can be format compliant. The key data is encrypted first, then returned to original bit streams, and encapsulated to IP packet to transmit. This method will not have any influences to DVB-RCS satellite communication system data encapsulation and can be integrated to the whole system seamlessly.

## 89.3 A Format-Compliant Selective Content Encryption Method

A format-compliant selective content encryption method is proposed in this chapter; only key content information is encrypted; therefore, the encryption efficiency is high. A variable modulus encryption method is also proposed to solve the ciphertext format-compliant problem.

### 89.3.1 *The Exaction of Key Information*

The exaction of key information is very important in our method. Main content information is DCT coefficients and motive vector (MV) codeword.

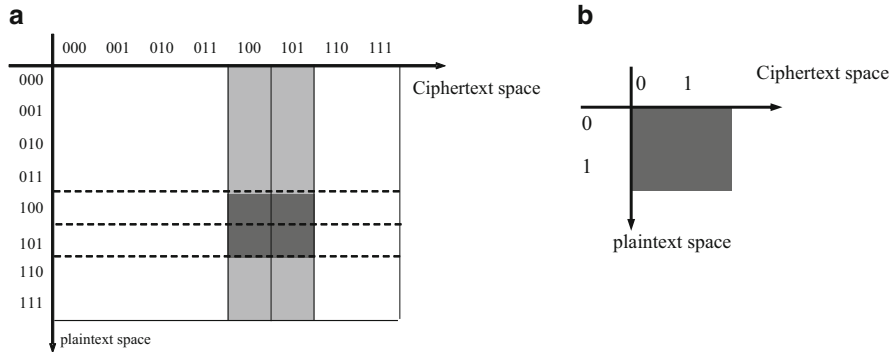
AC coefficients are related to image contour, and DC coefficients are related to average luminance and chrominance of each MB. Therefore DC and AC coefficients can be chosen as key information.

MV includes image's motion information and is used to do motion compensation. If the MV cannot be decoded correctly, the motion information cannot be recovered and the image quality will be reduced remarkably. Therefore, MV can be used as key information.

Except for DCT coefficients and MV, different coding standard of visual media has different content information. For example, the predict mode codeword and quantization information in H.263 standard can be treated as key information.

### 89.3.2 *The Encryption of VLC*

Keeping encrypted stream format compliant is the difficult issue of VLC encryption. In image and video compression standards, transformation and quantization process are followed by the variable length coding process to get higher compression ratio. The length of VLC codeword is variable and cannot occupy the whole codeword space, and direct encryption of VLC codeword will result in invalid VLC codeword and the altering of encrypted stream structure so that the ciphertext will not be compliant to the syntax of compression standard. Take 3-bit codeword for example; there are totally 8 different codeword values (000–111) corresponding to 3-bit codeword; however, only 2 VLC codewords are valid: 100 and 101. Because of the randomness of encryption operation  $E()$ , two valid plaintext codeword will map to the random position in ciphertext space, which includes the grey region and translucent grey region of Fig. 89.2a. If the valid plaintext codeword is mapped to the translucent grey region, the ciphertext will not be format compliant. Only mapping plaintext to the grey region can get valid ciphertext codeword, as shown in Fig. 89.2b.



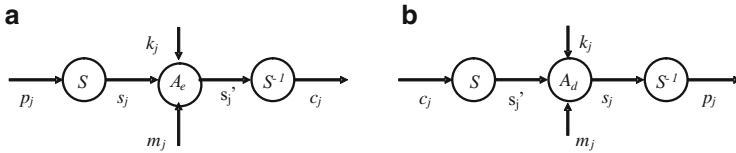
**Fig. 89.2** The valid field and invalid field of encryption mapping. (a) The mapping from plaintext space to ciphertext space. (b) Grey region-valid VLC region

**Table 89.1** MPEG4 VLC codeword

row	VLC	length(bit)	Last	Run	Level	modulus	Sequence number
1	10 s	3	0	0	1	2	1, 2
2	110 s	4	0	0	2	2	1, 2
3	1111 s	5	0	0	3	4	1, 2
4	1110 s		0	1	1		3, 4
...	.....	...	...	...	...	...	...

A variable modulus encryption method is proposed in this chapter in order to solve the problem of VLC codeword encryption. Firstly, valid equal-length VLC codeword table is constructed. Take AC coefficients for example; after run-length coding, the data is represented as an event of a combination of (Last, Run, Level). VLC codeword is found in the table according to (Last, Run, Level). In our method, VLC codewords having the same length are classified to the same group, and each group corresponds to an alphabet. Codewords in the same alphabet are sequenced, each codeword is allocated to a sequence number, and the number of codewords in the group is the modulus of the alphabet. By this way a new table is constructed, where two columns including modulus and sequence number are added.

For example, Table 89.1 shows the MPEG4 VLC for intra-luminance and -chrominance TCOEF, the “s” in the VLC codeword is the sign bit (1 means negative and 0 means positive), and the first row expressing event (Last = 0; Run = 0; Level = ±1) has 2 3-bit codeword: 100 and 101, and the modulus is 2; the sequence number of 101 is 1 and 102 is 2. The modulus of the second row is also 2, where the sequence number of 1101 is 1 and 1100 is 2. After the table is constructed, the plaintext can be encrypted. The theory is shown in Fig. 89.3.



**Fig. 89.3** The theory of variable modulus encryption. (a) Encryption. (b) Decryption

Assuming plaintext sequence  $P = p_1p_2 \dots p_j \dots$ ; the modulus sequence is  $M = m_1m_2 \dots m_j \dots$ ; the plain text  $P$  can be mapped to  $S = s_1s_2 \dots s_j \dots$ , ( $s_j \in [0, M_j - 1]$ ,  $j = 1, 2, \dots$ ) according to Eq. (89.1):

$$s_j = S(p_j) \tag{89.1}$$

For each  $s_j$  and its modulus  $m_j$ ,  $k_j$  is generated in  $[0, m_j - 1]$ ,  $j = 1, 2, \dots$ , and by Eq. (89.2) we get  $s'_j$ :

$$s'_j = (s_j + k_j) \bmod m_j = (s_j + k_j) \% m_j \tag{89.2}$$

According to Eq. (89.3) we get ciphertext  $C = \{c_j; j = 1, 2, \dots\}$ :

$$c_j = S^{-1}(s'_j) \tag{89.3}$$

In the decryption process, with the same mapping rule  $S()$ ,  $c_j$  can be mapped to  $s'_j$ , as shown in Eq. 89.4:

$$s'_j = S(c_j) \tag{89.4}$$

With the same key sequence  $k_j$  using in the encryption, and because  $0 \leq s_j$ ,  $k_j < m_j$ , and  $s_j \bmod m_j = s_j$ ,  $k_j \bmod m_j = k_j$ , we get  $s_j$  by Eq. (89.5):

$$(s' - k_j) \bmod m_j = (s_j + k_j - k_j) \bmod m_j = s_j \bmod m_j = s_j \tag{89.5}$$

Remapping  $s_j$  to  $p_j$  we get plaintext sequence  $P$ , as shown in Eq. (89.6):

$$p_j = S^{-1}(s_j) \tag{89.6}$$

For the discrete random variable  $k_j$ , when  $j$  is different, the sample space is different, and the modulus is variable; thus, the  $K = \{k_j; j = 1, 2, \dots\}$  is called as variable random sequence. For each mod  $m_j$  add operation, the value of  $m_j$  is variable, and the mod add operation is called as variable mod add operation.

For example, the first row in Table 89.1 is the alphabet of 3-bit codeword, and the valid VLC codewords are 100 and 101. According to the standard VLC codeword table there are two events: Event 1 (Last = 0, Run = 0, Level = 1) and Event 2 (Last = 0, Run = 0, Level = -1). If the  $j$ th event waiting for coding is Event 1 and the modulus of this alphabet is 2, then a random number  $R_j$  in  $[0, 1]$  is generated. If  $R_j = 0$ , we can get  $(0 + 0) \% 2 = 0$ , and then the encrypted VLC codeword is 100; if  $R_j = 1$ , then the encrypted VLC codeword is 101. In the authorized decryption side, with the same  $R_j$ , the plaintext Event 1 can be recovered. Otherwise, the unauthorized user cannot determine which one is the right result because  $R_j$  is unknown.

### 89.4 Experiment Results

Using our method to encrypt four MPEG4 testing video sequences, the results are shown in Fig. 89.4. The encrypted video is unrecognizable, and the unauthorized users cannot get any video information.

The proportion of key information in the whole compressed stream of the testing video sequences is shown in Table 89.2. We can see that the proportion of key information is no more than 20 % of the compressed stream, and the information



Fig. 89.4 Encryption results of MPEG4 map of experiment collocation

Table 89.2 The computing complexity of MPEG encryption

Test sequence	Key information/ compressed stream (%)	Time spent in the exaction of key information (ms/frame)	Time spent in encryption/coding (proportion)
Foreman	18.69	2.10	18.47
Mother_daughter	13.72	1.16	15.26
News	11.98	1.38	20.03
Tempete	18.24	2.60	22.27
Mobile	17.28	2.91	21.60
Hallmoniter	15.71	2.12	17.02

needed to be encrypted only accounts for a small part of the whole compressed stream. We can also see that the time spent in encryption is only equal to 20 % of encoding time. By this way the computing complexity is reduced and the high encryption efficiency is guaranteed. Therefore our method is suitable for real-time application.

## 89.5 Conclusion

A high efficient selective content encryption method suitable for DVB-RCS satellite communication system is presented in this chapter, the key content information in the compressed stream is extracted and encrypted, and the variable modulus encryption method is proposed to solve the problem of variable length code encryption; thereby, the encrypted stream can be format compliant. The experimental results prove the effectiveness of our method.

**Acknowledgments** This work is supported by National Natural Science Foundation of China (No. 41101416), National Basic Research Program of China (No. 2011CB302204), and Open Research Fund of The Academy of Satellite Application (No. 20121689).

## References

1. Pillai, P., & Yim-Fun Hu. (2006). Design and analysis of secure transmission of IP over DVB-S/RCS satellite systems. *Proceedings of the International Conference on Wireless and Optical Communications Networks* (pp. 1–5). Santiago, Chile: Springer
2. H. Cruickshank, S. Iyengar, S. Combes, L. Duquerroy, G. Fairhurst, & M. Mazzella. (2007). Security requirements for IP over satellite DVB networks. *Proceedings of the Sixteenth Mobile and Wireless Communications Summit, IEEE* (pp. 1–6). Piscataway, NJ
3. Qi wang, & Shengwu Wang (2009). Securing your satellite network and its contents. *Satellite and Network*, 90(12), 42–45.
4. Peng, C. (2010). *Research on key security technologies in space networks*. Changshai: National University of Defense Technology (in Chinese).
5. Guevara Noubir, Laurent von Allmen. (1999). Security issues in internet protocols over satellite links. *Proceedings of the IEEE Vehicular Technology Conference, IEEE* (pp. 2726–2730). Piscataway, NJ
6. Jonah, P. (2007). Performance implications of instantiating IPsec over BGP enabled RFC 4364 VPNs[C]. *Proceedings of IEEE Military Communications Conference, IEEE Piscataway, NJ* (pp. 1–7)
7. Xie, D. (2010). Discussion on long-distance encrypt technology of satellite HDTV. *Radio and TV Broadcast Engineering*, 37(10), 147–150 (in Chinese).
8. Yang, Z. (2005). An overview of encryption scheme for digital video. *Geomatics and Information Science of Wuhan University*, 30(7), 570–574 (in Chinese).



# Chapter 90

## Network Design of a Low-Power Parking Guidance System

Ming Xia, Yabo Dong, Qingzhang Chen, Kai Wang, and Rongjie Wu

**Abstract** A parking guidance system can help a driver quickly find an available parking space. Most currently available parking guidance systems require wire deployment in installation, thus entailing high installation costs. In this chapter, we discuss the network design of a low-power parking guidance system. We developed a tiered communication architecture including Wireless Sensor Network (WSN), General Packet Radio Service (GPRS) network and Internet to realize wireless parking space availability data transmission, and thus installation complexity can be greatly reduced. In order to reduce the battery replacement frequency of the WSN, we designed a power-minimized Medium Access Control (MAC) protocol. The proposed MAC protocol divides one network working cycle into four dedicated intervals to realize robust network organization and energy-efficient data delivery. Experimental results showed that the proposed MAC protocol can extend the battery lifetime of the WSN to more than ten years. Based on the collected parking space availability data, we built a portable parking guidance terminal to let drivers locate available parking spaces conveniently.

---

M. Xia (✉)

College of Computer Science and Technology, Zhejiang University, Hangzhou 310027, China

College of Computer Science and Technology, Zhejiang University of Technology,  
Hangzhou 310023, China

e-mail: [xiaming@zjut.edu.cn](mailto:xiaming@zjut.edu.cn)

Y. Dong

College of Computer Science and Technology, Zhejiang University, Hangzhou 310027, China

Q. Chen • K. Wang • R. Wu

College of Computer Science and Technology, Zhejiang University of Technology,  
Hangzhou 310023, China

## 90.1 Introduction

Searching for an available parking space in urban areas is becoming more and more annoying to drivers. This is partly caused by the rapid increase of the number of cars in large cities, and partly by the lack of parking space availability information. In order to alleviate the problem, parking guidance systems, which detect and guide drivers to available parking spaces, were developed in recent years. However, many currently available parking guidance systems require communication wire to transmit parking space availability information [1], thus entailing high costs in system deployment. Researchers proposed to use Wireless Sensor Networks (WSNs) to realize full wireless deployments [2]. A WSN typically consists of a large number of low-power, low-cost wireless sensor nodes to collect, process and transmit data collaboratively. As a result, sensor nodes can be battery powered, and the deployment complexity can be greatly reduced. Current researches on applying WSNs to parking guidance systems have discussed sensor design [3], vehicle detection algorithm design [4] and parking space searching policy design [5]. Nevertheless, we found that communication strategies including communication architecture and protocols are also critical issues in system design. As a result, we will focus on the design of efficient communication architecture and protocols for parking space availability information transmission in this chapter. First, we will introduce our tiered communication architecture including WSN, General Packet Radio Service (GPRS) network and Internet to realize wireless and wide-area parking space availability information transmission. Then, we will present the design of a power-minimized WSN Medium Access Control (MAC) protocol to maximize the battery lifetime of parking space sensors. We chose battery lifetime maximization as our major optimization object because parking space sensors are installed on the ground and battery replacement is relatively difficult. In order to reduce maintenance costs, long battery lifetime is a must. After that, we will discuss system implementation on customized hardware and a portable parking guidance terminal based on smartphone to enable drivers to acquire parking space availability information at any place and any time. At last, the experimental deployment of the system at a parking lot verifies the effectiveness of the proposed network design.

## 90.2 Network Design

We developed a tiered communication architecture which integrates WSN, GPRS network, and Internet to meet the deployment requirement of parking space availability monitoring and parking guidance, as shown in Fig. 90.1.

*WSN*: The communication between parking space sensors and gateway nodes employs WSN technology. In order to maximize the battery lifetime of parking space sensors, we designed a power-minimized MAC protocol.

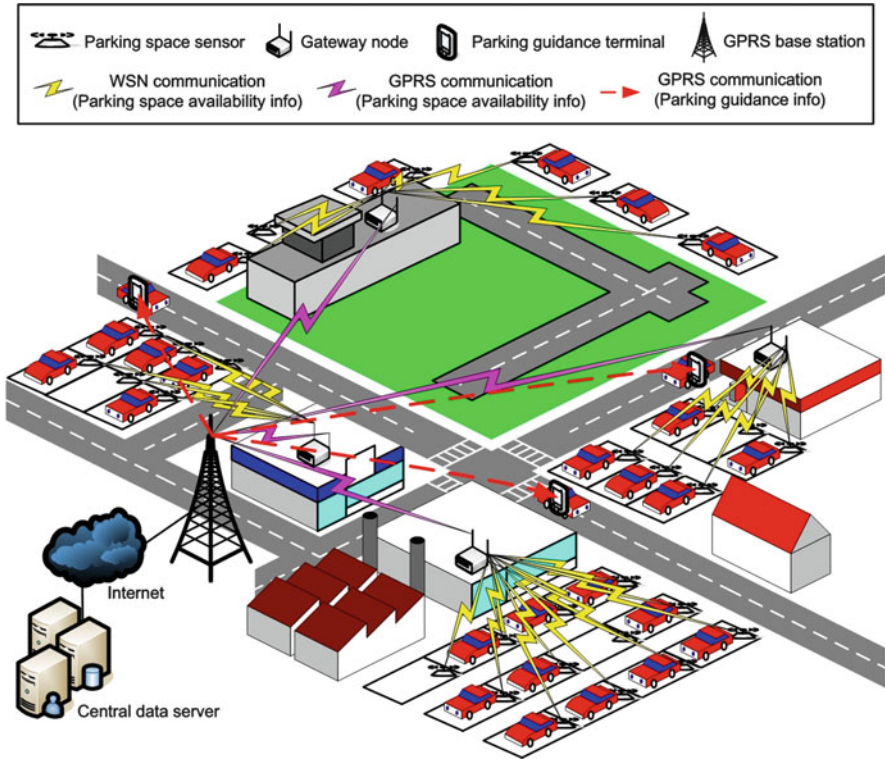


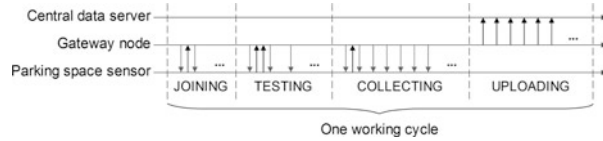
Fig. 90.1 Communication architecture

*GPRS network and Internet:* The communication between gateway nodes and the central data server uses GPRS network and Internet to ensure the large coverage area. At the same time, central data server will send the parking guidance information to the portable parking guidance terminal through GPRS network and Internet.

We will then elaborate on the details of the MAC protocol design. One of the major tasks of our MAC protocol for WSNs is to reduce the power consumption of sensor nodes. The design principles of WSN MAC protocols can be roughly categorized as either contention-based or Time Division Multiple Access (TDMA)-based. We chose to employ TDMA in our MAC protocol design because it does not suffer from collisions and thus can frequently achieve lower power consumption.

Generally, the gateway node is responsible for the establishment of the network, and the parking space sensors will select and join the “best” network established by one gateway node. The gateway node divides one working cycle into four intervals, including JOINING, TESTING, COLLECTING and UPLOADING, as shown in Fig. 90.2. In different intervals, the gateway node broadcasts different types of beacons periodically to notify parking space sensors.

**Fig. 90.2** Four intervals of the proposed MAC protocol



The proposed MAC protocol works as follows:

- (1) A gateway node scans all usable channels one by one, and chooses one silent channel to establish the network.
- (2) A parking space sensor scans all usable channels one by one when powered on. It records all beacons captured in scanning and compares the signal strength and will try to select the gateway node with the strongest signal strength to transmit parking space availability information. In order to notify the selected gateway node, the parking space sensor waits for the JOINING beacon from the selected gateway node and immediately replies a *JOINING\_REQUEST* after the beacon. Here we use a simple Carrier Sense Multiple Access/Collision Avoidance (CSMA/CA) protocol to avoid collision.
- (3) The gateway node that receives the *JOINING\_REQUEST* will reply a *JOINING\_CONFIRMATION* to the parking space sensor if it is able to adopt more sensors. In the *JOINING\_CONFIRMATION*, the gateway node tells the parking space sensor the time slot assigned to it in the TESTING interval. If the gateway node is unable to adopt more sensors, it will reply a *JOINING\_REJECTION* to the parking space sensor and the sensor will repeat step (2) and select another gateway node.
- (4) A parking space sensor that receives the *JOINING\_CONFIRMATION* will go sleep and wait for the time slot assigned to it to wake up in the TESTING interval. It will start testing (i.e., transmitting several packets to the gateway node and calculating the transmission success rate) immediately once it receives the TESTING beacon broadcasted by the gateway node. We designed a dedicated TESTING interval to ensure the communication quality between the parking space sensor and the selected gateway node because the received signal strength frequently cannot accurately reflect the link quality. If the transmission success rate is high enough, the parking space sensor will send a *REGISTERING\_REQUEST* to the selected gateway node. Otherwise, it will go back to step (2) and select another gateway node.
- (5) A gateway node that receives the *REGISTERING\_REQUEST* will reply a *REGISTERING\_CONFIRMATION* to the parking space sensor telling it the time slot assigned in the COLLECTING interval.
- (6) A parking space sensor that receives the *REGISTERING\_CONFIRMATION* will go sleep and wait for the time slot that is assigned to it to wake up in the COLLECTING interval. Once it wakes up and receives a COLLECTING beacon, it will transmit parking space availability information to the gateway node immediately. The gateway node will cache the received parking space availability information for uploading.

- (7) In the UPLOADING interval, the gateway node uploads the received parking space availability information to the central data server via GPRS network.
- (8) In the next round, a parking space sensor will only wake up in its time slot in the COLLECTING interval, and the COLLECTING beacon will be the time synchronization signal to avoid clock drifting. In order to conserve energy, the sensor will transmit data only if the state of the parking space (occupied/available) changed. Otherwise, the sensor only reports its state at a low frequency. If the sensor encounters continuous data transmission error, it will go back to step (2) to re-join a network.

### 90.3 Hardware Design

In this section, we will discuss the design of parking space sensors, gateway nodes and parking guidance terminal.

*Parking space sensor.* Our parking space sensor is directly installed on the ground of a parking space, and thus we designed a special robust and water-resistant enclosure for it. The design of the enclosure and its photo are given in Fig. 90.3. We chose an STMicroelectronics STM32F103 as microcontroller and an Atmel AT86RF212 as radio chip, which works on 700/800/900 Mhz. We chose a Honeywell HMC5883L geomagnetic sensor to monitor if the parking space is occupied. If there is car on the sensor, the earth's magnetic field will be changed, and we can monitor this event through measuring the output of the sensor.

*Gateway node.* The gateway node also adopts an STMicroelectronics STM32F103 as its microcontroller, and it has two radio chips, one is the 700/800/900 Mhz Atmel AT86RF212 radio chip for the communication between the parking space sensors and the gateway node and the other is a GPRS modem for the communication between the gateway node and the central data server. The gateway node uses a rechargeable battery as its power supply, because the gateway node may be installed on a street lamp and directly uses the power from the street lamp. Unfortunately, the street lamp is frequently powered on only at night, and thus we designed a recharging circuit to charge the battery at night, and the battery will provide power supply to the gateway node in day time. The gateway node is

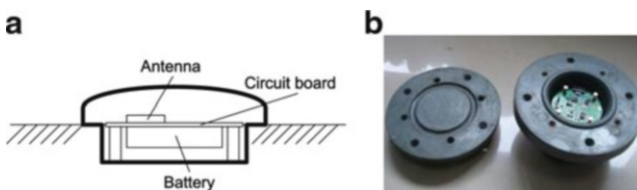


Fig. 90.3 Parking space sensor. (a) Sensor design concept (b) The photo of the sensor

Fig. 90.4 Gateway node

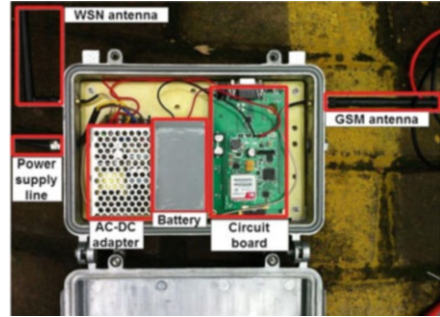


Fig. 90.5 Parking guidance terminal



encapsulated in a water-resistant enclosure. The inner structure of a gateway node is given in Fig. 90.4.

*Parking guidance terminal.* The parking guidance terminal is built based on an Android smartphone, as shown in Fig. 90.5, in which the small circle on the screen indicates the current position of the car, and the blue arrow on the screen indicates the position of the detected available parking space. We used a Global Positioning System (GPS) receiver to mark the positions of parking spaces in system deployment. Because the position measuring error is only several meters, we believe that user experience will not be affected. The parking guidance terminal generally works as follows: (1) when a driver arrives at the destination, he just presses the “search” button on the screen; (2) the terminal communicates with the central data server to find nearby available parking spaces and then tells the driver; (3) the driver selects his favourite available parking space and the terminal will guide the driver to the chosen available parking space.

### 90.4 System Deployment and Evaluation

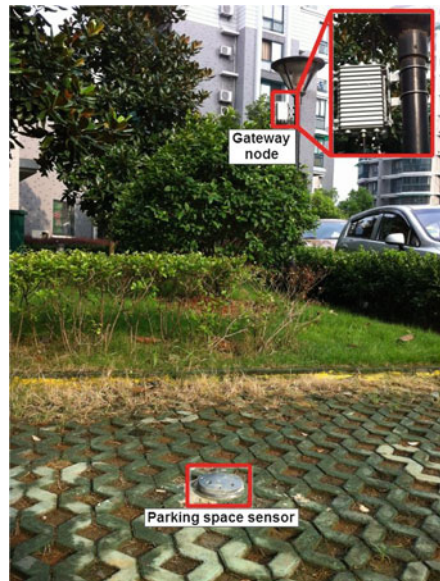
We have deployed an evaluation system covering about 12 parking spaces at a parking lot. Figure 90.6 shows the deployment of a parking space sensor and a gateway node. In deployment, we let parking space sensors transmit 100 packets in the TESTING interval and set the transmission success rate threshold to 80 %. Under these settings, parking space sensors typically spent no more than three working cycles to establish stable connections to gateway nodes.

Because of the extremely low active time of our parking space sensors after network joining, the battery lifetime can be extended to several years in deployment. We estimate parking space sensor’s battery lifetime based on its typical working parameters. Table 90.1 gives the parameters.

We can estimate the battery lifetime ( $T_B$ ) of parking space sensors according to (90.1), and the result is about 10.8 years.

$$T_B = W_B T_D / (I_A T_A + I_S (T_D - T_A)) \tag{90.1}$$

**Fig. 90.6** System deployment



**Table 90.1** Typical working parameters of parking space sensors

Item	Value
Active current ( $I_A$ )	40 [mA]
Maximum active time ( $T_A$ )	0.05 [s]
Sleep current ( $I_S$ )	0.03 [mA]
Working cycle ( $T_D$ )	60 [s]
Battery capacity ( $W_B$ )	6000 [mAh]

## 90.5 Conclusion

In this chapter, we focused on the network design of a low-power parking guidance system. We developed a tiered communication architecture including WSN, GPRS network and Internet, and thus the deployment complexity can be greatly reduced. In order to maximize the battery lifetime of parking space sensors to reduce battery replacement costs in system maintenance, we designed a power-minimized TDMA-based MAC protocol for the WSN. Experimental results showed that the proposed MAC protocol can extend the battery lifetime of parking space sensors to more than 10 years.

**Acknowledgments** This work is supported by the Research Program of Department of Science and Technology of Zhejiang Province (2012C33073), the Key Scientific and Technology Project of Zhejiang Province (ZD2009011), and the Collaborative Industry-University Research Project of Hangzhou (20112731E54).

## References

1. Yao, G. Z., Wang, J. Q., Li, Z. S., Ran, X. J. (2010) The design of parking guidance and information system based on CAN[C]. *Proceedings of the 2010 International Conference on Intelligent Control and Information Processing IEEE, Dalian, China* (pp. 171–174).
2. Bi, Y. Z., Sun, L. M., Zhu, H. S., Yan, T. X., & Luo, Z. J. (2006). A parking management system based on wireless sensor network[J]. *ACTA AUTOMATICA SINICA*, 32(6), 968–977.
3. Idris, M. Y. I., Tamil, E. M., Noor, N. M., Razak, Z., & Fong, K. W. (2009). Parking guidance system utilizing wireless sensor network and ultrasonic sensor[J]. *Information Technology Journal*, 8(2), 138–146.
4. Yoo, S., Chong, P. K., Kim, T., Kang, J., Kim, D., Shin, C., Sung, K., Jang, B. (2008). PGS: Parking guidance system based on wireless sensor network[C]. *Proceedings of the 3rd International Symposium on Wireless Pervasive Computing IEEE, Santorini, Greece* (pp. 218–222).
5. Wang, H. W., He, W. B. (2011). A Reservation-based smart parking system[C]. *Proceedings of the 2011 International Conference on Computer Communications Workshops. IEEE, Shanghai, China* (pp. 690–695).



# Chapter 91

## Strategy of Domain and Cross-Domain Access Control Based on Trust in Cloud Computing Environment

Bo Li, Ming Tian, Yongsheng Zhang, and Shenjuan Lv

**Abstract** Under the current cloud computing environment, a reasonable and practicable access control strategy is needed, which is a guarantee to protect cloud computing suppliers to provide services and many cloud users access to services. In this paper, based on analysis of many cloud computing safety features, trust management is introduced into the cloud computing service access control, within the domain of a trust-based access control strategy, in domain, presents a trust-based access control policy. Credible value will be given through the comprehensive treatment of the entity, and then AAC (authentication and authorization center) authorizes the appropriate access rights to achieve the control of the monomer in the domain. Combined with the characteristics of the existing cloud computing environment, in multiple management domains, this paper proposes a role mapping, with the role mapping relationship between the domain, which can make the inter-domain access to resources and security shared access between different domains, in order to avoid the problem of permission penetration and privilege escalation, this paper presents the mirror role based on role mapping, ultimately solves the problem.

---

B. Li

Academic Affairs Office Shandong Polytechnic, Jinan 250104, China

M. Tian (✉) • Y. Zhang • S. Lv

School of Information Science and Engineering, Shandong Normal University, Jinan 250014, China

Shandong Provincial Key Laboratory for Novel Distributed Computer Software Technology, Jinan 250014, China

e-mail: [Tiancius@163.com](mailto:Tiancius@163.com)

## 91.1 Introduction

Cloud computing era poses a huge shock to the traditional information industry, changes the traditional IT applications, and has created an enormous change to our lives. Individual or enterprise users can access to IT services only by a cloud service provider who provides a simple operation interface; users no longer need to upgrade and maintain hardware and software. Cloud computing brings great convenience to our lives, but it also presents some security risks [1]. According to the understanding of the definitions and concepts of cloud computing, the mode of cloud computing operation is to provide the user data and the corresponding computing tasks to the server. Storage of user's data, as well as operations such as handling and protection of user data, is in the "cloud" to complete. In this way, it will inevitably make the user's data in a potentially unsafe state of destruction and theft and also have more detailed personal information exposed on the network, which is a very large exposure. Judging from today's cloud computing development, the security of user's data, user privacy information protection, data stored and cloud computing to their own security and stability, and many other regulatory aspects of cloud computing issues directly relate to cloud computing user acceptance; thus, security becomes the most important factor affecting the development of cloud computing business [2]. A reasonable and practicable access control strategy is needed, which is a guarantee to protect cloud computing suppliers to provide services and many cloud users access to services.

## 91.2 Related Works

Cloud computing acts as a new information service mode, which brings new security risks and challenges, but has no essential difference with traditional IT security information service requirements [3,4]. It remains the core requirement of data and application of confidentiality, integrity, availability, and protection of privacy, and key technologies to meet these security requirements are access control technologies [5,6].

The document [7] analyzes dynamic demand for access control in a cloud computing environment and makes role-based access control (RBAC) model to a cloud computing environment to meet the need of complex access control management of cloud computing and dynamic management for access control and increase the maintainability. But the RBAC model is based on identification and close, whose access control mechanisms in centralized closed network environments, and does not apply to large-scale, distributed and open network, especially unable to meet the security needs of cloud computing environments [8]. In addition, the RBAC model, in assigning roles to users, only verifies the authenticity of the identity of a user, without taking into account the user's credibility.

RBAC model is used for allocating roles for access authorization, users who actually use permission do not be supervised and controlled, and insufficient against the RBAC model to extend it, Blaze's "trust management", based on trust introduced the concept of access control mechanisms, proposed a trust-based access control model TRBAC (Trust Role-Based Access Control Model) [9]. But some of the above study did not take into account the characteristics of the cloud computing security management domain, and did not explicitly give access control method for cloud computing environments.

### 91.3 The Trust in the Cloud

Human society is a complex system. The interactions between the entities in the system depend on trust relationships between each other. Cloud computing researchers now introduce the trust mechanism from the human society into the cloud computing environment as a basis of the exchange between entities in the cloud computing environment [10].

Trust in real life is a subjective concept, depending on the person's experience, and trust relationships would be difficult to reference to a cloud computing environment by using empirical measure.

We can use trust value to determine the degree of trust. The degree of trust allows the definition of security strategy to be more clear, and for different trust systems, we can define different security strategies [11].

#### 91.3.1 Access Control Trust Relationship in a Domain

In accessing other entities that are in the same domain security management, trust can be directly introduced to the access control model for secure operation.

**Definition 1 (Domain Trust Value):** In the same domain, an entity and another entity complete an interaction, there will be an assessment for the entity to another entity. Use  $T$  to denote  $-1 \leq T \leq 1$ . Negative values are not satisfactory; will reduce the trust, in contrast, expressing satisfaction with the integrity; and will enhance the trust. Entity  $n_j$  after the completion of the interaction of  $n_i$   $k$  times which gives trust values can formalize for  $T(n_i, n_j)^k$ .

**Definition 2 (Service Satisfaction Degree):** In the same domain, an entity and another entity complete an interaction; another entity to the entity's overall service satisfaction is denoted by  $S$ . After  $k$  times services entity  $n_j$  to  $n_{is}$  overall satisfaction with the formula as follows:

$$S(n_i, n_j)^k = \beta \times S(n_i, n_j)^{k-1} + (1 - \beta) \times T(n_i, n_j)^k \quad (91.1)$$

**Definition 3 (Direct Trust Degree):** An entity's direct trust degree is related to domain trust value, the higher assessed value in a domain is, the higher entity credibility is, also the direct trust value is high, notation DTD. For two entities that never interact, DTD value is usually set to zero. In the domain entity  $n_i$  and  $n_j$ , direct trust degree after the completion of the  $k$  times interaction formula is as follows:

$$DTD(n_i, n_j)^k = \alpha \times DTD(n_i, n_j)^{k-1} + (1 - \alpha) \times T(n_i, n_j)^k \quad (91.2)$$

**Definition 4 (Credit):** The credibility of entities in the domain has to interact with all other entities within the domain in order to obtain satisfaction, expressed in  $Rp$ . Entity  $n_i$  in domain  $A$  can be expressed as to the credibility of  $Rp(n_i, A)$ ; the specific formula is as follows:

$$Rp(n_i) = \frac{\sum_{j=1, j \neq i}^k S(n_i, n_j) \times Rp(n_i, A)}{k} \quad (91.3)$$

**Definition 5 (Domain Trust Degree):** In the domain an entity's trust degree is the credibility degree in the field of domain, directly made up by direct trust degree and credit, represented by symbol TD.

The  $i$ -th entity  $n_i$  trust degree in domain  $A$  can be formalized representation for TD ( $n_i, A$ ); the formula is as follows:

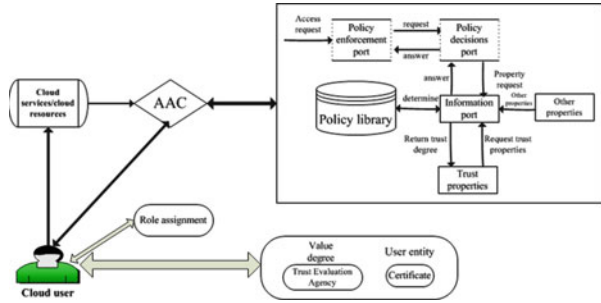
$$TD(n_i, A) = \gamma \times \frac{\sum_{j=1, j \neq i}^k DTD(n_i, n_j)^j}{k} + (1 - \gamma) \times Rp(n_i, A) \quad (91.4)$$

Above all,  $\alpha, \beta, \gamma = 0$ , the value of weight parameter associated with the local security policy, which is stored in the local domain authentication and authorization center and, through the definition above, decides the proper policy access control to the cloud users and the service providers.

### 91.3.2 Access Control Strategy in the Same Domain

We introduce the trust degree which is the basic property of cloud users and cloud services or resources into role-based access control; the certification center AAC (authentication and authorization center) is responsible for access control authentication, authorization, and trust management.

Fig. 91.1 Access control strategy in the same domain



In the same domain, each time a user requests access to cloud services or cloud resources, AAC will examine cloud users’ trust degree and ensure that the trust degree reaches its threshold, and then cloud users’ requests for access will be permitted. The process is shown in Fig. 91.1.

Domain Access Control Process

1. Under the role-based access control, firstly the user needs to request role assignments before he/she wants to access control requests, thereby indirectly obtaining the appropriate access control permissions; in this policy, users can use access permissions or not as determined by trust management stage.
2. Users send access control requests to the AAC; the request information includes user ID, password, and the ID of access to resources or services. AAC firstly certifies user’s authentication information, once passing through, trust management will authorize user’s rights correspondingly. Authorization process includes the following:
  - Policy library initializes a security policy in this domain.
  - Policy implementation which sends user’s access request is passed to the policy decision port.
  - Policy decision side passes the requests to the policy information port.
  - Policy information port obtains user’s trust degree and other property information, returns to policy decision port.
  - In accordance to the user’s information and the current security policy, policy information port makes decisions and returns the policy to enforcement port.
  - Policy enforcement port feedbacks the results to the user entity.
3. User performs the appropriate access control permissions and accesses cloud services or cloud resources.
4. Following the requests, providers of cloud services or cloud resources make assessment of the user and feedbacks to AAC.

### 91.4 Cross-Domain Access Control Policy

Because users often need to access different cloud services or cloud resources in different domain, safe and effective cross-domain access control policy is necessary. This paper proposes a role-based access control model and presents a new role mapping through a domain relationship, reaching to the result of resources sharing between domains.

Firstly we define two different security management domains which are Domain A and Domain B. The two roles Role A and Role B are in the separate domain.

Role A and Role B can access their cloud resources or enjoy the cloud service in their own domain. Suppose logical domains a and b are two partnership units, Domain A is an enterprise, and Domain B is a scientific research institute; two domains are using role-based access control management; now, both the enterprise and research institute will develop a research project, which needs to achieve the shared resources. At this time, the Domain A user wants to access the information of Domain B, which is cross-domain access. Through the relationship of role mapping, Domain A user by role exchanging can access resources of Domain B.

Cross-domain role mapping is shown in Fig. 91.2; the dotted line represents the mapping meaning.  $R_A$  represents role collection within the safe management Domain A.  $R_B$  represents role collection within the safe management Domain B.  $R_A R_B$  stands for the role mapping relation from  $R_A$  to  $R_B$ . With mapping relationship any  $R_A R_B$  ordered pair  $(R_{An}, R_{Bm})$  from Domain A's role  $R_{An}$  maps to the Domain B's role  $R_{Bm}$ , Domain B will assign their own domain role  $R_{Bm}$  to domain's role  $R_{An}$ , and then the domain A's role  $R_{An}$  has the same rights with  $R_{Bm}$  to access to resources in the domain B, so  $R_{An}$  can achieve access to some resources on the domain b. This mapping can be written as  $R_{An} \rightarrow R_{Bm}$ .

Based on the role mapping, we enable cross-domain resource sharing to become more convenient. Meanwhile, it may also bring the problem of permission penetration and privilege escalation. For example, Domain A's role  $R_{A1}$  maps Domain B's role  $R_{B3}$ ; at the same time,  $R_{B3}$  maps Domain C's role  $R_{C1}$ , and then it is inevitable that  $R_{A1}$  inherits the rights of  $R_{C1}$ . This situation is not what we want to see.

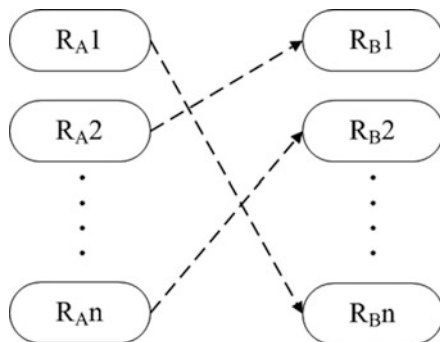
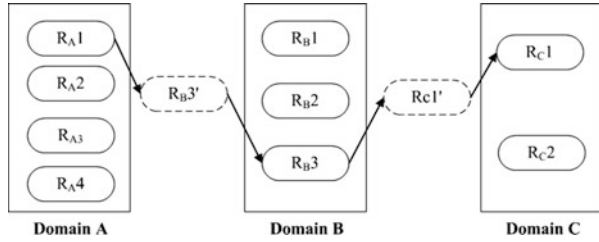


Fig. 91.2 Cross-domain role mapping relationship

**Fig. 91.3** Domain A, B, and C mapping example with mirror



In order to avoid this situation, we introduce a concept which is called the mirror role, and the specific circumstance is shown in Fig. 91.3. From Fig. 91.3 we will know both  $R_{B3'}$  and  $R_{C1'}$  are mirror roles which come into being. In such conditions, we can avoid the problem of permission penetration and privilege escalation. Although  $R_{A1}$  can obtain the same operating authority as  $R_{B3}$  in Domain B and have the resource access qualifications,  $R_{A1}$  cannot map  $R_{C1}$  which is in Domain C through the  $R_{B3}$  and cannot get the corresponding authority of  $R_{C1}$  in Domain C as well. Because the identity that  $R_{A1}$  gets is the mirror role  $R_{B3'}$  only,  $R_{B3'}$  cannot map  $R_{C1'}$  because of the interrupting of  $R_{B3}$ ; in other words, there is no mapping relationship between  $R_{B3'}$  and  $R_{C1'}$ . Thus it can be seen that the mapping through the mirror role restricts the role of a bridge played by some roles in the process of role mapping, so executing the role mapping transfer finally makes some roles acquiring authority which they should not have. Therefore, the access control based on the mirror role mapping avoids the problem of permission penetration and privilege escalation well which appears in the role mapping.

### 91.5 Conclusion

In this paper, we discuss the current access control issues in the environment of cloud computing and present an in-domain and cross-domain access control policy based on trust. In the same domain through the calculation of user's trust degree, which reaches to the allocation of their according privileges, accomplishes the result of an access to resources or the corresponding service. Cross-domain we provide a role mapping to achieve resource sharing, but it may cause the problem of permission penetration and privilege escalation, in order to avoid the problem, we present the mirror role which based on role mapping, ultimately solve the problem. Through the access control strategy of in-domains and cross-domains, it can achieve the purpose of the security of cloud computing environment, users, and platform.

**Acknowledgements** This research was supported by the Project of Shandong Province Higher Educational Science and Technology Program under Grant No. J12LN61 and Grant No. J13LN64. In addition, the authors would like to thank the reviewers for their valuable comments and suggestions.

## References

1. Feng, D. G., Zhang, M., Zhang, Y., & Xu, Z. (2011). Study on cloud computing security [J]. *Journal of Software*, 22(1), 71–83. In Chinese.
2. Li, W., Ping, L., & Pan, X. (2010). Use trust management module to achieve effective security mechanisms in cloud environment [C]. *IEEE International Conference on Electronics and Information Engineering (ICEIE2010), Kyoto, Japan* (pp. 14–19).
3. Wang, S., Zhang, L., & Li, H. (2010). Evaluation approach of subjective trust based on cloud model [J]. *Journal of Software*, 21(6), 1341–1352. In Chinese.
4. Takabi, H., Amini, M., Jalili, R. (2007). Trust-based user-role assignment in role-based access control [C]. *IEEE Proceedings of the ACS/IEEE International Conference on Computer Systems and Applications 2007, Amman, Jordan* (pp. 807–814).
5. Jie Zhao, Nanfeng Xiao, Junrui Zhong. (2009). The behavior trust control based on Bayesian network and behavior log mining [J]. *Journal of South China University of Technology (Natural Science Edition)*, 37(5), 94–100 (In Chinese).
6. Hur, J., & Noh, D. K. (2011). Attribute-based access control with efficient revocation in data outsourcing systems [J]. *IEEE Transactions on Parallel and Distributed Systems*, 22(7), 1214–1221.
7. Chuang Lin, Fujun Feng, Junshan Li. (2007). Access control technology under the new network environment [J]. *Journal of Software* 18(4), 955–966 (In Chinese).
8. Shouxin Wang, Li Zhang, Hesong Li. (2010). A subjective trust evaluation method based on cloud model [J]. *Journal of Software*, 21(6), 1341–1352 (In Chinese).
9. Wu Liu, Haixin Duan, Hong Zhang, Ping Ren, Jianping Wu. (2011). TRBAC: Trust based access control model [J]. *Journal of Computer Research and Development*, 48(8), 1414–1420 (In Chinese).
10. Guangwei Zhang, Jianchu Kang, Hesong Li. (2007). Research on subjective trust management model based on cloud model [J]. *Journal of System Simulation*, 19(14), 3310–3317 (In Chinese).
11. Chunhua Hu, Xinxing Luo, Sichun Wang, Yao Liu. (2011). Approach of service evaluation based on trust reasoning for cloud computing [J]. *Journal on Communications*, 32(12):72–81



# Chapter 92

## Detecting Unhealthy Cloud System Status

Zhidong Chen, Buyang Cao, and Yuanyuan Liu

**Abstract** In this paper, in order to detect the unhealthy status in the cloud system, a Basic Detection Strategy and a Threshold Strategy based on mathematic theory and statistical knowledge is proposed to solve this problem. By introducing unhealthy status percentage parameter  $\alpha$ , both Basic Detection Strategy and Threshold Strategy are combined to detect and monitor the unhealthy cloud system status. For illustration, an eBay company example is utilized to show the feasibility of Basic Detection Strategy and Threshold Strategy. Empirical results show that Basic Detection Strategy with setting a suitable value to  $\alpha$  can pinpoint most of unhealthy status in the cloud system, however, for some special unhealthy status, it must adopt the Threshold Strategy to pinpoint. The combination of Basic Detection Strategy and Threshold Strategy can effectively detect and pinpoint the unhealthy status in the cloud system and help staff to improve the performance of cloud system.

### 92.1 Introduction

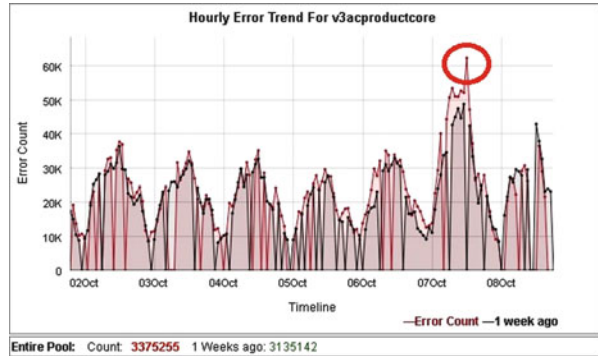
With the development of the cloud computing, more and more enterprises are willing to adapt the strategy of deploying services on Cloud Platforms and are keen to improve the utilization of resource and reduce costs. Cloud Computing System shoulders an important mission to provide a healthy system environment for the different server applications and to ensure that each application service request can timely access required resources (CPU, memory, disk space, network and so on) [1–3]. In an ideal scenario, Cloud Computing System is able to provide sufficient resources for application requests to consume while these application

---

Z. Chen (✉) • B. Cao  
School of Software Engineering, Tongji University, Shanghai 201804, China  
e-mail: [chenfang3376@gmail.com](mailto:chenfang3376@gmail.com)

Y. Liu  
Site Reliability Engineering, eBay Engineering and Research Center (Shanghai),  
Shanghai 201210, China

**Fig. 92.1** The data sample of 2 weeks



requests yield the desired results and return them to the clients (service invokers) [4, 5]. Furthermore, there should be no exception thrown while the requests are being processed. Although it is not official and not quite accurate, we define the Cloud Computing System to be in a healthy status if the scenarios described above are applicable. Unfortunately, these ideal scenarios may never exist. In order to provide customers with better and more robust services, we need to identify exceptions in time and be able to take corresponding actions effectively.

The status detection problem comes from one of the world largest e-commerce companies. The company has several huge data centers worldwide where the private cloud platforms are hosted. The company has designated technical staff members or system administrators who monitor the system status, process exceptions, and resolve system problems to ensure good services required by large amount of customers. To monitor the system status and find abnormalities effectively is crucial for this company. In order to find exceptions or abnormalities, the company had defined the rules of determining the system status (healthy or unhealthy) and conducting their manual interventions when the Cloud Computing System malfunction. For an efficient detection of system's unhealthy status we need to convert these rules established by this company into an actionable model.

A data center of the company (eBay) keeps the log files recording the number of errors occurring at different time periods. Figure 92.1 depicts the number of errors recorded in the log files of a data center over a given time period. The  $x$ -axis represents the time while  $y$ -axis indicates the number of recorded errors at different time points. Upon different scenarios, the company sets the rules below to identify the healthy status of the system:

- At some time points the curve suddenly rises or falls (we call them *spiking points*) and the changing range exceeds the predefined normal range, then these moments potentially are considered as unhealthy or in unhealthy status.
- When the spiking point in rule 1 is compared with the one point at the same moment last week and their trends are similar, and then it is considered to be healthy at the moment.

- The number of the recorded errors at a certain time point is above or below the specified threshold, and then the system could be thought being in unhealthy status at this moment.

Albeit the rules can be applied to monitor the system status manually, it is an extremely inefficient way to monitor the system status manually. Here we will propose a model-based methodology to identify the unhealthy status of the system effectively. The solution is developed by incorporating a set of algorithms based upon three rules mentioned. With the help of this approach the unhealthy status of the system can be detected in time and the corresponding message will be pushed to the staff members of the data center in time for possible actions. We first introduce the following two important definitions.

**Definition 1.** a recorded point  $P_n(x_n, y_n)$  is called a *Peak*: if  $(x_{n-1} < x_n < x_{n+1})$  &&  $(y_{n-1} < y_n && y_n > y_{n+1})$  where  $P_{n-1}(x_{n-1}, y_{n-1})$ , and  $P_{n+1}(x_{n+1}, y_{n+1})$  are left and right adjacent points of  $P_n(x_n, y_n)$ .

**Definition 2.** a recorded point  $P_n(x_n, y_n)$  is called a *Valley* if  $(x_{n-1} < x_n < x_{n+1})$  &&  $(y_{n-1} > y_n && y_n < y_{n+1})$  where  $P_{n-1}(x_{n-1}, y_{n-1})$ , and  $P_{n+1}(x_{n+1}, y_{n+1})$  are left and right adjacent points of  $P_n(x_n, y_n)$ .

This paper is organized as follows: The second section proposes two status detection algorithms. The computational experimental results are presented in the third section to demonstrate the effectiveness of the algorithms. The paper is concluded with the summary.

## 92.2 Status Detection Algorithms

The system status detection approach proposed here employs the recorded number of errors (the information is usually stored in the log files of the system) over the given time periods as shown in Fig. 92.1. As it is mentioned above, there are two curves mirroring the number of recorded errors over two time periods with the same length, e.g., this week and last week. The technical staff member usually determines the system healthy status by comparing the values for these two time periods. The criterion or rule of defining healthy or unhealthy status of the system is formed based upon the operational experience of the underlying e-commerce company. According to the rules and his own working experience, the technical staff member generally focuses on “peak” and “valley” points defined above to find unhealthy status of the system. A “peak” or “valley” point is usually caused by the exceptions when the Cloud Computing System is at an unhealthy status. In this case the number of recorded errors goes up or down sharply. The points where “peaks” or “valleys” occurs are called spiking points (for the differentiation purpose, one is called upward spiking where a peak occurs while the other is named as downward spiking where a valley appears). We might pinpoint the times when the unhealthy status of the system could occur by checking the spiking points. Nevertheless, it is

tedious and ineffective to find the potential unhealthy status of the system and take actions by checking the diagram manually. We are going to propose a method that is able to detect the unhealthy status automatically.

### 92.2.1 The Basic Detection Strategy

Basic idea: in order to find a spike (a sudden up or down in the number of recorded errors) that may represent an unhealthy status, we first compute the slope  $k$  of two adjacent points so that we can collect a new sample including much slope  $k$ , and then for the new sample we will make statistics, meanwhile, though setting a percent that much larger slope  $k$  are filtered out of the new sample and consist another sub-sample [6–8]. With the help of the measurement described above, we are able to initially acquire some suspected spikes or potential unhealthy points according to the sub-sample. After the data is obtained, we will compare them to the ones of the same time points of last week to find out the tendency. If their tendencies are the same, then they are considered to be in healthy status otherwise they are in unhealthy status. An unhealthy status therefore may be identified.

Approach detail description: for a given set of recorded number of errors occurring in a data center, it can be plotted as shown in Fig. 92.1. The  $y$ -values of a curve may appear up and/or down abnormally and the variances are pretty bigger that generate different tendencies. The situations with sharp up or down  $y$ -values may reveal that the system encounters some exceptions at certain times.

According to the definitions discussed above it is obvious that the sharp changes in  $y$ -value form peaks and valleys of the curve or functions. For the convenient purpose these peaks or valleys will be named as local extreme values (or extreme values for short) of the function. Let one point be  $P_n(x_n, y_n)$  of the function, its left adjacent point be  $P_{n-1}(x_{n-1}, y_{n-1})$ , and its right adjacent point be  $P_{n+1}(x_{n+1}, y_{n+1})$ , respectively. If the product of slope  $K_n$  for linear segment  $P_{n-1}P_n$  and slope  $K_{n+1}$  for linear segment  $P_nP_{n+1}$  is less than zero, namely:

$$\frac{y_{n+1} - y_n}{x_{n+1} - x_n} \times \frac{y_n - y_{n-1}}{x_n - x_{n-1}} < 0 \quad (92.1)$$

Then  $(y_{n+1} - y_n) \times (y_n - y_{n-1}) < 0$ , because the number of errors is recorded with the equal time interval, that is  $x_{n+1} - x_n = x_n - x_{n-1}$ . Based on the definitions for peaks and valleys we conclude  $P_n$  is the extreme value since its  $y$ -value is either bigger or smaller than its adjacent ones [9–11]. Using this formula we are able to define a set of extreme value points and name it as EVS (Extreme Value Set). Suspected spike points that may indicate abnormal status usually appear in the EVS.

Apparently in this way we may get a lot of extreme values for the given set of data while most of their up or down trends are more tempered that can be considered as healthy status. Therefore we need to focus on these extreme value

points whose  $y$ -values vary sharply while filtering out the points whose  $y$ -values don't change sharply.

The following methodology will be applied to EVS (Extreme Value Set) sorted in non-decreasing order of  $y$ -value changes. According to the above discussions, it is conceivable that non-health spike points usually vary severely in terms of their  $y$ -values. Using the definitions presented above we conclude that the value of  $|y_n - y_{n-1}|$  or  $|y_{n+1} - y_n|$  for point  $P_n$  is greater than the difference between  $y$ -values of two healthy adjacent points. Furthermore if a threshold  $T$  is set properly, then  $|y_n - y_{n-1}| < T$  or  $|y_{n+1} - y_n| < T$  indicates that the  $y$ -value of  $P_n$  has not changed that much and the system should be in healthy status. Based upon the real applications, the probability of system exceptions occurring is relatively small and therefore unhealthy spike points should occur with low probabilities. The probability of unhealthy spike points appearing can be defined as follows:

$$P(|y_n - y_{n-1}| > T) < \alpha \tag{92.2}$$

Where  $\alpha$  is small positive number. Then:

$$t = [\text{Total (EVS)} - \text{Total (EVS)}^{\alpha}] \tag{92.3}$$

Total (EVS) represent the cardinality of  $|y_n - y_{n-1}|$  in EVS, variable  $t$  is the time index satisfying:

$$T = |y_t - y_{t-1}| (t < n) \tag{92.4}$$

In this paper the value of  $\alpha$  is determined based upon the operational experiences of technical staff members of the data center.

Provided that the  $\alpha$  value is given then the value of  $T$  can be obtained via Eqs. (92.2), (92.3), and (92.4). Furthermore, we define a point  $P_n$  for which  $(|y_n - y_{n-1}|)$  is less than  $T$  to be a healthy point. Otherwise they are possibly unhealthy.

After the value of  $T$  having been given, the elements in EVS can be classified as follows.

For a given point  $P_n$  if  $|y_n - y_{n-1}| \geq T$  or  $|y_{n+1} - y_n| \geq T$ , then  $P_n$  will be added into the set called EPUHPS (Extremely Possible Unhealthy Point Set).

EPUHPS helps narrow the space to be investigated to find out the real unhealthy points. The same period of the historical data is also applied to conduct the analyzing procedure of seeking unhealthy points. The basic idea behind the analysis is to identify if any point in EPUHPS possesses the same pattern as its historical records. If the patterns are similar it is considered to be healthy otherwise it is unhealthy, where the same two time points of current and last week have the similar pattern if they both either are peaks or valleys.

The following method shows how an element in EPUHPS is identified as an unhealthy one. In Fig. 92.1 we plot two curves:  $l_1$  and  $l_2$ , where  $l_1$  the data (number of recorded errors) is curve of current week and  $l_2$  is the one of last week. Each  $P_n$

that is in EPUHPS of  $l_1$  curve is either a “peak” or a “valley”. If its trend is different from the trend (“valley” or “peak”) of the same period in  $l_2$ , then  $P_n$  will be added into Unhealthy Point Set (UHPS), that is, it is an unhealthy point. However if the trend is similar, we will have to compare their local variance  $s_p$  (the variance computed based upon the  $y$ -values of the three points:  $P_n, P_{n-1}, P_{n+1}$  of  $l_1$ , and the  $y$ -values of the corresponding  $P_n, P_{n-1}, P_{n+1}$  of  $l_2$ ) and the global variance  $s_t$  (calculated based upon the  $y$ -values of all the points of  $l_1$  and the corresponding ones of  $l_2$ ) described below respectively. If  $|s_t - s_p| < \xi$  ( $\xi$  is a very small positive number determined by the system administrator or technical staff members of the data center), then the point  $P_n$  is a healthy point; otherwise  $P_n$  will be added into the UHPS.

$s_p$  and  $s_t$  are calculated as follows: let  $V = \{P_{l_11}(x_{l_11}, y_{l_11}), \dots, P_{l_1n}(x_{l_1n}, y_{l_1n})\}$  be the set of the  $l_1$  points and  $V = \{P_{l_21}(x_{l_21}, y_{l_21}), \dots, P_{l_2n}(x_{l_2n}, y_{l_2n})\}$  be the set of the points of  $l_2$  for the give time periods with the same length (1 week, for instance), then

$$s_t = \frac{\sum_{i=1}^n (y_{l_1i} - y_{l_2i})}{n} \tag{92.5}$$

Let  $y_1, y_2, y_3$  be the  $y$ -values of points  $P_n, P_{n-1}, P_{n+1}$  of  $l_1$ , and  $y'_1, y'_2, y'_3$  be the  $y$ -values of the corresponding points of curve  $l_2$ , then

$$s_p = \frac{(y_1 - y'_1)^2 + (y_2 - y'_2)^2 + (y_3 - y'_3)^2}{3} \tag{92.6}$$

The results of the experiments demonstrate the effective filtering ability for narrowing the searching space of possible unhealthy status.

In addition to the spike points detected by the basic approach, there are some special cases to be considered. For instance, some point’s  $y$ -value is very large and way above the normal, but the trend is relatively smooth. In order to find this type of spikes, we propose the following Threshold Strategy.

In this method threshold value  $T$  is set by the system administrators or technical staff members of the data center depending on their experience or the given rules. The  $y$ -value of the observed point  $P_n(x_n, y_n)$  will be compared with  $T$ . If the former is bigger, then the observed point is considered as an unhealthy one and the corresponding point is added to UHPS.

The set of unhealthy points is therefore formed via methods described above and the corresponding alerts will be disseminated to the interested subscribers (system administrators or technical staff members) for taking proper actions.

## 92.3 Experiment Results

The experimental data was collected from the data center of one of the largest e-commerce companies of the world. Three experiments conducted: experiments A, B, and C. The basic approach are performed in experiment A, and in experiment B, the combination of basic approach and Threshold strategy is applied. In both experiments we set up  $\alpha = 0.05$  in the experiment A, B. In the experiment C, we set  $\alpha = 0.1$  and apply the combination of Basic Detection Strategy and the Threshold Strategy. The purpose of these experiments is to investigate the impacts of the combination of the detection methods and the parameter in addition to the effectiveness of the proposed methods. The results can provide the guidance of establishing business rules in detecting the system status.

Based on the obtained results yielded by different methods, we will evaluate their performances and provide the guideline of applying these methods in detecting the system status.

The actual unhealthy spike points are listed in Table 92.1:

### Experiment A

In this experiment, we will use the sample in Fig. 92.1 and adapt the Basic Detection Strategy, where  $\alpha = 0.05$ .

The result shown in the Table 92.2 demonstrates that the Basic Detection Method is able to find unhealthy spike points partially. It detects about 45 % reported unhealthy spiking points shown in the Table 92.1 under the current parameter setting.

### Experiment B

In this experiment, we will use the sample in Fig. 92.1 and apply the combination of Basic Detection Strategy and the Threshold Strategy, where  $\alpha = 0.05$  the threshold as 50,000 for the Threshold strategy.

The result in the Table 92.3 shows the improvement in detecting unhealthy status of the system as it is able to find 54% of reported unhealthy spiking points.

### Experiment C

In this computational experiment we set  $\alpha = 0.1$  and the threshold to be 50,000 for the combination of Basic Detection Strategy and the Threshold Strategy.

The outcome in the Table 92.4 demonstrates the superiority of the combined strategies with the parameter setting. It is able to obtain much better results than those in experiments A and B. More than 90 % of the unhealthy spiking points can be detected that in turn can provide necessary alerts in time to handle the exceptions.

Based on the experimental results, the combination of the Basic Detection Strategy and the Threshold Strategy is proven to be an effective procedure to detect unhealthy spiking points with a proper setting of  $\alpha$ . The methods discussed in this paper provide a solid base to detect the unhealthy status of a Cloud Computing System.

**Table 92.1** Actual unhealthy spike points in the Fig. 92.1

Date/time	Error count
10/02/2012 10:00:00 AM	0
10/02/2012 02:00:00 PM	0
10/03/2012 08:00:00 AM	31,633
10/04/2012 01:00:00 PM	0
10/04/2012 02:00:00 PM	28,601
10/06/2012 08:00:00 AM	35,103
10/06/2012 10:00:00 AM	0
10/06/2012 11:00:00 AM	33,916
10/07/2012 04:00:00 AM	0
10/07/2012 12:00:00 PM	62,455
10/08/2012 01:00:00 PM	36,509

**Table 92.2** Result in experiment A

Date/time	Error count
10/02/2012 02:00:00 PM	0
10/04/2012 01:00:00 PM	0
10/06/2012 08:00:00 AM	35,103
10/07/2012 04:00:00 AM	0
10/08/2012 01:00:00 PM	36,509

**Table 92.3** Result in experiment B

Date/time	Error count
10/02/2012 02:00:00 PM	0
10/04/2012 01:00:00 PM	0
10/06/2012 08:00:00 AM	35,103
10/07/2012 04:00:00 AM	0
10/07/2012 12:00:00 PM	62,455
10/08/2012 01:00:00 PM	36,509

**Table 92.4** Result in experiment C

Date/time	Error count
10/02/2012 10:00:00 AM	0
10/02/2012 02:00:00 PM	0
10/04/2012 01:00:00 PM	0
10/06/2012 07:00:00 AM	0
10/06/2012 08:00:00 AM	35,103
10/06/2012 10:00:00 AM	0
10/06/2012 11:00:00 AM	33,916
10/07/2012 04:00:00 AM	0
10/07/2012 12:00:00 PM	62,455
10/08/2012 01:00:00 PM	36,509



## 92.4 Conclusion

In this paper we first introduce the business background of detecting unhealthy status of a Cloud Computing system. We conduct the brief analysis on the data recording the number of errors during the system operation periods to lay the foundation of detecting algorithm development. A basic detection strategy and a threshold-based method are proposed, which can help finding the spiking points where the system could be unhealthy. The computational experiments are conducted to demonstrate the effectiveness of the proposed methods and impacts of various parameter settings. A system administrator or technical staff member is able to adjust the parameter upon his experience/desire to find unhealthy spiking points. Together with the proposed methods the computational results reveal the direction of establishing business rules for detecting unhealthy status effectively.

We are planning to collect more real datasets and perform more computational experiments to seek the further improvements of the algorithms. The topic of triggering an efficient business process to handle exceptions when unhealthy status is detected is also one of our future researches.

## References

1. Saripalli Prasad, Kiran, G. V. R., Shankar R. Ravi, Narware Harish, Bindal Nith. (2011) Load prediction and hot spot detection models for autonomic cloud computing. In: *4th IEEE/ACM International Conference on Cloud and Utility Computing (UCC 2011)*. IEEE Computer Society, Los Alamitos (pp. 397–402).
2. Xiaojun Yu, Qiaoyan Wen. (2010). A view about cloud data security from data life circle. In: *International Conference on Computational Intelligence and Software Engineering*. Peking University Press, Beijing (pp. 203–208).
3. Donglin Chen, Mingming Ma, Qiuyun Lv. (2012). Study on transaction management system in cloud service market. In: *2012 International Conference on Technology and Management*. Springer Verlag, Germany (pp. 479–483).
4. Zhengping Wu, Nailu Chu, Peng Su. (2012). Improving cloud service reliability. In: *2012 I.E. International Conference on Services Computing*. IEEE Computer Society, Los Alamitos (pp. 90–97).
5. Jianwei Yin, Yanming Ye, Bin Wu, Zuoning Chen. (2011). Cloud computing oriented network operating system and service platform. In: *2011 I.E. International Conference on Pervasive Computing and Communications Workshops*. IEEE, Piscataway (pp. 111–116).
6. Sunahara, Y. (1982). Treatment of irregular data. I. Probability models and statistics. *Systems & Control*, 26(4), 228–236.
7. Mitchell, M. (2003). Constructing analysis of variance. *Journal of Computers in Mathematics and Science Teaching*, 21(4), 381–410.
8. Hommes, S., State, R., Engel, T. (2012). A distance-based method to detect anomalous attributes in log files. In: *2012 IEEE/IFIP Network Operations and Management Symposium*. IEEE, Piscataway, NJ, USA (pp. 498–501).
9. Frei Adrian, Rennhard Marc. (2008). Histogram matrix: Log file visualization for anomaly detection. In: *3rd International Conference on Availability, Security, and Reliability*. IEEE, Piscataway, NJ, USA (pp. 610–617).

10. Stermsek, G., Strembeck, M., Neumann, G. (2007). A user profile derivation approach based on log-file analysis. In: *2007 International Conference on Information and Knowledge Engineering*. Las Vegas, NV, USA: CSREA Press (pp. 258–264).
11. Cheng, Y.-H., & Huang, C.-H. (2008). A design and implementation of a Web server log file analyzer. *WSEAS Transactions on Information Science and Applications*, 5(1), 8–13.

# Chapter 93

## Scoring System of Simulation Training Platform Based on Expert System

Wei Nie, Ying Wu, and Dabin Hu

**Abstract** In order to reduce the cost of operation training and improve efficiency of examination, the development of simulation training platform has achieved very good results. An intelligent scoring system based on expert system plays the role of the teacher and gives the student a just assessment. It uses the professional theory and practical experience as the evaluation criteria and analyzes the operator's operation process to realize the automatic scoring through the program algorithm. The application of scoring system evaluates the operation level of students and gives students guiding opinions and error analysis.

### 93.1 Introduction

Simulation training [1, 2] is now used to all walks of life because it is more efficient and affordable, compared with traditional training methods. In recent years, simulation training is a necessary training tool in many companies. They make simulation training as an important part of job training and a way of identification of technical and skills contest.

Simulation training platform establishes a virtual ship operation environment to make the operator with the feel of real boat scene. It can improve the rapid response capability of ship operation and capacity of safety operation and reduce operation accident of virtual equipment [3].

How would you assess the level of actual operator? The scoring system uses the professional theory and practical experience as the evaluation criteria and analyzes the operator's operation process to realize the automatic scoring through the program algorithm. This paper is based on the practical experience of the automatic

---

W. Nie (✉) • Y. Wu • D. Hu  
College of Naval Architecture and Power, Naval University of Engineering,  
Wuhan 430033, China  
e-mail: [niewei213@163.com](mailto:niewei213@163.com)

grading system exploitation, and the platform of simulation training is ship power system. The scoring system [4, 5] uses programming techniques to assess automatically operation ability of candidate, which is based on operation experience of ship power system of technical personnel that are restored in the database as the assessment standard. The scoring system has been applied to practical ship power simulation system and has received a good effect. It not only has improved the examination efficiency and fairness of the assessment but also has saved a lot of funding. The basic module and function of composing the system have been discussed. The key technologies to achieve the system have been described.

### **93.2 The Principle of Expert System Theory in the Scoring System**

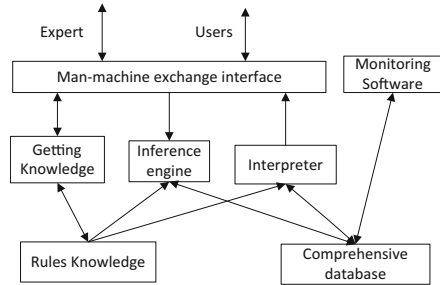
The expert system [6] is one of the main areas of artificial intelligence research. The expert system is also known as knowledge-based systems. The expert system is designed to simulate the work of the expert thinking and ways of tackling problems in certain areas by an intelligent computer program. It takes advantage of the human expert knowledge and problem-solving approach to solve the problems in the field.

In the traditional way of training and examination, the ability of candidate operation relies on the coaches' assessment. The training and examination of simulation training platform also rely on the coaches. After the candidates finish the operation of the devices, the coach judges the level of operation of the candidates and finally gives grades. The professional knowledge and practical experience of coaches play a decisive role in this process. It is an invisible criterion in the assessment process. The scoring system is used to replace the role of coaches and implements the scoring method of coaches through a computer. The realization of scoring system is combined with expert systems theory and data analysis method. The expert system theory makes all aspects of specialist expertise and practical experience as the standard of the automatic scoring. Certain reasoning mechanism can recognize the operation steps of the candidates and make reasonable judgment through the computer program. Data analysis is the basis of the scoring system which obtained candidates operation time steps, instrumentation and indicator of changes in the situation.

The expert system mainly consists of rules knowledge base, comprehensive database, inference engine, the interpreter, knowledge acquisition, and man-machine interface. Figure 93.1 shows the expert system structure of the scoring system of simulation training platform.

Rules knowledge is the key of quality of the expert system, including the quality and quantity of knowledge in database. In general, the expert system knowledge base and expert systems program are independent to each other. The user can change the content of knowledge base to improve the performance of the expert system. The inference engine matches rules in the knowledge base to access to new

**Fig. 93.1** The expert system structure of the scoring system



conclusions and get the results of problem solving. The inference engine of the scoring system realizes by programming language. The inference engine is the organization control mechanism of the expert system. Reasonable inference engine can make use of the rules of the knowledge base to solve practical problems.

### 93.3 The Knowledge Base of Expert System

Appropriate knowledge representation could convert the expert knowledge to the handle expression of computer system. The knowledge base of the scoring system is expert knowledge and practical experience of technical staff of simulation training platform operation, which is stored in the database.

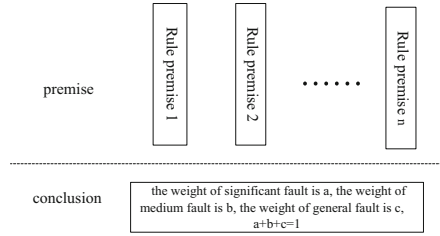
#### 93.3.1 Knowledge Acquisition

Knowledge acquisition [7, 8] is one of the main works of the expert system development, which is the basic technique of artificial intelligence and knowledge engineering. The most important sources are expert knowledge and practical experience. In the scoring system, field experts and technical personnel work together to complete the acquisition of knowledge. The process of development mainly divided into the following steps.

First, the developer must have the complete mastery of simulation training platform, including the method of operation, its composition, functions, and training. They could apply books and operating manual to practice and exchange with skilled operators.

Second, the obtained knowledge was systematized. The developer must have actual operating experience and repeat discussions with experts and professors to do further research on the knowledge of science and rationality.

**Fig. 93.2** The basic form of knowledge representation of the scoring system



### 93.3.2 Knowledge Representation

Knowledge representation of knowledge is a formal and symbolic process, which uses a computer language to encode the domain knowledge. It has an important impact on the performance of the system. Knowledge representation methods relate to the design of various data structures. The purpose of knowledge representation is to be able to use this knowledge to reason and make decisions through effective representation of knowledge. Figure 93.2 shows the basic form of knowledge representation of the scoring system.

In this paper, the knowledge representation is production rules. The basic form of rule representation is as follows:

If P then Q

Where P is the premise of the rule and Q is the conclusion. For example, if the valve of A has not turned on before you the switch B, This situation was recorded as a medium error. Where, the premise of the rule is a compound condition that constitutes by two simple conditions. In the mark-reducing method, the conclusion can be divided into several levels from low to high; the higher the level, the more the points. If the degree of operation mistake is between the medium fault and the significant fault, you can introduce the idea of fuzzy mathematics.

The knowledge representation is determined mainly by the following aspects: knowledge-use efficiency, understandability, and the degree difficulty of knowledge maintenance. The same knowledge can be expressed through a variety of methods, but the effect will be very different. Every kind of knowledge representation has its own characteristics, which also have their own advantages and disadvantages to different areas. Thus, it can combine several representations to achieve a particular field of knowledge representation.

The form of production rules knowledge is simple to understand and explain. The rules are independent to the extraction and formalization of knowledge. But constraints and interactions of production lead to the low inference efficiency. In the scoring system, each production rule is one of the scoring criteria. The knowledge base of each item consists of a certain rule.

### 93.4 The Scoring System of Simulation Training Platform

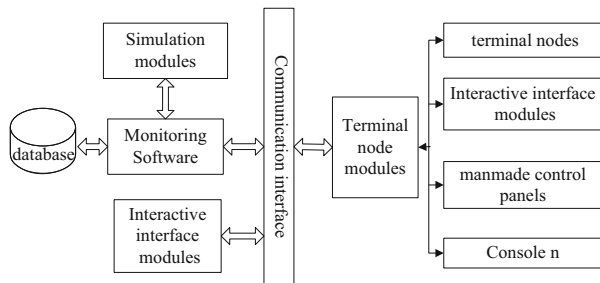
In the training of simulation training platform, the teacher makes the judgment to the operational level of students after examination. The teacher’s specialized knowledge and practical experience have played the decisive role in this process. Now, we have built a system in which the scoring plays the role of the teacher.

Simulation training platform achieves the purpose of training students through restoring the operating environment and feel of the real equipment. The simulation training platform consists of simulation modules, monitoring software, communication interface, terminal node modules, database, and so on. Figure 93.3 shows the basic principle of the scoring system of simulation training platform.

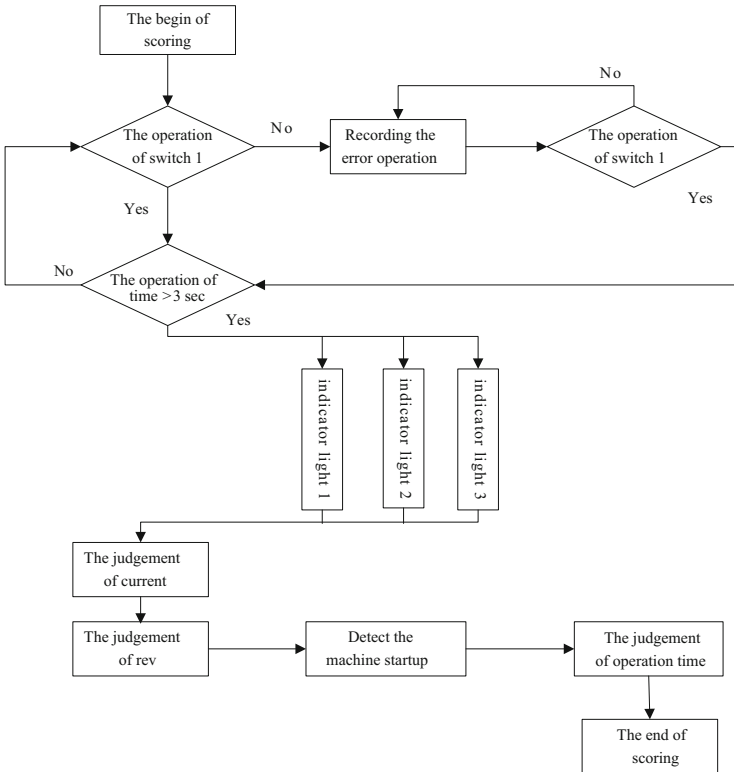
Simulation modules can simultaneously perform multiple simulation models, including the dynamic and steady state simulation of the main equipment of simulation training platform, the operation simulation of normal and fault conditions of the power system, and the data interconnection with other systems. Monitoring software can exchange the data with other software and obtain the data recording the student’s operation information from the terminal node modules. The function of database is to record the process of students’ operation information.

The scoring system is mainly made up of the database and software programming. The database stored all kinds of data information, including scoring criteria of expert knowledge, operation information, and basic personal information and examination results. Those data could exchange with other software and obtain the data recording the operation information from the consoles. Through analyzing those data from the database, the system can identify the operation steps of students. The scoring rules stored in the database can give a reasonable grading of operation level by programming technology.

The reasoning implementation of the scoring system is forward reasoning, which is based on the known facts as the starting point of reasoning. The basic idea of forward reasoning is that the computer identify the current applicable knowledge from the initial facts, and then constitute a set of applicable knowledge and some conflict resolution strategies. Figure 93.4 shows the reasoning mechanism of single armature start-up operation.



**Fig. 93.3** The basic principle of the scoring system of simulation training platform



**Fig. 93.4** The reasoning mechanism of single armature start-up operation

In the reasoning process of the scoring system, the key is the entire production rule. A rule expresses that if the premise is satisfied, you can launch the conclusion that the conclusions under the operation will be performed. In order to promote the operation of the scoring system of simulation training platform, the establishment of a knowledge base must follow the next regulations.

1. All rules constitute the rules knowledge base of scoring system operation judge, which contains all possible operational errors. Each rule can consist of certain sliver rules. If any of the operation steps meets this rule, the conclusion of this rule will be executed.
2. When an operation error may satisfy the two rules processing and trigger by the two rules, it should get the greater point as the coefficient of operation error.
3. The weights of the scoring system are initially drawn up according to the actual staff of long-term experience and the advice of experts, which are stored in the database as the initial default value. According to the reasonableness of the results of the scoring, those weights can be dynamic changes in the process of the scoring system debugging, making the weights more reasonable.



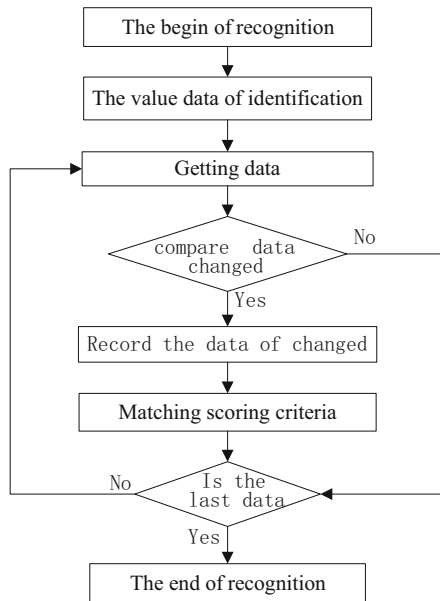
4. In the scoring rules, there are many compound checking rules. It considers mainly the following two aspects: First, whether is the time of the operation right which is accordance with the operation instructions. Second, whether is the steps correct which includes the operation of button and the button in the correct order. If the student gives a correct operation after drain operation, it is the operating part of the operating mistakes.

### 93.5 The Example of Program Algorithm of Scoring System

After the end of the examination, the operation information of the candidates would be stored in the database. Automatic scoring system identifies the operation of candidates through access database data and saves it to an array variable. The most critical of the data analysis is the algorithm of the recognition process. Each record of discrete variables in the database corresponds to an operation of the candidates. It can uniquely identify candidate operation by detecting the changes in the database, such as the operating button or switch. It is an important foundation of the recognition algorithm. Figure 93.5 is a simplified schematic of the implementation process.

Analysis of the data is that procedure of the candidates was got from the recorded data. The steps identified were mainly realized through certain procedure algorithm.

The candidate’s operation identification mainly includes two aspects. First, combing with the time variable and the value of the operating variables, operating



**Fig. 93.5** A simplified schematic of the implementation process

time of the button or switch was achieved by similar methods of code program. Second, the problem is how to determine the change direction of the button or switch operation.

## 93.6 Conclusion

The development of the scoring system of simulation training platform is a great help for training and teaching. The design and implementation of the scoring system involve many fields of knowledge. It is a new way that an expert system theory was applied to the scoring system of simulation training platform. The emphasis of system development is the establishment of a knowledge base. The established methods and guidelines of the knowledge base were particularly introduced. In practical applications, the automatic scoring system also achieved good results.

## References

1. Qingfu Kong, Jingyang Shong, Fanming Zeng. (2010). Technological condition and developmental trend of marine engine simulation training equipment. *Ship Science and Technology*, 32(2), 138–140.
2. Wei Nie, & Ying Wu. (2010) Study of automatic scoring system for simulation training. *2010 3rd International Conference on Advanced Computer Theory and Engineering* (pp. 403–406). Chengdu: IEEE Conference Publications.
3. Abhishek Kumar, Seung, Kyum Choi. (2011). Tolerance allocation of assemblies using fuzzy comprehensive evaluation and decision support process. *The International Journal of Advanced Manufacturing Technology*, 55(4), 379–391.
4. Xingtao Zhao. (2008) Improvement of automatic scoring of the power plant simulator training. *Automation of Electric Power Systems*, 18(3), 74–75.
5. Guojiang Bao. (2008). *Research on the ship maneuvering evaluation system based on ship handling simulator*. Dalian: Dalian Maritime University.
6. Shenghua Cai, Zhuxiao Liang. (2008). Student evaluation of large power plant simulation system. *Journal of System Simulation*, 20(21), 5989–5992.
7. Hong Zhang, Feng Xiu, Biguang Jin. (2005). Simulating test of ship navigation safety evaluation using ship handling simulator. *Ship Science and Technology*, 26(5), 567–571.
8. Yongsheng Fan, Fangzhen Cheng. (2000). Study on the scoring system of power plant training simulator. *Journal of System Simulation*, 12(3), 282–28.

# Chapter 94

## Analysis of Distributed File Systems on Virtualized Cloud Computing Environment

Tiezhu Zhao, Zusheng Zhang, and Huaqiang Yuan

**Abstract** Although various performance characteristics of distributed file system have been documented, the potential performance efficiency of distributed file system on virtualized cloud computing infrastructure is not clear. This chapter focuses on the performance of Hadoop Distributed File System (HDFS) on virtualized Hadoop. We construct a virtualized Hadoop platform and perform a series of experiments to investigate the performance of HDFS on the virtualized Hadoop cluster. Experimental results verify the efficiency of distributed file system on virtualized Hadoop to process the mass-intensive application.

### 94.1 Introduction

Distributed file systems can effectively solve the problems of the mass data storage and I/O bottlenecks in the mass distributed storage system and become the research hotspot of the storage industry and academia. Distributed file systems are key building blocks for cloud computing applications. Therefore, the industry is witnessing distributed file systems for large data center storage [1]. The performance of distributed file system directly affects the efficiency of the whole distributed computing environment. Therefore, the performance of distributed file system is the key research issue.

Hadoop is a highly scalable compute and storage platform for implementing the Google MapReduce algorithms in a scalable fashion on commodity hardware. The core Hadoop project solves two problems with big data: fast, reliable storage (HDFS) and batch processing (MapReduce) [2]. The HDFS cluster consists of a single NameNode, a master server that manages the file system namespace and regulates access to files by clients. In addition, there are a number of DataNodes,

---

T. Zhao (✉) • Z. Zhang • H. Yuan  
Engineering and Technology Institute, Dongguan University of Technology,  
Dongguan 523808, China  
e-mail: [tzzhao83@163.com](mailto:tzzhao83@163.com)

usually one per node in the cluster, which manage storage attached to the nodes that they run on. HDFS exposes a file system namespace and allows user data to be stored in files.

Although the HDFS have been widely studied for several years, there are relatively few studies on HDFS with the virtualized Hadoop platform. The potential impact to application performance of distributed file system on the virtualized Hadoop platform is not clearly understood. In this chapter, we focus on the performance of HDFS on virtualized cloud computing environment. The main contribution of this chapter can be summarized as follows (1) A virtualized Hadoop platform is constructed to investigate the performance of HDFS on the virtualized Hadoop cluster; (2) We verify the performance characteristics of HDFS in the context of different application scenarios.

The remainder of this chapter is organized as follows. We begin by introducing related work in Sect. 94.2. The virtualized Hadoop platform is proposed in Sect. 94.3. We investigate the performance of HDFS on the virtualized Hadoop cluster, discuss the experiment results in Sect. 94.4, and conclude the chapter in Sect. 94.5.

## 94.2 Related Work

Existing research for distributed file system can be classified into four categories: (1) Performance analysis of distributed file system with the specific application scenario. The use of clustered file systems as a backend for Hadoop storage has been studied previously. The performance of distributed file systems such as Lustre, PVFS and GPFS with Hadoop has been compared to that of HDFS [3]. Most of these investigations have shown that non-HDFS file systems perform more poorly than HDFS, although with various optimizations and tuning efforts, a clustered file system can reach parity with HDFS [4]. (2) Metadata management and query optimization. Metadata management is critical in scaling the overall performance of large-scale data storage systems and a large-scale distributed file system must provide a fast and scalable metadata lookup service. Wang et al. proposed a two-level metadata management method to achieve higher availability of the parallel file system while maintaining good performance [5]. (3) Performance parameter analysis and tuning. Yu et al. indicated that excessively wide striping can cause performance. To mitigate striping overhead and benefit collective IO, authors proposed two techniques: split writing and hierarchical striping to gain better IO performance [6]. Yu et al. presented an extensive characterization, tuning, and optimization of parallel I/O on the Cray XT supercomputer (named jaguar) and characterized the performance and scalability for different levels of storage hierarchy [7]. (4) Optimizing data distribution strategy and data access strategies. Li et al. modeled the whole storage system's architecture based on closed Fork-Join queue model and proposed an approximate parameters analysis method to build performance model [8]. Yu et al. adopted a user-level perspective to empirically reveal the implications of storage organization to parallel programs running on

Jaguar and discovered that the file distribution pattern can impact the aggregated I/O bandwidth [9]. Piernas et al. adopted a novel user-space implementation of active storage for Lustre and the user-space approach has proved to be faster, more flexible, portable, and readily deployable than the kernel-space version [10].

### 94.3 Virtualized Hadoop Platform

#### 94.3.1 Execution Engine of Hadoop

To get an idea of how data flows between the client interacting with HDFS, the NameNode, and the DataNode, consider Fig. 94.1, which shows the execution engine of Hadoop. The I/O flow of execution engine consists of four main modules: Application Launcher module, JobTracker module, TaskTracker module, Hadoop Distributed File System (HDFS) module. All the four modules cooperate with each other to implement MapReduce application, using HDFS for storage.

The execution engine workflow is shown as follows: (1) Application Launcher submits a job to JobTracker and saves jars of the job in the HDFS file system. (2) JobTracker monitors all jobs' execution status and makes scheduling decision. JobTracker divides the job into several map/reduce tasks and schedules tasks to TaskTracker for execution. According to the implementation situation, JobTracker reports execution status to JobTracker. (3) TaskTracker runs on every node and manages the status of all tasks which run on that node. TaskTracker fetches task jar from the HDFS file system and saves the data into the local file system. (4) HDFS is responsible for storing data, which come from Application Launcher and TaskTracker.

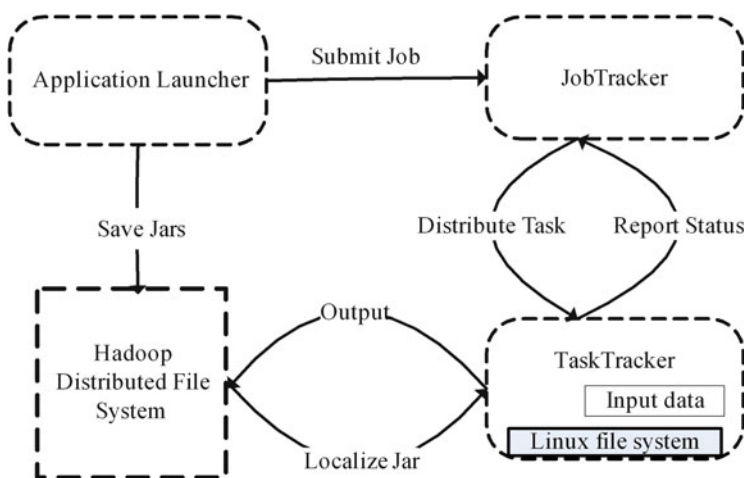


Fig. 94.1 Execution engine of Hadoop

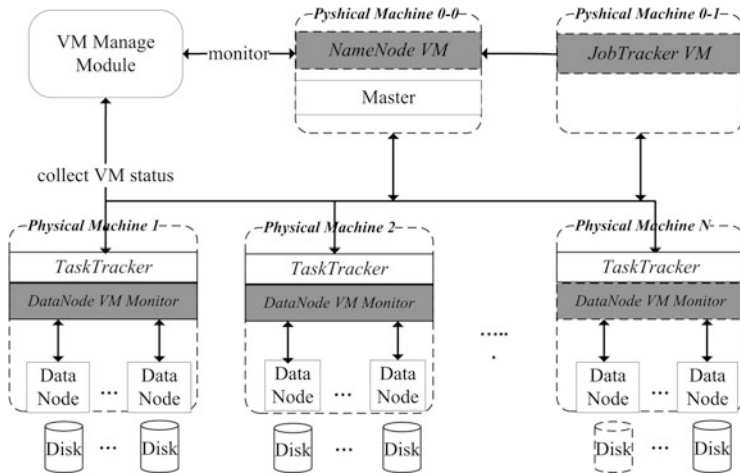


Fig. 94.2 Architecture of virtualized Hadoop platform

### 94.3.2 Architecture of Virtualized Hadoop Platform

Figure 94.2 illustrates the architecture of the virtualized Hadoop cluster platform. In this platform, VM manage module, which manages all virtual resources, is the key component. It is responsible for monitoring the virtual machine resource status and tuning the configuration parameters of the virtualized Hadoop cluster. DataNode VM monitor is responsible for monitoring the utilization status of DataNode, including the utilization of CPU, memory, disk, and network status.

By using the virtualization technology, one physical machine can be shared by several virtual machines. For the performance consideration, NameNode VM and JobTracer VM are deployed in two physical machines, respectively. Each physical machine can deploy multiple DataNode VMs, which are monitored by DataNode VM monitor.

## 94.4 Experiment Analysis

### 94.4.1 Experiment Setup and Scenarios Design

The experiment was performed on the Hadoop cluster with nine physical nodes and the configuration is shown in Table 94.1. The experiment Hadoop cluster is constructed as: two nodes for servers (one for NameNode, one for JobTracer) and the remaining nodes are DataNodes.

**Table 94.1** Configuration of experiment environment

Component	Detailed description
Hadoop version	Hadoop 0.20.1
Physical machine OS	Red Hat Enterprise Linux 5
Virtual machine OS	CentOS 5.5
NameNode	6 core Xeon X5650 processors at 2.66 GHz and 48 GB Memory
DataNode	6 core Xeon X5650 processors at 2.66 GHz and 48 GB Memory
Network	10-Gigabit TCP/IP Ethernet
Xen version	Xen 3.4.4
Benchmark	TeraSort, TestDFSIO

Our experiment chooses two typical benchmarks to evaluate the performance of HDFS on virtualized Hadoop platform: (1) TeraSort benchmark, represents one typical use case, is probably the most well-known Hadoop benchmark that combines testing the HDFS and MapReduce layers of the Hadoop cluster. (2) TestDFSIO benchmark, an I/O-intensive Hadoop standard benchmark, is used to test the backend file system performance of HDFS and Lustre in the context of a Hadoop job. It is helpful for tasks such as stress testing HDFS and Lustre, to discover performance bottlenecks in your network, and to give you a first impression of how fast your cluster is in terms of I/O.

The experiment mainly considers three aspects: read/write, datasize scale, and DataNode scale. The experiment cases are designed as follows:

*Case 1:* Analyze the read/write running time corresponding to different dataset size using TeraSort benchmark. The detailed configuration is as follows: 7/28 virtual DataNodes (vDataNodes) with the configuration of 2 vCPUs and 2,048 MB vMemory; datasize = 100 MB, 400 MB, ..., 2,200 MB; blocksize = 64 MB; replication = 3.

*Case 2:* Analyze the read/write throughput corresponding to the different number of vDataNodes using TestDFSIO benchmark. The detailed configuration is as follows: num. of vDataNodes=1, 5, 9, 13, 17, 21, 25, 29, 33; each vDataNode with 2 vCPUs and 2,048 MB vMemory; datasize = 5 GB; blocksize = 64 MB; replication = 3.

## 94.4.2 Experiment Result Analysis

For all experiment scenarios, the writing/reading performance is tested three times and a median value will be compared to the other to avoid outliers. In case 1, the running time is selected as the performance metric. Figure 94.3 plots the read/write running time for different datasizes, which vary from 100 to 2,200 MB.

As shown in Fig. 94.3, we note that as the number of datasize scales, the running time increases quickly. When the datasize is small, the time difference is relatively small. However, when the data size exceeds 1,300 MB, the running time difference

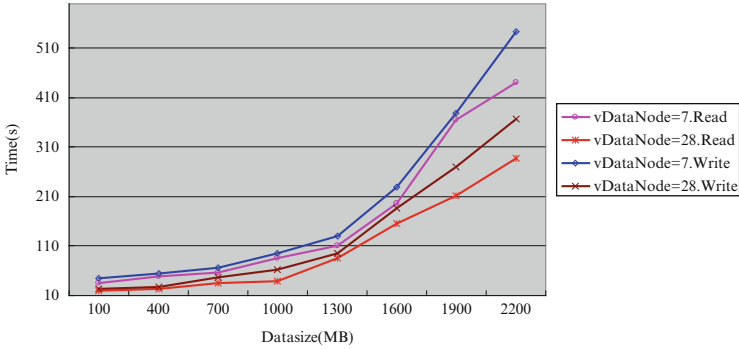


Fig. 94.3 Running time comparison of the different datatypes

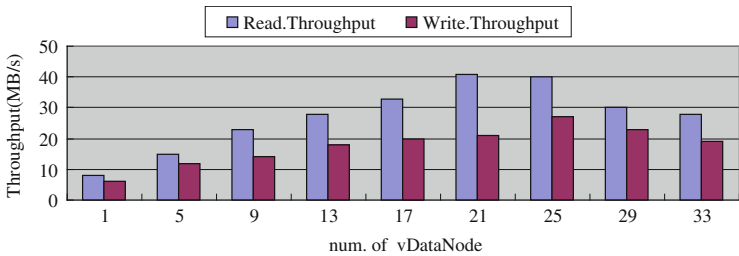


Fig. 94.4 Throughput comparison of the different number of vDataNodes

increases quickly. The read performance of HDFS on the virtualized Hadoop platform is better than the write performance. The performance of HDFS is better when the number of vDataNodes is 28.

In case 2, the throughput (MB/s) is selected as the performance metric. Figure 94.4 shows the performance (throughput) changes with the increase of the number of vDataNodes.

As illustrated in Fig. 94.4, the read throughput is better than write throughput. The read throughput continues to rise with the increase of the number of vDataNodes, and arrives at the maximum value when then vDataNode number is 21. Later, the read throughput is declining as the number of vDataNodes increases. The main reason is the competition for physical resources when the number of vDataNodes is too much (>21). The change of the write throughput is similar to the read throughput. The results clearly indicate that the distributed file system can maintain better efficiency on the virtualized cloud computing environment.



## 94.5 Conclusion

In this chapter, we study the performance and efficiency of distributed file system on the virtualized computing environment. We first introduce the I/O flow characteristic of HDFS and construct a virtualized Hadoop platform. Then, we perform a series of experiment to investigate the performance of HDFS on the virtualized Hadoop cluster platform. The experiment result shows that the distributed file system can maintain better efficiency on the virtualized cloud computing environment. It is necessary to study the performance of distributed file system in some specific application environments.

**Acknowledgements** This work is supported by the Natural Science Foundation of Guangdong Province, China (Grant No. S2012040007746), the Scientific Research Foundation for Doctors of DGUT(ZJ130604), the National Natural Science Foundation of China (Grant No. 61170216, 10805019, 61272200).

## References

1. Cheng, K., & Wang, N. (2012). The feasibility research of cloud storage based on global file system. In *Proceeding of 2012 9th International Conference on Fuzzy Systems and Knowledge Discovery* (pp. 2507–2511). Piscataway, NJ: IEEE.
2. Konstantin, S., Hairong, K., Sanjay, R., et al. (2010). The Hadoop distributed file system. In *Proceedings of the 2010 I.E. 26th Symposium on Mass Storage Systems and Technologies* (pp. 1–10). Washington, DC: IEEE Computer Society.
3. Sun Microsystems Inc. (2010) *Using Lustre with Apache Hadoop*. White Paper. pp. 1–25.
4. Xyratex Inc. (2011). *Map/reduce on Lustre*. White Paper. pp. 1–16.
5. Wang, F., Yue, Y. L., Feng, D., et al. (2007). High availability storage system based on two-level metadata management. In *Proceedings of the 2007 Japan–China Joint Workshop on Frontier of Computer Science and Technology* (pp. 41–48). Washington, DC: IEEE Computer Society.
6. Yu, W., Vetter, J. S., Canon, R. S., et al. (2007). Exploiting lustre file joining for effective collective IO. In *Proceeding of the Seventh IEEE International Symposium on Cluster Computing and the Grid* (pp. 267–274). Washington, DC: IEEE Computer Society.
7. Yu, W., Vetter, J. S., & Oral, H. S. (2008). Performance characterization and optimization of parallel I/O on the Cray XT. In *Proceeding of the 2008 I.E. International Symposium on Parallel and Distributed Processing* (pp. 1–11). Piscataway, NJ: IEEE.
8. Li, H. Y., Liu, Y., & Cao, Q. (2008). Approximate parameters analysis of a closed fork-join queue model in an object-based storage system. In *Proceeding of the Eighth International Symposium on Optical Storage and 2008 International Workshop on Information Data Storage* (pp. 1–8). Bellingham, WA: SPIE.
9. Yu, W., Oral, H. S., Canon, R. S., et al. (2008). Empirical analysis of a large-scale hierarchical storage system. In *Euro-Par 2008, LNCS, 5168* (pp. 130–140). Berlin: Springer.
10. Piernas, J., Nieplocha, J., & Felix, E. J. (2007). Evaluation of active storage strategies for the lustre parallel file system. In *Proceeding of the 2007 ACM/IEEE Conference on Supercomputing* (pp. 1–8). New York, NY: ACM.

# Chapter 95

## A Decision Support System with Dynamic Probability Adjustment for Fault Diagnosis in Critical Systems

Qiang Chen and Yun Xue

**Abstract** In order to locate and remove the faults in the critical systems where the faults occur, this paper proposes a three-layer decision support system for fault diagnosis, in which both static information and dynamic information of the system are used to find out suspicious components. In the process of locating the faults, a bipartite graph is applied to describe the relation between the symptom and the components, on the basis of which a method is proposed to calculate the value of fault evidence of a component. Then, the components whose values are larger are chosen as the result. Meanwhile, the decision support system adjusts the data of the bipartite graph according to the actual situation in order to improve the effectiveness of the diagnosis. The experiment shows that the fault diagnosis process in the decision support system can locate the fault more effectively.

### 95.1 Introduction

The traditional small application system becomes large and critical, which is widely used in military, defense, finance, and other important areas. Due to the limitation of time, personnel, and technical conditions, the practical application of the system is impossible to be developed without defect. Once these important applications failed and cannot be fixed quickly, that may result in immeasurable loss and serious consequences.

The traditional fault diagnosis is achieved by experts, which may take a long time, and the accuracy of the results is closely related to the individual ability.

---

Q. Chen (✉)

School of Management, Huazhong University of Science and Technology,  
Wuhan 430074, China  
e-mail: [kurt\\_cq@163.com](mailto:kurt_cq@163.com)

Y. Xue

Naval Academy of Armament, Beijing 100055, China

Meanwhile, the ascendants may change frequently, so the personal experience in this analysis is difficult to be shared with the others. Therefore, the traditional analysis process which emphasizes individual ability obviously cannot meet the actual requirements.

This paper presents a fault diagnosis decision support system, which applies the bipartite graph to build a system fault diagnosis model to analyze the possible reason to the failure. What is more, the model parameters can be adjusted dynamically according to the actual result of the fault. The result of the experiment indicates that the algorithm in this paper can get a better diagnostic effectiveness.

## 95.2 Related Work

Fault diagnosis is an analysis process to obtain the final cause of the failure based on the system through the fault symptoms and related information [1]. The fault diagnosis process can be taken as a black box, and its input is the information of the failed system which is organized in a particular way, and its output is the possible reason.

The existing diagnosis methods usually make use of the certain types of information of the failed system. Gunjan Khanna creates a real-time causality diagram by analyzing the information interaction between the components and analyzes the possible cause of the fault with the legal sequence of interactions defined according to the rules [2]. Wen generates a structured system model by calculating the program dependence graph and then uses the program slice to determine the reason of the fault [3].

Steinder used heuristic method to create a set of fault hypothesis which can explain all the events received and calculated the assumptions set of all the assumptions confidence. Then assumption with the largest confidence is selected as the results [4]. Zheng proposed a three-layer belief network model and analyzed the relationship between network events [5]. Zhang improves the accuracy of fault diagnosis by analyzing the relationship of the alarm information in the adjacent window of time [6].

There will be inconsistencies in the diagnosis result with the actual system operation, by reasoning based on the static information or symptoms. For example, the diagnosis result may indicate that the component A is failed, but A is not involved in the actual operation and cannot cause the generation of the fault. So the result greatly reduces the efficiency of diagnosis. If we only use the run-time information for diagnosis, the set of components which may result in the system failure can be generated, but it is difficult to analyze which is of the largest possibility, and it will also reduce the efficiency of diagnosis. Thus, we should use the two types of information for diagnosis, and the result can meet the actual needs.

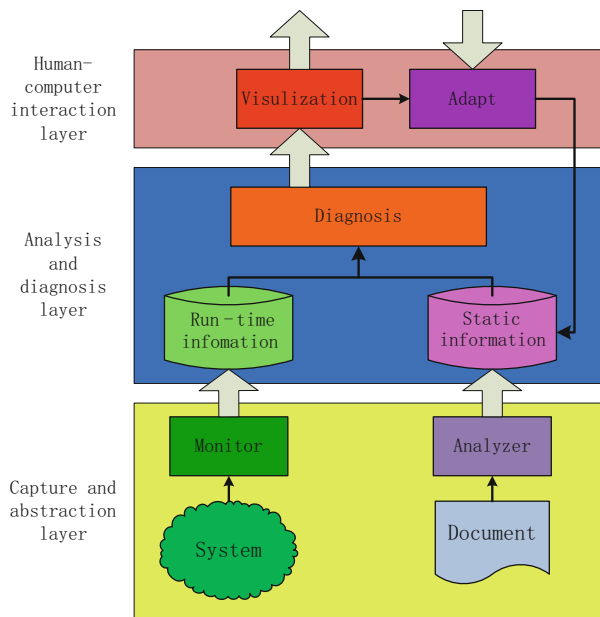
According to the above analysis, we provide a support system for fault diagnosis, which makes use of the static information and run-time information for diagnosis, and can adjust the result in the diagnosis to improve the efficiency.

### 95.3 Decision Support System for Fault Diagnosis

#### 95.3.1 System Framework

As shown in Fig. 95.1, the decision support system consists of three layers. From the bottom to the top, the diagnostic information is processed through the three levels to generate the result.

The capture and abstraction layer injects probes to capture the related information and extracts the static information from related documents. Based on this information, the analysis and diagnosis layer generates the result graphically displayed in the human-computer interaction layer. After each diagnosis, the decision support system should adjust the information for diagnosis according to the actual cause of the fault, in order to improve the efficiency of subsequent diagnosis.



**Fig. 95.1** Framework of the decision support system

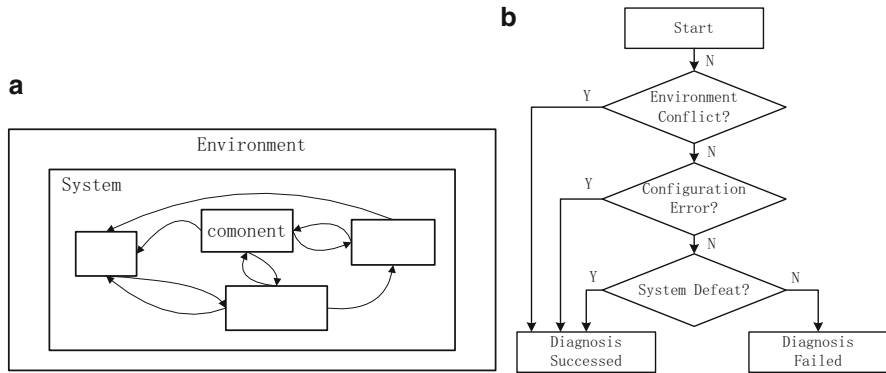


Fig. 95.2 (a) Schematic diagram of a failed system. (b) Diagnosis process

### 95.3.2 The Fault Diagnosis Strategy

As shown in Fig. 95.2a, the analysis of a failed system on the actual environment consists of the internal system environment, system boundary, and the system itself. Any problem in each part may result in failure. For example, the system is developed in a lab environment; there are some differences between the simulation environment and the actual environment, which are likely to cause the system to fail. If the configuration of each subsystem is incorrect, it will also result in failure. At the same time, each component of the system may also trigger fault. Therefore, the diagnosis strategy provided in this paper is to analyze the system from the outside to the inside.

Diagnosis strategy is shown in Fig. 95.2b. When a fault occurs, we first decide whether the system is incompatible with the environment. If the answer is no, then we will judge the correctness of the system configuration. Based on the exclusion of configuration problems, we will analyze each component and the interaction between them to find out the cause of the failure, and then the final diagnosis results are given.

In this paper, we focus on the third part of the diagnosis, and we provide an algorithm of fault location to achieve the diagnosis in the next section.

### 95.3.3 Fault Location Algorithm

Because of the spread of fault, the fault which occurs in one component may cause other component to fail which will generate different symptoms. At the same time, there are multiple mapping relationships between the symptom and the cause of the fault, in which multiple faults show the same symptoms or one fault results in

some different symptoms. The bipartite graph is competent for describing multiple mapping between the symptom and the cause, so we use the bipartite graph as the analysis model.

**95.3.3.1 Analysis Step**

The fault diagnosis process consists of the following steps:

First, the execution path  $t_r$  and bipartite graph model  $G_B$  of the failed system are extracted. Bipartite graph model can be defined as  $G_B = \{C, S, P_E, S_o\}$ .  $C$  stands for the component set  $C = \{(c_i, p_{ci})\}$ , in which  $c_i$  is the  $i$ th component and  $p_{ci}$  is the probability of the component to be failed.  $S$  is the symptom set including all the possible symptoms in the system.  $P_E$  is the set of the relationship between component and symptom, and  $P_E = \{p_{ci, sj}\} = \{p(s_j|c_i) \mid s_j \in S, c_i \in C\}$ , in which  $p(s_j|c_i)$  is the probability of component  $c_i$  to generate symptom  $s_j$ . If  $p_{ci, sj} > 0$ , component  $c_i$  can generate symptom  $s_j$ .  $E(c_i) = \{s_j \mid p_{ci, sj} > 0\}$  is the set of symptoms which can be generated by  $c_i$ .  $S_o \in S$  is the set of observed symptoms.

Second, defeat exists in the execution paths  $t_r$ , so we can focus on the bipartite graph  $G_B$  slice  $G_{B_r}$  according to the  $t_r$ .

The next step is to find out a set of components which can explain all the observed symptoms. In this paper, a heuristic greedy algorithm is applied to find out the suspicious component. In this process, we should calculate the evidentiary value of the component to select the suspicious component; therefore, the algorithm of calculating the component failure evidentiary value is important.

**95.3.3.2 Algorithm of Calculating the Component Failure Evidentiary Value**

Actually, the failure evidentiary value of the component  $c_i$  is affected by the following:

1. In the observed symptom set  $S_o$ , the more the symptoms of the component  $c_i$ , the greater the evidentiary value of the component; thus, the evidentiary value can be defined as

$$W_{eo}(c_i, S_o) = \sum_{s_j \in E(c_i) \wedge s_j \in S_o} p(s_j, c_i) \tag{95.1}$$

2. Symptom  $s_k$  can be explained by component  $c_i$  but has not been observed. Such symptoms can appear in the following cases. First, other failure symptoms that make  $s_k$  cannot be observed, so these symptoms should enhance the evidentiary value of the  $c_i$ .  $P_L(s_k)$  is defined as the probability of  $s_k$  that occurs without observation, and the evidentiary value can be defined as

$$W_{euc}(c_i, S_O) = \sum_{s_k \in E(c_i) \wedge s_k \in S - S_O} p_L(s_k) \cdot p(s_k, c_i) \tag{95.2}$$

Another case is that these symptoms without observation have not been produced actually. The more such symptoms, the smaller the probability of failure of  $c_i$ ; thus the evidentiary value can be defined as

$$W_{eue}(c_i, S_O) = \frac{1}{\sum_{s_m \in E(c_i) \wedge s_m \in S - S_O} 1 - p(s_m, c_i)} \tag{95.3}$$

3. In the set of the observed symptoms which can be explained by  $c_i$ , there are some false symptoms because of the incorrect threshold and inaccurate monitoring methods and so on. The observed false symptoms may affect the evidentiary value of  $c_i$ . So, the value can be calculated as

$$W_{es}(c_i, S_O) = \sum_{s_n \in E(c_i) \wedge s_n \in S_O} p_S(s_n) \cdot p(s_n, c_i) \tag{95.4}$$

According to the analysis above, the evidentiary value  $W(c_i, S_O)$  of  $c_i$  is defined as follows:

$$\begin{aligned} W(c_i, S_O) &= \frac{W_{eo}(c_i, S_O) + W_{euc}(c_i, S_O)}{W_{eue}(c_i, S_O) + W_{es}(c_i, S_O)} \\ &= \frac{\sum_{s_j \in E(c_i) \wedge s_j \in S_O} p(s_j, c_i) + \sum_{s_k \in S_U} p_L(s_k) \cdot p(s_k, c_i)}{\sum_{s_m \in E(c_i) \wedge s_m \in S - S_O} 1 - p(s_m, c_i) + \sum_{s_n \in E(c_i) \wedge s_n \in S_O} p_S(s_n) \cdot p(s_n, c_i)} \end{aligned} \tag{95.5}$$

Then, the fault diagnosis algorithm can be provided as follows:

Algorithm GFDA

Input: bipartite graph model  $G_B = \{C, S, P_E\}$  and the observed symptom set  $S_O$ ;

Output: a set of suspicious components  $H$ ;

GFDA ( $G_B, S_O$ ) {

$H = \phi$ ;  $C_e = \phi$ ;  $S_D = \phi$ ;

while( $S_O \neq \phi$ ) {

    analyze each  $s_j \in S_O$ , compute the  $C_e = \{c_i \mid c_i \in C, s_j \in E(c_i)\}$ ; calculate the evidentiary value  $W(c_i, S_O)$  of each  $c_i$  in the set of  $C_e$ ; choose the component  $cm$  with the greatest value;

$H = H \cup \{c_m\}$ ;  $S_O = S_O - \{s_n \mid s_n \in E(c_m), s_n \in S_O\}$ ;

$S_D = S_D \cup \{s_n \mid s_n \in E(c_m), s_n \in S_O\}$ ;

$F_e = \phi$ ;

  }

  OutPut( $H$ );

}

### 95.3.3.3 Algorithm Complexity Analysis

The GFDA outer cycle removes a symptom from the observed symptom set  $S_o$  at each time until all symptoms are removed. Thus, the outer cycle needed loop  $|S|$  times at most. Calculating the failure evidentiary value of the component  $c_i$  executes  $|C_e|$  times, of which the maximum times is  $|C|$ . Therefore, the complexity of GFDA is  $O(|S| \times |C|)$ .

### 95.3.4 Probability Adjustment Strategy

In the diagnosis process, some probabilities are used to calculate the component failure evidentiary value. These probabilities are obtained by analyzing the system historical data and configuration information based on the expert experience, so there is uncertainty in estimating these probabilities. Thus, it is very necessary to adjust the probability according to the actual diagnosis result.

$P_{c_i}$  is the probability of  $c_i$  to be failed, and it is in the range  $[0, 1]$ . If  $P_{c_i}$  is 0, it means that  $c_i$  does not fail. If  $P_{c_i}$  is 1, it means that  $c_i$  will fail. For the component with a higher failure probability, its growth rate based on a successful confirmation should be less than the growth of the low probability. Therefore,  $P_{c_i}$  can be adjusted as follows:

$$p_{c_i} = \alpha(1 - p_{c_i})^2 \quad (0.1 \leq \alpha \leq 0.2) \quad (95.6)$$

## 95.4 Experiment

In this experiment, we randomly generated a system with  $n$  components and the set  $S$  and  $P_E$ , according to the model  $GB = \{C, S, P_E, S_o\}$ . Then, the fault set is selected, and the set  $S_o$  is also selected from the set  $S$  randomly. In order to simulate the symptom loss, some symptoms in  $S_o$  are removed randomly. In addition, some of the symptom is not contained in  $S_o$  initially and will be added into  $S_o$  to simulate the mendacious symptoms. The number of this removed or added symptom is decided randomly. Finally, the algorithms proposed by Steinde, Zhang, and our paper will be performed ten times to compare their effectiveness, in which the number of component is increased gradually.

We use two marks to assess the fault location algorithm: detection rate and the false-positive rate, which may indicate the effectiveness of the algorithm. And Fig. 95.3 shows the result of the experiment.

Figure 95.3a shows that the detection rate of our algorithm is higher. What is more, as the number of component is increased, there are fluctuations in the other two algorithms. But our algorithm can still maintain a high level due to the adjustment strategy which can increase the detection rate as the experiment is



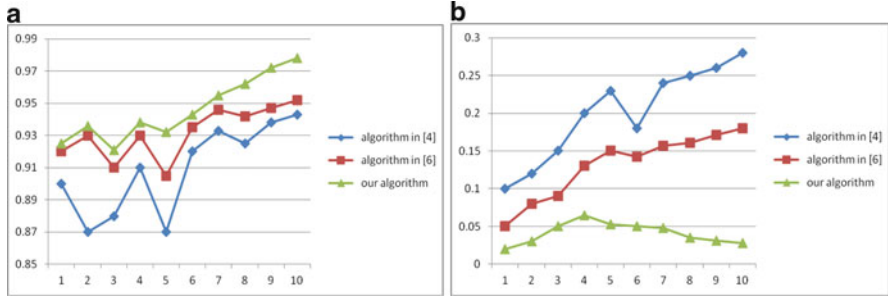


Fig. 95.3 (a) Detection rate of algorithms. (b) False-positive rate of algorithms

repeated. As shown in Fig. 95.3b, our algorithm is the best in the false-positive rate. Although incensement of the false-positive rate of the other two algorithms is not rapid in the later experiments, our algorithm can reduce the rate as the repeat of the experiment process.

In conclusion, the result of experiment indicates that the algorithm in this paper can get a better diagnostic effectiveness especially in the multiplication of the experiment.

### 95.5 Conclusion

This study constructed a decision support system with three layers to assist the maintainer to fix the failure. We used the bipartite graph to describe the correspondence relationship between the components and the symptoms. Then the set of suspicious components which can explain all the observed symptoms was generated by calculating the probability of component to be failed. The probability adjustment mechanism was introduced in the support system in order to improve the efficiency of fault diagnosis.

We focus on the faults caused by the components in the system. These faults caused by the environment or configuration should be studied further to improve the comprehensiveness of the fault diagnosis decision support system.

### References

1. Małgorzata Steindera, & Adarshpal S. Sethi. (2004). A survey of fault localization techniques in computer networks. *Science of Computer Programming*, 53(2), 165–194.
2. Khanna, G., Cheng, M. Y., Varadharajan, P., Bagchi, S., Correia, M. P., & Verissimo, P. J. (2007). Automated rule-based diagnosis through a distributed monitor system. *IEEE Transactions on Dependable and Secure Computing*, 4(4), 266–279.

3. Wanzhi Wen, Bixin Li, Xiaobing Sun, & Cuicui Liu. (2013). Technique of software fault localization based on hierarchical slicing spectrum. *Journal of Software*, 24(5), 977–992 (In Chinese).
4. Steinder, M., & Sethi, A. S. (2004). Probabilistic fault diagnosis in communication systems through incremental hypothesis updating. *Computer Networks*, 45(4), 537–562.
5. Zheng, Q., Hu, W., Qian, Y., Yao, M., Wang, X., & Chen, J. (2008). A novel approach for network event correlation based on set covering. *The Proceeding of Fifth International Conference on Fuzzy Systems and Knowledge Discovery*, 3(3), 122–126.
6. Zhang, C., Liao, J., Li, T., & Zhu, X. (2012). Probabilistic fault localization with sliding windows. *Science China Information Sciences*, 55(5), 1186–1200.

# Chapter 96

## Design and Implementation of an SD Interface to Multiple-Target Interface Bridge

Guoyong Li, Leibo Liu, Shouyi Yin, Dajiang Liu, and Shaojun Wei

**Abstract** The design and implementation of an SD card controller circuit architecture for multiple-target interface, suitable for communication function extension of existing electronic device for UBICOMP, are presented in this paper. The SD to multiple targets bridge includes an SD memory controller, a ping-pong FIFO, and a target selectable interface, such as UART, SPI, parallel, and NAND Flash IO. The bridge follows SD memory card v2.0 specification so that it is fully flexible in terms of portable device without any special drivers. The ping-pong FIFO increases the throughput of this system, and the availability of UART, SPI, parallel, and NAND flash interfaces provides flexibility for implementation of applications that requires the conversion of data to feed the SD bus. A tidy NAND flash is also implemented in the multiple-target interface for FTL of NAND flash. The new design has been verified and implemented in FPGA. It has also been synthesized and will be taped out through a 0.18  $\mu\text{m}$  CMOS technology. Experiment reveals that the proposed architecture presents superior performance in platform-independent, interface-scalability and integrality compared with existing works.

### 96.1 Introduction

Ubiquitous computing (UBICOMP) [1] is the trend towards increasingly ubiquitous connected computing devices in the environment. Information processing of UBICOMP in human-computer interaction has been thoroughly integrated into everyday objects and activities. Meanwhile, peer-to-peer (P2P) communication [2]

---

G. Li

Research Center for Mobile Computing, Tsinghua University, Beijing 100084, China

L. Liu (✉) • S. Yin • D. Liu • S. Wei

Research Center for Mobile Computing, Tsinghua University, Beijing 100084, China

Institute of Microelectronics, Tsinghua University, Beijing 100084, China

e-mail: [liulb@tsinghua.edu.cn](mailto:liulb@tsinghua.edu.cn)

is needed by electronic devices in UBICOMP system. However, most existing electronic terminals are not available for P2P communication. Therefore, how to extend these terminals with function of P2P communication handily is very important. Security Digital (SD) protocol is an industry-leading memory card storage standard that simplifies the use and extends the life of consumer electronics, e.g., mobile phones, for millions of people every day. Most of modern computers and mobile terminals are equipped with SD slots, such as mobile phone and PDA. This universality makes the SD-based function extensions, especially P2P communication, become a new and hot research area, where SD interface does not only serve as a memory access protocol but also becomes a channels of function extensions of existing mobile terminals.

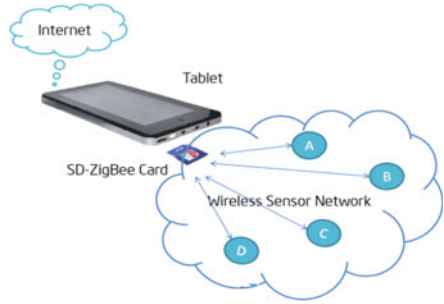
The SD Memory card slave controller is designed to reside within an SD memory card. It serves as a bridge between the SD bus and user logic that provides the actual function of the card. Using the standard SD bus protocol, computers and mobile terminals could access various application devices such as blue tooth, WiFi, and ZigBee. Thus, this bridge makes computing devices have the ability of P2P communication and UBICOMP of existing terminals becomes possible with this extending bridge.

Currently, most SD controllers support standard SD interface on one side and 8-bit flash IO on the other side, such as the one introduced by inCOMM [3]. So these controllers are most used in memory storage card. iWave company shows an SDIO to UART bridge controller [4], which could convert data only from UART to SDIO interface. Although it supports SDIO Interrupt feature, SDIO aware host controller is needed in the mobile terminals, which is much different from the normal SD memory host controller in hardware and software. Altera company also gives a SD slave controller IP core [5], SD/SDIO/MMC Slave Controller, whose user interfaces are mainly master parallel interfaces, and the common interfaces, such as UART and SPI, are not included. In addition, microcontrollers would be added to these SD controllers to implement the flash translation layer (FTL) for NAND flash management.

Taking all the above considerations into account, the goal of our proposed SD slave controller is a platform-independent, interface-extensible and function (FTL)-complete bridge which could convert the data from the common interfaces of existing mobile terminal's transceiver to standard SD interface. This bridge contains an SD slave controller with a ping-pong FIFO in it and a multiple-target interface. The ping-pong FIFO increases the throughput of this bridge and the multiple-target interface is designed to extend interfaces for various electronic devices. Moreover, in order to implement the function of SD memory, a tidy NAND Flash controller, including a flash translation layer (FTL), is designed in hardware in the multiple-target interface. It will present better performance in integrality compared with FTL implemented in software in microcontroller. One application on ZigBee of this bridge is shown in Fig. 96.1, where an SD-ZigBee card is implemented with this bridge.

The architecture of the remainder of this paper is as follows. The next section gives a design consideration of our bridge, including the overview of hardware

**Fig. 96.1** Demonstration of SD-ZigBee card with our proposed bridge



architecture, the design of the SD slave controller and the multiple-target interface. Then the implementation of the bridge, performance analysis, and demonstration are presented in Sect. 96.3. Finally, conclusion is discussed in Sect. 96.4.

## 96.2 Design Consideration of Proposed Bridge

The architecture of proposed SD to multiple-target interface bridge, as shown in Fig. 96.2, consists of an SD card slave controller meeting SD memory card v2.0 specification [6] and a multiple-target interface for various electronic devices. In this section, structure and functions of main components will be described in details.

### 96.2.1 Platform Independent SD Card Slave Controller

The SD card slave controller is designed to comply with SD Physical Layer Specification Version 2.0. It supports the standard interface of 9-pin and is designed to operate at a maximum frequency of 50 MHz. The bus interfaces supported are SD 1-bit and 4-bit modes. The controller handles hot insertion, removal of card. As shown in the left side of Fig. 96.2, the controller consists of a command handler, a data handler, a card status state machine, and a ping-pong FIFO.

With this SD memory controller, any terminal that could access SD memory card is able to access the application cards based on the proposed bridge, such as SD-ZigBee and SD-WiFi, without any driver or hardware change. Compared with this, an SDIO device needs an SDIO aware host and a custom driver to support its operation. Therefore, this proposed bridge is platform independent and could be applied to various OS, such as Android, WinCE, and Symbian.

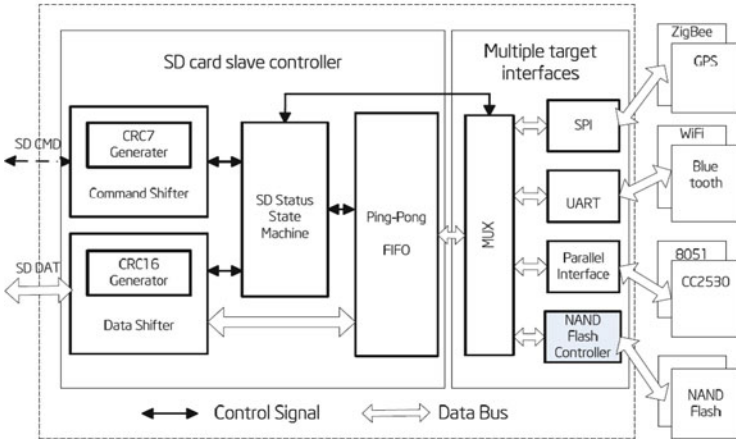


Fig. 96.2 Architecture of proposed SD to multiple targets bridge

96.2.1.1 Command Handler

The command handler receives the 48-bit command and transfers it from serial to parallel command. All commands are protected by cyclic redundancy check (CRC) bits. If the CRC check fails, the handler will not respond and the command will not be executed. Meanwhile, handler will not respond and change its status if an illegal command is received, such as a command that is not defined, or not supported by the card controller. The state machine diagram is shown in Fig. 96.3.

In the beginning, the handler is in “IDLE” state. If a command arrives, it transfers to “RAED CMD” state and starts to read and then checks the command. If the received command has a response, the handler transfers to “SEND RESPONSE” state and starts to send a response to host. Otherwise, it will return to “IDLE” state directly.

The CRC7 check is used for all commands, all response except type R3, and the CSD and CID registers. The CRC7 is a 7-bit value and is computed as follows:

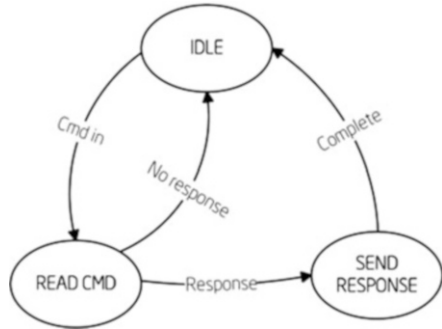
$$G(x) = x^7 + x^3 + 1. \text{ (Generator polynomial)}$$

$$M(x) = (\text{1st bit}) * x^n + (\text{2nd bit}) * x^{(n-1)} + \dots + (\text{last bit}) * x^0.$$

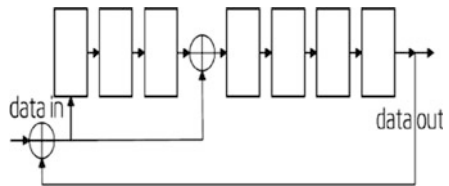
$$\text{CRC}[6 \dots 0] = \text{Remainder} [(M(x) * x^7) / G(x)].$$

As shown in Fig. 96.4, CRC7 Generator/Checker is composed of seven shift registers and two adders. Command or response is shifted into the circuit serially. When all the bits are shifted, the CRC7 check code is presented from the seven registers.

**Fig. 96.3** Command state machine diagram



**Fig. 96.4** CRC7 generator/checker



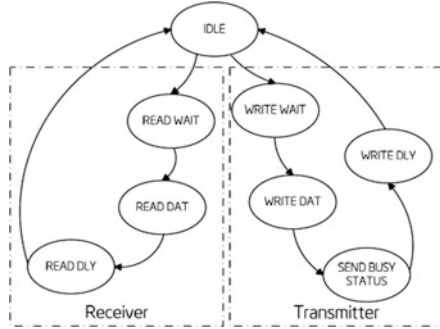
**96.2.1.2 Data Handler**

The data handler handles data transactions in the controller, including a transmitter and a receiver. If a data read command is received, the receiver will get data from FIFO, convert them into serial with additional CRC check and send it to SD host. If a data write command is received, the handler will receive serial data from SD interface and store them in the data buffer for further processing. If all write buffers are full, and as long as the card is in Programming State, transmitter will send busy status to host to indicate host the slave card is busy. The state diagram of data handler is shown in Fig. 96.5. Data access also needs CRC check and it uses a 16-bit CRC generator and checker for a more accurate data check. The CRC16 circuit is similar to CRC7 module in command handler.

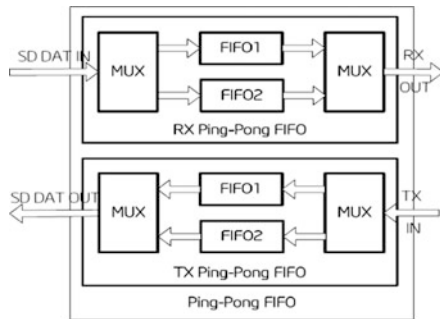
**96.2.1.3 SD Card State Machine**

The card state machine handles the card state described in SD 2.0 Physical specification. It includes Inactive State, Idle State, Ready State, Identification State, Stand-By State, Transfer State, Sending-Data State, Receiving-Data State, Programming State, and Disconnect State, where the state transitions depend on the received command.

**Fig. 96.5** Data state machine diagram



**Fig. 96.6** Ping-pong FIFO operation multiple-target interface

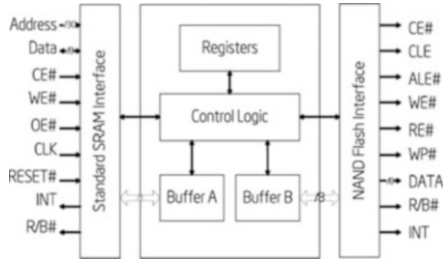


### 96.2.1.4 Ping-Pong FIFOs

The card slave controller has two different clock domains, one is SD clock whose frequency range is 0 to 50 MHz and the other is programming clock with maximal frequency 50 MHz. To handle the data stream of different clock domains, we use ping-pong FIFOs [7]. The treatment scheme of ping-pong operation is as below: The input data stream is divided into two FIFOs via input data multiplexer (MUX). In the first buffer cycle, the input data fills in FIFO1; In the second buffer cycle, the input data stream fills in FIFO2, meanwhile, the data stored in FIFO1 in the first buffer cycle outputs to processing unit via an output data multiplexer. In the third buffer cycle, the input FIFO and output FIFO are exchanged. The cycle changes like this flow and goes round and round. With the rhythmical switching of input data MUX and output data MUX, the data via buffer could be sent to processing unit without a pause. The diagram of ping-pong FIFO is shown in Fig. 96.6.



**Fig. 96.7** NAND flash controller functional block diagram



## 96.2.2 Multiple-Target Interface

### 96.2.2.1 Target Selectable Interface

In order to provide flexibility for implementation of applications that require the conversion of data to feed SD bus, a multiple-target interface is proposed, including SPI, UART, parallel interface, and NAND Flash interface. As shown in the right part of Fig. 96.2, this multiple-target interface module is composed of an SPI host controller, a UART controller, a standard Parallel interface, a tidy NAND Flash interface, and a multiplexer. With the MUX, one of the target interfaces is selected at a time. We allocate a different address for each target interface. Generally, we allocate one sector (512 bytes) for each non-memory target such as ZigBee device [8], WiFi device, and blue tooth device. Through analyzing the address from SD memory host, this module decides which target to access.

### 96.2.2.2 Tidy Flash Controller Including FTL

The NAND Flash controller comprises a control logic module with two buffers and two different interfaces. One side of the tidy NAND Flash controller is a standard SRAM interface so that SD controller could access NAND flash as a SRAM conveniently, the other side is a standardized NAND Flash device interface which is compliant to the open NAND Flash interface specification (ONFi). The function block diagram is shown in Fig. 96.7.

Since there is not an MCU in our bridge, the bad block management and address mapping is performed in hardware in the control logic module. Using a RAM, logic address to physic mapping is implemented [9], which is shown in Fig. 96.8. A block-level mapping method is adopted for the sake of memory size. The selected NAND flash has totally 8,192 physical blocks, where 7,680 blocks are mapped to user data area and the remaining 512 blocks are mapped to free pool for bad block management. The bad block mark and logic block number (LBN) are written in the

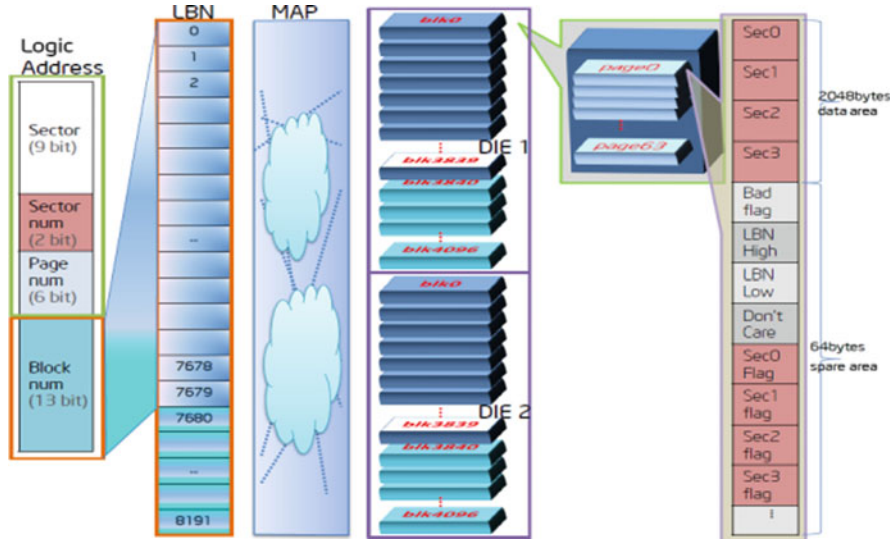
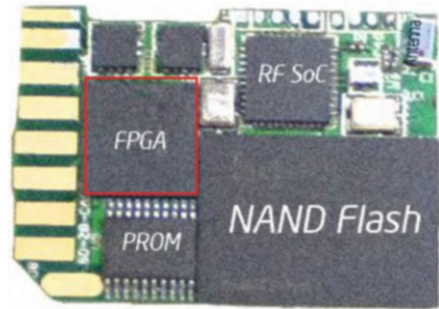


Fig. 96.8 The address map of NAND flash

Fig. 96.9 Circuit implementation of verification system



extra space of pages of every physical block. Thus, this proposed bridge could be adopted to implement the storage function without extra MCU and presents good performance in integrity.

### 96.3 Implementation and Analysis

#### 96.3.1 Implementation

This bridge has been prototyped on an FPGA-based verification system [10], which consists of a Xilinx Spartan-3E FPGA, a Hynix 8 Gb NAND Flash, and a CC2530 ZigBee SoC. The circuit implementation of the verification system is shown below (Fig. 96.9).

**Table 96.1** Comparison between the proposed bridge and other controllers

Interface	IN371AC [3]	iW-SDIO slave controller [4]	EP560 [5]	SDZ537 [11]	Proposed bridge
SD	√	√	√	√	√
SDIO			√	√	
UART		√		√	√
SPI					√
Parallel master interface			√		√
Flash IO	√				√

The entire bridge architecture has been FPGA verified and implemented. It also walked through a 0.18 μm CMOS technology and would be taped out. Synthesis experiments show our bridge reaches 50 MHz with the implementation size of 10 mm<sup>2</sup>. The equivalent gate count is about 80,000 and the total memory is 47 Kb.

### 96.3.2 Performance Analysis

Verification on this subsystem is based on simulation except that the wireless communication interface is verified on FPGA. The programming frequency is set at 50 MHz and SD clock frequency is set at 50 MHz. Interface-scalability compared with that of other controllers is shown in Table 96.1.

The results in Table 96.1 give comparisons between our proposed bridge and other SD controllers. The proposed bridge could support UART, SPI, and parallel and NAND flash IO for connection to various target devices, while others can just support part interface. Therefore, it shows superior performance in interface-scalability.

### 96.3.3 Demonstration

As Fig. 96.1 shows, with an FPGA implemented bridge, a ZigBee SoC CC2530 and other basic components, an SD-ZigBee card is made. Communication interface between our bridge and CC2530 is parallel interface. With other four sensor nodes (A, B, C, D), we could establish a wireless sensor network. Via the general SD memory interface, a portable tablet with an upper machine application software could read environment information instantaneously, which could be recorded in magnanimous NAND Flash storage at the same time. Some comparisons are shown in Table 96.2 below.

**Table 96.2** Comparisons between SD-ZigBee card and SDZ 537

Parameter	SD-ZigBee card (proposed bridge)	SDZ 537 [11]
Embedded MCU	×	√
SDIO aware host	×	√
Custom drivers	×	√
Power voltage	2.7–3.6 V	2.8–3.6 V
Operation current	40–50 mA (max)	24–29mA (average)
Communication distance	40 m	30–100 m
Data rate	250 kbps	250 kbps

From the results shown in the Table 96.2, for the approximate power (Voltage and Current) and performance (Distance and Data Rate), an application card equipped with this bridge could implement the NAND flash storage function without extra MCU and an SD memory host controller could access it without any hardware (SDIO aware) or driver change. The integrality and platform independent features would reduce the difficulty and period of system design. Other SD-based communication devices could be implemented easily with our proposed bridge such as SD-WiFi card and SD-Blue Tooth card.

## 96.4 Conclusion

This paper presented an SD interface to multiple-target interface bridge. The bridge was designed to adapt to different electronic devices and could extend the function of existing ubiquitous computing devices. A tidy NAND Flash controller was also implemented in hardware so that an 8 Gb NAND Flash could be connected directly to build up an SD memory card. The implementation and demonstration showed that it is very convenient to design various SD-based communication devices using our proposed bridge, and thus many existing terminals could be extended with P2P communication functions, which will make the UBICOMP of existing electronic device possible. Experiment reveals that the proposed bridge presents good performance in platform-independent, interface-scalability and integrality.

## References

1. Lyytinen, K., & Yoo, Y. (2002). Ubiquitous computing. *Communications of the ACM*, 45(12), 63.
2. Bruda, S. D., et al. (2012). A peer-to-peer architecture for remote service discovery. *Procedia Computer Science*, 10, 976–983.
3. inCOMM Technologies Co. Ltd. (2009). Flash card controller.
4. iWave Company. (2012). *iW-SDIO slave controller*. SDIO slave controller datasheet, Rel. 1.4.
5. Eureka Technology Inc. (2009). Ep560 datasheet.
6. S. C. Association. (2006). SD specification, version 2.0.

7. Huang, P., He, H., & Xu, D. (2008). Asymmetric asynchronous FIFO design in navigation receiver. *Journal of Projectiles, Rockets, Missiles and Guidance*, 1, 77.
8. Z. Alliance. (2005). *Zigbee specification*. ZigBee document 053474r06, version 1.
9. Kim, J., Kim, J., Noh, S., Min, S., & Cho, Y. (2002). A space-efficient flash translation layer for compact flash systems. *IEEE Transactions on Consumer Electronics*, 48(2), 366–375.
10. Dajiang Liu, Shouyi Yin, Jianfeng Chen, Hui Gao, Shaojun Wei. (2011). A portable environmental monitoring system based on WSN for off-the-shelf sensors. *International Conference on Computational Problem-Solving (ICCP)*, 2011 (pp. 569–572).
11. L. Spectec Computer Co. (2009). SDZ-537 microSD ZigBee card datasheet.

# Chapter 97

## Cloud Storage Management Technology for Small File Based on Two-Dimensional Packing Algorithm

Zhiyun Zheng, Shaofeng Zhao, Xingjin Zhang,  
Zhenfei Wang, and Liping Lu

**Abstract** In order to improve storage efficiency of small files in the cloud storage systems based on HDFS (Hadoop Distributed File System), this paper proposed a merging process approach based on a two-dimensional packing algorithm, called TDPHDFS (two-dimensional packing for HDFS). In it the correlations between file size and arrival time are comprehensively considered to assist the small files to be merged into large ones. The simulation results demonstrate that the storage efficiency of small files is improved, while the stability remains the same, yet less resource is consumed. The TDPHDFS algorithm can effectively reduce the performance penalty in both storage space and memory consuming while managing massive small files.

### 97.1 Introduction

HDFS (Hadoop Distributed File System) is a distributed file system model with highly fault-tolerant performance [1], which can be deployed on ordinary machine or virtual machine. It supports Java runtime environment and provides high-throughput data access. The metadata of the file system is placed in memory by NameNode [2]; if there are a large number of small files, the system will undoubtedly reduce the storage efficiency and storage capacity of the entire storage system. Therefore, how to solve the problem of small file storage efficiency has been researched as a popular topic of cloud storage.

---

Z. Zheng • S. Zhao • X. Zhang • Z. Wang (✉)  
School of Information Engineering, ZhengZhou University, ZhengZhou 450001, China  
e-mail: [iezfwang@zzu.edu.cn](mailto:iezfwang@zzu.edu.cn)

L. Lu  
Department of Information and Engineering, Henan College of Finance and Taxation,  
ZhengZhou 450001, China

This paper is focused on merging small files before they were uploaded to HDFS, and it proposes a method of merging small files by draw packing algorithm. The paper is organized as follows. In Sect. 2 we describe the relevant aspects of HDFS. In Sect. 3 we give an algorithm to merge small files. In Sect. 4 we describe our experiments and evaluation, and in Sect. 5 we present conclusions and future works.

## 97.2 HDFS Basic Framework and Related Research

A typical HDFS cluster is composed of a single NameNode and multiple DataNodes. The NameNode main function is maintaining file system namespace and managing all files and directories in the file system tree and the whole tree [3].

As the file system work node, the DataNode stores the actual data block, retrieves data blocks according to the need, and responds to the command which includes create, delete, and copy the data block.

The NameNode is the single point and stores the metadata information [4]. When the number of small files increases to a certain extent, the NameNode resource consumption will become a bottleneck of the system performance.

Currently, the mainstream idea of research is consolidating or combining small files into large files; there are usually two methods to be used: one is using Hadoop Archive (HAR) technology to merge small files (file size is less than 10 MB) [5]; the other is approaching a certain method of the small files' combination for specific applications.

The research results by using archiving technique to process small files are as follows: Mackey takes advantage of HAR technology to achieve the merging of small files, thereby improving the efficiency of HDFS metadata storage [6]. Yu Si and Gui Xiaolin uses the SequenceFile to combine small files, integrate the multi-attribute decision theory and experiment, and then obtain an optimal file merging scheme [7].

The research results by approaching a certain method for specific application are as follows: Liu combines small files into large files by using the WebGIS access mode features and establishes a global index for them; the storage efficiency of small files is improved [8]. Liu L optimizes the concurrent access of small files in distributed storage system [9].

## 97.3 Small Files Merging Algorithm

In order to solve the problem of low efficiency for store small files in HDFS, this paper proposes TDPHDFS (two-dimensional packing for HDFS) algorithm during pre-merging small files with two stages. In the first stage, sort these files and generate the optimal solution file sequence. In the second stage, use file stream to merge small files to ensure the efficiency of the cost of the merger, and then submit the merged file to the DataNode by client.

### 97.3.1 Two-Dimensional Packing Problem

Bin-packing (BP) problem [10] is a combinatorial optimization problem having very strong application background, and its solution is extremely difficult. In it, objects of different volumes must be packed into a finite number of bins or containers each of volume in a way that minimizes the number of bins used [11]. Treat each block of HDFS as a box, and small file waiting to be merged as the object; drawing on the idea of bin-packing algorithm can maximize the utilization of the HDFS storage space.

One representative part of the two-dimensional packing problem is 2 BP (two-dimensional bin-packing problem) [12], described as follows: given  $n$  rectangular items,  $F = \{1, 2, \dots, n\}$  and the infinite plurality of rectangular boxes with same size. The width of the  $i$ th item is  $w_i$ , and the height is  $h_i$  ( $i = 1, 2, \dots, n$ ). The rectangular box's width is  $W$ , and the height is  $H$ . All items are put into the box with the least number.

### 97.3.2 TDPHDFS Algorithm

With the idea of dynamic delay, TDPHDFS does not require users to upload small files immediately, only packing when there are enough (total file size is greater than a threshold value such as 64 MB) small files in the buffer. Considering the waiting time of the file in the buffer, TDPHDFS increases the file's priority level as the waiting time increases at speed  $A$  (set 2).

TDPHDFS proposes the following definition: there is only one box and the size of box is set to 63 MB; 1 MB space is reserved as the compressed file index information.

For case of description, Table 97.1 describes the parameters used in TDPHDFS. The TDPHDFS algorithm's description is as follows:

1. Initialization process defines a two-dimensional array named FileArray[MD5][key]; the key is initialized as file's size.
2. Read the size of all files in buffer; update the FileArray[MD5][key]; if the array already exists, the same MD5 value file, the corresponding weight is multiplied by 2.
3. Sort the file by weights from the largest to smallest. If total weights are greater than 63MB,  $\sum_{i=1}^n \text{FileArray}[i][2] \geq 63M$ , start packing or else exit the program, and wait for the next packing scheduling.
4. Begin packing from the maximum weight of the file. If file  $f_i$  is boxed and the total size of the files which are boxed already is less than 63 MB, then store  $f_{i+1}$  and delete the  $f_i$  records from the array. Loop step 4.



**Table 97.1** The parameters used in TDPHDFS

Symbol	Description
$f_i$	The $i$ th file to packing
$F$	A collection on files in buffer $F = \{f_1, f_2, \dots, f_n\}$
$n$	The number of files in buffer
$a$	The rate change of file priority, set 2
FileArray[MD5][key]	File information; the first column indicates the MD5 value of the file, and the second column indicates the weight of the file
Key	The weights of the file
Flevel	The priority of the file
$y$	The file number waiting for merge

5. Calculate the actual size of the files in the box; if it is less than 62 MB (packing threshold), go to step 2.
6. Calculate the number of files to be merged, and set  $y$ . Generate the index of ranked files, and attach the index to the post-merger file; combine them.
7. Upload merged file, and go to step 2.

In the fifth step, the reason for TDPHDFS finding and packing files again is because in step 2, there will be some small files that are not loaded while weights are increased. After priority loading these small files, there will be some virtual space occupied. So it is necessary to confirm the actual size after packing. The best time complexity of the proposed algorithms is  $T(n) = O(n \times \log n)$ , and the worst time complexity is  $T(n) = O(n^2)$ .

## 97.4 Simulation and Result Analysis

### 97.4.1 The Configuration of Simulation

In order to verify the algorithm performance on improving the storage utilization and reducing resource consumption, the following experiments are designed, detailed experimental environment configuration is shown in Table 97.2.

### 97.4.2 Simulation Analysis

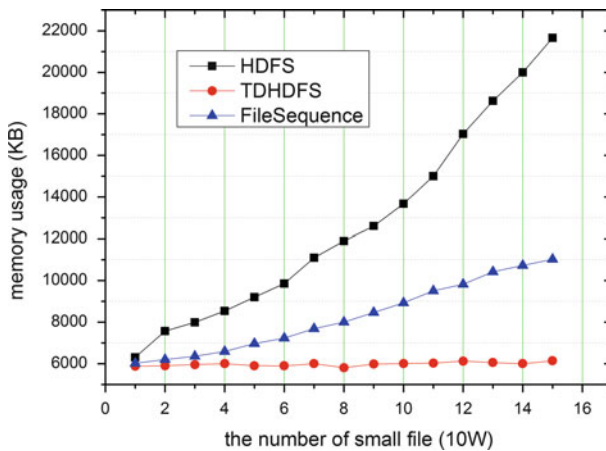
In the experiment, replication is set to 1; when all the files are uploaded to HDFS, use “du” command to show the size of all files in HDFS. The results are shown in Table 97.3. When we use DS1, the total size of files in HDFS is 488 G (upload directly), which is over 200 times the space cost than when using TDPHDFS. The DS2 files’ space cost is almost a differential of 60–1.

**Table 97.2** Experimental environment

Components	Description
Hadoop version	0.2.3
NameNode number	1
DataNode number	7
OS	Ubuntu 10.04
CPU	NameNode: Intel I3-2130 3.4 G DataNode: Intel Core Duo T2410 2.0 G
Memory	NameNode, 4 G; DataNode, 2 G
Net situation	LAN

**Table 97.3** The total size of files after upload

Dataset	Description	Upload way	Total size
DS1	8571 files, total size 1.49 G	Use TDPHDFS	2.3 G
		Directly	488 G
DS2	1000 1MB files, total size 1 GB	Use TDPHDFS	1.1 G
		Directly	62.5 G

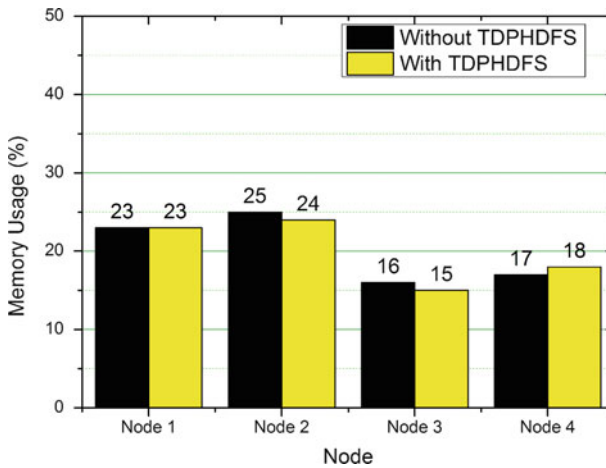
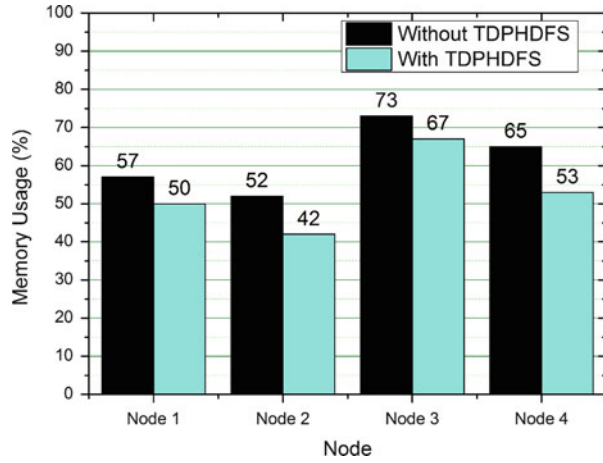


**Fig. 97.1** Memory usage in NameNode with three methods

As it is shown in Fig. 97.1, when the number of small files increased, the memory usage in NameNode increases, and compared with the FileSequence, the TDPHDFS can reduce this consumption more effectively. As the number of small files increased, TDPHDFS can control the memory usage cost under a very low level.

A comparison of the memory usage in 4 DataNodes is shown in Figs. 97.2 and 97.3 with or without the use of TDPHDFS algorithm. The average memory

**Fig. 97.2** Memory usage comparisons during upload files



**Fig. 97.3** Memory usage comparisons while HDFS is idle

occupancy rate is 56 % when uploading files to HDFS, and if the server is idle, the average memory occupancy rate could be 21 %. By using TDPHDFS, the average memory occupancy rate could be 42 and 19.5 % separately. This test shows that with TDPHDFS, the pretreatment could reduce the consumption of system resources.

## 97.5 Conclusion

In this paper, we have analyzed the HDFS basic framework and proposed a TDPHDFS algorithm to solve low storage performance when storing small files in HDFS. This algorithm uses improved packing algorithm to merge small files to achieve the maximum usage of HDFS' default block. There is a greater improvement in time and space to store small files with TDPHDFS compared with the origin HDFS.

## References

1. Borthakur. *HDFS architecture guide*. Hadoop Apache Project. Retrieved from [http://hadoop.apache.org/common/docs/current/hdfs\\_design.pdf](http://hadoop.apache.org/common/docs/current/hdfs_design.pdf).
2. Shengjun Xue, Wu-Bin Pan, & Wei Fang. (2012). A novel approach in improving I/O performance of small meteorological files on HDFS. *Applied Mechanics and Materials* 17 (11), 1759–1765.
3. Ming Chen, Wei Chen, Likun Liu, & Zheng Zhang. (2007). An analytical framework and its applications for studying brick storage reliability. *IEEE*, 7(26), 242–252.
4. Zhanjie Wang, & Lijun Zhang. (2012). Mix-P2P architecture of distributed storage system based on HDFS. *Advanced Materials Research* 382(7), 92–95.
5. Shvachko, K., Huang, H., Radia, S., & Chansler, R. (2010). The hadoop distributed file system. *MSST, IEEE 26th Symposium*, 4(1), 1–10.
6. Sehrish, M. S., & Wang, J. (2009). Improving metadata management for small files in HDFS. In *Cluster computing and workshops, 2009* (pp. 1–4). CLUSTER'09. IEEE International Conference, IEEE.
7. Yu, S., Gui, X., Huang, R., & Zhuang, W. (2011). Improving the storage efficiency of small files in cloud storage. *Journal of Xi'an Jiaotong University.*, 10(1109), 65–72.
8. Yang, C. T., Huang, K. L., Liu, J. C., Chen, W. S., Hsu, W. H. (2012). On construction of cloud IaaS using KVM and open nebula for video services. In *Parallel processing workshops (ICPPW), 2012 41st International Conference*, IEEE.
9. Thusoo, A., Sarma, J. S., Jain, N., Shao, Z., Chakka, P., Anthony, S., et al. (2009). Hive: A warehousing solution over a map-reduce framework. *Proceedings of the VLDB Endowment*, 2 (2), 1626–1629.
10. Lodi, A., Martello, S., & Monaci, M. (2002). Two-dimensional packing problems: A survey. *European Journal of Operational Research*, 141(2), 241–252.
11. Côté, J. F., Gendreau, M., & Potvin, J. Y. (2013). An exact algorithm for the two-dimensional orthogonal packing problem with unloading constraints. *Interuniversitaire Research Centre on Enterprise Networks*, 6(1), 1–32.
12. Andrea Lodi, Silvano Martello, Michele Monaci, Claudio Cicconetti, Luciano Lenzini, Enzo Mingozzi et al. (2011). Efficient two-dimensional packing algorithms for mobile WiMAX. *Management Science*, 57(12), 2130–2144

# Chapter 98

## Advertising Media Selection and Delivery Decision-Making Using Influence Diagram

Xiaoxuan Hu and Fan Jiang

**Abstract** The influence diagram (ID) is introduced into advertising media selection and delivery strategy making by reducing uncertainty in the process of decision-making. This paper conducts a survey and selects relevant variables including product category, advertising budget, target audience, media selection, authority, and coverage. The topology layer of the ID model is constructed by distinguishing the causal relationship among variables, and the parameter layer is defined through the judgment of conditional probability. Empirical results show that scientific assessments of the various expected utility values in the decision-making program are put by probabilistic reasoning. Based on it, the larger profit can be obtained under a smaller cost with the principle of expected maximization. Therefore, the model does an effective job and provides reference for decision-makers.

### 98.1 Introduction

The huge role of advertising is more and more valued by many enterprises as an important part of corporate marketing activities. At the same time, the effect of advertising is declining seriously because many audiences respond to it with intensive indifference and resistance. Thus, how to select the suitable media and make a right decision are the primary decision-making problems in the corporate advertising works. However, due to the presence of many uncertainties, it greatly increases the risk and difficulty of decision-making. Therefore, the use of appropriate tools and methods, the quantitative analysis for uncertainties, and having scientific assessment of the various decision-making programs are all important issues.

---

X. Hu • F. Jiang (✉)

School of Management, Hefei University of Technology, Hefei 230009, China  
e-mail: [huxiaoxuan@vip.sina.com](mailto:huxiaoxuan@vip.sina.com); [caocaoyatou66@163.com](mailto:caocaoyatou66@163.com)

By now, a large number of researchers do qualitative work to have the advertising media selection and decision-making from the point of view of expert knowledge and experience. Very small part does quantitative analyzing to get optimal strategy, such as, mathematical programming, multi-attribute evaluation method. For example, we can use the dual simplex method of linear programming problem to select the best combination of advertising media. Buratto made use of the linear programming model to obtain the maximum profit media solutions and did research about the advertising channel selection in the market segments [1]. Viscolani applied nonlinear programming method to solve the general advertising decisions assuming that each market segment demand function was linear and the function of media advertising cost was quadratic [2]. Jha adopted the same objective programming method to maximize the total profit in all markets, taking into account the different characteristics of different markets [3]. Ngai got the best website for online advertising by making use of AHP [4]. Other researchers applied constrained goal programming model, multiple criteria decision-making model, or double goal programming model to solve the problem of advertising decision-making [5–7]. In a word, these methods are certain specific solutions to advertising decision-making, however, when facing other similar scenarios, it's necessary to re-plan all processes, thus, resulting in large amount of calculation and failure to do the universal advertising decision-making.

In this paper, we will introduce influence diagram (ID) to construct the model of advertising media selection and delivery strategy, analyze each factor which will effect decision-making qualitatively and quantitatively using powerful ability to describe problems and inference probability, and ultimately, put a sequence for various decision-making programs by the calculation of ID, thus providing a reference for policy makers.

## 98.2 Decision-Making Modeling

Professor Howard proposed the graph model based on uncertain information expression and solved complex decision problems in 1984 [8], consisting of the dependencies between variables, conditional independence relationships, and decision-makers' preference information. The theory and application are developing more and more rapidly so that it could build related model for any statistical analysis problem [9–11]. The generation process of ID is a hierarchical process and the specific process is shown in Fig. 98.1.

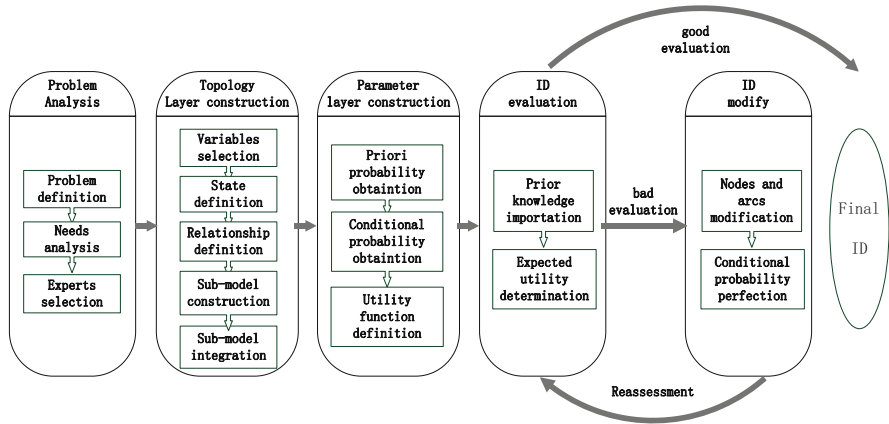


Fig. 98.1 The generation process of ID

### 98.2.1 Selection of Variables and the Corresponding State Space

To build the graph of ID, determining all variables related is key, usually finished by the knowledge and expertise of experts. There are a lot of variables during the process of corporate advertising media selection and delivery; however, not every variable has an important influence on the final decision-making. So we will conduct a questionnaire survey, discriminate the importance among variables, and delete some irrelevant variables. Then, we distribute these questionnaires to experts in the field and ask them to give their points ranging 1–5, “1” representing unimportance of variables, while “5” representing a large extent. Ultimately, the associated variables can be shown in Table 98.1.

### 98.2.2 Construction of Topology Layer for Advertising ID

According to the variables previously obtained, we begin to build the topology layer based on consultancies of the experts. In order to reflect accurately the conditional relationship among these variables, we repeatedly ask the experts, such as persons of experience in the advertising industry, the media people, and business owners. The specific information of questionnaire can be seen in Fig. 98.2.

Based on the above questionnaire, we could draw influence relationships among variables properly, resulting in two network diagrams: media selection and delivery strategy, as shown in Figs. 98.3 and 98.4, respectively.

**Table 98.1** The state spaces of each variable

Variables	State space	Remark
Product category	Hotel, food, clothing, tobacco and alcohol, computer	For many categories of all kinds of products on the market, it's complex to be covered in the "Product Category" variable, just to pick a few representative kinds
Advertising budget	Low	Deterministic variable with the lowest cost, hoping to get a larger income by the lowest budget
Target audience	Old man, middle-aged, youth	It's a simple division in order to obtain the conditional probability, assuming that they have the purchasing power
Product life cycle	Introduction, growth, maturity	The recession is not the introduction of state space because of a sharp decline in product sales in this period, and the profit will fall significantly so that companies no longer put into advertising the products, but consider what strategy out of the market to adopt as soon as possible
Media selection	TV, radio, Internet, outdoor, newspaper	It's a decision node
Delivery strategy	Centralized, continuous, intermittent	It's a decision node
Authority	High, middle, low	It's a random variable, indicating that the type of media has a different degree of authority
Coverage	High, middle, low	It's a random variable, representing that the type of media has a different degree of coverage
Audience awareness	High, middle, low	It's a random variable. The probability value would change with the difference among delivery strategies in a certain stage for different products
Audience preference	High, middle, low	Different delivery strategies will lead to different audiences' preference
Expected sub-income 1	High, middle, low	It's influenced by the target audience, audience preference, and the corresponding coverage
Expected sub-income 2	High, middle, low	Both audience awareness and audience preferences will affect its conditional probability value
Total revenue	High, middle, low	The final state will be decided by two sub-incomes
Media costs	High, low	The media costs are different due to the diversity type of media
Delivery costs	High, low	Choosing different delivery strategies will result in different delivery costs
Total costs	High, low	Both the media costs and delivery costs determine its value



Q1: What are the key factors (ie, a priori node) in the choice of media clearly?  
 A1: Only clearing out product categories, budget of advertising, target audience and product life cycle, we could be able to select the media category.  
 Q2: After selecting the established media, which factors will impact the decision-making program?  
 A2: Based on historical experience and data, we can mainly divide into the following several aspects—media authority, media coverage and media costs. At the same time, the type of audience will affect the final income.  
 Q3: What criteria should be determined when evaluating the result of decision-making?  
 A3: In order to facilitate our research, the economic benefits of advertising could be considered, such as, the expected revenue and cost.

Fig. 98.2 A simple questionnaire

Fig. 98.3 Media selection module

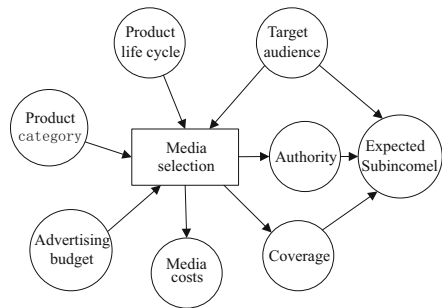
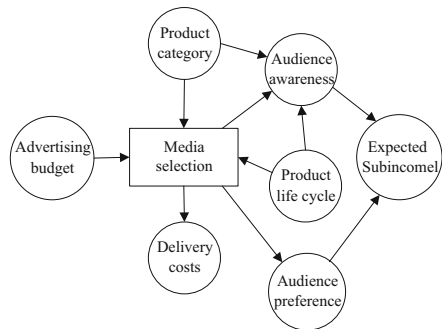


Fig. 98.4 Delivery strategy module



The overall model should be built by the joint of some intermediate nodes, connecting the main and common nodes in two sub-modules, such as *Total income* and *Total cost*. Thus, the integrated ID could be built on the basis of sorting out the logical relationship among variables, as shown in Fig. 98.5.

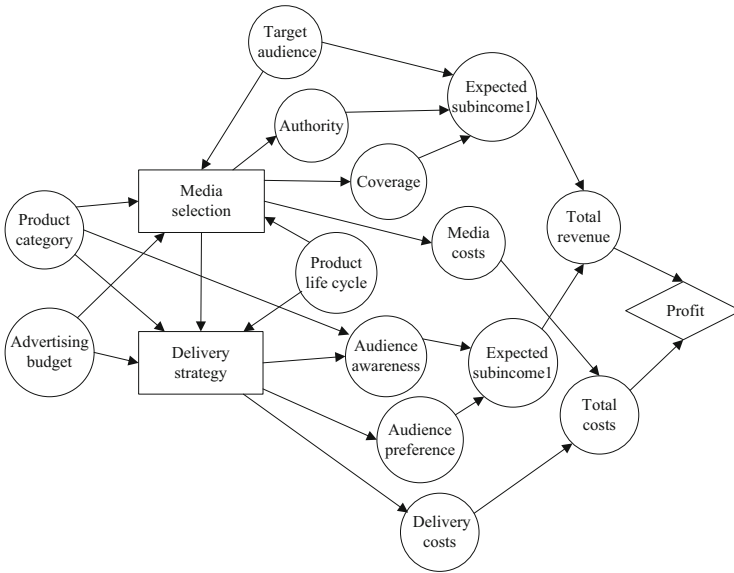


Fig. 98.5 The integrated ID

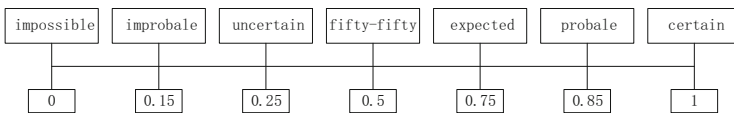


Fig. 98.6 The probability assessment scale

### 98.2.3 Construction of Parameter Layer for Advertising ID

The next step is to quantify conditional relationship among all variables, using the knowledge and experience of experts and applying historical statistic data usually. Since there exists some difference between different enterprises, and different types of data are not homogeneous, so it is difficult to obtain accurate a priori probability and conditional probability in modeling. Therefore, the conditional probability is solved by using the first method, consulting experts in the field. We will employ the easiest but most effective way for description of the probability—the probability assessment scale [12]—as shown in Fig. 98.6.

The key of building ID properly is to establish an appropriate function. Accordingly, we will also invite experts in the field and some experienced staffs to describe functions. In order to be able to get the opinions of experts, we will make the conditional probabilities in the ID into questionnaire, completed by all participating

**Table 98.2** The conditional probabilities of “Authority”

Media	Authority		
	High	Middle	Low
TV	0.7	0.2	0.1
Radio	0.3	0.4	0.3
Internet	0.3	0.4	0.3
Outdoor	0.5	0.3	0.2
Newspaper	0.6	0.3	0.1

experts; finally, all assignments will be updated and stated averagely so that the conditional probability tables are obtained, as seen in Table 98.2, and the others are similar to Table 98.2.

For whole ID, since the date of some input node is difficult to get, we temporarily simplify the processing to have the same probability for each state of all input nodes, when determining some a priori probabilities.

### 98.3 Evaluation of Advertising ID

Nowadays, the research of ID has been quite mature and the algorithms are also very complete. A number of software have been developed to solve the problem of the calculation and optimization of ID in foreign countries, such as Analytica, Netica, Pulcinella, Hugin, BayesialLab, Smile, SmileX, and GeNIe. In this paper, the GeNIe2.0 was applied to help us solve some calculation, which is powerful and has a wealth of graphic elements and a variety of decision-making tools and user-friendly interface.

The integrated ID is derived with some a priori probabilities and conditional probabilities based on the above questionnaire, applying GeNIe2.0 software. The next step is to verify the ID. For example, there is a kind of wine, which is a mature brand and whose target audience is middle-aged person, then we input some priori information based on it—target audience, product category and product cycle and update the ID, thus, the expected utility values are as shown in Table 98.3.

From Table 98.3, the largest expected utility *EU* equals 36.697281, meaning that the enterprise should select the program of TV and intermittent delivery strategy when there is the smallest advertising budget. By entering different data, the corresponding maximum expected utility values are obtained by virtue of the ID, whose classification results are consistent with the actual situation, reflecting the excellent decision-making function. In the process of choosing kinds of advertising media by actual corporate, we usually try to combine with other media when there is sufficient funds in order to bring greater profits. Thus, it is possible to sort the expected utility values through the ID, and further work may also be needed to select the programs with larger utility values and to study the specific advertising media composition problem.

**Table 98.3** Expected utility values

	Centralized	Continuous	Intermittent
TV	26.593488	31.972454	36.697281
Radio	21.706331	27.103728	24.704282
Internet	25.752049	31.811325	30.957273
Outdoor	23.511136	28.949377	29.166374
Newspaper	27.636540	34.238384	32.237819

## 98.4 Conclusion

Advertising media selection and delivery strategy making are important ways to promote products in a corporation. This paper proposes the model of ID to analyze and calculate the process of decision-making, reducing complexity and uncertainty to a large extent. Applying this model has the following advantages: (1) representing intuitively the relationships among corresponding factors in the process of choosing advertising media and increasing data correlation and logic, overcoming some uncertainty in advertising program; (2) estimating quantitatively the relationships by virtue of the form of conditional probability; (3) reducing some subjectivity and macro considerations comparing to some enterprise experience decision-making and avoiding prejudices made by some leaders; and (4) making the decision more convenient and more realistic with collection of two-stage decision-making process in the ID model. Therefore, our results indicate that the approach is a promising one.

**Acknowledgements** This research is supported by the National Natural Science Foundation of China (No. 71071045, 71131002, 71001032) and Humanities and Social Science Projects of Ministry of Education of China (13YJC630051).

## References

1. Buratto, A., Grosset, L., et al. (2006). Advertising channel selection in a segmented market [J]. *Automatica*, 42(8), 1343–1347.
2. Viscolani, B. (2009). Advertising decisions for a segmented market [J]. *Optimization*, 58(4), 469–477.
3. Jha, P. C., et al. (2011). Optimal media planning for multi-products in segmented market [J]. *Applied Mathematics and Computation*, 217(16), 6802–6818.
4. Ngai, E. W. T. (2003). Selection of web sites for online advertising using the AHP [J]. *Information & Management*, 40(4), 233–242.
5. Bhattacharya, U. K. (2009). A chance constraints goal programming model for the advertising planning problem [J]. *European Journal of Operational Research*, 192(2), 382–395.
6. Kwak, N. K., et al. (2005). An MCDM model for media selection in the dual consumer/industrial market [J]. *European Journal of Operational Research*, 166(1), 255–265.

7. Perez Gladish, B., et al. (2010). Planning a TV advertising campaign: A crisp multiobjective programming model from fuzzy basic data [J]. *Omega*, 38(1–2), 84–94.
8. Howard, R. A., & Matheson, J. E. (2005). Influence diagrams [J]. *Decision Analysis*, 2(3), 127–143.
9. Smith, J. Q. (1989). Influence diagrams for statistical modeling [J]. *The Annals of Statistics*, 17(2), 654–672.
10. Cobb, B. R., & Shenoy, P. P. (2008). Decision making with hybrid influence diagrams using mixtures of truncated exponentials [J]. *European Journal of Operational Research*, 186(1), 261–275.
11. Kjaerulff UB, Madsen AL (2013) Bayesian Networks and influence diagrams: A guide to construction and analysis [M]. *Springer* (pp. 70–95).
12. Witteman, C., & Renooij, S. (2003). Evaluation of a verbal–numerical probability scale [J]. *International Journal of Approximate Reasoning*, 33(2), 117–131.

# Chapter 99

## The Application of Trusted Computing Technology in the Cloud Security

Bo Li, Shenjuan Lv, Yongsheng Zhang, and Ming Tian

**Abstract** For the lack of safety and reliability of the information in the cloud computing environment, in order to create a more flexible and adaptable security mechanism, the combination of cloud computing and credible concept is a major research direction in today's security. Based on the view mentioned above, this paper strengthens the research of trust computing technology to solve the security issues in cloud and cloud-based trust transfer, on the basis of the practical work of the experts and scholars on the trust transfer technology, and expands the theoretical model of the trust chain. This paper uses the stochastic process algebra and Petri nets as a modeling tool to build two trust chain models, demonstrates the credibility of certain behavioral characteristics of the chain, analyzes several constraints of credible chain, and provides a valuable reference for engineering practice of the credible chain.

### 99.1 Introduction

Cloud computing obtains much of the industry's attention since its birth, and with the increasing globalization of the world economy and the development of information technology, cloud-based applications can be promoted in all areas of society [1]. At the same time, the advantages of cloud computing can reflect in all walks of life in society, from business to science to many key sectors of national security,

---

B. Li (✉)

Academic Affairs Office Shandong Polytechnic, Jinan 250104, China

e-mail: [lvshenjuan2011@163.com](mailto:lvshenjuan2011@163.com)

S. Lv • Y. Zhang • M. Tian

School of Information Science and Engineering, Shandong Normal University,

Jinan 250014, China

Shandong Provincial Key Laboratory for Novel Distributed Computer Software Technology,

Jinan 250014, China

and can see the trail of cloud computing. However, with the development of information technology, people give more importance on the information security and cloud data security, and privacy protection is generally considered to be an important aspect of cloud computing. So far, however, the customer does not have any means to prove the integrity and confidentiality of the applications and the upload of their own data, resulting in the issue of trust between the client and the cloud service provider, which will become the biggest obstacle to hinder the development of the cloud computing world. Currently, cloud computing technology is facing many security technology crises [2, 3], such as counterfeiting electronic signature, electronic signature repudiation, Trojan attacks, and viruses' damage, which seriously threatens the Internet cloud computing trust and cloud security. For a range of issues, the combination of trust computing technology and cloud computing and of the application of trust computing technology in the cloud security is the direction to solve this problem; therefore, it's important to strengthen the trust technology [4].

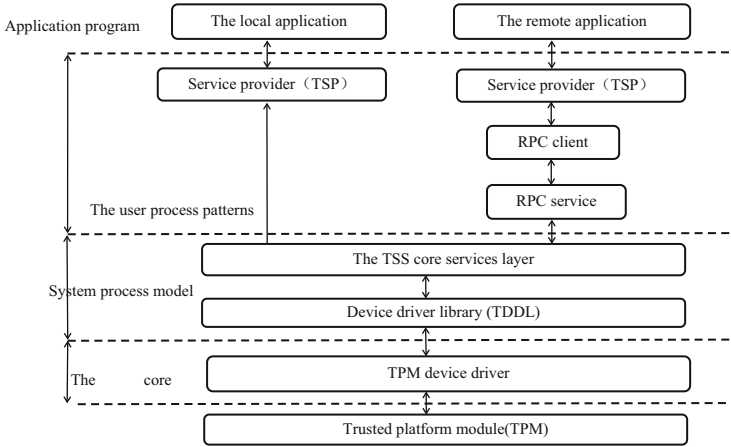
## 99.2 Related Works

In recent years, many experts and scholars use different theories and methods in cloud security issues and achieve certain results.

Today, the most active organization on cloud security is Cloud Security Alliance (CSA). CSA as an industry-recognized research on Cloud Security Alliance do more in-depth cloud security research. According to the CSA on cloud security research, other organizations of international numbers, such as CAM (Common Assurance Metric beyond the cloud) [5], Microsoft and Green League, and other institutions also do some research on the field of cloud security and put forward the framework of computing security and cloud security technology solutions research. But overall, the international research on cloud security issues is not deep enough. In connection with cloud security issues, this paper will begin with trust computing technology and strengthen the research of trust computing technology to solve the security issues in cloud and cloud-based trust transfer as the research object, at the conclusion on the basis of the practical work of the experts and scholars in the trust transfer technology; expand the theoretical model of the trust chain; analyze several constraints of credible chain; and provide a valuable reference for engineering practice of the credible chain [6].

## 99.3 Trusted Computing Technology

Nowadays, implementation of the trusted computing platform is trusted PC; the module TPM (Trusted Platform Module) is an SOC chip, which is the root of trust of the trusted computing platform. TSS (TCG Software Stack) is the TPM support



**Fig. 99.1** Trusted computing system. TPM: Trusted Platform Module, it is a security chip. OS: Operating system, it is the system software collection which manages hardware resources of computer, controls the other program's running and provides interactive operation interface for users

software of a trusted computing platform [7]; its main function is to provide the environment of application software development compatible with heterogeneous trusted platform module. TNC (Trusted Network Connect) ensures the integrity of the website visitor through the network to access, collect, and verify the requester's integrity information. We evaluate this information according to certain security policy and then decide whether to allow the requester and the network connection [8, 9]. Trusted computing system is shown in Fig. 99.1.

## 99.4 The Research of Trusted Computing Technology in Cloud Security

### 99.4.1 The Data Model

**Definition 1:** The activity of atomic component is a four-tuple as shown in Eq. 99.1:

$$Activity = (\alpha, I_p, r, t) \tag{99.1}$$

Among them,  $\alpha \in Act, p \in P$ , when component  $p$  executes behavior  $\alpha$ ,  $I_p$  expresses the integrity property;  $I_p = true$  represents component  $p$  is complete,



$I_p = false$  represents that component  $p$  is incomplete,  $r$  represents the integrity metric size of component  $p$ , and  $t$  represents the duration of the activity.

**Definition 2:** The operator refers to the relationship between atomic components, such as synchronization, concurrency, and competition, and structure composite components by operator. Operators are expressed as in formula 99.2:

$$P ::= (\alpha, I_p, r, t) \cdot P \mid P + Q \mid P_L Q \mid Nil \mid A \stackrel{def}{=} P \quad (99.2)$$

Among them,  $(\alpha, I_p, r, t) \cdot P$  expresses performing  $(\alpha, I_p, r, t)$  before operation  $P$ ;  $P + Q$  expresses performing activities of  $P$  and  $Q$ ;  $P_L Q$  expresses components  $p$  and  $Q$  performing together;  $L$  expresses communication between  $P$  and  $Q$ ;  $Nil$  expresses computing platform deep in a deadlock situation and stop the transfer of trust; at present, the platform is in an incredible state;  $A \stackrel{def}{=} P$  expresses performing activity  $P$  in a circle.

**Definition 3:** Define two trust transfer rules in trust transfer process as follows:

*Rule 1* The prefix operator transfer rules

$(\alpha, I_p, r, t) \cdot P$  occurred prefix operation and migrated to  $P$  after  $(\alpha, I_p, r, t)$  as shown in 3:

$$(\alpha, I_p, r, t) \cdot P \xrightarrow{(\alpha, I_p, r, t)} P \quad (99.3)$$

The state transition of trusted chain is shown in formula 99.4:

$$I_{TC} \rightarrow I_{TC} \oplus (I_p, r) \quad (99.4)$$

Among them,  $I_{TC}$  represents the state of trusted chain,  $I_{TC} = true$  expresses platform is in a credible state, and  $I_{TC} = false$  expresses the transmission of trust is stopped and platform is in an incredible state. The initial value of  $I_{TC}$  is true.  $\oplus$  expresses extending trusted chain with integrity metric size  $r$ . The calculation method is shown in formula 99.5:

$$I_{TC} = \begin{cases} I_{TC} & \text{and } I_p \text{ (under } r \text{)} \\ I_{TC} & \text{(otherwise)} \end{cases} \quad (99.5)$$

The trust transfer rule is shown in formula 99.6:

$$\frac{I_{TC} = true}{(\alpha, I_p, r, t) \cdot P \xrightarrow{(\alpha, I_p, r, t)} P} I_{TC} \rightarrow I_{TC} \oplus (I_p, r) \quad (99.6)$$

**Rule 2** Select operator transfer rules

$P + Q$  migrate to  $P'$  under the premise of  $P \xrightarrow{(\alpha, I_p, r, t)} P'$ . Its trust transfer rule is shown in 7:

$$\frac{I_{TC} = true \& P \xrightarrow{(\alpha, I_p, r, t)} P'}{P + Q \xrightarrow{(\alpha, I_p, r, t)}} P' I_{TC} \rightarrow I_{TC} \oplus (I_p, r) \quad (99.7)$$

Under the premise of  $Q \xrightarrow{(\alpha, I_Q, r, t)} Q'$ , we measure the integrity of  $Q$  with integrity metric size  $r$ ; its trust transfer rule is shown in formula 99.8:

$$\frac{I_{TC} = true \& Q \xrightarrow{(\alpha, I_Q, r, t)} Q'}{P + Q \xrightarrow{(\alpha, I_Q, r, t)}} Q' I_{TC} \rightarrow I_{TC} \oplus (I_Q, r) \quad (99.8)$$

## 99.4.2 Semantic Model Design

In accordance with the transmission process of trust in TPM, OS(Operating System) and the user mode process, this paper establishes semantic model of TPM, OS and the user mode process, after which, we link each model, making them become trusted chain model.

### 99.4.2.1 The Semantic Model of TPM

In the process of trust transfer, activities which need to be completed by TPM are extending and extended; if the two atomic behaviors are completed, then platform is in an idle state. The semantic model of TPM is expressed in formula 99.9:

$$TPM \stackrel{def}{=} (extending, I_{TPM}, r, t) \cdot (extended, I_{TPM}, r, t) \cdot (TPM\_idle, I_{TPM}, r, t) \cdot TPM \quad (99.9)$$

### 99.4.2.2 The Semantic Model of OS

OS is an important part in the process of trust transfer. First we should complete integrity measuring of OS and then extend the results to trusted chain by extending. Its semantic model is expressed in formula 99.10:

$$OS \stackrel{def}{=} (run, I_{OS}, r, t) \cdot (mean, I_{OS}, r, t) \cdot (extending, I_{OS}, r, t) \cdot (OS\_idle, I_{OS}, r, t) \cdot OS \quad (99.10)$$

### 99.4.2.3 The Semantic Model of the User Mode Process

The user mode process starts to run after it gains scheduling and finally gets into the waiting queue. The semantic model of the user mode process is expressed in formula 99.11:

$$U \stackrel{def}{=} (schedule, I_U, r, t) \cdot (run, I_U, r, t) \cdot (U\_idle, I_U, r, t) \cdot U \quad (99.11)$$

### 99.4.2.4 The Semantic Model of Trusted Chain

Based on the above semantic model, we can build the semantic model of trusted chain. It is expressed in formula 99.12 as follows:

$$TC \stackrel{def}{=} (U_1 U_2 \cdots U_n)_{\{schedule\}} OS_{\{extending\}} TPM \quad (99.12)$$

In this model, the interaction between user and OS is realized by schedule; the interaction between OS and TPM is realized by extending.

## 99.5 The Simulation Results and Analysis

In this paper, we compare the performance of the algorithm in two aspects through the simulation experiments:

1. The analysis of credibility's accuracy.
2. Efficiency analysis of computing reliability.

In the same situation, if the value of untrust is bigger, the confidence probability is smaller and the accuracy is lower; if the value of pratio is bigger, the shorter the trust transfer time and the higher the efficiency.

The experimental environment is Pentium(R) Dual-Core 2.1 GHZ CPU, 2 G memory, 500 G hard disk and Windows 7 ultimate operating system; experimental data analysis software environment is MATLAB 7.11.

If the value of untrust is not equal to zero, set a different value of pratio and untrust, comparing transmission efficiency and accuracy probability of credibility under different conditions. The experimental results are shown in Fig. 99.2.

From Fig. 99.2, we can find when values of pratio are equal; if the value of untrust is bigger, the less the number of completed activities are and the shorter trust transfer time is; when values of untrust are equal, if the value of pratio is bigger, the less the number of completed activities are, and the shorter trust transfer time is.

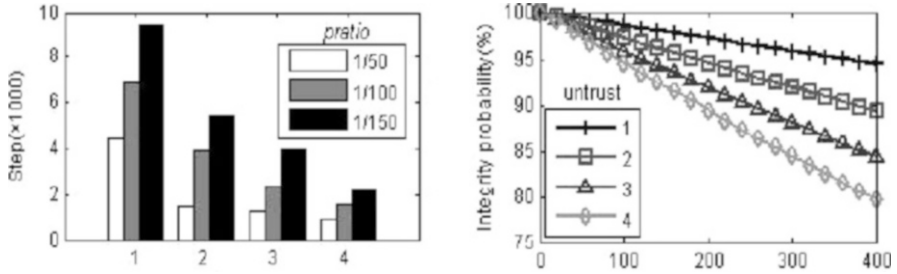


Fig. 99.2 The results of the experiment credibility efficiency and reliability accuracy experiment

### 99.6 Conclusion

This paper defined some transfer rules, and on the basis of trust, we designed a new semantic model for TPM, OS, the user mode processes and trusted chain; overcame the shortcomings of the original transfer rules and semantic model; and strengthened the research and application of trusted computing in the cloud security, solving some problems related to trust decision in the cloud security effectively. Simulation results showed that the model has certain advantages in the accuracy and efficiency of trust degree. In addition, the research work of cloud security based on trusted technology is still in its primary stage. For the direction, we also need to do a lot of research work.

**Acknowledgements** This research was supported by the Project of Shandong Province Higher Educational Science and Technology Program under Grant No. J12LN61. It was also supported by the Project of Shandong Province Higher Educational Science and Technology Program under Grant No. J13LN64. In addition, the authors would like to thank the reviewers for their valuable comments and suggestions.

### References

1. Xiao-yong Li, Xiao-lin Gui, & Qian Mao. (2009). The monitoring dynamic trust model based on the adaptive behavior. *Journal of Computer*, 32(04), 664–674 (In Chinese).
2. Deng-guo Feng, & Yu Qin. (2008). Trusted computing environment method research. *Journal of computer*, 31(09), 1640–1652(2008) (In Chinese).
3. Deng-guo Feng, & Yu Qin. (2010). A property proof protocol based on TCM. *China Science: Information Science*, 40(02), 189–199 (In Chinese).
4. Yu Qin, & Deng-guo Feng. (2009). Remote attestation based on the component properties. *Journal of Software* 20(06), 1625–1641 (In Chinese).
5. Zi-wen Liu, & Deng-guo Feng. (2010). Dynamic integrity measurement architecture based on trusted computing. *Journal of Electronics and Information*, 32(04), 875–879 (In Chinese).
6. Huan-guo Zhang, Lu Chen, & Li-qiang Zhang. (2010). Research on trusted network connection. *Journal of computer*, 33(4),706–717 (In Chinese).

7. Xing Zhang, Qiang Huang, & Chang-xiang Shen. (2010). An analysis method of transmitting trust chain based on non-interference model. *Journal of Computer*, 11(1), 74–81 (In Chinese).
8. Run-lian Zhang, Xiao-nian Wu, Sheng-yuan Zhou. (2009). A Trust Model Based on Entity Behavior Risk Assessment. *Journal of Computer*. 32(04), 688-698 (In Chinese)
9. Jie Zhao, Nan-feng Xiao, & Jun-rui Zhong. (2009). The behavior trust control based on Bayesian network and behavior log mining. *Journal of South China University of Technology (Natural Science Edition)*, 37(05), 94–100 (In Chinese).

# Chapter 100

## The Application Level of E-commerce in Enterprises in China

Yinghan H. Tang

**Abstract** Based on the process of corporate value formation and performance system, this chapter extracts key factors indicating E-commerce application level in enterprises and has established a set of E-commerce measurement indicator system. In addition, this chapter uses Delphi method and Analytic hierarchy process to identify the coefficients of various factors. By applying this model to measure the E-commerce application level in 23 heterogeneous enterprises in Chinese domestic market, this study proves that the proposed model can yield a relative accurate measurement of E-commerce application level in enterprises. The results also indicate that there are strong individual differences among different enterprises in China. The E-commerce application level in individual enterprises is affected by corporate strategy, informatization level, E-commerce application performance, and human resources. The nature and the size of enterprises have significant correlation with E-commerce application level. The study also finds that the big-sized enterprises will become stagnant when they develop to a certain level, which is known as a “trap”.

### 100.1 Introduction

In the era of network economy, E-commerce has gradually penetrated into every economic and social aspect. It is becoming obvious that E-commerce is useful in saving costs, improving efficiency, increasing market share and enhancing corporate competitiveness advantages. But there exist great differences in E-commerce application among different enterprises due to some factors of the individual enterprises themselves. Then how to compare the E-commerce application level among enterprises? So far, there is no universally accepted measure method. Many

---

Y.H. Tang (✉)

Management School of Shenzhen Polytechnic, Shenzhen, Guangdong 518055, China  
e-mail: [tangyh@szpt.edu.cn](mailto:tangyh@szpt.edu.cn)

enterprises think they are E-commercialized when part of their business is done online. The answer to this question lies in some indicators to measure the E-commerce level in enterprises so as to compare the differences among them. Is the next questions are how to measure the E-commerce application level in enterprises?

## **100.2 The Definition of Enterprise E-commerce Application Level**

E-commerce application level of enterprises reflects the situation when the enterprise applies E-commerce at a certain time point. Different definitions are given from different perspectives of study in this field. From the viewpoint of corporate value creating process, E-commerce application level is a process to manipulate modern network technology, information technology, and computer technology to manage and control people, possessions, materials, and information so as to create value for the enterprise [1]. It is a dynamic process. But from the perspective of corporate production, it involves taking advantage of E-commerce technology to efficiently exploit and utilize various resources, to improve operation, management, and production, to lower cost and enhance quality as well as to strengthen the innovation ability and competitiveness. No matter from which perspective to define E-commerce application level in enterprises, one point is clear: it is the degree that E-commerce application has on the enterprise final performance and the actual outcome that drives companies to adopt E-commerce.

From the above discussion, it can be concluded that the E-commerce application level in enterprises is a measurement tool to estimate E-commerce level in the operation and development process of enterprises as well as a contribution degree that E-commerce has to the development of enterprises. In the following paragraphs, E-commerce indicator is used to replace the phrase “the E-commerce application level.”

The studies on how to measure E-commerce started from researches on the measurements of information economics. Mark Lupe is said to be the first in this field. In 1960s, Mark Lupe first used The Final Demand Method (also called the Expenditure Approach, the Final Product Approach) to measure American knowledge industry; the second scholar in this field is Borat who studied the measurement of American information economy. Many international organizations and research institutions, such as OECD, APED, and IDC, have made studies on E-commerce indicator system. Starting from the 1990s [2], EU began researching on how to analyze and compare E-commerce level between countries and to predict its future trend by means of indicator system and methodologies. UNCTD have released many editions of E-commerce and its development report since 2001. UK, USA, and Japan also published their respective E-commerce study reports successively [3]. Among these studies, the one made by a Japanese scholar who proposed the

Social Informatization Index is widely adopted because this measurement approach is easy to use.

Apart from official organizations, some consultancies and investigation firms as well as the academic circle also get involved in this field. For instance, the two biggest Internet investigation firms, Forrester Research and Jupiter Communications, Stanford University and University of Minnesota have all established their own E-commerce centers or projects to do relative researches. In 2002, PWC Consulting cooperated with Carnegie Mellon University to study measurement methods of E-commerce level and have proposed an E-commerce maturity model [4].

The network economic research center in Peking University has also made a series of studies on network economy and released a report involving IT development, E-commerce, E-government, etc. In a word, the measurement studies made in China are about informatization or network economy. Few studies are done exclusively on E-commerce application level in enterprises.

In conclusion, in order to develop E-commerce and drive it forward, a lot of countries have already made studies on E-commerce application and development as to how to set data collection, to make quantitative analysis, to evaluate, and to make predictions. But none of these studies have adopted a universally accepted measurement approach. Some studies already done both home and aboard are about how to measure information, rather than E-commerce; others do measure E-commerce level, but they are merely at a national or regional level. There are no studies in a strict sense that have measured it at a micro and medium level. However, these studies have great significances in providing references for further researches about the measurement of E-commerce application level in enterprises.

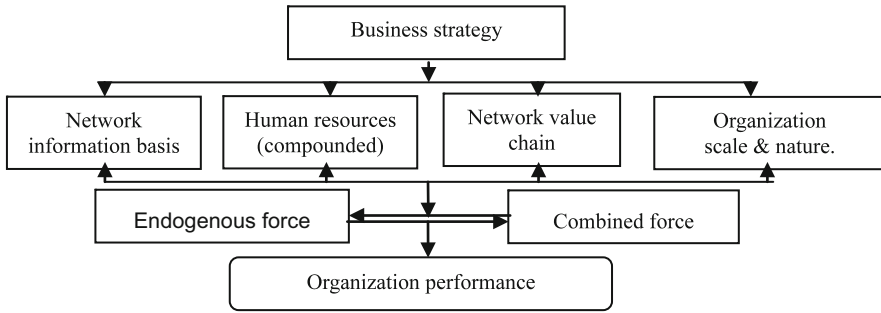
### **100.3 Constructing the Index System and Measurement Approaches of E-commerce Application Level in Enterprises**

#### ***100.3.1 Selection and Determination of Major Indicators***

There are various factors that have an impact on E-commerce application level. Many of them are interconnected and highly correlated, making it difficult to choose. According to the operation performance flow in E-commerce enterprises, E-commerce can be divided to six fundamental parts, e.g., [5] Nature and scale of the enterprise, business strategy, network informatization level, human resources, operation performance, E-business, and customer management. Then the key indicators are chosen from the six parts just discussed. (see Fig. 100.1) [6].

From the process shown in Fig. 100.1, the measurement indicators of E-commerce application level in organizations can be identified from Table 100.1.





**Fig. 100.1** The process showing impact of E-commerce application on organization performance

### ***100.3.2 Determining the Indicator Weight of E-commerce Application Level in Organizations***

In the comprehensive evaluation of many indicators, the determination of weight is the most fundamental but important task because it affects the outcome of comprehensive evaluation. Currently, there are many ways to determine indicator weight [7]. Generally speaking, there are two categories. One is called subjective method of weighting. The second is called objective method of weighting. Subjective method of weighting determines, from a qualitative perspective, the corresponding weight value of indicators according to their meaning and effects they show. It involves Delphi Method (also called Experts Grading Method) and the Analytic Hierarchy Process (AHP). Objective method of weighting belongs to a quantitative analysis, including Entropy method, Principal Component Analysis, and Factor analysis method. Generally speaking, Delphi Method is commonly used. Though it is somewhat subjective, it can usually reflect the real situation.

The studies discussed above on the evaluation index system of E-commerce application level have explained from various perspectives definitions and features of E-commerce application in organizations. In order to test the rationality of evaluation index systems and to explore the rules of E-commerce application level, the author of this chapter tries to construct an evaluating model of E-commerce application level in organizations.

### ***100.3.3 Evaluating Model Constructing***

This chapter uses the Simple linear weighting method (that is part of the comprehensive scoring method) to construct the index model of E-commerce application level in enterprises.

**Table 100.1** Weight value and standard value of indicators

Target layer	Criterion layer	Evaluating indicator	Unit	Standard value	Index property	Index weight	
E-commerce application level in enterprises	E-commerce strategy	Business strategy	-	1	+	0.0549	
		Goals	-	1	+	0.0395	
		Network project teamwork	-	1	+	0.0424	
	E-commerce application performance	System ready status	%	70	+	0.0204	
		System elasticity	%	60	+	0.0385	
	E-commerce business situation	E-commerce application performance	Ratio of annual expenses to train employee e-business technology and skills to total education training expense	%	8	+	0.0268
			Ratio of information infrastructure input expense to fixed assets	%	12	+	0.0313
		E-commerce business situation	Ratio of value added by E-commerce to total value created	%	50	+	0.0328
			Rate of profit	%	24	-	0.0125
			Contribution rate of network assets	%	25	+	0.0281
			Rate of network customer	%	35	+	0.0312
			Rate of network procurement	%	75	+	0.0291
			Rate of network booking	%	80	+	0.0391
			Sales rate online	%	50	+	0.0442
			Sales rate offline	%	30	-	0.0338
Network informatization status	Customer relationship management	%	100	+	0.0301		
	Ratio of business online to offline	%	100	+	0.0321		
	Ratio of E-commerce investment to infrastructure, upgrading and reconstruction	%	25	+	0.0220		
	Number of computers per 100 people	%	75	+	0.0231		
E-commerce business situation	Number of telephones per 100 people	%	90	+	0.0126		
	Rate of Internet access to computers	%	80	+	0.0313		
	Rate of online information release	%	80	+	0.0304		

(continued)

Table 100.1 (continued)

Target layer	Criterion layer	Evaluating indicator	Unit value	Standard	Index property	Index weight
		Ratio of E-commerce personnel to total employees	%	50	+	0.0314
		Average band width of Internet access per capita	*1	20	+	0.0231
		Modes of Internet access	*2	1	+	0.019
		Total capacity of Web server	-	2	+	0.0202
		Average of daily online time/work time of front line staff	%	60	+	0.0391
		Degree of getting information through Internet	%	65	+	0.0218
		Degree of exchanging information through Internet	%	90	+	0.0314
	HR condition	Number of professional technicians per 100 person	%	6	+	0.0307
		Number of students per 100 people	%	13	+	0.0261
		Ratio of the amount of labor done online to that of offline	%	100	+	0.0270
	Enterprise condition	Industry/sector	-	1	+	0.0218
		Enterprise size	-	1	+	0.0222

Note 1: Business strategy, business goals, network project teamwork, and ways of Internet connection are given coefficient by levels

Note 2: The reference standard value comes from conclusions drawn from statistics and analysis of organizations [1]

Note 3: \*1 = M/100 people, \*2 = ADSL/fiber(0.6/1)

$$I = \sum p \cdot \omega_i \tag{100.1}$$

In formula (100.1), I refers to E-commerce index; p means the value of the *i*th evaluation index after dimensionless treatment.  $\omega$  means the *i*th weight of evaluation index. In order to make it easy to calculate p, after comparing the real value and standard value in enterprises, this model choose the value after dimensionless treatment. That is,  $p = pi/\tau_i$  where  $\tau_i$  is a standard value . Then formula (100.1) can be changed into:

$$I = \sum p \cdot \omega_i = \sum \left(\frac{pi}{\tau_i}\right) \cdot \omega_i \tag{100.2}$$

This is the E-commerce index of enterprises. It can be used to measure the E-commerce application level in enterprises and also can be made to compare the E-commerce application level in one enterprise with that of others.

### 100.4 A Case Study of Applying E-commerce Application Level and Its Analysis

To apply the above index system model to practice, we randomly choose 23 enterprises (mainly located in Shenzhen) for interview to get the data we want (Table 100.2).

From Table 100.2, it is obvious to see that there are striking differences in E-commerce application level of enterprises. In some enterprises, the E-commerce application level is relatively quite high. In those enterprises, E-commerce is adopted in every business transactions and has significant impact on organization performance; while in some other enterprises, E-commerce application level is quite low. Mean 50.15 indicates that the overall E-commerce application level of Chinese enterprises is rather low.

Table 100.3 indicates the differences in E-commerce application level among enterprises of different industries. The manufacture industry shows generally low E-commerce application level while the information software industry has an E-commerce application level that is obviously above the average. From the viewpoint of enterprise size, those super large and medium-sized enterprises are reported to have high level of E-commerce application. Nevertheless, the small-sized ones have relatively low levels. This investigation finds out that there is a “trap” in E-commerce application level of enterprises, that is, when the application level

**Table 100.2** E-commerce application index of Chinese enterprises

	Mean	Maximum	Minimum	Std deviation	No. of samples
E-commerce application index	50.15	110.15	8.00	46.52	23

Source: Data comes from the result of questionnaire with managers in 23 enterprises

**Table 100.3** Distribution of the E-commerce application index of enterprises

		Mean	Maximum	Minimum	No. of samples
Nature of the enterprise	Manufacture	31.19	76.43	8.00	6
	Information software service	72.56	110.15	59.67	5
	Business, trade	55.36	90.21	50.16	6
	Service	41.49	95.13	17.32	6
Scale of the enterprise	Super large	74.06	82.14	65.98	2
	Large	51.11	65.22	35.26	3
	Medium	61.04	75.2	38.43	6
	Small	19.38	88.86	8.00	12

Note: there are 23 enterprises in this survey. Scale of enterprises is divided according to the number of people in an enterprise, e.g., 500 people or more is super large; 100–499 is large; 30–99 is medium; 30 or fewer is small

in the super large and large-sized enterprises develops to a certain degree, it will become stagnant, while once the small-sized enterprises achieve a breakthrough, their E-commerce will develop into its maximum extreme.

## 100.5 Conclusions and Suggestions

Firstly, there exist sharp differences in E-commerce application among Chinese enterprises. Large organizations generally have a higher level of E-commerce application than small and medium-sized enterprises. A main reason to this lies in the fact that large enterprises tend to invest more to infrastructure facilities, such as network information. Moreover, large enterprises have relatively more advantages over human resources. However, when the application level comes to a certain stage, it will fall into stagnation. Therefore, to look for a new breakthrough in E-commerce application is an urgent issue that most Chinese E-commerce enterprises face to tackle.

Secondly, the development strategy of E-commerce specialization has substantial effect on E-commerce application in enterprises. No matter what industry the enterprise belongs to, the degree of importance it attaches to E-commerce decides the level of E-commerce application. Without a long-term development strategy, some small-sized enterprises get lost in the wave of the current E-commerce tide. Though the quality of human resources occupies only a small proportion of weight, it is an inherent factor that is critical to E-commerce application level index of enterprises because it drives their endogenous growth of E-commerce. Besides, some enterprises are difficult to survive just because they have no sufficient compounded talents in the field of E-commerce.

Finally, the nature of industry decides the E-commerce application level of enterprises. For manufacturing industry, their need for E-commerce application is not as strong as other industries because it is still a traditional industry and its modes of production are also traditional. However, in the long run, it should lay some emphasis on E-commerce as E-commerce will help it for transformation and upgrading, and also for better competitiveness. Many indicators of the E-commerce application level index have close relationship with the online and offline businesses, but the final enterprise performance directly decides whether the enterprise will continue with E-commerce and reinvest in E-commerce or not. However, this kind of “circle” needs adjusting based on the development strategy of the organization. After all, E-commerce application in enterprises is still in a process of exploration. Our investigation shows that the current combination of online and offline business mode is a main approach to apply E-commerce in enterprises. This mixed management model is a product of traditional industry combined with new technology and new services. The new development of E-commerce application in enterprises is an innovation of E-commerce application model for organizations.

## References

1. Huang, J. Q., Shui, M., & Cai, W. J. (2012). E-commerce ready level model and its case study. *Statistics and Decision*, 2012(18), 100–103. In Chinese.
2. OECD (1998) Measuring the ICT sector. <http://www.oecd.org/>, retrieved on May 21, 2011
3. OECD (2002) Measuring the information economy [R/OL]. <http://www.oecd.org>, retrieved on March 20, 2012.
4. Committee for Information, Computer and Communications Policy, OECD. (1997) Measuring electronic commerce. <http://www.oecd.org/>, retrieved on Oct. 30, 2012.
5. Yang, H. H., & Li, P. (2011). Analysis on new industrialization evaluation index and measurement. *Economic Management*, 10, 121–125. In Chinese.
6. Zhao, J., Zhu, Z., & Wang, F. (2010). E-commerce performance evaluation model of enterprises based on process. *Manage Eng J*, 2010(1), 17–24. In Chinese.
7. Chen, X. D., & Fu, L. S. (1999). Study on the measurement of Chinese industry informalization level. *Social Science Management*, 6, 21–31. In Chinese.

# Chapter 101

## Toward a Trinity Model of Digital Education Resources Construction and Management

Yong Huang and Qingchun Hu

**Abstract** This chapter aims to solve the problem of how to construct and manage digital educational resources effectively. It puts forward a trinity mode based on system architecture, workflow, and technology system. The trinity model consists of “Pre-Stage, Mid-Stage, Post-stage,” “Theory, Practice, Regulation” and “Approach, Tool, Rule.” By combining the trinity mode with case studies, the issues concerning construction and management of digital educational resources are to be analyzed, including topic selection, relationship between quantity and quality, implementation. Over the past year, results have showed that the trinity model could shorten more than 50 % of the development cycle of the project. The model could greatly help improve the construction and management of digital educational resources.

### 101.1 Introduction

With popularity of the mobile storage technology and equipment, the digital teaching and learning would play a prominent role in the future. A number of e-learning materials have been made. The traditional teaching and e-learning materials will coexist in current classrooms [1, 2]. There has been researches focusing on the effect of digital-integrated education on national innovation systems [3, 4]. Our study is based on the project of “To Carry Out the Construction of Digital Curriculum and Learning Environment Change” supported by the Shanghai Municipal Education Commission [5]. In the past 10 years, we have done a lot to

---

Y. Huang (✉)

Shanghai Audio-Video Education Center, Shanghai Distance Education Group,  
Shanghai 200086, China  
e-mail: [hyong@shtvu.edu.cn](mailto:hyong@shtvu.edu.cn)

Q. Hu

School of Information Science and Engineering, East China University of Science and  
Technology, Shanghai 200237, China

build Shanghai Education Resource Center (<http://www.sherc.net>). Now, we need to summarize and reflect on our past 10 years' work. We find that the following three aspects should be improved.

The first aspect goes with topics selection. Most resources in Shanghai Education Resource Center are not fully used. The contents in the textbooks are updated frequently, such as the learning contents in Chinese and English for instance have been updated by more than 50 %. However, some digital materials in the Shanghai Education Resource Center are not updated subsequently [5].

The second aspect is the relationship between quantity and quality. The huge number of resource in Shanghai Education Resource Center makes users difficult to search for their favorite resources. And also, the poor quality of some resources reduces the usability of resources.

Thirdly, there is a point at implementation and publishing. With the popularity of electronic equipment, digital teaching and learning will become a mainstream in the future [6, 7]. The importance of efficient construction and management of high-quality digital education resources grows with increasing government investment on educational resources [8, 9]. It is worthy to have a study on how to improve the quality and the usability of the digital education resources. It is quite complex, for it involves many factors, such as storages [10], materials, applications [11], teachers' views [12], marketing reports, copyright [13], and so on [14].

For the past 2 years, our focus has gradually shifted from the construction of the repository to the development of digital educational resources. Till now, we have issued 10 digital textbooks and more than 200 APPs (Nada Online <http://www.ndapk.com>). Our work of digital education resources construction and management usually start from the perspective of education or engineering, but we are often at a loss in this way.

Currently, we had a lot of successful cases, but there still exist a lot of problems and cases of failures. It is time to summarize and study on more effective methods to construct and manage the resources. On basis of the past years' research and development activities, this chapter discusses a digital education resources construction management model—a Trinity Model.

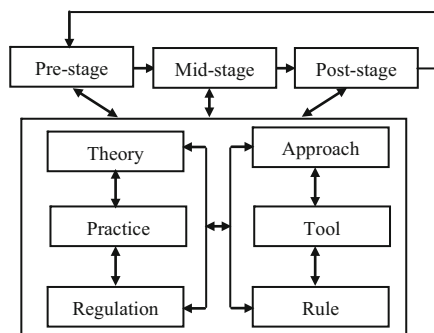
## 101.2 The Components and Meanings of the Trinity Model

The word “trinity” is commonly used to describe three individuals, three things, or aspects in a tight inseparable whole. The concept of trinity is to be used, and more appropriate in our practices of constructing and managing digital education resources.

In our practice, digital education resources construction and management consist of three parts: workflow, system architecture, technology.

Furthermore, the construction and management of digital education resources can be seen as a big project. Workflow in the project could be divided into “Pre-stage, Mid-stage, and Post-stage,” these three stages are whole inseparable.



**Fig. 101.1** A trinity model

The system architecture of construction and management can be divided into a threesome “Theory, Practice, Regulation.” The technology used in the construction and management can be divided into an inseparable “Approach, Tool, Rule.”

The workflow, system architecture, and technology combine to form a unity, named as a trinity development model together (Fig. 101.1).

### 101.2.1 Pre-stage

Pre-stage is the first stage of the project. It includes mainly literature review of the relative theory, preparation of topics, the top-level designs, data collection, feasibility analysis, copyright issues, regulations, methods, and tools. This stage should start from the users’ demand to consider which topics to be designed. Only those topics what can really improve the users’ learning interest, should be worthy to be done. The key points during the first stage include the following:

Pre-stage is a continual and iterative process, focusing on data accumulation, analysis and research. The work during the pre-stage builds library that benefits the following work, including cases library, theoretical library, methods library, rules library, tools library, Apps library. All these six libraries provide decision-making managers with effective help.

Pre-stage is a decision-making process. The aim of decision-making is to reach a certain goal. It is an analysis and judgment process, which adopts scientific methods, selecting a feasible method among two or more. The decision-making process chooses the best in the process of the optimal solution through analysis and comparison.

### 101.2.2 Post-stage

There is no little thing after finishing resource construction. In fact, work during post-stage is also very important and rich. It includes collecting students’ or

teachers' feedback; revision and updating; products issued online; customer service; research activities; information dissemination; developing new topics; and project evaluation. The work in post-stage is to improve the education resource, such as three-round assessment activities, discussion, performance, regular meetings of assessment by experts, and feedback mechanisms.

### ***101.2.3 Mid-stage***

Mid-stage is in the mid-term in workflow, including content, technology, evaluation, application, feedback, releases, and upgrade. The development process involves a lot of methods, tools, and rules, including technical level, education level, government level, and resources on ecology. If the members in a project team do not communicate and collaborate well, it could lower the development efficiency and the quality.

We need to systematically deal with pre-stage, mid-stage, and post-stage of the project, and deal them with theory, practice, regulations, techniques (methods, tools, rules) in order to improve usability of the digital resources.

### ***101.2.4 The Relationships***

Communications among the three stages is very important.

From Pre-stage to Mid-stage, there should be documents of requirements specification, bidding, experts' assessments, project contracts, etc.

The documents from Mid-stage to Post-stage involve experts' position paper, usability testing reports, project inspection reports, products backup, etc.

Documents from Post-stage to Pre-stage should be evaluation reports, new project proposals, seminars, etc.

The advantage of this model lies in two aspects.

1. It helps to find the reason of inefficiency. In order to develop digital resources better, beside a focus on Mid-stage of a project, we need to have an in-depth research and attention to the relationship among these three stages. Sometimes, one project is inefficient because of ignoring the key role of the Post-stage to Pre-stage. And, there is no good technology system and relevant theories at hand.
2. It helps to find the right way to solve the problem. The practice of optimization methods organized into libraries is very valuable and reused. A comparative study of different tools for decision-making and different methods for developing efficiently has a significant impact. With establishment and update of rules library and methods library, it is easy to form a coordination mechanism.

### 101.3 Implementation of Trinity Model

We have developed a lot of digital resources. And some are still in development. The implementation of “Theory, Practice, Regulation” and “Approach, Tool, Rule” reflects past practices, as shown in Tables 101.1 and 101.2.

We cannot reach the essence of the technology if we view technology only as a tool or means. The technology is composed of a variety of methods; tools and rules of the system work together for a purpose. In the past 2 years’ practice, we frequently encounter three typical problems: topics, quality and quantity, implementation.

#### 101.3.1 Selecting Topics

Some APP topics, such as learning English APPs, are expected to be popular; however, the download record is poor. On the other hand, sports APPs, which seem to be less-popular topics, came out with many downloads. This phenomenon can be analyzed with Long Tail theory, which is a new theory appearing in the Internet Age proposed by Americans Chris Anderson [15]. APPs are virtual products, as Google Adwords, iTunes. They are consistent with the Long Tail theory. All of their payment and delivery cost is close to zero. It can be said that the virtual products’ sales are inherently suitable for the Long Tail theory. The products of online retail giant Amazon are all-inclusive, rather than just a few commodities that can create high profits. Result shows that the Amazon model is successful, and the

**Table 101.1** Implementation of “Theory, Practice, Regulation”

	Theory	Practice	Regulation
Pre-stage	T1: the long tail theory	P1: topics	R1: topics regulation
Mid-stage	T2: game theory	P2: quality and quantity	R2: elevate regulation
Post-stage	T3: computational thinking	P3: implementation	R3: publish regulation

**Table 101.2** Implementation of “Approach, Tools, Rules”

	Practice	Approach	Tool	Rule
Pre-stage	P1: selecting the topics	A1: brainstorm	T1: converter	R1: if average down per day >5 then upgrade APP
Mid-stage	P2: improving the quality and quantity	A2: project	T2: iBook author	R2: if the score of efficient <3 then redesign
Post-stage	P3: implementation	A3: group	T3: feedback tool	R3: if the score of user <3 then redesign

profit is not ideal if we ignore the Long Tail and focus only on a few bestsellers products.

In market analysis, we should concentrate on specific target markets, to create a new product and service. Our suggestions are listed below:

- (a) Checking the products we have developed and analyzing the available resources. We need to know which of them is a success or failure? What is the reason? Is there any possibility of “redevelopment” for those products?
- (b) Analyzing market demand, to clarify the nature of the issues. What are the specific issues that need to be addressed? What are the key problems to be resolved? Why is it necessary? What are the difficulties? What are the existing resources? Are there any especially good methods to solve the problem? What is the significance to solve the problem? Whether there exists a similar product?
- (c) Spending more time on the topics’ selection. It is the basis of the publishing business that decides whether the follow-up work is a success or failure. The successful topics can help the digital publishing institutions form and establish a good follow-up development cycle.
- (d) Establishing rules for developing activities. For instance, if downloads are over five items per day in a whole month, then these products should be worthy to be updated and improved.

### ***101.3.2 Improving Quality and Quantity***

In practice, quality and quantity are often on contrary. In general, quality is difficult to guarantee if focus is on pursuit of the quantity. In this case, most products are with a small number of download. The Game Theory helps to solve this contradiction. One solution is pursuit of quality with the rise of the number; pursuit of quantity with the improvement of quality.

We have a wealth of APPs resources online. To our disappointment, we have only a small number of download with poor usability. The phenomena make us to refine our current products and publications; establish a three-round assessment system for digital publishing.

### ***101.3.3 Implementations***

It is better to understand the problem and seek prompt solution to the problem based on the concept of “Computational Thinking” [16, 17]. Computational thinking means the method to solve problems using the computer technology. We do not study digital terminal when we face to the digital terminal implementation of digital teaching and learning, just like the astronomy is not a research of the telescope. We need to solve the problem like computer scientists do, to understand the transformations of digitized teaching and learning.

Our suggestions include the following: improving digital educational resources evaluation criteria; collecting data continually for future development after issue of a product. After an assessment activity, a summary report is needed for next-stage revision and improvement.

Based on our actual situation and through the trinity model, we argue: on road of digital publishing, teacher training, and supplementary teaching resource pack.

### **101.3.3.1 On Road of Digital Publishing**

From professional views in the industry, the most important factor in “digital transformation of teaching and learning” is teaching and learning materials. Currently, at least 70 % of domestic publishing companies in China rely on textbooks and tutorial resources to survive. There would be crises in the current digital age. In addition, the textbooks are varied in different provinces, and their update speed is not faster. There are many issues to be addressed if the “digital transformation of teaching and learning” project plan is launched all around our country by the government. With regard to e-textbooks, how much does it cost and how to resolve copyright issues involve the three departments: education, publishing company and the market. The form and content of the “digital transformation of teaching and learning” would pose a challenge to education and publishing industry, but there exists great opportunities.

### **101.3.3.2 Teacher Training**

In the teaching system, for the intervention of the new media, there is an urgent need to train teachers. There are many good platforms in teacher training and research activities. We can collect a large number of practice cases, such as the digitized teaching and learning application in different schools.

The reform of information technology in education is irresistible. Children would be left behind if he/she is less trained by “transformation of digital teaching and learning resource.” The research towards transformation includes physiological and psychological impact of its students, the applications model in education and teaching and education management.

In 2013, Chinese government plans to launch a number of e-learning materials projects for primary and secondary schools in the city of Shanghai. The massive digital resources accumulated over the past decade in Shanghai Distance Education Group. In the project, all these resource are grouped by the “cloud platform” for sharing among all schools in Shanghai.

Using the trinity model, we propose to speed up the research on digitized transformation of teaching and learning office institutional settings, such as the establishment of data centers, research centers, training centers, technical centers, and service center.

### 101.3.3.3 Supplementary Teaching Resource Pack

The key to the effective implementation of digital teaching and learning should have two sides, including developing resources and teaching mode. We need to consider how to provide supplementary teaching resource pack. First of all, we should make full use of educational resources and our publishing resources rather than let teachers search on internet. These resources are valuable and useful. It is very important to consider how to conduct secondary development, and how to deliver them to the hands of teachers and students.

## 101.4 Conclusion

In early 2011, in the project “Reform of Digital Teaching and Learning,” we had developed digital resources for three courses involving mathematics, information literacy, and the English language. It took 6 months to discuss and organize. The result, however, is not satisfactory.

It fails in five parts. Requirement analysis is not enough in the Pre-stage, such as selecting tools, designing ideas, and copyright issues. There is less collaboration among developers. There is less education theory to guide the process. And there are less evaluation and feedback mechanisms.

In late 2011, we analyzed our work from the three levels: workflow, system architecture and technology. And the trinity model was proposed and used at that time. In 2011, in our projects, it generally cost 6 months to finish an e-textbook or an APP project without the trinity model. Now, we can develop it within 2 months under the model. The number of download and the amount of our APPs products in 2012 rise ten times more than in 2011. The result shows that the model could shorten the development cycle of a project by more than 50 %. The model could greatly help improve the construction and management of digital educational resources.

The trinity model makes digital education resources construction and management more effective. The Pre-stage and Post-stage work of digital education resources construction in practice is most likely to be ignored, although they are very important. We usually deal with the digital education resources construction and management with conventional thinking and theoretical support in practice. But lack of system construction and support results in a lot of disorder and inefficient activities. It will be more economic, scientific, and logical to make a decision in construction and management of the digital educational resources, based on the construction of theoretical library, method library, rule library, tool library, and case library.

**Acknowledgements** This research is supported by Digital Educational Resources Ecological Construction and Sharing Mode under Shanghai Municipal Education Commission by National Educational Programs Grant No. NOESP dca110194.

## References

1. Nussbaum, M., & Diaz, A (2013). Classroom logistics: Integrating digital and non-digital resources. *Computers & Education*, 69(0), 493–495.
2. Chen, N.-S., et al. (2011). Augmenting paper-based reading activity with direct access to digital materials and scaffolded questioning. *Computers & Education*, 57(2), 1705–1715.
3. Wiseman, A. W., & Anderson, E. (2012). ICT-integrated education & national innovation systems in the Gulf Cooperation Council (GCC) countries. *Computers & Education*, 59(2), 607–618.
4. Allegra, E., et al. (2011). Cross-border co-operation and education in digital investigations: A European perspective. *Digital Investigation*, 8(2), 106–113.
5. Shanghai's long-term Education Reform and Development Plan (2010~2020). Available from: [http://www.360doc.com/content/11/03/13/21/5344705\\_100838514.shtml](http://www.360doc.com/content/11/03/13/21/5344705_100838514.shtml) (2010) (In Chinese)
6. Loveless, A., & Underwood, J. (2010). Learning in digital worlds: A view from CAL09. *Computers & Education*, 54(3), 611–612.
7. Thompson, P. (2013). The digital natives as learners: technology use patterns and approaches to learning. *Computers & Education*, 65, 12–33.
8. He, K. (2009). The status and strategies on the construction of digital learning resources. *E-education Research*, 10, 5–9 (in Chinese).
9. Jing, Y. J., & Li, X. (2011). The study and practice of a pattern on the community of regional education information resources infrastructure. *China Educational Technology*, 1, 83–86 (in Chinese).
10. Fu, X., et al. (2011). On data integration, warehousing and software reuse in the construction of digital campus: A review on performance. *Procedia Engineering*, 15, 3109–3113.
11. Tohidi, H. (2011). Human resources management main role in information technology project management. *Procedia Computer Science*, 3, 925–929.
12. Petko, D. (2012). Teachers' pedagogical beliefs and their use of digital media in classrooms: sharpening the focus of the 'will, skill, tool' model and integrating teachers' constructivist orientations. *Computers & Education*, 58(4), 1351–1359.
13. Hunter, B. (2013). The effect of digital publishing on technical services in university libraries. *The Journal of Academic Librarianship*, 39(1), 84–93.
14. Kreijns, K., et al. (2013). What stimulates teachers to integrate ICT in their pedagogical practices? The use of digital learning materials in education. *Computers in Human Behavior*, 29(1), 217–225.
15. Anderson, C. (2009). *The long tail: Why the future of business is selling less of more* (pp. 67–80). New York: Hyperion Books.
16. Wing, J. M. (2007). Computational thinking. *Communications of ACM*, 49(3), 33–35.
17. Karp, R. M. (2011). Understanding science through the computational lens. *Journal of Computer Science and Technology*, 26(4), 569–577.

# Chapter 102

## Geographic Information System in the Cloud Computing Environment

Yichun Peng and Yunpeng Wang

**Abstract** Cloud computing has become a very popular vocabulary in recent years. The combination of cloud computing and GIS (geographic information system) can improve the performance of GIS. By analyzing the technology of cloud computing, this paper introduces the concept of GIS based on cloud computing; based on the current major GIS application development trends, key technologies of cloud GIS are proposed; finally four application modes of cloud GIS are presented. Cloud GIS can improve stability and efficiency services to end users by optimized network resource allocation of underlying data and services.

### 102.1 Introduction

At present, in the field of GIS, there exist two questions: firstly, in data aspect, because of widely data sources, which could cause the following question, coordinates and formats are not interchangeable, the data is not compatible, semantics are not uniform, and difficult to share, difficult to interoperate and so on, moreover, how to store, manage, update and analyze these massive data is also difficult to achieve. Secondly, in application, with the development of society, people's demand for geographical information service also continues to grow; however,

---

Y. Peng (✉)

Guangzhou Institute of Geochemistry Chinese Academy of Science,  
Guangzhou 510640, China

City College of Dongguan University of Technology, Dongguan 523106, China

University of Chinese Academy of Sciences, Beijing 100049, China

e-mail: [yichunpeng678@hotmail.com](mailto:yichunpeng678@hotmail.com)

Y. Wang

Guangzhou Institute of Geochemistry Chinese Academy of Science,  
Guangzhou 510640, China

e-mail: [wangyp@gig.ac.cn](mailto:wangyp@gig.ac.cn)



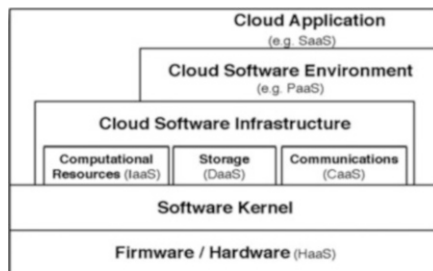
data production unit is relatively less and a single system is also hard to own all the resources and the ability of processing, therefore, users unable to get their required data from a single source too. In addition, large-scale concurrent access and high cost of upgrading are the main factors that hinder the development of GIS. Service-based “cloud computing” has the advantages of massive data storage, large-scale computing, and in-depth data mining, which is very suitable for GIS development; in addition, GIS service provider can deploy flexibly GIS applications on cloud computing platforms and can dynamically adjust the system’s software and hardware requirements. GIS end user may gain the service on demand. Therefore, GIS based on cloud computing can not only simplify system deployment and management and reduce the cost of investment, operation, and maintenance but also improve the flexibility of GIS applications and infrastructures. In this paper, the concepts of GIS based on cloud computing are introduced; based on the current major GIS application development trends, key technologies of cloud GIS are proposed; finally four application modes of cloud GIS are presented.

## 102.2 Cloud Computing Technologies

Since Google puts forward to the concept of cloud computing in 2006, Amazon’s “Elastic Computer Cloud” service, IBM’s “Blue Cloud” plan, Microsoft’s Internet operating system “Midori,” Sun’s “Black Box” plan, SAP, Yahoo, and some other large companies have developed their own “Cloud” plans or have launched their “Cloud” products. In early 2008, IBM cooperated with the Wuxi municipal government and established Wuxi Software Park Cloud Computing Center, which began commercial applications of cloud computing in China. A number of industries and localities also have launched some cloud computing plans, such as Rising’s “Cloud Security” plan, Beijing’s “Auspicious Cloud Computing,” Shanghai’s “Yunhai Plan,” Suzhou’s “Fengyun Online,” Guangzhou’s “Tianyun Plan,” China Mobile’s “Tianyun Plan,” China Unicom’s “Woyun Plan,” and China Telecommunication’s “Nebula Plan.”

What is cloud computing? Nowadays, cloud computing is still an evolving paradigm. Its definitions, use cases, underlying technologies, issues, risks, and benefits will be refined in a spirited debate by the public and private sectors. These definitions, attributes, and characteristics will evolve and change over time. The NIST definition of cloud computing is as follows [1]: “Cloud computing is a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction. This cloud model promotes availability and is composed of five essential characteristics, three service models, and four deployment models.” The cloud computing industry represents a large ecosystem of many models, vendors, and market niches. This definition attempts to encompass all of the various cloud approaches.

**Fig. 102.1** Five-layer model of cloud computing



In 2008, Lamia Youseff, a doctoral student of the University of California Santa Barbara, Maria Butrico, and Dilma Da Silva, researchers of New York IBM T.J. Watson Research Center, published a research report entitled “Toward a Unified Ontology of Cloud Computing.” This report established a five-layer model [2] as shown in Fig. 102.1.

At present, there are mainly three kinds of cloud computing service model, IaaS (Infrastructure as a Service), PaaS (Platform as a Service), and SaaS (Software as a Service) [3].

IaaS: one service of the closest underplayed in cloud computing, its output products are some resources of computing, storage, and network, such as VM, Storage, CDN and DNS. SaaS and PaaS will be created on the IaaS. IaaS provides cloud computing and cloud storage services for high availability, flexible expansion, on-demand billing, easy to use, low cost, and other advantages favored by many domestic and foreign companies. At abroad, such as Amazon AWS, IBM Smart Cloud, Microsoft Windows Azure, Rackspace, and NASA open source products: Open Stack, in addition, VMware, BlueLock, CSC, GoGrid, Savvis also launched its own IaaS technology or products; domestic such as Aliyun, Grand Cloud, and HUAWEI Single Cloud.

PaaS: in addition to providing computing, storage, and network infrastructure hardware resources, and PaaS also provides the basic framework for software development. Application developers must develop and host applications according to language and specification of platform, independent software vendors, or other third parties for vertical industries to create new solutions, but do not have to purchase the development, control the quality, or build the server. At present, there are some cloud computing platforms such as Google Application Engine, Sina Application Engine, Salesforce.com’s Force.com, and Microsoft’s Azure.

SaaS: is a software layout model, is a completely innovative software application model, its application is designed for network delivery, and is convenient for the user to host, develop, and access through the Internet; its output is information system, such as OA, CRM, ERP, and CMS. SaaS providers put up the network infrastructure and software hardware operating platform of which requirement in realizing enterprises informatization and will be responsible for implementation and maintenance of the system; enterprises do not need to purchase software and hardware, build a computer room, or recruit IT staff, but can use information system

via the Internet. The main products are alesforce.com, NetSuite, Google Gmail, Zimbra, Zoho, the IBM Lotus Live and SPSCommerce.net, Ali software, etc.

With the deepening of cloud computing, database technology has been changing from the traditional relational database memory data grid to NoSQL, Database technology transformation has produced the fourth kind of cloud computing mode: DaaS, Data as a Service.

DaaS is a strong complement to SaaS model, a service model of web-based virtual storage; it can be for business users and business intelligence users to simplify the process of information retrieval, the user according to the actual storage capacity to pay. The benefit of which database migrate into cloud is data integration, usually in large enterprises; database needs to be shared across different departments; cloud services can be integrated into a single custody DBMS, so DaaS can reduce the problem of interior database expansion. The main products are Amazon's SimpleDB, VMware's vFabric the Data Director, Google's AppEngine and China Telecom Shanghai branch in collaboration with EMC "e cloud," etc.

IaaS is the foundation of cloud computing; DaaS is based on IaaS; SaaS can be deployed on PaaS or deployed directly on IaaS; and PaaS can be built on IaaS or be directly built on the physical resources. SaaS, PaaS, and IaaS combined with DaaS can build a complete cloud computing environment.

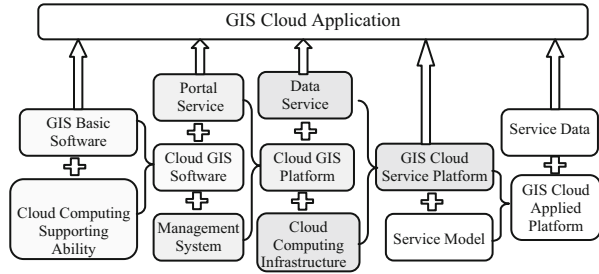
### 102.3 The Key Technology of Cloud GIS

Cloud GIS is the result of the combination of cloud computing and GIS; it is no longer a single GIS software platform, but provides storage, software, and content which can be virtually flexible, deployed, or rented, as long as through the PC desktop, mobile phone, and a web browser; the user can access on-demand data, map, spatial analysis, and Geoprocessing services which are provided by cloud GIS. At present, there have some GIS products based on cloud computing at home and abroad. At abroad, there are, such as, Google's Google Earth, Google Moon and Google Mars, ESRI's ArcGIS Online and ArcGIS10.1 etc. At home, there are, such as, SuperMap's SuperMap GIS 6R, MapGIS's MapGIS K9 SP3, GeoStar's GeoCloud etc.

Corresponding to the four-service models of cloud computing, cloud GIS also has some related concepts and technologies: cloud computing infrastructure belongs to IaaS, GIS application, cloud GIS platform, GIS cloud service platform, and GIS cloud application platform; all that belongs to PaaS; cloud GIS software belongs to SaaS; data services belong to DaaS; the composition system of cloud GIS is shown in Fig. 102.2 [3, 4].

Cloud GIS must run through data, software, and development; the user can really get the GIS resources at any time, so cloud GIS has its own unique technology in cloud computing platform and data interoperability, GIS spatial data storage, management, analysis and processing, and terminal access.

**Fig. 102.2** The composition system of cloud GIS



### 102.3.1 Technology of Cloud Computing Platforms and Data Interoperability

Cloud GIS to realize cross operating system (Linux/Unix/AIX/Windows), cross GIS platform, support and synchronize with a variety of hardware architectures, can meet the private cloud and public cloud environments as a set of unified architecture and realize integrated connectivity and interaction of private cloud and public cloud, and its application fields including desktop, LAN, and Internet. It can support single release, automatic synchronization, frequency statistics and automatic optimization, support cloud internal data interoperability, private and public cloud interoperability, and interoperation between cloud center. In order to achieve the above operation, we must establish interoperation and integration standards of cloud computing, cloud computing service interface standards and application development standards, cloud computing interface standards between the different levels, cloud computing service catalog management, seamless migration between different cloud portability standards, cloud computing standards of business indicators, cloud computing architecture management standards, and cloud computing security and privacy standards, etc. Cloud computing technology use the multilevel structure framework, from top to bottom: business logic layer, application layer, distributed file and operating system layer, virtualization layer, hardware layer and data center infrastructure layer.

### 102.3.2 Technology of GIS Spatial Data Storage, Management, Processing, and Analysis

Cloud GIS uses virtual storage technology to establish a highly efficient, seam-less, multi-source, multi-scale, multi-spatio-temporal data model, which realize massive spatial information storage, management [5, 6].

**Storage technology:** the development of spatial data storage from the file system to a distributed file system and cloud storage system based on Internet technology entirely; spatial database has developed from enterprise database to distributed spatial database and will support for BigTable, HBase, and NoSQL to store and manage data in the future and support uniform access to it through the spatial database interface standard and REST interface. There are two main technologies at present: Google's non-open source GFS (Google File System) and Hadoop's open source HDFS (Hadoop Distributed File System). Most of the IT companies, including Yahoo's and Intel's "cloud" plans, adopt the data storage technology of HDFS. The future development will focus on large-scale data storage, data encryption, and security guarantee and continue to improve the rate of I/O.

**Data management:** uses virtualization technology to realize the unified management of spatial database; support for rapid migration and automatic synchronization of data between systems, departments, levels, etc.; has off-line application and online update technology; and achieves distributed, multilevel, and supporting multi-terminal spatial data security process. This method that the table divided by column and then stores it to storage is usually used for the data management. There are two main technologies at present: Google's BigTable data management technology and Hadoop's open source data management module that is similar to BigTable.

**Data processing:** a task-oriented asynchronous spatial data processing architecture, supports concurrent processing and process control in large clusters, supports long-time running and long transaction processing, supports for mobile terminal handling large spatial databases, has those functions of visualized design for processing flow and monitoring real-time running status of system, can cross-platform, cross-regional integrate spatial data processing, and immediately release the processing result.

**Spatial analysis:** has unified spatial analysis framework and rich spatial analysis model, establishes a standardized analysis model library, and supports for rapid construction and automatic operation of spatial analysis process and the immediate release of the analysis results.

### ***102.3.3 Technology of Terminal Access***

The final purpose of GIS platform based on cloud computing is to let the user access all the GIS functions through a browser. To meet smooth transmission of the massive data in a different network, it is requested that the system has unified kernel and interface, and its services can be accessed by various types of desktop, web, and mobile client, ultimately achieving those effects such as data synchronization access, consistent processing results, and elegant user experience.

## 102.4 The Application Mode of Cloud GIS

According to the four models of cloud computing, cloud GIS also put forward four kinds of GIS application services, namely [7, 8], GIS Software as a Service (namely, SaaS), Geographical Information Platform as a Service (namely, PaaS), Geographic Information Infrastructure as a Service (namely, IaaS), and Geographic Information Content as a Service (namely, DaaS). Figure 102.3 shows the cloud application structure of ArcGIS [9].

### 102.4.1 GIS Software as a Service

Also known as the “cloud of geographic information services,” this service refers to the use of the Internet to provide online geographic information processing services, including map publishing, data format conversion, spatial analysis, and other services. Generally use the development of multi-tier architecture: the upper adopts SOA architecture pattern; geographic information services are packaged into standard web services and are incorporated into the management and use of the SOA system, its contents, including service interface, service registration, service search, and service access; the middle layer realizes the billing of the user using geographic information, takes charge of the load balance and map tile service; data layer is the underlying data service provided by the GIS server. This service mode, the GIS system developers, can build their own solutions, such that they can develop their own GIS solutions by using GIS platform firm’s PaaS services; use cloud

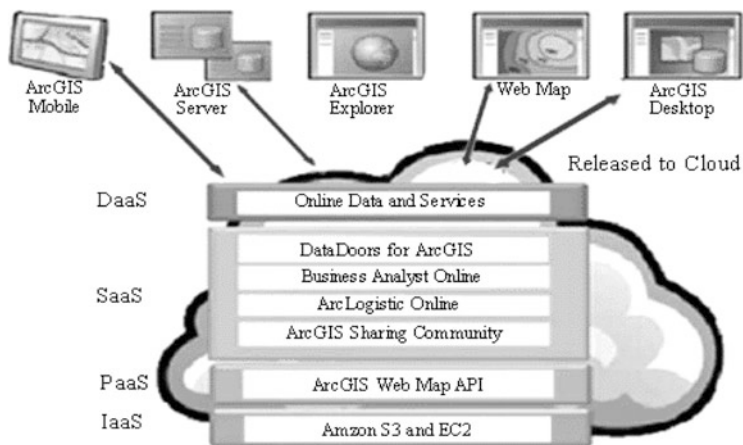


Fig. 102.3 The cloud application structure of ArcGIS

computing model to provide services for their customers, such as ESRI Business Analyst Online; it also allows user to combine GIS technology with thematic data, consumer data, and business data to achieve on-demand analysis and report and map web services, but Business Analyst Online is maintained by ESRI; users do not have to worry about data management and technical updates.

### ***102.4.2 Geographical Information Platform as a Service***

This service provides the whole GIS development environment to the user as a service. From ArcGIS9.2, ESRI began to promote its ArcGIS Online, and to provide GIS services by a series of API, GIS developers can develop GIS software in the Google App Engine platform and run on Google's cloud computing infrastructure. At home, GIS platform providers are basically only using the cloud computing model to provide services for their customers and partners, for example, GIS platform provider can create PaaS service from their own GIS platform, then, when its partners develop their various GIS, they do not need to buy the GIS platform license but only need to rent the platform even; don't install or deploy the GIS platform in local server but be able to directly carry out the development of GIS on the Internet; all that will bring great convenience and save a lot of cost for GIS partners. The cloud services platform of SuperMap GIS consist of geographic information cloud services, navigation product and spatial data processing. This platform is open to the third-party, they can add some applications and provide SaaS services to those end users.

### ***102.4.3 Geographic Information Infrastructure as a Service***

Geographic information infrastructure as a service is the basis of the "cloud" model, which is the foundation of geographical information software as a service and geographic information content as a service; therefore, this infrastructure environment and service mode are indispensable parts of cloud GIS geographic information services deployed in the cloud; then cloud GIS users can pay a monthly fee to rent the commercial cloud computing platform software and hardware resources. At present, these are the main companies: Amazon, IBM, and some of telecommunications, which provide the cloud infrastructure. Amazon provides two rent modes: Elastic Compute Cloud (namely, EC2) and Simple Storage Service (namely, S3). ESRI's ArcGIS cloud applications are built in Amazon's EC2 and S3; it provides cloud map slice services which can be uploaded to the cloud and establishes the data center in the clouds; users can put the map cache in the data center of Amazon cloud.

### 102.4.4 Geographic Information Content as a Service

Geographic information content as a service provides data, map information, and a simple query service for GIS end-user online; the user can access the content on demand, without the need to establish and maintain data, which is the lowest level of cloud GIS application. ESRI’s ArcGIS Online Map and GIS Server is a typical SaaS model, users can configure service on demand, and be able to quickly produce thematic maps, access seamless based map, etc. In addition, Baidu Map, Google Maps, Bing Maps, Yahoo Maps, etc., generally provide API for developers to use their cloud services; this API is a set of JavaScript or Flash language application programming interface, which can help users to build some function-rich, interactive map applications in their website.

### 102.5 Applications

Figure 102.4 shows the Geographic Information Society Service Platform of Jilin province based on SuperMap GIS cloud services platform.

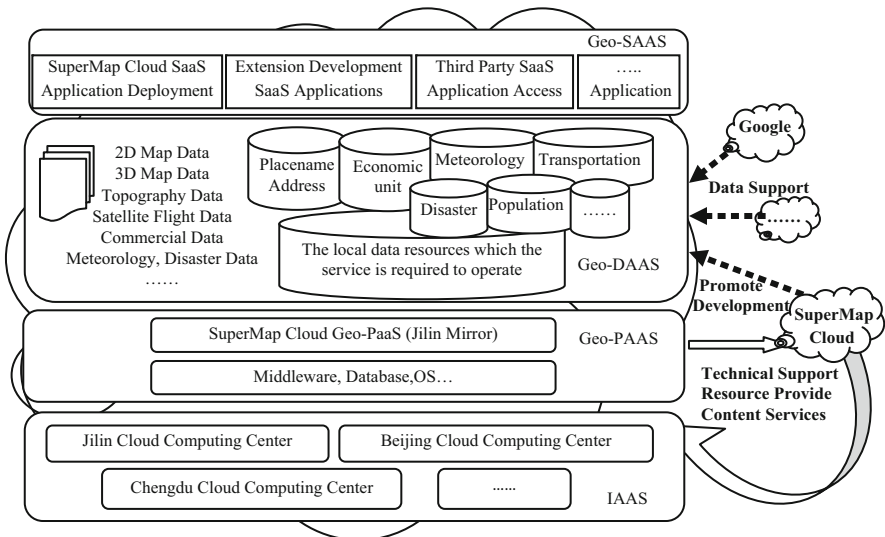


Fig. 102.4 The geographic information society service platform of Jilin province



## 102.6 Conclusion

With the development of cloud computing technology and the gradual and further application of GIS, cloud computing and GIS will be further fused, which can improve stability and efficiency services to end users by optimized network resource allocation of underlying data and services. But the cloud computing-based GIS in a real sense is required to a deep study; cloud GIS still has a long way to go and GIS rushes up high in the clouds also still to face great challenges. However, the arrival of the cloud era was an irresistible trend of development; cloud computing-based GIS is bound to be one of the GIS main development trends in the future.

**Acknowledgements** This work was supported by the Scientific and Technological Projects of Guangdong (No. 2009B010800042) and in part by Projects of Science and Technology of Dongguan (No. 201110825100119).

## References

1. Peter Mell, & Timothy Grance. (2011). *The NIST Definition of Cloud Computing*. <http://csrc.nist.gov/groups/SNS/cloud-computing/index.html>.
2. Michael Armbrust, Armando Fox, Rean Griffith, Anthony D. Joseph, Randy Katz, Andy Konwinski et al. (2009). *Above the clouds: A Berkeley view of cloud computing*. Berkley: University of California.
3. Liqian Dai, & Na Chen. (2009). The development of GIS in the times of cloud computing. *Journal of Anhui Agricultural Science*, 37(31), 15556–15557, 15572 (in Chinese).
4. Liu Yang. (2011). Research on GIS application model based on cloud computing. Henan University (in Chinese).
5. Alexander Lenk, Markus Klems, Jens Nimis, Stefan Tai, Thomas Sandholm. (2009). What is Inside the Cloud? An Architectural Map of the Cloud Landscape, Software Engineering Challenges of Cloud Computing, CLOUD '09. ICSE Workshop on, 23–31 (2009).
6. Yonggang Wang, Sheng Wang, & Daliang Zhou. (2009). Retrieving and indexing spatial data in the cloud computing environment. *Cloud Computing*, 12(4), 322–331 (in Chinese).
7. Er qi Wang. (2011). Cloud computing and GIS technology innovation. *New Economy Weekly*, 10(2), 83–87 (in Chinese).
8. Esri China Information Technology Co., Ltd. ArcGIS and Cloud Computing Technology. Esri technical white paper (2010) (in Chinese).
9. Fang Lei, Yao Shenjun, Liu Ting, & Liu Renyi. (2010). A cloud computing application in land resources information management. *IEEE*, 388–393

**Part V**  
**Embedded Systems**

# Chapter 103

## Memory Controller Design Based on Quadruple Modular Redundant Architecture

Yuanyuan Cui, Wei Li, and Xunying Zhang

**Abstract** For space application to improve the reliability of the memory operation, quadruple modular redundant (QMR) architecture is used in all registers of the memory controller. The QMR architecture in this paper can correct one-bit faults, detect two-bit faults, and also tolerate single event transient (SET). By modifying finite state machine (FSM) of the memory controller, when one uncorrectable fault is checked, the memory operation can be terminated in time and return the error information. Compared with triple modular redundancy (TMR), although the area overhead is increased by  $47,530.59297 \mu\text{m}^2$ , the single event upset (SEU) failure rate is lower by 6 orders of magnitude. Experimental results show that when 1 bit-flip or 2 bit-flips are injected in QMR registers, they can be corrected or detected in time, respectively. Memory controller using QMR architecture increases the area overhead, but the advantage is the higher reliability valuable for safety system.

### 103.1 Introduction

Dependability issues are the most important for space application. Due to the continuous increase in the integration level of electronic systems, an acceptable degree of reliability is increasingly difficult to be guaranteed. It is necessary to design fault-tolerant memory controller which is the important interface between the processor and the memory. To protect against SEU errors, the registers of the key modules in the system can often be implemented using TMR, for example, the LEON-FT processor [1], the SCS750 single board [2], and the Virtex FPGAs [3].

---

Y. Cui (✉) • W. Li

Graduate Department, Xi'an Microelectronics Technology Institute, Xi'an 710054, China  
e-mail: [hebutcyy@126.com](mailto:hebutcyy@126.com)

X. Zhang

Research and Development Department, Xi'an Microelectronics Technology Institute,  
Xi'an 710054, China

For the TMR architecture, any 1 bit-flip can be corrected, but the faults with multiple bit-flips can generate false results. The probability of multiple bit-flips is increasing as integrated circuits continue to scale into the deep micron regime. The system using TMR will not meet the requirement of reliability.

To improve the dependability of access memory, the QMR architecture is adopted in this paper. In recent decades, QMR has been successfully used in a wide variety of fields such as aviation and railroad [4–6]. Compared with TMR, although the area overhead is increased by  $47,530.59297 \mu\text{m}^2$ , the SEU failure rate of memory controller using QMR is lower by 6 orders of magnitude. By using separate clock trees to QMR registers, memory controller has the ability to tolerate SET. For QMR, any correctable error will be removed automatically within a clock cycle. When a group of QMR have 2 bit-flips, the detected error cannot be corrected. So in every clock cycle, the error detection signals of all QMR need to be examined. Using SMIC 130 nm standard CMOS process to synthesize, the error detection circuit's delay is 1.10 ns, meeting the design timing. By modifying FSM, any memory operation having a non-corrected error can be terminated in time and return the error information. Experimental results show that when 1 bit-flip or 2 bit-flips are injected in QMR registers, they can be corrected or detected in time, respectively. Memory controller using QMR increases the area overhead, but the advantage is the higher reliability valuable for safety system.

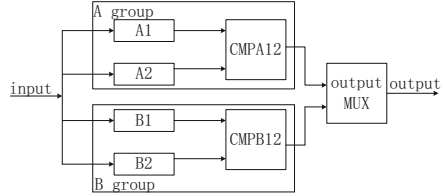
## 103.2 QMR Fault-Tolerant Technique

### 103.2.1 QMR Architecture

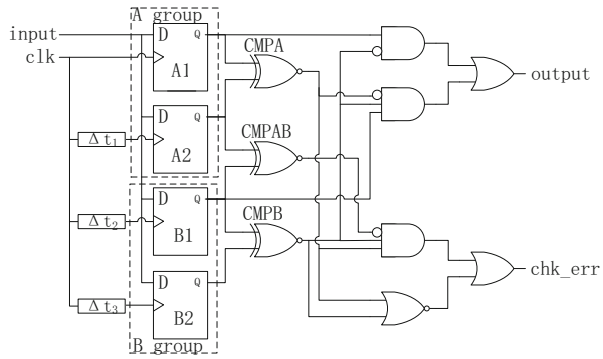
Figure 103.1 shows the QMR architecture. Every harden module needs to add three same ones. Every two modules constitute a group, namely, A group and B group. At the same time, one is a host group, another is a standby group. At the beginning, we assume the A group is the host group. If outputs of two modules from A group are the same, comparison circuit A12 will send a control signal to the output multiplexer and the output is from A group. If A12 always shows the results of A group are the same, the output is still from A group. When the results are different, B group will switch to the host group. In the same cycle, if the results from B group are the same, the output will be from B group; otherwise, the error detection signal will be set. Thus, QMR can correct and detect errors. The QMR is usually used in a subsystem which can self-repair. Because registers may be updated every cycle, they have the ability of restoring themselves naturally. So in Fig. 103.1, A1, A2, B1, and B2 can be registers.

Figure 103.2 shows the proposed circuit structure of the QMR. To simplify the design, we assume A group is still the host group, only if the results of the A group are different and that of the B group are the same; the output will be from the B group. When there are 2 bit-flips, the error cannot be corrected, but the error

**Fig. 103.1** QMR block diagram



**Fig. 103.2** QMR circuit structure



detection signal `chk_err` will be set. Compared with TMR, the delay of one “xor” logic is increased.

Due to SET faults, a signal from combinational circuit can have a glitch. When the glitch is clocked in a register, the register will have an error. To avoid it being captured into all four registers, each of the four lanes can have separate clock trees. Compared with the clock tree of register A1, phase differences are  $\Delta t_1$ ,  $\Delta t_2$ , and  $\Delta t_3$ , respectively. In the space environment, the width of the glitch is often 0.35~1.3 ns [7]; thus, the phase difference  $\Delta t$  is usually 1~1.5 ns. If  $\Delta t_1 = \Delta t$ ,  $\Delta t_2 = 2\Delta t$ , and  $\Delta t_3 = 3\Delta t$ , the glitch whose width is less than  $\Delta t$  can be captured only by one register; the error can be corrected. When the glitch can be captured only by two registers, the error will be detected.

Compared with TMR,  $3\Delta t$  is too large; system performance is reduced. A tradeoff scheme can be used; let  $\Delta t_1 = \Delta t$ ,  $\Delta t_2 = 0$ , and  $\Delta t_3 = 2\Delta t$ ; parts of glitches whose widths are less than  $\Delta t$  can be captured by A1 and B1 at the same time. This kind of errors cannot be corrected but can be detected. QMR with separate clock trees can protect against SEU errors and SET errors.

### 103.2.2 QMR Reliability

When the QMR registers have no error, or one bit-flip error, the output is correct; when two bits are flipped, the error information can be output, so this case is safe too. Let the reliability of single modular be  $R$ , so that QMR will be  $R_{QMR} = 6R^2 - 8R^3 + 3R^4$ . To make sure  $R_{QMR} > R$ ,  $R > 0.23$  must be satisfied. That means

when  $R > 0.23$ , the QMR architecture will just be valid. Reliability is a function of time. Commonly used reliability functions have exponent distribution function, normal distribution function, and Weibull distribution function. Exponent distribution function is often adopted, because exponent distribution is a single parameter distribution type and has broad applicability. So we suppose that reliability function obeys exponent distribution,  $R(t) = e^{-\lambda t}$ ;  $\lambda$  is SEU failure rate, and here, let  $\lambda = 10^{-6}$  error/(bit·day). Therefore,  $R_{QMR}(t) = 6e^{-2\lambda t} - 8e^{-3\lambda t} + 3e^{-4\lambda t}$ , the mean time between failures (MTBF) of QMR is

$$\begin{aligned} \text{MTBF} &= \int_0^\infty R_{QMR}(t)dt = \int_0^\infty 6e^{-2\lambda t} - 8e^{-3\lambda t} + 3e^{-4\lambda t} dt \\ &= 13/(12\lambda) > 1/\lambda = \int_0^\infty R(t)dt \end{aligned} \tag{103.1}$$

Formula (103.1) shows the integral result of  $R_{QMR}$  is greater than that of  $R(t)$  in  $0 \leq t < \infty$ . Because  $R(t)$  is a descending function, if  $t > -\ln 0.23/\lambda$ , then  $R(t)$  is less than 0.23, and  $R_{QMR}(t)$  is less than  $R(t)$ . In fact, due to limited battery life, undated life-span for the satellite system is impossible. The life of low Earth orbit satellites is about 3~10 years; that of geostationary Earth orbit and middle Earth orbit satellites is about 12~15 years. When  $t < -\ln 0.23/\lambda = 4,027$  years,  $R_{QMR}(t)$  is still greater than  $R(t)$ , so it is feasible to use the QMR architecture. For the TMR, when 3 registers have no error or one bit-flip, the output can be the correct value. Thus, TMR reliability is  $R_{TMR} = 3R^2 - 2R^3$ . To guarantee  $R_{TMR} > R$ ,  $R$  must be greater than 0.5. Let  $R(t) = e^{-\lambda t}$  and  $R_{TMR}(t) = 3e^{-2\lambda t} - 2e^{-3\lambda t}$ , the MTBF of TMR is

$$\begin{aligned} \text{MTBF} &= \int_0^\infty R_{TMR}(t)dt = \int_0^\infty 3e^{-2\lambda t} - 2e^{-3\lambda t} dt = 5/6\lambda < 1/\lambda \\ &= \int_0^\infty R(t)dt \end{aligned} \tag{103.2}$$

Formula (103.2) shows the integral result of  $R_{TMR}$  is less than that of  $R(t)$  in  $0 \leq t < \infty$ . But when  $t < \ln 2/\lambda = 1,899$  years, using TMR is also feasible. Figure 103.3 shows the relation of  $R(t)$ ,  $R_{TMR}(t)$ , and  $R_{QMR}(t)$ . The horizontal axis represents time, the vertical axis represents reliability. With the time growing,  $R(t)$ ,  $R_{QMR}(t)$ , and  $R_{TMR}(t)$  are all decreasing. When  $0 < t < \ln 2/\lambda$ ,  $R(t) > 0.5$ , so  $R_{QMR}$  and  $R_{TMR}$  are all greater than  $R(t)$ , namely, using TMR and QMR can all improve reliability. But when  $\ln 2/\lambda < t < -\ln 0.23/\lambda$ ,  $0.23 < R(t) < 0.5$ ,  $R_{QMR}(t) > R(t)$ , but  $R_{TMR} < R(t)$ , that is to say, using QMR is still feasible, but using TMR, the reliability is even lower than that of single modular. As Fig. 103.3 shows, the reliability of QMR is always greater than that of TMR.

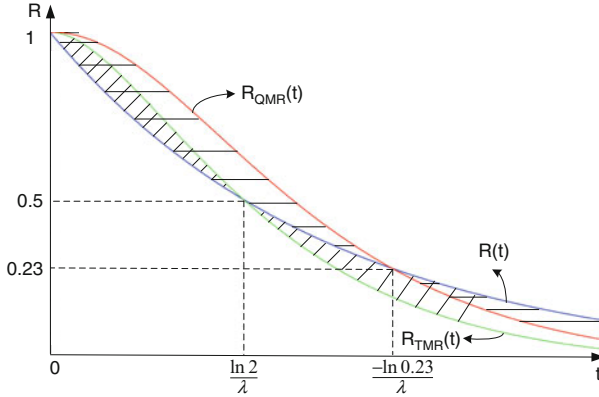


Fig. 103.3 Comparison of reliability of QMR and TMR

### 103.2.3 QMR SEU Failure Rate

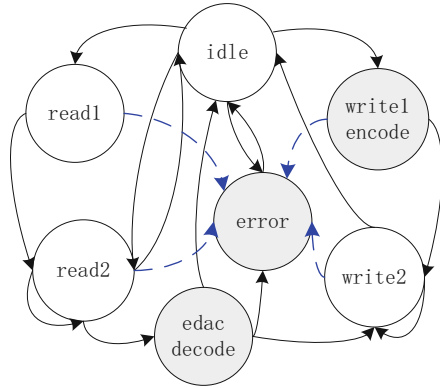
The failure rate function subjecting to exponential distributions is  $\lambda(t) = \lambda$ . To simplify the analysis, we assume SEU failure rate of every memory element is the same. Let SEU failure rate of single modular be  $\lambda$ ; thus, that of TMR is  $\lambda_{TMR} = 3\lambda^2 - 2\lambda^3$ , and that of QMR is  $\lambda_{QMR} = 4\lambda^3 - 3\lambda^4$ . The memory controller contains 464 registers; SEU failure rate of memory controller is  $\lambda_{memctrl} = 1 - (1 - \lambda)^{464} = 4.64e - 4$  error/(bit-day), which cannot meet the requirement of reliability. When using TMR,  $\lambda_{memctrl\_TMR} = 1 - (1 - \lambda_{TMR})^{464} = 1.39e - 9$  error/(bit-day); when using QMR,  $\lambda_{memctrl\_QMR} = 1 - (1 - \lambda_{QMR})^{464} = 1.86e - 15$  error/(bit-day). Compared with TMR, using QMR to all registers, using SMIC 130 nm standard CMOS process to synthesize, the area overhead of the memory controller is increased by 44,892.833344  $\mu m^2$ , but the SEU failure rate is lower by 6 orders of magnitude.

## 103.3 Design Implementation

Using QMR, there are two implementation modes. One is the memory controller as a subsystem; in Fig. 103.1, A1, A2, B1, and B2 are all memory controller modules. Another mode is all registers of the memory controller can be implemented using QMR. About area consumption, the front mode adds the area of combinational circuit, and the after mode adds the area of the comparison circuit and multiplexer of QMR for all registers. But SEU failure rate of the front mode is  $\lambda_{memctrl\_QMR} = 4\lambda_{memctrl}^3 - 3\lambda_{memctrl}^4 = 3.99e - 10$  error/(bit-day) and is greater by 5 orders of magnitude. So we select the after mode.

Figure 103.4 shows timing control of accessing memory operation. In an idle state, the request from the processor is still detected. When there is a read operation,

**Fig. 103.4** FSM with fault-tolerance design



switch to read1; the address will be given, chip select signal and read enable signal will be valid. In read2, the data from the external memory will be locked. To protect external memory using edac (error detection and correction) circuit, an edac state is added. In the edac state, if the decoder result is no error, the data will be propagated to the processor and will switch to idle. If there is a correctable error, switch to write2; the corrected data will be written back memory and will be sent to the processor. If there is a non-corrected error, switch to error state; memory error information will be propagated to processor. When there is a write operation, switch to write1; the address will be given, chip select signal will be valid, and the data and corresponding checksum will be prepared. In write2, write enable signal is still valid, and when write operation is completed, switch to idle state.

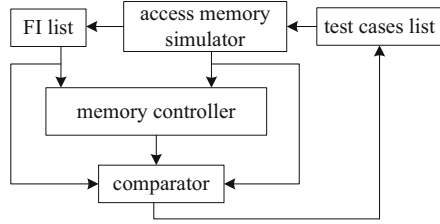
The QMR has error detection function; all error detection signals need to be checked every cycle. The memory controller contains 464 flip-flops, namely, 464 error detection signals do “or” operation and become a final error detection signal. This combinational circuit’s delay is 1.10 ns; for the system of clock frequency no more than 909 MHz, the delay meets design timing and the area overhead is 2,637.759626  $\mu\text{m}^2$ . By modifying FSM, when the final error detection signal is set, in whichever state the FSM is, it will switch to error state. In Fig. 103.4, broken lines indicate timing control modification. When some faults are checked, the operation of a memory can be terminated in time and can return the error information, improving the reliability of the accessing memory.

## 103.4 Experimental Results

To evaluate the reliability of our approach, it is necessary to inject faults in QMR registers and to code relevant test cases. Figure 103.5 shows simulated structure. Access memory simulator simulates operation timing from AHB bus. When a test case (an access memory operation) is running, the correlative fault can be injected, test case list and fault injection (FI) list are corresponding, and every group of QMR



**Fig. 103.5** Simulated structure



**Table 103.1** The simulation results

Category	Description	Bits	Case type		FI time	1 bit-flip cycles	2 bit-flips cycles
			Read	Write			
Rw_ctrl	Read and write control	79	9	70	read1/ write1	167	158
addr	Read/write address	32	0	32	write2	64	96
data	Read/write data	64	32	32	edac/write1	160	192
fsm_state	FSM state	4	0	4	write2	8	12
ahb_ctrl	Ahb control signals	14	0	14	write1	28	28
config	Config registers	102	0	102	write2	204	306
en_dec	Encode and decode info	169	160	9	edac/write1	498	658

registers has 4 bits FI signals RegN\_err[3:0] corresponding to A1, A2, B1, and B2 respectively. These signals are input of memory controller. When one FI signal is valid, the relevant register will have 1 bit-flip. Comparator receives the FI information and the number of cycles of the current operation; if the fault is 1 bit-flip, the ending timing of the corresponding test case is checked whether it matches the “okay” ending timing of AHB bus. If the fault is 2 bit-flips, the ending timing is checked whether it matches the “error” ending timing.

In this experiment, read access and write access use 3 cycles and 2 cycles, respectively, by configuration. Table 103.1 lists the various categories of registers and provides a description for each, as well as the number of bits of registers. One test case is coded for every group of QMR registers of every category. When the faults injected to the registers are only correlative with read access, test cases use read operation; otherwise, to save simulation time, write operations are used. The number of different types of test cases is given in the fourth and fifth columns. Each test case is executed twice; 1 bit-flip is injected for the first time, 2 bit-flips is injected second time, and FI time is shown in the sixth column. The 1 bit-flip can be self-repaired, the timing of access memory is not affected, and the cycles of running test cases are indicated in the seventh column. The 2 bit-flips are injected in the QMR registers at FI time; after one cycle, the test case would be finished; the simulated cycles are shown in the last column. The earlier the FI time is, the less the cycles of fault detection are. To correct the 2 bit-flips fault, except for the simulated cycles in the last column, accessing memory operation must be executed again, so more cycles are needed, but the compensation is higher reliability desired for the high-safety system.

## 103.5 Conclusion

Instead of TMR, all inner registers of memory controller are implemented using the QMR architecture; the area overhead is increased by  $47,530.59297 \mu\text{m}^2$ , but SEU failure rate of the memory controller is lower by 6 orders of magnitude. Although the error detection signal of every QMR needs to be examined every cycle, using SIMC 130 nm standard CMOS process to synthesize, this error detection circuit's critical path delay is 1.10 ns meeting design timing that is at most 909 MHz. By modifying FSM, when some faults are checked, the operation of a memory can be terminated in time, error information can be returned, and the reliability of the system can be improved. For the hardware system with high safety, if the overhead of area and power meet the design requirements, the QMR architecture can be used to reduce the failure rate of the key components.

## References

1. Gaisler, J. (2002). A portable and fault-tolerant microprocessor based on the SPARC v8 architecture. In *Proceedings International Conference on IEEE, USA* (pp. 409–415).
2. Longden, L., Thibodeau, C., Hillman, R., Layton, P., Williamson, G., & Dowd, M. (2002). Designing a single board computers for space using the most advanced processor and mitigation technologies. *European Space Components Conference*, Toulouse (pp. 313–316).
3. Carmichael, C. (2001). Triple module redundancy design techniques for virtex FPGAs. *Xilinx Application Note XAPP, 197*, 1–37.
4. Chen, G., Fan, D., & Wei, Z. (2010). All electronic computer interlocking system based on double 2-vote-2. *China Railway Science*, *31*(4), 138–144 (In Chinese).
5. Zhang, B., Lu, Y., Han, J., & Wei, Z. (2009). Reliability and security analysis of double 2-vote-2 redundancy system. *Journal of System Simulation*, *21*(1), 256–261 (In Chinese).
6. Zhang, J., Wang, H., & Jiang, D. (2006). Analysis of double 2-vote-2 fault-tolerant architecture used in computer-based interlocking system. *Computer and Telecommunication*, *15*(11), 46–49 (In Chinese).
7. Eaton, P., Benedetto, J., Mavis, D., Avery, K., Sibley, M., Gadlage, M., & Turflinger, T. (2004). Single event transient pulsewidth measurements using a variable temporal latch technique. *IEEE Transactions on Nuclear Science*, *51*(6), 3365–3368.

# Chapter 104

## Computer Power Management System Based on the Face Detection

Li Xie, Yong He, Yanfang Tian, and Tinghong Yang

**Abstract** In order to reduce the unnecessary power waste of computer system, the working principle of earlier Windows power management program and the new face recognition function of Windows 8 are analyzed in this paper. And the conflict between the convenience of use and the effects of energy conservation and environmental protection is given attention to. We put forward a new method based on the detection of frontal face in front of the monitor instead of the events of keyboard or mouse. Experimental results show that the method is a fast and effective one. Particularly, when user is leaving for a moment, this method is better than the work of Windows power management program. The results tell us that this method can save electrical energy about 4.28 % than windows power management program.

### 104.1 Introduction

With the development of society, there is a growing demand for energy. So the energy conservation and emission reduction will undoubtedly become a very noteworthy subject nowadays. However, due to the huge increase in the number of computer users, the lack of existing power management program, and many other factors, a great waste of energy is made on the use of the PC.

According to the data released in 2008 by the market research firm Gartner Inc., the global PC had been more than ten million, and the number still kept a steady

---

L. Xie (✉)  
Chongqing Electric Power College, Chongqing 400053, China  
e-mail: [43329588@qq.com](mailto:43329588@qq.com)

Y. He  
Chongqing Experimental High School, Chongqing 401320, China  
e-mail: [tumblerman@126.com](mailto:tumblerman@126.com)

Y. Tian • T. Yang  
Logistic Engineering University, Chongqing 401311, China

improvement in the growth rate of 12 % each year. If this improving speed continues, the number of global PC will reach 20 billion in 2014. For a so large number, a little waste on each computer will be unforgivable.

When using a computer, we often encounter emergencies which need a temporary leave such as answering phones, receiving express, and taking printed statements. For convenience, we generally do not turn off the computer. Occasionally once cannot result in much waste of energy, but the cumulative number of all the world's computers should not be ignored.

But the earlier system comes with power management program usually requires a relatively long reaction time (usually at least a few minutes) to enter a power-saving mode [1]. The traditional criterion to determine whether the system enters power saving mode is: "Is there keyboard click or mouse click?" It brings us some trouble to determine the response time. If the threshold of response time is set too long, the power saving effect is not obvious. But if it is too short when we just use the computer to read some documents and information, and do not use the keyboard and mouse for a long while, the system will in turn affect our work.

Recently, Microsoft's latest release of Windows 8 system provides a face recognition function. But this function is mainly focused on the management and login of user's accounts, and this way is found not safer than the early password security. Only needing a certain account user's photo, anyone can illegally log on to the system. In addition, the Windows 8 system's functions of detecting the user's arrival with the distance inductive sensor and automatically booting must obtain the support of sensor hardware. But for PC in current extensive use, this kind of sensor does not yet exist.

Therefore, according to the technical level of existing computer hardware, to develop a highly intelligent computer power-saving system will be of great significance for energy conservation.

## **104.2 Working Principle and System Structure**

### ***104.2.1 Working Principle***

"Is there a keyboard or mouse action" obviously has its limitations as a decision standard to enter the power-saving state, since it is difficult to obtain the compromise between the energy-saving effect and the normal using of users. In order to resolve this contradiction, we must find a way that gives a more accurate and clear description of these two states.

We find that the face is often positive facing on the monitor when someone is normally using a computer. Based on this premise, we can determine whether someone is using the computer according to if some positive face is in front of the monitor or not.

We use camera, a kind of commonly used external equipment, to collect the image in positive front of the monitor. Then, we detect faces through a face detection algorithm and control the computer's power-saving state according to the test results. In certain detection period (off-screen time, in seconds), the system will immediately turn off the monitor into the primary power-saving state when the computer does not sense the face because of the user's temporary leave. And once the user is back to the computer, it will immediately turn on the monitor and return to normal. Further, if no human face is detected for a very long time (standby time, in minutes), the system will go into a better state of power-saving standby. In this way, if there is a temporary or a long-term leave of user, the computer will give a quicker and more accurate judgment and turn off the monitor or enter system standby in time. All those come to our purpose of power saving.

### **104.2.2 Feasibility Analysis**

The face detection is a complex pattern recognition process. Its main difficulties are the inner face changing and the environmental impact on face to be detected. More details:

1. Very complex details in changing of faces, different physical characteristics such as face shape and color, and different expressions such as opening and closing of eyes and mouth
2. The blocking of other objects on the face, such as glasses, hair, head ornaments, and other external objects
3. The influence of light, such as image brightness, contrast variation, and shadow

All those set up obstacles for people to solve the problem of face detection [2].

In order to remove those difficulties, in this paper we take an existing approach with better robustness in the field of face detection, i.e., approximate Haar characteristics method, to classify the targets, better dealing with adverse factors such as light and color [3].

In addition, we find that the monitor power is about 30–100 W, the total power of a desktop computer is about 350 W, and the notebook is about 100 W. Compared with all those, the camera's working power is around 90 mW, which is negligible and does not increase the load on system with any operation. So our method is of a better feasibility.

### **104.2.3 System Structure**

This system mainly consists four parts: "system tray," "parameter setting module," "video processing module," and "hardware control module." The working principle is shown in Fig. 104.1.

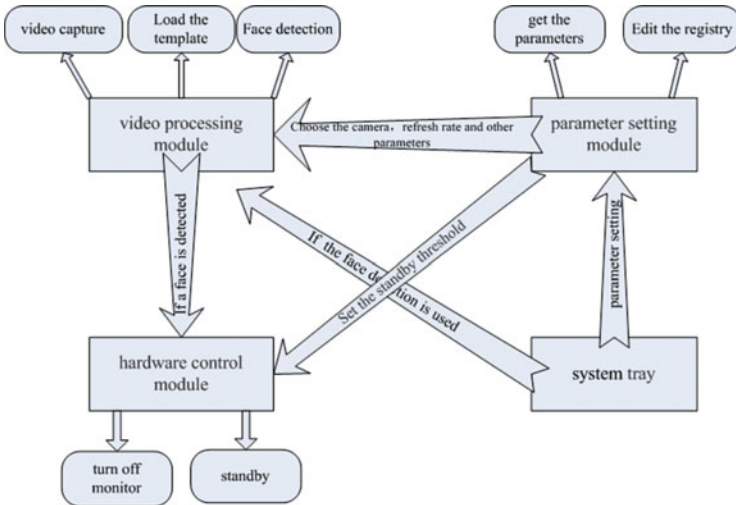


Fig. 104.1 Working principle of the system

The “system tray” is the input interface of all commands and parameters. In special cases, such as watching videos or listening to music, you can suspend or cancel the running of the system through this module. In most cases, the module is mainly used to set and modify the system parameters.

“Parameter setting module” is primarily responsible for the gaining and saving of system parameters, including the settings of power-saving parameters, video capture devices, and acquisition parameters. Through setting power-saving parameters, the system can control when the monitor should be turned off after the face detection fails and when the system enters standby mode. Through the setting of video capture devices and parameters, one can control the frequency of system detection and then control the sensitivity of face detection.

“Video processing module” is the core of this system, which is mainly responsible for detecting image in front of the monitor and examining whether there is user’s frontal face in image detected by face detection algorithm. If the system cannot detect the front face, and cannot reach the turning-off time threshold set by “parameter setting module” to tell “hardware control module” to turn off monitor, the “video processing module” will keep working before the system gets into deeper power-saving state. If the face is detected sometime later, the system will tell “hardware control module” to turn on the monitor. But if the face data is still not detected until it reaches the standby threshold, the “hardware control module” should be told to enter deeper power-saving state. The flow chart of video processing is shown in Fig. 104.2.

“Hardware control module” is for controlling monitor’s turning on and off, according to the system state informed by parameter setting module and the relevant parameters of video processing module. It also examines whether the system enters the standby and is responsible for the interaction with computer hardware.

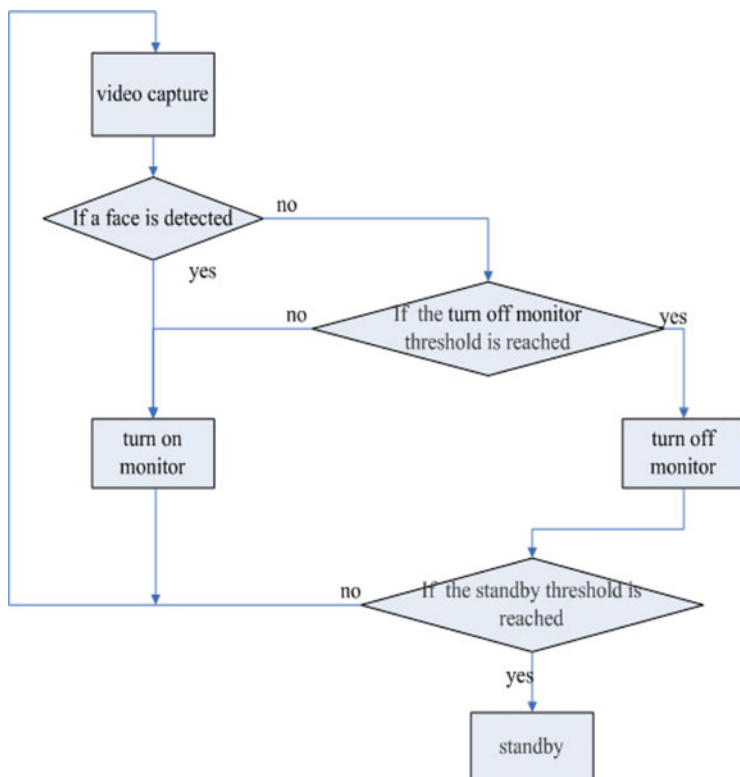


Fig. 104.2 Flow chart of video processing

### 104.2.4 System Implementation

In actual using, we develop the system with cross-platform programming language QT. Its core part is the face detection using a face detection algorithm supported by OpenCV (Open Source Computer Vision Library). Since it only works on examining whether there is a frontal face in all video images, but does not have to deal with face recognition, the algorithm is simple and is easy to handle with [4].

Face detection program completes three-part work, i.e., loading classifier, loading the image to be detected, as well as detecting and marking.

The target detecting classification of “haarcascade\_frontalface\_alt2.xml” file storage supported by OpenCV is used in this system [5]. The test results show that it basically meets the actual requirement, without the tedious steps of their own training classifiers.

The loading classifier program is as follows:

```

QString cascade = QCoreApplication::applicationDirPath().
    replace("/", "\\") +
    "\\haarcascade_frontalface_alt2.xml";
  
```

```

if( !cascade.load( cascadeName.toStdString() ) )
{
cerr << "ERROR: Could not load classifier cascade" << endl;
return -1;
}

```

“Loading the image to be detected” and “detecting and marking” are two periodically performed steps. They are mainly based on the refresh time of system parameters setting to periodically collect image in front of the screen through the camera and to give the appropriate backup for calling of system detection part.

```

IplImage* iplImg = cvQueryFrame( capture );
frame = iplImg;
if( frame.empty() ) return -1;
if( iplImg->origin == IPL_ORIGIN_TL )
frame.copyTo( frameCopy );
else
flip( frame, frameCopy, 0 );
int r=detectAndDraw( frameCopy, cascade, scale ,disp);
where the“detectAndDraw” function codes are:
int Facedetect::detectAndDraw( Mat&img, CascadeClassifier&
cascade, double scale, bool disp)
{
int i = 0;
vector<Rect> faces;
const static Scalar colors[] =
{
CV_RGB(0,0,255),
CV_RGB(0,128,255),
CV_RGB(0,255,255),
CV_RGB(0,255,0),
CV_RGB(255,128,0),
CV_RGB(255,255,0),
CV_RGB(255,0,0),
CV_RGB(255,0,255)
};
Mat gray, smallImg( cvRound( img.rows/scale),
cvRound( img.cols/scale), CV_8UC1 );
cvtColor( img, gray, CV_BGR2GRAY );
resize( gray, smallImg, smallImg.size(), 0, 0, INTER_LINEAR );
equalizeHist( smallImg, smallImg );
cascade.detectMultiScale( smallImg, faces, 1.1, 2, 0, Size
(20, 20));
int radius=0;

```



```

for( vector<Rect>::const_iterator r = faces.begin(); r !=
    faces.end(); r++, i++)
{
    Point center;
    Scalar color = colors[i%8];

    center.x = cvRound((r->x + r->width*0.5)*scale);
    center.y = cvRound((r->y + r->height*0.5)*scale);
    radius = cvRound((r->width + r->height)*0.25*scale);
    if (disp==true)
    {
        circle (img, center, radius, color, 3, 8, 0);
    }
}
if (disp==true)
{
    cv::imshow( " Face detection results ", img );
}
else
{
    cvDestroyWindow("Face detection results ");
}
return radius;
}

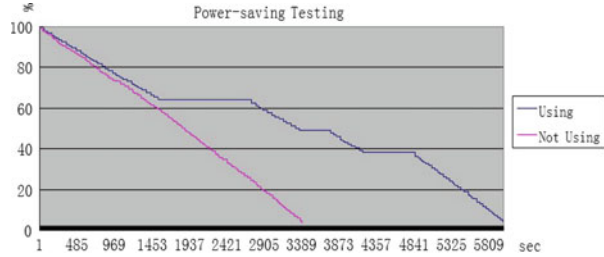
```

### 104.3 System Testing

In order to test the performance of the system, we use a laptop to do a power discharging test, respectively, under normal operating conditions (programming and Internet) to use the software and not to use until the system naturally turns off. The test data is shown in Fig. 104.3. Obviously, the power-saving effect is significantly superior to the effect of not using the software. This advantage comes mainly from power saving of closing the screen system standby while there is temporary leave.

The system chooses the power-saving strategy of turning off monitor to deal with temporary leave. In actual test, there are temporary leaves about six times in former 1,540 s, each time about 5 s to 5 min (due to the density of coordinate, the display is not obvious after screen-off, but the cumulative power discharging trend after multiple screen-off is more gentle). And at the time 1,540 s, the measured remaining power is, respectively, 64.39 and 60.11 %, i.e., there is a saving approximately 4.28 % of electricity.

**Fig. 104.3** Power-saving testing data



For user's longtime leaving, the system chooses standby. Shown at the 1,541–2,742, 3,371–3,771, and 4,189–4,861 s in Fig. 104.3, power consumption is almost negligible.

## 104.4 Conclusion

The “computer power management system based on face detection” works on the condition of existing computer hardware level. And in addition to the long system standby strategy of conventional power management system, when the user is temporarily away, this face detection system can more quickly give the accurate judgment, shut down the monitor, and save power. Therefore, using of face detection and recognition technology can solve the problems of slower response time and inaccurate judgment. It also can save unnecessary energy consumption of a computer system, and is of high practical value and promoting significance.

## References

1. Zhao, X., Chen, X., Guo, Y., & Yang, F. (2008). The research progress of operating system's power management. *Computer Research and Development*, 45(5), 817–824.
2. Chen, Z., & Jiang, M. (2012). Designing of face detection system based on the OpenCV. *Electronic Design Engineering*, 20(10), 82–185.
3. He, Y., & Li, G. (2012). Face detecting, tracking and feature points positioning system. *Electronic Design Engineering*, 20(8), 189–192.
4. Zhang, Y., & Li, Y. (2011). General face detection module designing based on the OpenCV. *Computer Engineering and Science*, 33(1), 97–101.
5. Tao, Y. (2012). Face recognition applications based on the OpenCV. *Computer Systems and Applications*, 21(3), 220–223.

# Chapter 105

## Twist Rotation Deformation of Titanium Sheet Metal in Laser Curve Bending Based on Finite Element Analysis

Peng Zhang, Qian Su, and Dong Luan

**Abstract** Laser sheet bending is a new metal forming process realized by thermal stresses resulted from the irradiation of laser beam scanning. Laser forming is a new type of sheet metal forming process. The sheet metal is formed by asymmetrical thermal stresses. The three-dimensional elastoplastic thermomechanical coupled finite element model of laser bending for Ti-6Al-4V plates was established with nonlinear finite element analysis software ANSYS. The bending properties of sheet metal with different processing parameters were simulated. The results show that the twist rotation deformation of sheet metal can be influenced by laser power, spot diameter, scanning velocity, scanning path curvature, and the distance between scanning path and free end.

### 105.1 Introduction

The laser is a kind of tool, and common light source cannot be compared with it because of high purity, high brightness, high coherence, and high directivity, so it is used widely. Laser processing technology is greatly used in the area of cutting, welding, and surface treatment with the features of high energy injection rate, low hot influence area, easy guidance, being not affected by electromagnetism, high speed machining, no tool wear, and noise pollution [1–3]. Based on the characteristic of thermal expansion and contraction, laser sheet bending is a new metal forming process realized by thermal stresses resulted from the irradiation of laser beam scanning. Compared with other conventional machine forming methods, laser sheet bending has many advantages such as on die molding, noncontact molding, no external force molding, hard-to-deformation thermal normal cumulative forming, and laser beam mode without specific requirements [4, 5].

---

P. Zhang (✉) • Q. Su • D. Luan  
School of Materials Science and Engineering, Harbin Institute of Technology at Weihai,  
Weihai, Shandong 264209, China  
e-mail: [pzhang@yeah.net](mailto:pzhang@yeah.net)

When trajectory of the laser beam is linear compared with sheet, V-shape parts are obtained; when trajectory is non-repeated or curve, composite curved special-shaped parts are obtained, such as cylindrical parts, disc-shaped parts, spherical part, and a variety of complex shapes three-dimensional shaped pieces [6]. The deformation process of sheet metal is very complicated, especially the residual stress of preorder deformation and the influence of geometry for deformation after unloading; with the curve scan path and because of the thermal effect of laser beam, the sheet also produces torsional deformation in addition to the bending deformation in author's previous studies [7]. Therefore, studying the torsional deformation of the sheet in the process of laser radiation, analyzing influence law of the sheet torsional deformation's laser processing parameters can lay a foundation for the implementation of the sheet laser thermal stress precision forming.

## 105.2 Finite Element Model of Laser Bending

Laser bending is a complex process of interaction of many factors, such as sheet performance parameters, geometric parameters, and laser processing parameters, and designing the processing technology by experimental method will spend much time and human and material resources. With the rapid development of computational mathematics and computer technology as well as the improvement and perfect of finite element algorithm, using the numerical simulation method to simulate practical production has been proved that it is an effective way and has a huge potential [8–12].

### 105.2.1 *Physical Model*

In view of the fact that laser bending process is complex and influenced by many factors, it is difficult to establish the model to accurately reflect the actual situation, so this paper makes hypothesis to simplify calculation model as follows: (1) The freedom of all nodes of one end of the sheet should be restricted to avoid the sheet occurring rigid body displacement, and this is consist with the actual situation; (2) the laser beam moves at equal speed and irradiates the surface of the sheet vertically; (3) the material is isotropic, and the thermal physics and mechanical properties change with temperature; (4) the thermal absorption coefficient of material is constant; and (5) the process of scanning is carried out under the melting point of the material.

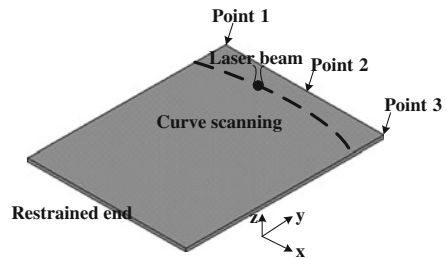
### 105.2.2 Movable Heat Source Based on Scanning Path

The three-dimensional elastoplastic thermomechanical coupled finite element model of laser bending is established with nonlinear finite element analysis software ANSYS, and the model uses three-dimensional eight-node hexahedral elements. Through ANSYS secondary development, subroutine of Gauss heat source based on scanning path is established, and laser thermal load is applied to the respective units in the form of heat flux. Scanning path uses parametric curve designing, and cubic B-spline curve can describe the scanning path easily and accurately.

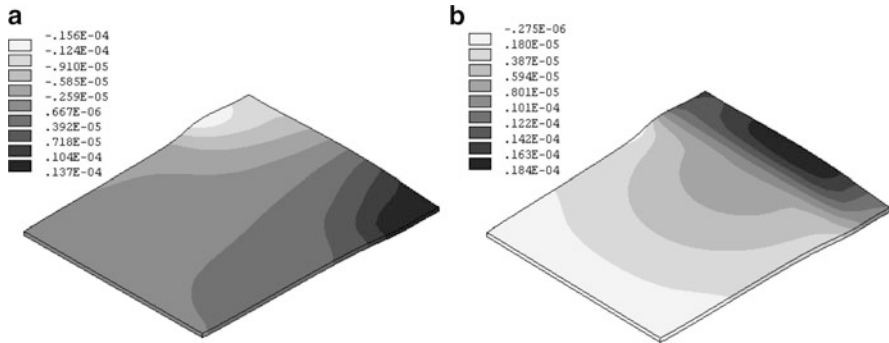
Figure 105.1 shows the scheme of laser curve scanning. Select the displacement of the free end of the sheet on the sideline three-point  $z$  to calibrate the influence of technical parameters for laser bending. The material of the sheet is Ti-6Al-4V, its performance parameter comes from literature [13], and geometry size is  $50 \times 40 \times 0.8$  mm.

### 105.3 The Twist of Sheets During Laser Curve Bending

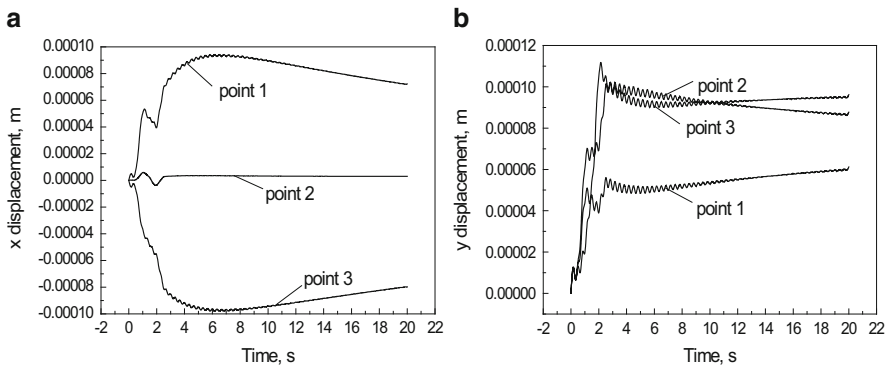
Figure 105.2 shows how the displacement field of sheets laser curve scanning distributes. Figure 105.3 shows how the movement of three points demarcated on sheets varies. It can be seen that the movement of Point 1 on  $x$ -direction always belongs to stretch deformation, and the movement of Point 3 on  $x$ -direction always belongs to compression deformation. And Point 2 hardly deforms on  $x$ -direction, only when laser beam goes by the region near Point 2 does it generate tiny deformation. There are different degree  $y$ -direction movements on the three points demarcated. Because the beginning of the laser beam scanning is asymmetric, the  $y$ -direction movement of Point 1 is minor than Point 2 and Point 3.



**Fig. 105.1** Sketch of laser scanning scheme



**Fig. 105.2** Displacement of laser curve scanning at  $t = 0.6$  s. (a) Displacement of  $x$  (b) Displacement of  $y$



**Fig. 105.3** Variation of displacement with time of laser curve scanning

### 105.3.1 Laser Power

Figure 105.4 shows how the laser power influences the twist of sheets. It can be seen that with the increasing of the laser power, the  $x$ -direction movement and  $y$ -direction movement of sheets increase. The main reason is that the thermal expansion increases when the input energy increases.

### 105.3.2 Spot Diameter

Figure 105.5 shows how the spot diameter influences the twist of sheets. It can be seen that with the increasing of the spot diameter, the  $x$ -direction movement of

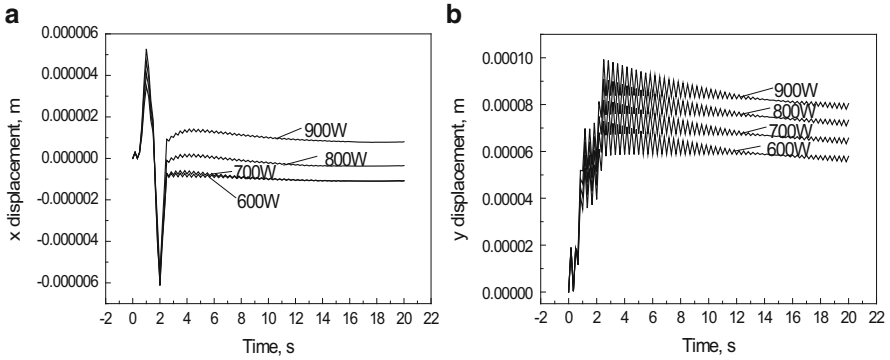


Fig. 105.4 Effect of laser power on twist rotation deformation

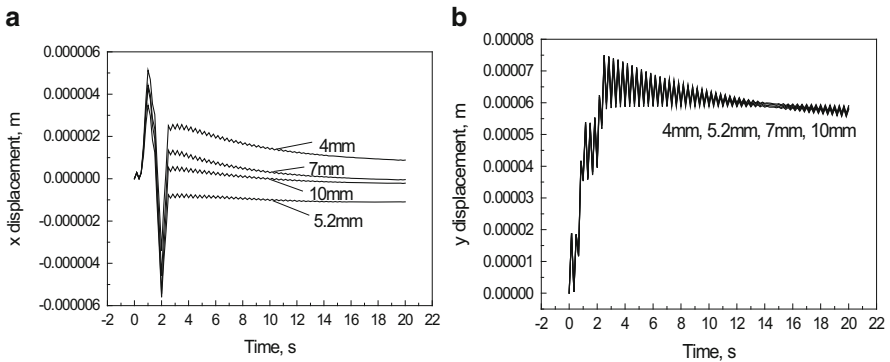


Fig. 105.5 Effect of spot diameter on twist rotation deformation

sheets decreases, while there is no significant change about y-direction movement of sheets. It can explain that y-direction movement of sheets is not sensitive to the spot diameter changes.

### 105.3.3 Scanning Speed

Figure 105.6 shows how the scanning speed influences the twist of sheets. It can be seen that with the increasing of the scanning speed, both the x-direction movement and the y-direction movement of sheets increase at first then decrease. When other progress parameters are certain, there is some scanning speed, which makes the x-direction movement and y-direction movement of sheets maximum.

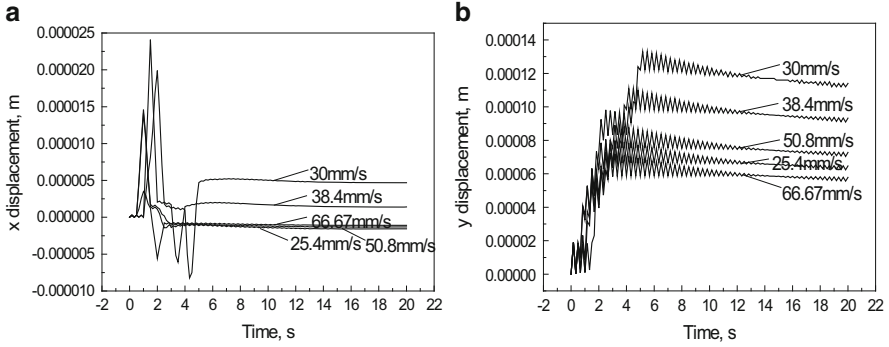


Fig. 105.6 Effect of scanning velocity on twist rotation deformation

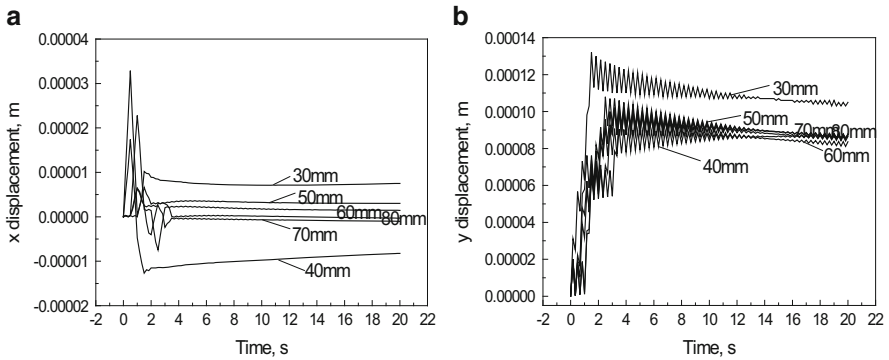


Fig. 105.7 Effect of scanning path curvature on twist rotation deformation

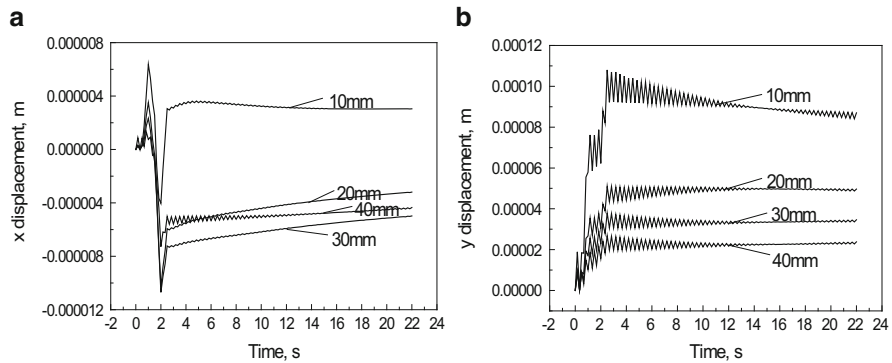
### 105.3.4 The Curvature of Scanning Paths

Figure 105.7 shows how the curvature of scanning paths influences the twist of sheets. It can be seen that with the increasing of the curvature of scanning paths, the *x*-direction movement of sheets decreases at first then increase, while the *y*-direction movement of sheets decreases at first then does not change significantly.

### 105.3.5 The Distance Between Scanning Paths and Free End

Figure 105.8 shows how the distance between scanning paths and free end influences the twist of sheets. It can be seen that with the increasing of the distance





**Fig. 105.8** Effect of distance between scanning path and free end on twist rotation deformation

between scanning paths and free end, the  $x$ -direction movement of sheets decreases at first then increase, while the  $y$ -direction movement of sheets decreases.

The laser power, the spot diameter, and the scanning speed are energy factors, and the increase of input energy leads to increasing in thermal expansion of sheets. The curvature of scanning paths has influence on the input energy and the size of heat-affected zone. The distance between scanning paths and free end has influence on the effect of the material between scanning line and free end on stiff constraint of heating zone.

## 105.4 Conclusion

1. Based upon the three-dimensional elastoplastic thermomechanical coupled finite element model of laser bending for Ti-6Al-4V plates, the bending properties of sheet metal with different processing parameters were simulated.
2. This study analyzed how the different processing parameters influence the twist of sheets. The increase of laser input energy would lead to an increase in thermal expansion of sheets, and the curvature of scanning paths can influence both the input of energy and the size of heat-affected zone, while the distance between scanning paths and free end can influence the effect of the material between scanning line and free end on stiff constraint of heating zone.

## References

1. Raya, T., & Umino, T. (1991). Present status of CO<sub>2</sub>-laser processing. *Journal of Materials Engineering*, 13(4), 299–360.
2. Geiger, M. (1993). Synergy of laser material processing and metal forming. *CIRP Annals*, 43(2), 563–570.

3. Washio, K., Takenaka, H., Okino, K., Aruga, S., Matsui, E., & Kyusho, Y. (1992). Welding and cutting car-body metal sheets with fiber delivered output from high-power ND-YAG lasers. *NEC Research and Development*, 33(1), 102–109.
4. Thomson, G. (1998). Improvements to laser forming through process control refinements. *Optics and Laser Technology*, 2(30), 141–146.
5. Gisario, A., Barletta, M., Conti, C., & Guarino, S. (2011). Springback control in sheet metal bending by laser-assisted bending: Experimental analysis, empirical and neural network modelling. *Optics and Lasers in Engineering*, 49(12), 1372–1383.
6. Hennige, T. (2000). Development of irradiation strategies for 3D-laser forming. *Journal of Materials Processing Technology*, 103(1), 102–108.
7. Zhang, P., Guo, B., Shan, D. B., & Ji, Z. (2007). FE simulation of laser curve bending of sheet metals. *Journal of Materials Processing Technology*, 184(1–3), 157–162.
8. Maji, K., Pratihari, D. K., & Nath, A. K. (2013). Analysis and synthesis of laser forming process using neural networks and neuro-fuzzy inference system. *Soft Computing*, 17(5), 849–865.
9. Ji, Z., & Wu, S. C. (1998). FEM simulations of the temperature field during the laser forming of sheet metal. *Journal of Materials Processing Technology*, 74(2–3), 89–95.
10. Wu, S. C., & Ji, Z. (2002). FEM simulation of the deformation field during the laser forming of sheet metal. *Journal of Materials Processing Technology*, 121(2–3), 269–272.
11. Kyrsanidi, A. K., Kermanidis, T. B., & Pantelakis, S. G. (1999). Numerical and experimental investigation of the laser forming process. *Journal of Materials Processing Technology*, 87(1–3), 281–290.
12. Chen, D. J., Wu, S. C., & Li, M. Q. (2004). Studies on laser forming of Ti-6Al-4V alloy sheet. *Journal of Materials Processing Technology*, 152(1), 62–65.
13. Ren, X. D., Zhang, R. K., Zhou, J. Z., Zhang, X. Q., & Lu, X. Z. (2006). Influence of laser parameters on laser-shock forming of Ti-6Al-4V alloy. *Chinese Journal of Nonferrous Metals*, 16(11), 1850–1854 (In Chinese).

# Chapter 106

## Voltage Transient Stability Analysis by Changing the Control Modes of the Wind Generator

Yu Shao, Feng Shi, and Xiang Li

**Abstract** The chapter studies voltage transient stability when the wind generator changes its control modes. The chapter studies the influence caused by connection with wind farms based on simulation and makes comparison between different control modes, then gives the conclusion. The chapter takes the real grid model and the result of the study has some means to the relative study.

### 106.1 Introduction

Problems on voltage stability have the direct bearing on the safe operation of the grid, and it hopes that the wind turbine itself could regulate the reactive power. Most studies are always using ideal models. The chapter used the real model of a regional grid to simulate transient stability under two control modes, which had great persuasion to the study of the modern grid.

The structure of the chapter is as follows: The classification of the control modes is described in Sect. 106.2. Section 106.3 presents the problem of voltage stability with the connection of wind farms. Section 106.4 presents the simulation. Section 106.5 concludes the whole chapter.

### 106.2 Classification of the Control Modes

The chapter compared the two control modes, such as the power factor constant mode and voltage constant mode. The characteristic of the power factor constant mode is that wind turbine follows the change of the power in transient process, regulating the

---

Y. Shao (✉) • F. Shi • X. Li  
Zhengzhou Electric Power Supply Company, Zhengzhou 450000, China  
e-mail: [672851649@qq.com](mailto:672851649@qq.com)

relative power to keep the power factor constant. The characteristic of voltage constant mode is that wind turbine follows the change of the voltage in transient process, regulating the relative power to keep the power factor constant [1].

## **106.3 Voltage Stability with the Connection of Wind Farms**

### ***106.3.1 Capacity Effect of Wind Farms***

The combination of the large and small wind farms takes different problems. There are two circumstances. One is that a large wind farm connects with EHV grid and the other is that several small wind farms connect with grid intensively. Small wind farm would affect the grid in several parts. Large wind farm would induce voltage vibration in the grid sometimes [2].

### ***106.3.2 Influence Caused by Characteristics of Wind Farm***

1. Type of wind turbine. The types of wind turbine widely used in the grid are constant speed turbine based on asynchronous motor and turbine based on DFIG (Double-Fed Induction Generator).
2. Pneumatic power control technology. The technology mainly includes fixed pitch control, non-fixed pitch control, and active stall control. Pitch angle control technology could strengthen voltage stability in a fault.
3. Level penetration. It means the proportion of the whole wind capacity connected to the grid with the load. High level means large capacity of wind was connected to the grid and the proportion of conventional generators was low [3].

### ***106.3.3 The Influence Caused by Reactive Power Compensator***

The asynchronous motor needs more reactive power when it is on operation and reactive power compensator is needed to be installed. The control of condenser bank is discrete and the need of reactive power raised by the change of active power on every compensator stage still be provided by the grid. Switches of condenser bank would cause jumps of voltage. Besides, the characteristics of condenser bank cut lots of reactive power when the voltage depressed, which made the need of reactive power raise in wind farm and worsen the voltage to collapse.

## 106.4 Simulation

### 106.4.1 Models of DFIG

Dynamic models of DFIG are composed of generator model, rotation control model, and reactive power model. The voltage equation based on d, q axes is shown in formula Eq. (106.1) [4]:

$$\begin{cases} U_{ds} = p\Psi_{ds} - \Psi_{qs} + r_s I_{ds} \\ U_{qs} = p\Psi_{qs} + \Psi_{ds} + r_s I_{qs} \\ U_{dr} = p\Psi_{dr} - s\Psi_{qr} + r_r I_{dr} \\ U_{qr} = p\Psi_{qr} + s\Psi_{dr} + r_r I_{qr} \end{cases} \quad (106.1)$$

In equations,  $U_{ds}$ ,  $U_{qs}$  are voltage on stator and  $U_{dr}$ ,  $U_{qr}$  are voltage on rotor.  $I_{ds}$ ,  $I_{qs}$  are current on stator and  $I_{dr}$ ,  $I_{qr}$  are current on rotor.  $\Psi_{ds}$ ,  $\Psi_{qs}$  are magnetic linkage on stator and  $\Psi_{dr}$ ,  $\Psi_{qr}$  are magnetic linkage on rotor.  $p = d/dt$  is differential operator. Let  $p\Psi_{ds} = p\Psi_{qs} = 0$ ,  $r_s = 0$ ; the dynamic equation is

$$\begin{bmatrix} U_{ds} \\ U_{qs} \\ U_{dr} \\ U_{qr} \end{bmatrix} = \begin{bmatrix} 0 & -X_{ss} & 0 & -x_m \\ X_{ss} & 0 & x_m & 0 \\ px_m & -sx_m & r_r + pX_{rr} & -sX_{rr} \\ sx_m & px_m & sX_{rr} & r_r + pX_{rr} \end{bmatrix} \begin{bmatrix} I_{ds} \\ I_{qs} \\ I_{dr} \\ I_{qr} \end{bmatrix} \quad (106.2)$$

Represent the dynamic equation on axes d and q:

$$\begin{cases} \frac{dE'_d}{dt} = -\frac{x_m}{X_{rr}} U_{qr} + sE'_q - \frac{1}{T'_{d0}} [E'_d + (X_{ss} - x') I_{qs}] \\ \frac{dE'_q}{dt} = -\frac{x_m}{X_{rr}} U_{dr} + sE'_d - \frac{1}{T'_{d0}} [E'_q + (X_{ss} - x') I_{ds}] \end{cases} \quad (106.3)$$

$$\begin{cases} E'_d = x' I_{qs} \\ |\dot{U}_s| - E'_q = x' I_{ds} \end{cases} \quad (106.4)$$

Under the influence of magnetic field on stator,  $U_{ds} = 0$  and  $M_e$  are decided by voltage and current on stator:  $M_e = |\dot{U}_s| I_{qs}$ . And  $|\dot{U}_s|$  represents amplitude of voltage.

The regulations of reactive power output are realized by current of axes d and q on rotor, called vector control. Current on rotor is regulated by  $U_{dr}$  and  $U_{qr}$ . In order to get the relation between voltage and magnetic linkage on rotor, equations of them are shown [5]:

$$\begin{cases} U_{dr} = p\Psi_{dr} - s\Psi_{qr} + r_r I_{dr} \\ U_{qr} = p\Psi_{qr} + s\Psi_{dr} + r_r I_{qr} \end{cases} \tag{106.5}$$

$$\begin{cases} \Psi_{dr} = X_{rr} I_{dr} + x_m I_{ds} \\ \Psi_{qr} = X_{rr} I_{qr} + x_m I_{qs} \end{cases} \tag{106.6}$$

When  $X' = x_{rr} - \frac{x_m^2}{X_{ss}}$ :

$$\begin{cases} U_{dr} = (r_r + X' p) I_{dr} - s X' I_{qr} \\ U_{qr} = (r_r + X' p) I_{qr} + s X' I_{dr} + r_r I_{qr} + \frac{s X_m}{X_{ss}} |\dot{U}_s| \end{cases} \tag{106.7}$$

Equation (7) is an electromagnetic transient equation. And the dynamic equation of rotor is  $\frac{ds}{dt} = \frac{1}{T_J} (M_m - M_e)$ .  $s$  is slip of generator.  $T_J$  is inertia time constant of rotor.  $M_m$  is machine torque.  $M_e$  is electromagnetic torque.

Electromagnetic transient equation showed that rotate speed could be changed by regulating electromagnetic torque. The control system could regulate the reactive power according to the aforementioned. The article used BPA software, and typical parameters of GE1.5MW were used in the simulation.

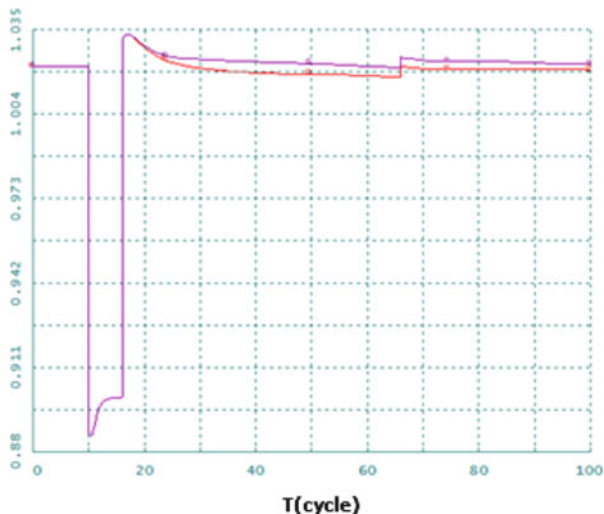
### 106.4.2 Simulation Under Different Control Modes

A small wind farm was connected to a 220 kV grid. The capacity of wind farm was 90 MW. The capacity per turbine was 1.5 MW. The wind farm was connected by a single line. A one-phase short-circuit fault happened on a 220 kV line in the grid. The fault happened at the 10th cycle; breakers at both ends of the line acted at the 16th cycle and reclosed at the 56th cycle successfully. Voltage of connection point and turbine are shown in Figs. 106.1 and 106.2. The reactive power of turbine is shown in Fig. 106.3. The red line represents for constant power factor control and the purple one represents for constant voltage control.

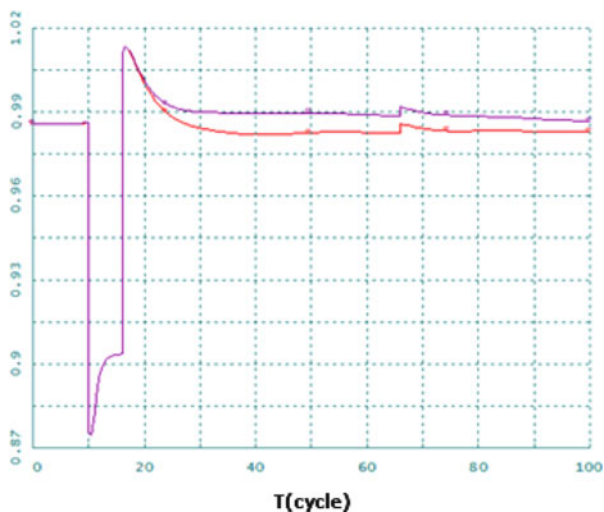
According to Figs. 106.1, 106.2, and 106.3, voltage of the grid could recover under two kinds of control. Differences are as follows: (1) For voltage of connection point and turbine, constant voltage control could make the level of voltage higher. (2) For reactive power of turbine, constant voltage control distributed more to the voltage by letting turbine made more reactive power output. In sum, constant voltage control is more positive to the voltage stability.

Reactive power made by turbines is far less than the shortage of the grid, and the reactive power made by automatic field forcing is 0.6 Mvar without other compensators, which distributes less to the grid voltage. So in process of a fault, automatic

**Fig. 106.1** Voltage of connection point



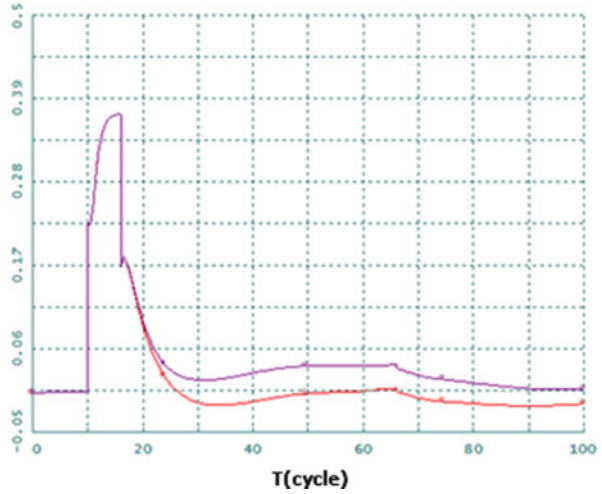
**Fig. 106.2** Voltage of the turbine



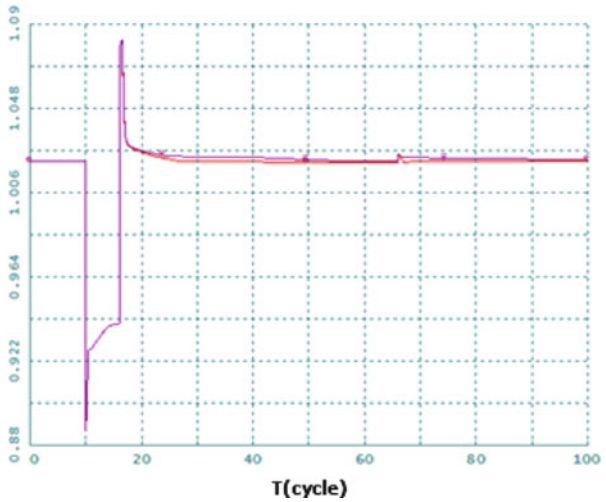
field forcing could not improve voltage of connection point and turbine obviously, and low voltage protection would cut turbines, which could cause vibration of the grid.

Installation of SVC could improve voltage transient stability. There would be a same simulation with the wind-farm-installed SVC, checking the effects of SVC. With the installation of SVC, the simulation waves are shown in Figs. 106.4, 106.5, 106.6, and 106.7. The red line represents for constant power factor control and the purple one represents for constant voltage control.

**Fig. 106.3** Reactive power of turbine



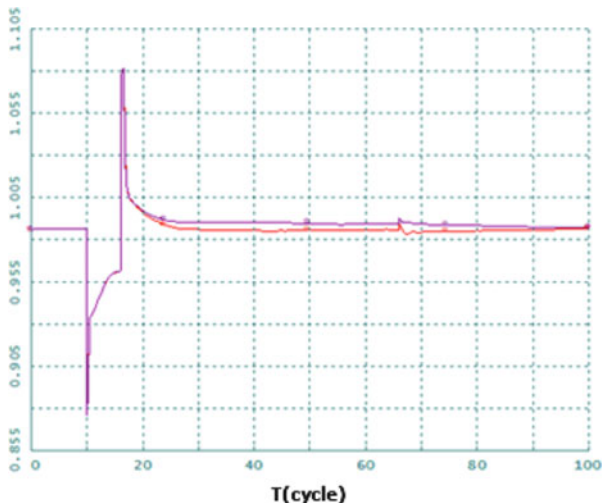
**Fig. 106.4** Voltage of connection point



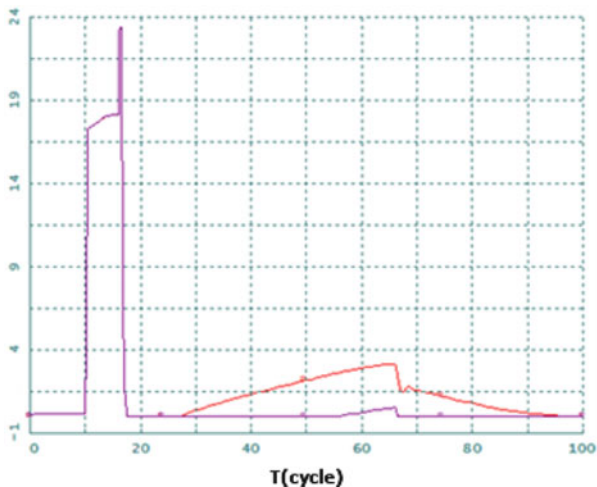
According to Figs. 106.4, 106.5, 106.6, and 106.7 and Tables 106.1 and 106.2, voltage of connection point and turbine is same under two types of control with SVC. The reason is that the reactive power generated by turbines is 0.3 Mvar, which is far less than 23 Mvar generated by SVC, and the recovery of voltage is mainly determined by SVC. When constant voltage control is used, the level of voltage is higher after recovering from the fault. Besides, time of voltage under 0.9 p.u is short and reduces the probability of cutting turbines.



**Fig. 106.5** Voltage of the turbine



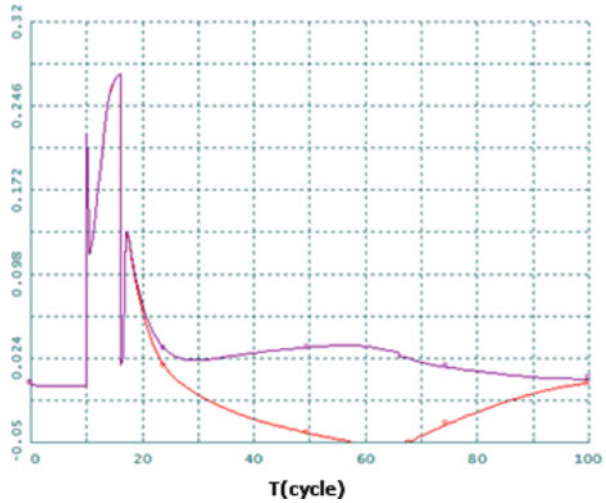
**Fig. 106.6** Reactive power of turbine



### 106.5 Conclusion

With the study on voltage transient stability, by changing the control modes of the wind generator, it gives the following conclusions: (1) In comparison with constant power factor control, wind farm with constant voltage control is positive to the voltage stability of the grid. (2) The installation of SVC could improve voltage transient stability effectively and the simulation of two control types is the same.

**Fig. 106.7** Reactive power of SVC



**Table 106.1** Voltage of the turbine

Control mode	The original		The minimum		The stable	
	Without SVC	With SVC	Without SVC	With SVC	Without SVC	With SVC
Constant voltage control	0.985	0.985	0.868	0.868	0.986	0.985
Constant power factor control	0.985	0.985	0.868	0.868	0.985	0.984

**Table 106.2** Voltage of connection point

Control mode	The original		The minimum		The stable	
	Without SVC	With SVC	Without SVC	With SVC	Without SVC	With SVC
Constant voltage control	1.022	1.022	0.89	0.89	1.023	1.022
Constant power factor control	1.022	1.022	0.89	0.89	1.019	1.02

## References

1. Hansena, A. D., & Michalke, G. (2007). Fault ride-through capability of DFIG wind turbines. *Renewable Energy*, 23(32), 1594–1610.
2. Xiang, D. W., & Li, R. (2006). Control of a doubly fed induction generator in a wind turbine during grid fault ride-through. *IEEE Transactions on Energy Conversion*, 21(3), 652–662.
3. Lopez, J., Sanchis, P., Roboam, X., et al. (2007). Dynamic behavior of the doubly fed induction generator during three-phase voltage dips. *IEEE Transactions on Energy Conversion*, 22(3), 709–717.

4. Lopez, J., Gubia, E., Olea, E., et al. (2009). Ride through of wind turbines with doubly fed induction generator under symmetrical voltage dips. *IEEE Transactions on Industrial Electronics*, 56(10), 4246–4253.
5. Bueno, C., & Carta, J. A. (2006). Wind powered pumped hydro storage systems, a means of increasing the penetration of renewable energy in the Canary Islands. *Renewable and Sustainable Energy Reviews*, 10(3), 312–340.

# Chapter 107

## The Generator Stator Fault Analysis Based on the Multi-loop Theory

Yu Shao, Feng Shi, and Xiang Li

**Abstract** Interturn short circuit is a common kind of fault in generator. The chapter takes multi-loop theory to analyze the theory of fault on generator stator and puts the math model of the generator. Changes of main parameters are analyzed separately when fault happens. According to the result, the chapter analyzes the influence on the main parameters caused by the fault of generator stator and summarizes the factors of parameter changes.

### 107.1 Introduction

Synchronous generators are the main force in thermal power generation, which are the most important in power system. The condition of synchronous generators decides the stability of the power system directly. The monitoring system of generators comprises local thermal insulation monitoring apparatus, partial discharge monitoring apparatus, palladium barrier leak detectors [1], etc.

In order to ensure the accuracy of operation parameters, the math models are used. For the reason that the magnetic field of generator is whirling in space and the structure of stator and rotor is very huge and the convert between magnetic energy and electric energy is abstract, it is possible to use the math model to study, which makes online monitoring meaningful [1].

The common inner faults are unsymmetrical faults, such as single-phase earth fault and phase-to-phase fault. When a fault happens, the harmonic rate is high in three phases, which can generate a different rotate speed and rotate orientation of the magnetic field. The winding EMF caused by unsymmetrical current is also confused. These are all marked features of inner faults of synchronous generators, which need an exact math model to deal with.

---

Y. Shao (✉) • F. Shi • X. Li  
Zhengzhou Electric Power Supply Company, Zhengzhou 450000, China  
e-mail: [672851649@qq.com](mailto:672851649@qq.com)

The structure of the chapter is as follows. The math models of inner faults in synchronous generators are described in Sect. 107.2. Section 107.3 presents the multi-loop theory. Section 107.4 presents equations of voltage and magnetic linkage. Section 107.5 presents simulation and analysis. Section 107.6 concludes the whole chapter.

## 107.2 Math Models of Inner Faults

Interturn short circuit, interphase short circuit, and open-phase short circuit are the main faults of synchronous generators. When faults are analyzed, current and voltage are set to be status-variable.

Math models are set to describe inner faults of stator windings based on multi-loop theory, analyzing the connections between stators and rotors. Multi-loop theory helps make voltage equations and magnetic linkage equations. Models of transformer and grid are made according to reality [2].

## 107.3 The Multi-loop Theory

The multi-loop theory makes winding to be a study unit, which breaks the ideal law that the motor seems to be the ideal one when analyzing inner faults of stator winding. Stator windings are seen as one-phase winding of three-phase winding, and rotor windings are dealt with in non-salient pole machine or salient pole machine. Windings are seen as an independent unit to be used in making magnetic equations.

In the procession of analysis, the motor is seen as many mutual sportive circuits. When the theory is used to study winding faults, influences caused by harmonic magnetic fields should be taken into account, so does the space structure of stator winding, connection shapes, etc. [3].

## 107.4 Equations of Voltage and Magnetic Linkage

### 107.4.1 Branch Equations

Any branch equations of voltage and magnetic linkage can be made according to multi-loop theory, such as branch Q:

$$\begin{cases} u_Q = pY_Q - r_Q i_Q \\ Y_Q = -\sum_{S=1}^{N_l} M_{QS} i_S + \sum_{i=1}^d \sum_{g=1}^{2P} M_{Qgi} i_{gi} + M_{Qr} i_r \end{cases} \quad (107.1)$$

In formula (107.1),  $u_Q, i_Q, Y_Q, r_Q$  represent for voltage, current, magnetic linkage, and resistance of branch Q in the stator, respectively.

$i_S, i_{gi}, i_f$  represent for current of branch S in the stator, current of damping branch “i” in pole “g,” and field current of generator.  $M_{QS}, M_{Q gi}, M_{Q f}$  represent for the mutual inductors.  $N_1$  represents for the number of branches in stator. d represents for the number of resistance branches, and P the number of poles. p is differential operator.  $Q=1,2,\dots N_1$ .

If there are no interturn faults of stator branches and voltage of three phases is  $u_A, u_B, u_C, u_Q$  would meet the relation that [4]

$$u_Q = \begin{cases} u_A & Q = 1, 2, \dots, a \\ u_B & Q = a + 1, a + 2, \dots, 2a \\ u_C & Q = 2a + 1, 2a + 2, \dots, 3a \end{cases} \quad (107.2)$$

When there is an interturn fault in the “i” branch of phase A in stator, the expression of  $u_Q$  is shown next:

$$u_Q = \begin{cases} u_A & Q = 1, 2, \dots, i - 1, i + 1, \dots, a \\ u_B & Q = a + 1, a + 2, \dots, 2a \\ u_C & Q = 2a + 1, 2a + 2, \dots, 3a \\ u_A - u_k & Q = i \\ u_j & Q = 3a + 1 \\ u_k - u_j & Q = 3a + 2 \end{cases} \quad (107.3)$$

### 107.4.2 Loop Equations

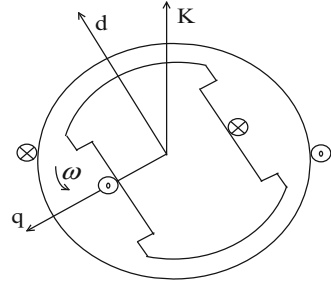
According to the above, the “i” loop equation of “g” pole is shown in formula (107.4):

$$\begin{cases} 0 = p\Psi_{gi} + r_{gi} i_{gi} - r_C (i_{g,i-1} + i_{g,i+1}) \\ \Psi_{gi} = -\sum_{S=1}^{N_1} M_{gi,S} i_S + \sum_{j=1}^d \sum_{k=1}^{2P} M_{gi,jk} i_{jk} + M_{gif} i_f \end{cases} \quad (107.4)$$

In formula (107.4),  $i_{gi}, \Psi_{gi}, r_{gi}$  represent for voltage, current, magnetic linkage, and resistance of loop “i” in pole “i,” respectively.

$i_{g,i-1}, i_{g,i+1}, r_C$  represent for voltage, current of loop “i-1,” and loop “i+1” in pole “g,” respectively.  $M_{gi,S}, M_{gi,jk}, M_{gi f}$  represent for mutual inductors between the resistance loop and other branches. And  $i=1, i_{g,i-1} = i_{g-1,d}, i=d, i_{g,i+1} = i_{g+1,1}$ .

**Fig. 107.1** Forward direction of current in stator and rotor



### 107.4.3 Equations of Field Winding

According to the forward direction in Fig. 107.1, equations of voltage and magnetic linkage in field winding are shown in formula (107.5):

$$\begin{cases} u_f = p\Psi_f + r_f i_f \\ \Psi_f = -\sum_{S=1}^{N_1} M_{fS} i_S + \sum_{i=1}^d \sum_{g=1}^{2P} M_{f gi} i_{gi} + L_f i_f \end{cases} \quad (107.5)$$

In formula (107.5),  $\Psi_f$ ,  $r_f$ ,  $u_f$ ,  $i_f$  represent for magnetic linkage, resistance, voltage, and current of field winding, respectively.  $i_{gi}$ ,  $i_S$  represent for current of the loop “ $i$ ” and stator winding “ $s$ ” in pole “ $g$ .”  $M_{fS}$ ,  $M_{f gi}$ ,  $L_f$  represent for mutual inductance and self-inductance of field windings [5].

### 107.5 Simulation and Analysis

The inner interturn faults of stators are familiar in normal operation, especially in synchronous generators with large capacity. The windings of synchronous generators with large capacity are in the same channel, which increase the probability of interturn faults.

In can be seen that the analysis of reason and theory of interturn faults has great value, which proves the matter to seek the discipline of the interturn faults. The simulation used PSCAD software to make models of generator and simple grid, simulating the operation of full-load generation. The simulation focused on one-winding and two-winding faults and looked for the discipline of faults through interturn faults. A 10 kW generator model was used in the simulation, which was made of PMVR, field system, etc.

**Table 107.1** Parameters of generator

Parameters	Range
Kinetic energy of generator	$E_{WMS} > 0$
Original active power (pu)	$0 \leq P \leq 1.0$
Original reactive power (pu)	$0 \leq Q \leq 1.0$
Active power on a node (pu)	$\Sigma P = 1.0$
Reactive power on a node (pu)	$\Sigma Q = 1.0$
Direct axis transient reactance	$X'd > 0, X_d > X'd > X''d$
Quadrature axis transient reactance	$X_q = X'q > X''q (T'q0 = 0)$
Damping reactor	$XL \leq X''d$

### ***107.5.1 Simulation of Interturn Faults in Stator on Full-Load Generator***

The chapter uses the typical parameters of generator in PSCAD. And parameters are shown in Table 107.1.

The model of magnetic field was the IEEE-1968 model, which could represent the typical generator. The full-load operation means generator operates with rated load, which is the most familiar working condition.

One of the features when the full-load generator is on operation is that the current is heavy. The long-time operation of full-load generator has great harm to the isolate winding, which makes it meaningful to study interturn faults under the condition when the generation is full load.

### ***107.5.2 One-Winding Fault of Phase A on Stator***

When one-winding short circuit happens on phase A of the generator, the simulation occurs, as shown in Figs. 107.2 and 107.3.

As shown in Figs. 107.2 to 107.3, current of three phases is heavier when the generator is on normal operation.

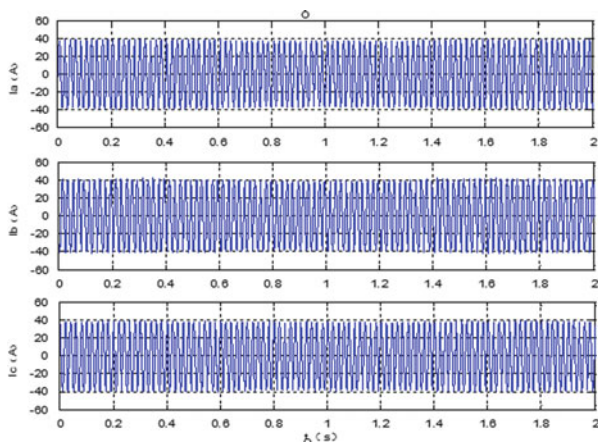
There would be an obvious change when a fault happens. On the contrary, field current has an obvious change and the amplitude is about 0.1 A, for which the field current is chosen to be the criterion of faults.

### ***107.5.3 Two-Winding Fault of Phase A on Stator***

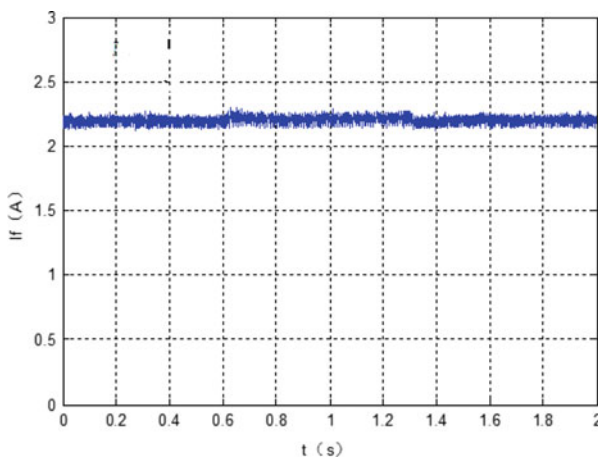
When two-winding short circuit happens on phase A of the generator, the simulation occurs, as shown in Figs. 107.4 and 107.5.



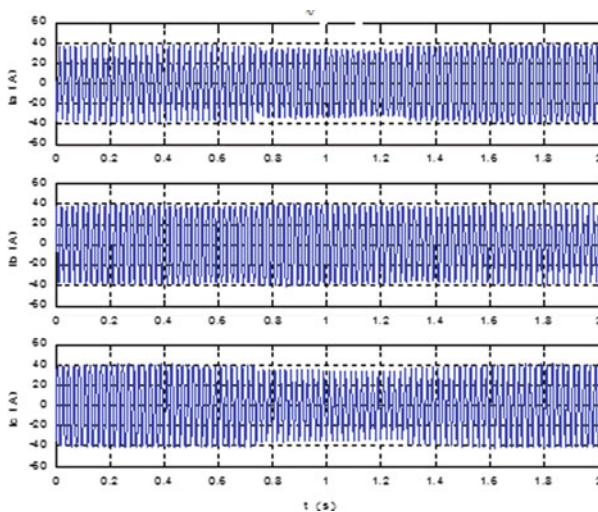
**Fig. 107.2** Current wave of three phases when one-winding short circuit happens



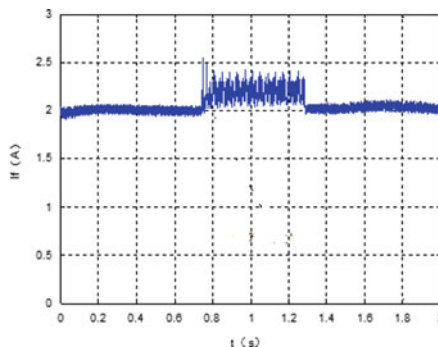
**Fig. 107.3** Field current wave when one-winding short circuit happens



**Fig. 107.4** Current wave of three phases when two-winding short circuit happens



**Fig. 107.5** Field current wave when two-winding short circuit happens



As shown in Figs. 107.4 to 107.5, current of three phases are heavier when the generator is on two-winding short circuit. The change of field current is obvious and the amplitude is about 0.4 A.

## 107.6 Conclusion

The chapter used PSCAD software to prove multi-loop theory could be used to study inner faults of motors, checking the validity that current variation and voltage variation of stator could be used to analyze the stator faults.

The math models of stator windings are with commonality when used to analyze faults, which are adapted to be used in analysis of non-salient pole machine and salient pole machine. The whole simulation has some meanings to the analysis of real motor faults.

## References

1. Ishida, K., & Dokai, K. (1992). Development of a 500kV transmission lines arrester and its characteristics. *IEEE Transactions on Power Delivery*, 7(3), 1265–1274.
2. Sakshaug, E. C. (1989). Metal oxide arresters on distribution systems fundamental consideration. *IEEE Transmission on Power Delivery*, 4(4), 2076–2089.
3. Flisowski, Z., Mazzetti, C., & Wlodek, R. (2004). New approach to the selection of effective measures for lightning protection of structures containing sensitive equipment. *Journal of Electrostatics*, 60(4), 287–295.
4. Mathier, L., Perreault, L., & Bobe, B. (1992). The use of geometric and gamma-related distributions for frequency analysis of water deficit. *Stochastic Hydrology and Hydraulics*, 6(4), 239–254.
5. Frank, J., & Masse, J. (1951). The Kolmogorov-Smirnov test for goodness of fit. *Journal of the American Statistical Association*, 46(253), 68–78.

# Chapter 108

## An Improved Edge Flag Algorithm Suitable for Hardware Implementation

Lixiang Wang and Tiejun Xiao

**Abstract** The traditional edge marking algorithm cannot fill the elongated polygon and a polygon with local points correctly. After doing a lot of research and analysis about polygon fill algorithms, this paper presents a new improved algorithm, which is suitable for hardware implementation, to meet the need for high-quality graphic display in the embedded system. The new algorithm makes full use of the characteristic that the local point or elongated point is accessed repeatedly when it meets local points and elongated points. We can define a measurement variable named FLAG, which is used to mark the boundary point of the polygon. The flag of the present point will add one when it is accessed. This method can conveniently and simply distinguish singular points and elongated points from ordinary points. What's more, the improved algorithm solves the previously mentioned problems effectively. In the new algorithm, we only use the addition operation so it is easy to be implemented by the hardware.

### 108.1 Introduction

With the wide application of embedded systems, polygon filling has been extensively used in the field of embedded systems gradually. The traditional display of the embedded graphics depended largely on the microprocessor. However, to satisfy the high-quality and high-efficiency requirements, using only software to do the complex process of polygon filling has become a bottleneck in the graphics display, so it is a feasible scheme to meet the high demand with the hardware. The problem of polygon filling [1] is one of the basic issues in computer graphics. Polygon fill algorithm mainly includes edge marking algorithm [2, 3], scan line fill algorithm [4], and seed fill algorithm [5], it has to call

---

L. Wang (✉) • T. Xiao  
School of Computer Science and Telecommunication Engineering, Jiangsu University,  
Zhenjiang 212013, China  
e-mail: [lxwang2012@163.com](mailto:lxwang2012@163.com)

stack time and time again, so the efficiency of the filling is reduced greatly. For the scan line fill algorithm, it has to build edge tables and rank them when filling the polygon, so the efficiency is also restricted. When scan line fill algorithm and edge marking algorithm are implemented in software, their efficiency is quite good; however, the edge marking algorithm doesn't need to build edge tables and maintain them, so its efficiency is greater by one to two orders of magnitude than that of the scan line fill algorithm. But the traditional edge marking algorithm will not fill normally when the polygon has elongated points and singular points. So a lot of papers have already been submitted to provide many solutions to improve it, such as Xiaohua Wang who proposed an improved edge marking algorithm in 2004 that successfully resolved the problem of the singular points. Unfortunately the introduction of the edge table, which is used in the scan line fill algorithm, made the data structure complex. Guodong Ye [6] presented a designed edge marking algorithm aiming at elongated polygon, but the efficiency of this algorithm isn't high while using the operation of intersection. This paper presents an improved algorithm that uses a measurement variable as the flag of boundary, whose initial value is 0. In the process of rasterizing the boundary, the flag of the current point will add one when it is accessed. The method distinguishes the normal boundary points from singular points and elongated points is very easy and it not only keeps the advantage of the traditional edge marking algorithm but also solves the anomalies when filling the special point effectively.

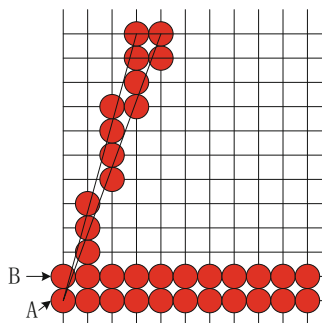
## 108.2 The Traditional Edge Marking Algorithm

The traditional edge marking algorithm is very easy to be understood as follows: first, marking the boundary and, second, filling the polygon, pixel by pixel in accordance with the order from the left point to the right for each scanning line intersecting with the boundary of the polygon. It is judged whether the pixel point is in the interior of the polygon or not. If the pixel point is in the interior of the polygon, fill it with filling color; otherwise, fill it with background color. But an abnormal phenomenon will appear when the polygon has local points and elongated points. In this process, a Boolean variable INSIDER marks whether the point is in the polygon or not. If the initial value of the INSIDER is FALSE, then the value of the INSIDER will be negated when the scanning line encounters the boundary point, and if the value of the INSIDER is TRUE, then the color of the point is set with filling color.

### 108.2.1 Singular Point

The general method to judge whether a local point is a singular point is as follows: assuming an intersecting point of A, if the two sides of another vertex are located on

**Fig. 108.1** The map of the local point and elongated points



**Fig. 108.2** Filling of local point in polygon



the same side of the intersection of A, the intersection is a singular point. Otherwise, it isn't a local point. In Fig. 108.1, point A is a singular point; when the scanning line encounters point A, the value of the INSIDER is TRUE, and then the pixels after point A will be filled, resulting in abnormal filling. Such a polygon (10, 10), (50, 120), (90, 10), (50, 70) which has local points and is filled by traditional edge marking algorithm is shown in Fig. 108.2. And so, we make mistake for it strongly.

### 108.2.2 Elongated Point

The problem of the elongated polygon is that when the slope of the two edges of the polygon is so close, the polygonal area is too thin and narrow to fit in this region at any point. In Fig. 108.1, point B is an elongated point. Then the pixels will be filled after point B, resulting in abnormal filling. Such an elongated polygon (50, 25), (50, 225), (90, 225), (90, 125), (55, 125) is shown in Fig. 108.3. The points near the highest point of the polygon should not be filled, but they are filled by traditional edge marking algorithm. Obviously, we also make mistake for it.

**Fig. 108.3** Filling of elongated polygon

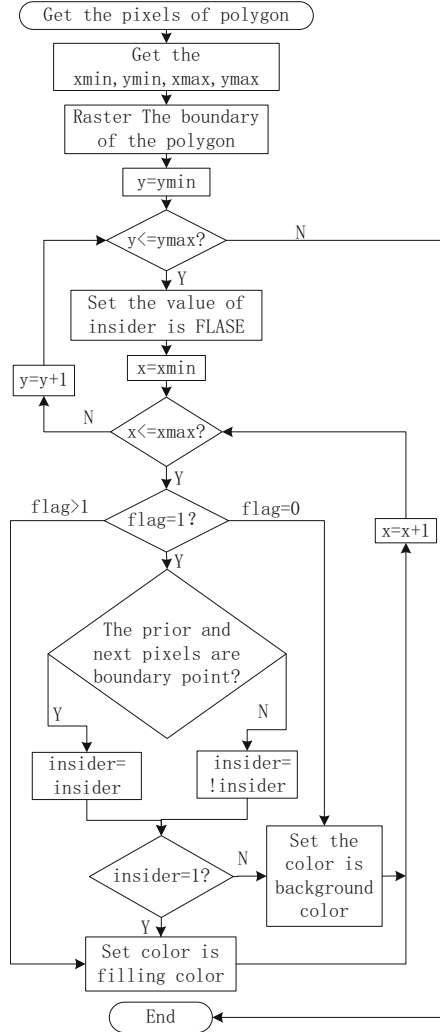


### 108.3 The Improved Edge Marking Algorithm and Flow Diagram

Based on the above descriptive analysis, the improved edge marking algorithm is explained in this section. The reason of the abnormal filling phenomenon is that the traditional algorithm can't make the distinction between the ordinary points and the local points and elongated points. It presents an easy method to distinguish them. And what's more, in the system we use a memory to store the flags of the boundary points.

As shown in Fig. 108.4, the main processes are as follows: Firstly, we can get the pixels of the polygon. Secondly, when scanning the boundaries of the polygon, we can get the maximum and minimum values of the transverse and longitudinal coordinates of the vertices of the polygon; they are defined as  $x_{min}$ ,  $x_{max}$ ,  $y_{min}$ , and  $y_{max}$ . Thirdly, in the process of rasterizing, read the flag of the memory, get the flag of the current point, and add one before storing it again in the memory. So when we begin to scan the polygon, we must initialize the scanning line- $scan\_y$  with  $y_{min}$ , and  $scan\_y$  increases by 1 from  $y_{min}$  to  $y_{max}$ . If  $y$  is less than  $y_{max}$ , then set the value of the INSIDER to FALSE and then initialize  $x$  with  $x_{min}$ . In the next step, if the value of the flag is greater than one, then set the color of the point with filling color and the value of the INSIDER remains unchanged; otherwise, if the value of the flag of the current pixel is 1, then judge the values of the prior and next pixels. If both of them are 1, then the value of the INSIDER remains unchanged; otherwise, invert the value of the INSIDER. If the value of the INSIDER is TRUE, then set the color of the current pixel with filling color; otherwise, set the color of the current pixel with background color. Take point A as an example. When the scanning line encounters A, because the value of flag of A is two, greater than one, we can set the color of A with filling color and the value of the INSIDER remains unchanged; the points at the back of A are not filled.

**Fig. 108.4** The flow chart of the improved algorithm



### 108.4 Result

The improved algorithm has been verified by the software. The software tool is VC++ 6.0. The filling effect of the algorithm, which has been improved, is shown in Figs. 108.5 and 108.6.

From the effect of the polygon filling with improved algorithm, we can see that the improved algorithm is correct. Take point A as an example. When the scanning line meets with A and B, as shown in Fig. 108.1, the value of the INSIDER remains unchanged, so the colors of the point at their back are filled with background color and their colors are filled with filling color because their flags are greater than 1.

**Fig. 108.5** Filling of elongated polygon with improved algorithm



**Fig. 108.6** Filling of local point in polygon with improved algorithm



### 108.5 Conclusion

The traditional edge marking algorithm is simple and easy to be implemented by hardware, and it has the advantages of needless to build, keep and order the edge tables. What we can see from the implementation of this new improved algorithm is that not only does it avoid the complex data structure called edge table, which is always used in the scan line fill algorithm, but also it avoids abundant intersection operation. All of them can take up a lot of CPU time, and in this algorithm, we only use the addition operation, so a conclusion can be made that this improved algorithm put forward by this paper is suitable for hardware implementation.



## References

1. Wu, Z., et al. (2003). Singular point distinguishing algorithm for area filling. *Journal of Computer-Aided Design & Computer Graphics*, 15(8), 979–983 (In Chinese).
2. Wang, X., & Yan, B. (2004). An improved algorithm of edge marking fill. *Computer Applications*, 24(6), 182–183 (In Chinese).
3. Hao, X. (2006). The common problems and solutions of boundary labeling method in the course of realization. *Journal of Xi'an University of Engineering Science and Technology*, 20(10), 643–645 (In Chinese).
4. Zhang, Z., Liu, X., Zhang, Z., et al. (2009). Improved method for polygon scan conversion. *Computer Engineering and Applications*, 45(4), 193–195 (In Chinese).
5. Zhang, Z., Ma, S., & Li, W. (2009). New regional filling algorithm based on seed. *Computer Engineering and Applications*, 45(6), 201–202 (In Chinese).
6. Ye, G., Lin, G., Zhu, C. (2009). A designed edge marking fill algorithm for elongated polygon. *2009 First International Workshop on Database Technology and Applications* (pp 22–24).

# Chapter 109

## A Handheld Controller with Embedded Real-Time Video Transmission Based on TCP/IP Protocol

Mingjie Dong, Wusheng Chou, and Yihan Liu

**Abstract** Cross-platform video transmission is of vital importance in industrial applications. In this paper, we introduce a method for transmitting video from the computer with Windows system to the ARM11 board with embedded Linux system using the Ethernet based on the TCP/IP protocol. The ARM11 board is used as the server to receive video information using its Qt GUI, while the computer on the bank is used as the client that receives video information from the remote-operated underwater vehicle showing with its MFC (Microsoft Foundation Classes) interface and then sends the video information to the handheld controller. The image gained from the computer MFC is JPG format, and after coding, the images are transmitted to the server on the handheld controller continuously. Then the Qt GUI receives the data and decodes the JPG images before displaying them on the screen. The transmission is based on TCP/IP protocol and an image parsing protocol made by us. After testing, the video image can successfully conduct real-time transmission and can meet the industry application.

### 109.1 Background Information

The project comes from the national 863 project—miniature underwater submarines. Because of the complex underwater environment of the nuclear power plant, the ROV (remote operated vehicle) must be controlled on the bank, and the information of the reactor pool gained by the camera attached to the ROV will be transmitted to the control box made of an IPC (Industrial Personal Computer) with Windows operation system on the bank. Considering the big size of the control box, we need to develop a portable handheld controller which can replace the control box to a certain extent. The video image information will be transmitted to the

---

M. Dong (✉) • W. Chou • Y. Liu  
Intelligent Technology and Robotics Research Center, School of Mechanical Engineering and Automation, Beijing University of Aeronautics and Astronautics, Beijing 100191, China  
e-mail: [buaadmj@gmail.com](mailto:buaadmj@gmail.com)

system box from the camera on the ROV, and then the control box will transmit the video image information to the handheld controller directly. The latter is just what we will talk on this paper.

The video transmission is based on the TCP/IP protocol [1]. The control box will be the client, while the handheld controller will be the server during the video transmission. The transmission is through Ethernet using network cable, so the transmission is very fast and can be real time after a certain optimization.

### 109.2 Introduction of the Hardware Platform

The handheld controller is made of an ARM11 development board with an S3C6410 platform. The S3C6410, whose CPU can be up to 667 MHz, is a 32-bit RISC microprocessor which is designed by Samsung to provide a cost-effective, low-power capabilities and high-performance processor solution for mobile devices [2]. During the research, a Linux operation system with kernel number 2.6.38 is transplanted to it. After cross-compiling using arm-Linux-gcc cross-compiler with version number 4.5.1, the program written on the host machine can run successfully on the development board [3]. Given its excellent property, the development board is right for video transmission (Fig. 109.1).

### 109.3 Software Application Platform Building

The software application platform is composed of two different parts, the software on the Windows system and the software on the embedded Linux system.

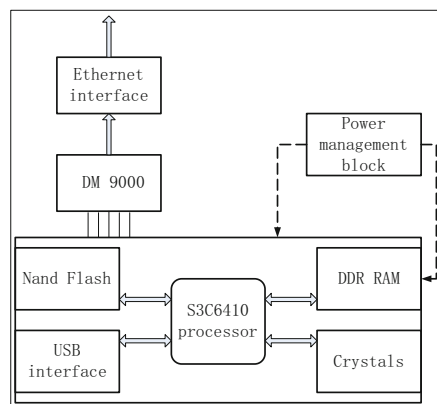
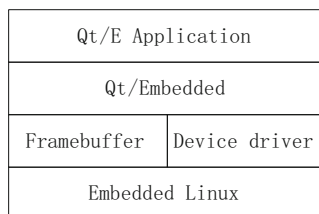


Fig. 109.1 The hardware platform of the handheld controller

**Fig. 109.2** Qt/E mechanism structure diagram



### 109.3.1 Software on the Windows System

The software on the Windows system refers to the platform using MFC built on the control box with the Windows system. This part shows the video image transmitted from the camera attached on the ROV. The MFC program uses the SDK functions of the image capture card to display the images. Meanwhile, it is used as the client of the network video transmission through socket based on TCP/IP protocol.

### 109.3.2 Software on the Embedded Linux System

In the ARM11 development board, a Linux system is transplanted with the kernel version number 2.6.38, and then we successfully setup the cross-compiler arm-linux-gcc with the version number 4.5.1 on it. With all above settled, we set up Qt/E (QT embedded) on the development board in order to use its convenient GUI to decode the JPG image and display the video information on the screen.

Qt/E is an open-source software development kit provided by the software developer Trolltech, with C++ program language as its development tools. Qt/Embedded has a lot of simple class libraries and interface modification tools. Especially, it has unique signals and slots mechanism [4]. It is across platforms and can be easily transplanted. The simple Qt/E mechanism is showed in Fig. 109.2.

## 109.4 Application of TCP/IP Protocol

Based on the TCP/IP protocol, we use socket as the application programming interface. The socket programming is based on the system call of the socket. At first, we use the function `socket ()` to create a new socket, and then we connect the socket address with the socket we just created using `bind ()`. In order to build the socket connection between the client and the server, we should use `connect ()` and `accept ()` functions. The function `connect ()` is used to build connection, while the function `accept ()` is used by the server, waiting for the connection request from the client. Meanwhile, the server uses the function `listen ()` to monitor whether there is a

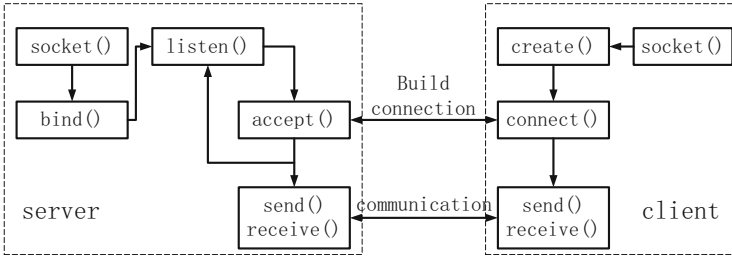


Fig. 109.3 The socket’s working procedure

client request. Once a connection between the client side and the sever side is built successfully, we can exchange data from the client and the server [5]. The socket’s working procedure is in Fig. 109.3 [6].

### 109.4.1 Socket on the Server Side

The server is built on the development board with embedded Linux operation system using the Qt/E GUI. We use QTcpServer, a Qt class, to build a sever class. Then, we use the function `listen ()` to monitor the port we set at the beginning [7]. Because of the high encapsulation of Qt, we can save a lot of trouble. Once the server side gets the connection request from the client, the server will accept the request and the Qt function `readyread()` will be called, and then in our program, the function `dataReceived()` written by us will run according to our settled signals and slots mechanism. The data from the client will be gained successfully.

### 109.4.2 Socket on the Client Side

On the client side, as we use the MFC in the visual C++6.0 environment, we should use the Windows socket API to build the socket and build the connection with the server in the Linux operation system. At first, we use `socket (AF_INET, SOCK_STREAM, 0)` [8] to create a new socket based on the TCP/IP protocol. Then we give the address variables a certain value to make sure the socket address is connected with the socket we created. Lastly, the client will send a connection request to the server using the function `connect ()`. When the server accepts the request, the connection is finished.

## 109.5 Protocol Development of the Video Transmission

For the data to be transmitted in the form of binary, so after capturing the images information from the camera attached to the ROV, we will change the image format from JPG to binary streams. In order to get the images streams coded in the client side and decoded in the server side securely and to accept the whole image streams in the development board smoothly, we should set up our own video transmission protocol.

### 109.5.1 Image Acquisition and Coding on the Client Side

Every time we capture an image, the image will be saved in the form of binary with a temporary name given by us. Then we make a protocol for the whole image transmission, using the format "AA:file name\*length:EE" in order to identify that we get a whole image all the times. The file name is the temporary name we give after an image is captured, while the length is its length of binary form. Using a while structure loop, the images will be sent to the server continuously.

### 109.5.2 Image Receiving and Decoding on the Server Side

In the server written by us using Qt GUI, a lot of data will be received. At first, we estimate whether a set of data the server received is a whole image through the image transmission protocol we made. If the data meet the format "AA:file name\*length:EE", the video transmission protocol made by us, we can make sure that the data the server received is a whole image transmitted from the client, and then the image will be displayed on the screen. Just as before, using a while structure loop, we decode and display the images continuously. Because of the fast flow of the continuous images using network cable and the excellent property of the development board, we can watch the video information smoothly (Fig. 109.4).

## 109.6 Experimental Verification

After building the environment on both the Windows system and the embedded Linux system, we programmed in the client side and the server side, respectively, based on the video transmission protocol made by us and the TCP/IP protocol. Then we test the program, and the result is very satisfying. The video information captured from the client side can be transmitted to the server side on the ARM11 development board smoothly.

Fig. 109.4 The video transmission procedure

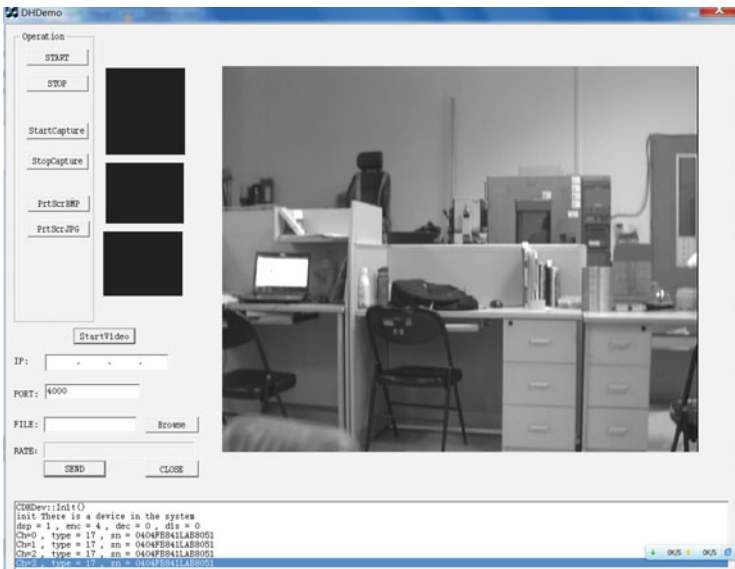
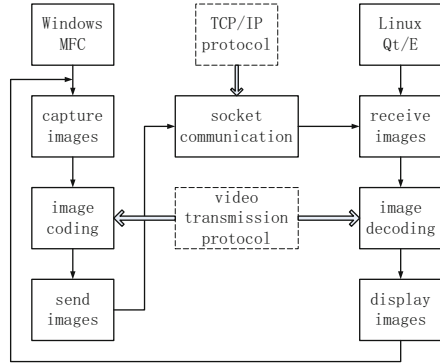


Fig. 109.5 The video information running in the client side on the control box

In order to test whether the video transmission is stable and to gain the transmission speed, we set up the function `frame_test()` in the Qt GUI. By calling the Qt class libraries with the function, we find that the transmission is very stable and the speed can be up to 24 frames per second. It is real time and can meet the industry application perfectly. The effect picture of the video transmission is shown in Figs. 109.5 and 109.6.



**Fig. 109.6** The video information running in the server side on the handheld controller

## 109.7 Conclusion

Summarily, this paper describes an effective method to transmit video image information across the different operation system based on the TCP/IP protocol. The images transmission protocol we made is simple and useful, while the cost of the whole handheld controller is very low and the video transmission is real time; it can easily meet the needs of the project.

## References

1. Wilder, F. (1993). *A guide to the TCP/IP protocol suite* (pp. 3–21). Boston, MA: Artech House.
2. Friendly ARM. (2011). *Tiny6410 hardware manual*. Retrieved August 5, 2011, from <http://www.arm9.net>
3. Liu, Z. (2012). Application of embedded database in room environment supervisory system. *Modern Electronics Technique*, 35(10), 2–3.
4. Bangwei, Y., & Deng, H. (2011). Control interface of embedded digital monitoring system based on Qt/embedded. *Part and Application*, 35(24), 1–2.
5. Anand Kumar, M., & Karthikeyan, S. (2011). Security model for TCP/IP protocol suite. *Journal of Advance in Information Technology*, 2(2), 1–5.
6. Zhang, Y.-G., Liu, C.-C., Liu, W., & He, F.-Z. (2006). Remote monitoring and control system based on socket and multithread. *Control Engineering of China*, 13(2), 2–3.



7. Blanchette, J., & Summerfield, M. (2008). *C++ GUI programming with Qt4* (2nd ed., pp. 295–303). Beijing, China: Publishing House of Electronics Industry.
8. Tian, L. (2012). Network communication technology based on Winsock. *Hubei: Software Guide*, 11(1), 1–4.

# Chapter 110

## Evaluating the Energy Consumption of InfiniBand Switch Based on Time Series

Huifeng Wang, Zhanhuai Li, Xiaonan Zhao, Qinlu He, and Jian Sun

**Abstract** Recently, energy consumption has emerged as a critical factor in designing storage system. In order to test the energy consumption of InfiniBand switch (IB switch), we establish an energy consumption model for IB switch and formulate the test cases. Using the method, you can obtain the energy consumption of the IB switch scientifically and efficiently. Empirical results illustrate the correctness of the energy consumption and reflect the distribution laws of the energy consumption of the IB switch clearly. The scheme can solve the problem of testing and analyzing the energy consumption of the IB switch efficiently. It has positive practice significance to reduce the cost of storage system.

### 110.1 Introduction

With the development of modern information technology, energy consumption accounts for a significant fraction of cost of data centers. As demonstrated by the successful emergence of the Green500 list [1], energy consumption has become as important as performance. Storage subsystems alone represent roughly 10–25 % of the power consumed by the data center [1]. Energy consumption of storage system can become a greater problem. So it is very meaningful to study on the energy consumption in order to reduce the cost of storage system.

In recent years, high-speed communication protocols and their supporting hardware have emerged, which can provide very low latency and very high bandwidth. As a result, they have become increasingly popular in the areas of high-performance computing (HPC) and enterprise computing where communication is critical to application performance [2]. InfiniBand switch (IB switch) is one of the

---

H. Wang (✉) • Z. Li • X. Zhao • Q. He • J. Sun  
School of Computer Science and Technology, Northwestern Polytechnical University,  
Xi'an Shaanxi 710129, China  
e-mail: [wanghuifeng12@163.com](mailto:wanghuifeng12@163.com)

central devices, which is capable to interconnect a large amount of nodes (large port count). The entire data transfer between servers and storage devices through IB switch. The energy consumption of IB switch is one of the most important sources the storage system consumes. Calabretta et al. [3] study on the relationship between energy consumption and latency. Fu et al. [4] propose a novel router architecture to allow each of its modules to adjust frequency according to traffic loads in order to gain energy efficiency. Vishwanath et al. [4, 5] explore buffer usage of routers and develop a buffer adapting algorithm for letting SRAM and DRAM buffers sleep while not being used.

We can see that the existing studies focus mainly on how to reduce the energy consumption of the storage system. There is little study on how to test the energy consumption of switch. This paper provides a solution to test the energy consumption of the IB switch and make a detailed analysis about the test results. Everyone can use the solution to obtain the distribution of the energy consumption of the IB switch. The test scheme we designed makes the test case ordered. The former test is the condition of the latter test case. The order not only can finish the test the energy consumption of the IB switch but also can make the test data easily and scientifically. The research on energy consumption of InfiniBand switch has positive practice significance on the construction of storage system. In order to describe easily, we abbreviate the InfiniBand switch as IB switch.

The rest of this paper is organized as follows. In Sect. 110.2, we introduce the test content and test methodology. In Sect. 110.3, we describe our test environment and make a detailed analysis on the experiment data of energy consumption of IB switch. We conclude and describe our future work in Sect. 110.4.

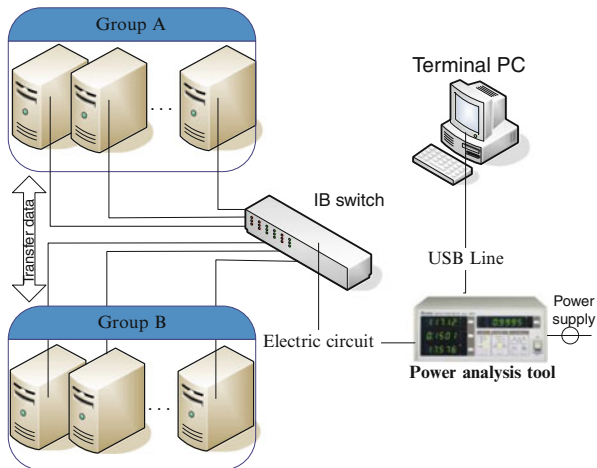
## 110.2 Methodology

In this section, we mainly introduce test contents and present the methodology we use. The paper mainly discusses the energy consumption distribution of the IB switch and investigates whether the workload on servers can make some effect on the energy consumption of IB switch. In order to collect the data correctly, we design a brief solution to test the energy consumption of the IB switch.

### 110.2.1 Test Topology Graph

We briefly introduce the topology graph about connection between the devices. Figure 110.1 shows that servers connect to the IB switch through IB line and the power the IB switch uses must pass the power analysis tool to provide power supply for IB switch. The power analysis tool has three ports to connect with other devices. The first port of the power analysis tool ties to power supply. The second port is connected to the terminal pc which is used to collect data of the energy consumption. The third port ties to the IB switch.

**Fig. 110.1** The topology graph of the test system



### 110.2.2 Energy Consumption Model

In order to make a quantitative analysis, we establish an energy consumption model for IB switch, which relates to the construction and working process of the IB switch. It describes as follows:

$$E_{sw} = E_{base} + E_{workload} + \sum E_{port} \tag{110.1}$$

The energy consumption of IB switch derives from three parts, which are the base energy consumption, the energy consumption of transferring data, and the energy consumption of IB cards. The three parts are described as  $E_{base}$ ,  $E_{workload}$ , and  $\sum E_{base}$ , respectively.

Although the model roughly describes the distribution of the energy consumption of IB switch, we need to study the proportion of every part to the total energy consumption of IB switch. The proportion is very helpful to evaluate the total energy consumption of IB switch.

### 110.2.3 Test Case

In order to make the test scientifically and efficiently, we design the test cases as described in Table 110.1. The design not only can save test time but also can make the test data more accurate. The former test case is the precondition and the basis of the latter test case. Using the results of former test case, we can only calculate the subtraction of the consecutive test cases. It need not test the starting value of every test case and also can make the change of two consecutive test cases to not too largely influence the experimental precision.

**Table 110.1** Test cases

Test case	Abbreviation	Description
1	Basis test	Test the energy consumption of the idle IB switch
2	Full-switched test	Test the energy consumption of IB switch in full-switched test
3	IB cards test	Test the energy consumption of the IB cards
4–8	Workloads test	Test the influence of workloads of servers to the energy consumption of IB switch

We design eight test cases to complete the test task. The test cases are in order. If you conduct the test cases in order, you can easily obtain the energy consumption of IB switch and make a detailed analysis.

Test case 1 is the basis test for the energy consumption of IB switch. The method is to power on the IB switch which is not connecting any IB card. The result of the test case is the energy consumption of the idle IB switch. Other test cases draw their conclusions through comparing with the result of the basis energy consumption.

The mission of test case 2 is to test the energy consumption of IB switch in full-switched. In test case 2, the internal ports transfer data mutually all the time. The result of the full-switched test reflects the energy consumption of switching data.

In test case 3, we connect 20 IB cards to the IB switch and stop the fully switching the data. The purpose of this test case is to collect the energy consumption of the IB cards.

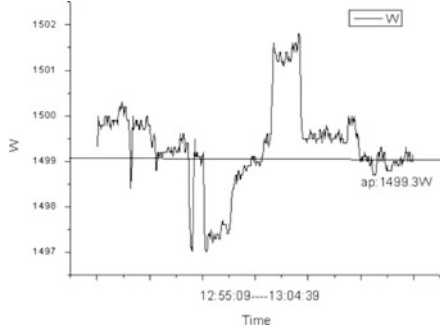
Test cases 4–8 are a series of test cases. Its task is to inspect the influence of workloads of servers to the energy consumption of IB switch. In these five test cases, we connect some servers to the IB switch through IB cards. These servers are divided into groups. At any test case, we make any group switch data between two servers through the IB switch.

Through conducting basis test and full-switched test, we can obtain the energy consumption of full-switched in the IB switch. We can gain the energy consumption of the IB cards by comparing test case 1 and test case 3. Through conducting test cases 4–8, we can obtain the discipline about the energy consumption of workloads to the IB switch.

### 110.3 Experiment and Analysis

The IB switch we study is used by Inspur Corporation. The switch has 324 ports. The server connected to the IB switch is the model of AS300N. The test equipment of energy is Chroma Digital Power Meter 66202. The tool can easily collect power of the tested object in real time. In order to describe easily, we abbreviate average power as ap.

**Fig. 110.2** The result of test case 1



As described in Sect. 110.2.3, the energy consumption of IB switch derives from three parts. We conduct the eight test cases to calculate the  $E_{base}$ ,  $E_{workload}$ , and  $\Sigma E_{base}$ , respectively, to obtain the distribution law of the energy consumption of the IB switch.

### 110.3.1 Test Case 1

We power on the IB switch which is not connecting any IB card to test the  $E_{base}$ . Figure 110.2 shows the result of test case 1, which is the energy consumption of the idle IB switch. We get the  $E_{sw1}$  which is the average power of test case 1. It is about 1499.3w. The  $E_{sw1}$  is equal to  $E_{base}$ . It is very simple but indispensable and very important because it is the standard basis of the other test cases. Other test cases draw their conclusions through comparing with the result of the basis energy consumption.

### 110.3.2 Test Case 2

The difference between test case 1 and test case 2 is that whether the IB switch is full-switched or not. We conduct test case 2 with the similar method of test case 1. The  $E_{sw2}$ , which is the ap of test case 2, is about 1757.6w. Based on the energy consumption model, we can get the largest energy consumption of switching data as follows:

$$E_{sw2} = E_{base} + E_{workload\_max} \tag{110.2}$$

$$E_{workload\_max} = E_{sw2} - E_{base} = E_{sw2} - E_{sw1} = 258.3w \tag{110.3}$$

Under normal circumstances the energy consumption is usually smaller than the value.

**Fig. 110.3** The result of test cases 4–8



### 110.3.3 Test Case 3

In test case 3, we connect 20 IB cards to the IB switch. We obtain that the  $E_{sw3}$ , which is the ap of test case 3, is about 1528.6w. So we can get the energy consumption of each IB card as follows:

$$E_{sw3} = E_{base} + \sum_1^{20} E_{port} = E_{sw1} + \sum_1^{20} E_{port} \tag{110.4}$$

$$E_{port} = \frac{E_{sw3} - E_{sw1}}{20} = 1.465w \tag{110.5}$$

We can see that the energy consumption of each IB card is a small proportion of the total energy consumption of IB switch. However, when we insert all the IB cards to the IB switch, the energy consumption of IB cards cannot be ignored.

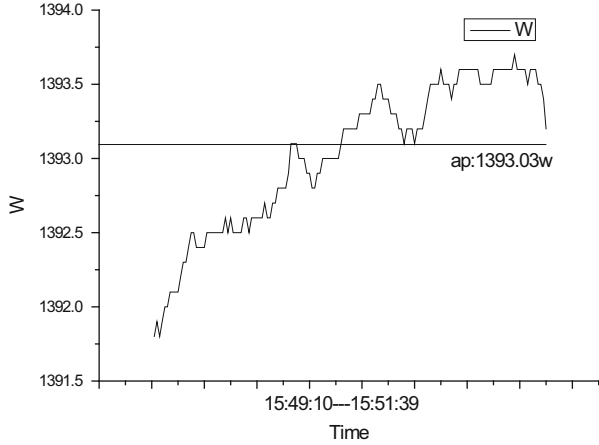
### 110.3.4 Test Cases 4–8

Figure 110.3 demonstrates the energy consumption of the IB switch when transferring data between the servers connected to the IB switch. In each experiment we add a group to transfer data. Each group has four servers and transfers data through the IB switch. We can find out the ap is a little changeable in the serials of experiments. The ap is about 1530w. It reflects that the energy consumption of transferring data is a small proportion of the total energy consumption of IB switch. In other words, transferring data has little impact on the energy consumption of the IB switch.

According to the empirical results, we can evaluate the energy consumption of IB switch. If n IB cards connect to the IB switch, we can obtain the  $E_{sw}$  as follows:

$$E_{sw} = E_{base} + \sum_1^n E_{port} = 1499.3 + 1.465n \tag{110.6}$$

**Fig. 110.4** The result of energy consumption IB switch provided by the Chinese Academy of Science



We also can evaluate the largest energy consumption of the IB switch as follows:

$$E_{sw} = E_{sw2} + \sum_1^{324} E_{port} = 1757.6 + 1.465 * 324 = 2232.26w \quad (110.7)$$

Using the method, we also test the IB switch provided by the Chinese Academy of Science. The energy consumption of the idle IB switch is about 1196.3w. When inserting six IB cards to the IB switch, the energy consumption grows to 1210w. As described in Fig. 110.4, when conducting full-switched test case, the ap of the IB switch is about 1393.03w. Empirical results show the same distribution law of the energy consumption of the IB switch. So adopting the solution to study on the energy consumption of IB switch has positive practice significance to reduce the cost of storage system.

### 110.4 Conclusion and Future Work

The series of experiments draw generally two conclusions. Firstly, the energy consumption of IB switch is mainly from three aspects, respectively, the energy consumption of normal run, the energy consumption of IB cards, and the energy consumption of communicating between internal ports. Secondly, the energy consumption of transferring data between servers connected to the IB switch is a small proportion of the total energy consumption of IB switch.

In the future, we will change the place where the IB ports connect to the IB switch to observe the law whether the port the server connect to the switch can have some effect on energy consumption. We plan to collect a long-term data of energy consumption of IB switch to observe the law of the IB switch energy consumption.



**Acknowledgements** This research program has been supported by the Ministry of Science and Technology of RPC (863 Program:2013AA01A215), the NPU Fundamental Research Foundation under Grant No.JC20120209, the National Key Technology Research and Development Program of the Ministry of Science and Technology of China under Grant No.2011BAH04B05, and the National Natural Science Foundation of China under Grant No.61033007.

## References

1. Prada, L., García, J., Calderón, A., García, J. D., & Carretero, J. (2013). A novel black-box simulation model methodology for predicting performance and energy consumption in commodity storage devices. *Simulation Modelling Practice and Theory*, *34*, 48–63.
2. Liu, J., Poff, D., & Abali, B. (2009). Evaluating high performance communication: A power perspective[C]. *Proceedings of the 23rd International Conference on Supercomputing* (pp. 326–337). Yorktown Heights, NY: ACM.
3. Calabretta, N., Luo, J., Lucente, S. D., & Dorren, H. (2012). Experimental assessment of low latency and large port count OPS for data center network interconnect[C]. *Transparent Optical Networks (ICTON), 2012 14th International Conference on* (pp. 1–4). Coventry: IEEE.
4. Fu, W., & Song, T. (2012). A frequency adjustment architecture for energy efficient router [C]. In *ACM SIGCOMM Computer Communication Review* (Vol. 42, pp. 107–108). Helsinki, Finland: ACM.
5. Vishwanath, A., Sivaraman, V., Russell, C., Zhao, Z., & Thottan, M. (2011). Adapting router buffers for energy efficiency[C]. *Proceedings of the Seventh Conference on Emerging Networking Experiments and Technologies* (pp. 1–12). New York: ACM.

# Chapter 111

## Real-Time Filtering Method Based on Neuron Filtering Mechanism and Its Application on Robot Speed Signals

Wa Gao, Fusheng Zha, Baoyu Song, Mantian Li, Pengfei Wang,  
Zhenyu Jiang, and Wei Guo

**Abstract** In order to implement the real-time filtering and tracking of robot signals with high efficiency, a novel real-time filtering method based on neuron filtering mechanism is developed in this paper. By considering the ubiquity of resonance in mammal and combining the mechanism of neural information processing, the derived details and the feasible parameter criterion under minimum error variance condition are given. For illustration, the application on quadruped robot is discussed. The quadruped robot feet speed signals are processed by developed real-time filtering method and Kalman filtering algorithm, respectively, and the computation time of both methods is tested. Experiment results show that the performance of developed real-time filtering method is better than that of Kalman filtering algorithm, not only in filtering and tracking performance but also in filtering speed. The novel real-time filtering method based on neuron filtering mechanism can effectively implement the real-time filtering and tracking with regard to robot signals.

### 111.1 Introduction

The mechanism of biological neuron filtering is being a subject of intense study since several years. Since the work of Hubel et al. [1], the filtering properties of single neurons in auditory cortex have been known and attract the interests of professionals [2, 3]. Recently, it is well known that biological neurons produce temporal filtering property and exhibit responses to vibration stimuli in neuron [4, 5]. It is a real-time

---

W. Gao • F. Zha • M. Li (✉) • P. Wang • Z. Jiang • W. Guo  
State Key Laboratory of Robotics and System, Harbin Institute of Technology,  
Harbin 150006, China  
e-mail: [skymoon.hit@gmail.com](mailto:skymoon.hit@gmail.com)

B. Song  
Department of Mechanical Design, Harbin Institute of Technology, Harbin 150006, China

information transferring process in mammal. Accordingly, it can inspire us to develop a novel filtering approach from the physiological standpoint.

Besides, vibration is ubiquitous in mammal and resonance phenomena can be widely found in vibrations such as the beating of a heart, the conveying of a sound, and the breathing of mammals [6, 7]. Resonance, which refers to the tendency of a system to oscillate with greater amplitude at some frequencies than at others, provides filtering ability by generating vibrations of a specific frequency or picking out specific frequencies from a complex vibration containing many frequencies [8, 9]. Hence, it is available to probe real-time filtering method starting from vibration.

In this paper, we are interested in developing a real-time filtering method by using biological neuron filtering mechanism. The paper is organized as follows. Firstly, the derived details and the feasible parameter criterion under minimum error variance condition are presented. Secondly, experiments of quadruped robot feet signals are given, and comparisons between developed real-time filtering method and Kalman filtering algorithm are discussed. Then, the paper is completed with some concluding remarks.

## 111.2 Real-Time Filtering Method Based on Neuron Filtering Mechanism

### 111.2.1 Methods

Generally, the single-degree-of-freedom vibration system can be seen as a typical and basic unit in vibration domain, and its mathematical expression is

$$f'' + \alpha f' + \beta f = F \tag{111.1}$$

where  $f$  is the system displacement output,  $F$  is the external excitation, and  $\alpha, \beta$  are the system parameters, respectively. Denote  $F = \beta u$ , Eq. 111.1 can be rewritten as

$$f'' + \alpha f' + \beta f - \beta u = 0 \tag{111.2}$$

where  $u$  represents the system input. We can solve Eq. 111.2 by the following form:

$$\begin{cases} \dot{f}' = x \\ \dot{x}' = \beta u - \beta f - \alpha x \end{cases} \tag{111.3}$$

where  $x$  represents the system velocity. Equation 111.3 represents the state function of single-degree-of-freedom vibration system. Define  $n \in \mathbb{N}$ , and  $\mathbb{N}$  represents the set of integers. Denote the sampling instant  $t_n$  and the sampling interval  $\Delta t = t_n - t_{n-1}$ . When the sampling interval  $\Delta t$  approaches zero, Eq. 111.3 can be discretized as Eq. 111.4.

$$\begin{cases} f_{n+1} = x_n \\ x_{n+1} = \beta u_n - \beta f_n - \alpha x_n \end{cases} \quad (111.4)$$

Then, we can obtain the numerical solution of Eq. 111.4 as follows:

$$\begin{cases} f_{n+1} = f_n + \Delta f_n \\ x_{n+1} = x_n + \Delta x_n \end{cases} \quad (111.5)$$

where

$$\begin{cases} \Delta f_n = (K_1 + K_2)/2 \\ \Delta x_n = (L_1 + L_2)/2 \end{cases} \text{ and } \begin{cases} K_1 = x_n \\ K_2 = x_n + hL_1 \\ L_1 = \beta u_n - \beta f_n - \alpha x_n \\ L_2 = (1 - ah)L_1 - \beta hK_1 \end{cases} \quad (111.6)$$

where  $h > 0$  is the step of numerical solution.

Information in neuron is transferred by synapses. The generations of temporal filters are shown among different synapses, and there exist information exchange processes while neural information transferring [10, 11]. As referred in ref. [12], we assume that Eq. 111.5 is a neural information delivery system,  $\Delta f_n$  and  $\Delta x_n$  are information segments transferred by different synapses. Hence,  $\Delta f_n$  and  $\Delta x_n$  shall exchange to complete the information transferring process. Thus,

$$\begin{cases} f_{n+1} = f_n + \Delta x_n \\ x_{n+1} = x_n + \Delta f_n \end{cases} \quad (111.7)$$

Define

$$\begin{cases} a = -\beta h^2/2 \\ b = h - \alpha h^2/2 \end{cases} \quad (111.8)$$

Substituting Eqs. 111.6 and 111.8 into Eq. 111.7 yields

$$\begin{cases} f_{n+1} = (1 - \beta b)f_n + (a - ab)x_n + \beta b u_n \\ x_{n+1} = a f_n + (1 + b)x_n - a u_n \end{cases} \quad (111.9)$$

Equation 111.9 is the kernel filtering function of developed filtering method.  $\alpha$ ,  $\beta$ ,  $a$ , and  $b$  are the parameters. Its derivation mainly depends on the principle of neuron filtering mechanism, and it is quite different from common filtering theory.

### 111.2.2 Preferences Under Minimum Error Variance Condition

Considering the filtering process and the observing process, we have

$$\begin{cases} X_{n+1} = \Phi_n X_n + B_n U_n \\ Z_{n+1} = H_{n+1} X_{n+1} + V_{n+1} \end{cases} \tag{111.10}$$

Denote

$$\Phi = \begin{bmatrix} 1 - \beta b & a - ab \\ a & 1 + b \end{bmatrix}, B = \begin{bmatrix} \beta b \\ -a \end{bmatrix} \tag{111.11}$$

where  $X_{n+1} = [f_{n+1} \ x_{n+1}]^T$  is the state sequence,  $Z_{n+1}$  is the observation sequence,  $\Phi$  is the state transfer matrix,  $H = [1 \ 0]$  is the observation matrix, and  $V$  is the white noise caused by observation equipment with the expectation  $E[V] = 0$  and the variance value  $\sigma[V] = r$ . Then, we can obtain the system step prediction equation:

$$\begin{cases} \hat{X}_{n+1,n} = \Phi_{n+1,n} \hat{X}_n + B_{n+1,n} U_n \\ \hat{Z}_{n+1,n} = H_{n+1} \hat{X}_{n+1,n} \end{cases} \tag{111.12}$$

where  $\hat{X}_{n+1,n}$  and  $\hat{Z}_{n+1,n}$  are the step predicted state value and the step observed value, respectively. Define the observed error as  $\tilde{Z}_{n+1,n}$ , and  $\tilde{Z}_{n+1,n} = Z_{n+1} - \hat{Z}_{n+1,n}$ . It can be used to upgrade the step predicted state value  $\hat{X}_{n+1,n}$ . Thus, we have

$$\hat{X}_{n+1} = \hat{X}_{n+1,n} + \tilde{Z}_{n+1,n} = \hat{X}_{n+1,n} + G_{n+1} (Z_{n+1} - H_{n+1} \hat{X}_{n+1,n}) \tag{111.13}$$

where  $G_{n+1}$  is the undetermined filtering gain matrix.

According to the linear minimum variance estimation method, the system error variance matrix function can be defined as follows:

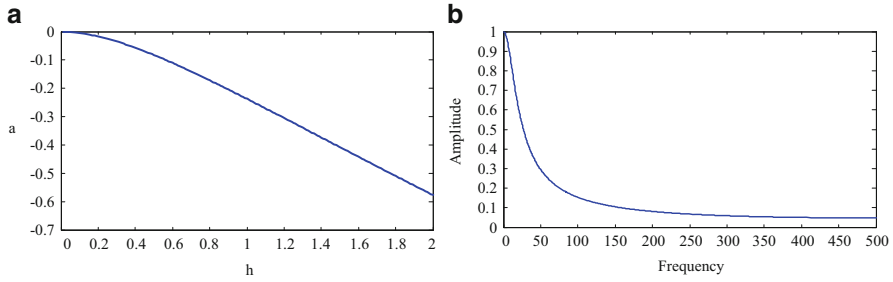
$$J_{n+1} = E[\tilde{X}_{n+1} \tilde{X}_{n+1}^T] \tag{111.14}$$

where  $\tilde{X}_{n+1}$  is the state error, and  $\tilde{X}_{n+1} = X_{n+1} - \hat{X}_{n+1}$ . Hence,

$$J_{n+1} = [I - G_{n+1} H_{n+1}] J_{n+1,n} [I - G_{n+1} H_{n+1}]^T + G_{n+1} J_{n+1} G_{n+1}^T \tag{111.15}$$

where  $J_{n+1,n} = \Phi_{n+1,n} J_n \Phi_{n+1,n}^T$  is the step error variance matrix.

Minimizing the system error variance matrix  $J_{n+1}$ , we can obtain the optimized gain matrix  $G = J_{n+1} H_{n+1}^T r^{-1}$ . When the filtering process is stable, the error variance is constant value, i.e.,  $J_{n+1} = J$ . Thus,



**Fig. 111.1** (a) The relationship between parameter  $a$  and  $h$ .  $h$  is the step of numerical method. (b) The amplitude-frequency figure of developed real-time filtering method when the parameter combination is selected as Table 111.1

**Table 111.1** The parameter combination

$a$	$b$	$\alpha$	$\beta$
-0.002	0.003	10	350

$$J = (I - J_r^{-1})\Phi J\Phi^T \tag{111.16}$$

Substituting Eqs. 111.8 and 111.11 into 111.16, we can yield the feasible parameter criterion as follows:

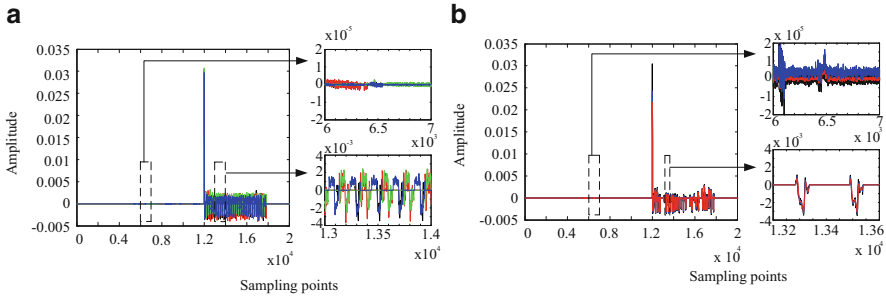
$$2a^3h^2 + 4a^2(h + 1) + 2ah^2(h + 2) + h^4 = 0 \tag{111.17}$$

where  $h$  is the step of numerical solution. The relationship between parameter  $a$  and  $h$  can be seen clearly in Fig. 111.1a. According to the value of  $h$ , the acceptable parameter combinations under minimum error variance condition can be obtained by Eqs. 111.8 and 111.17, and  $h$  depends on different scenarios.

The data in Table 111.1 is one of the feasible parameter combinations which meet Eqs. 111.8 and 111.17, and the corresponding amplitude-frequency figure is shown in Fig. 111.1b. The parameter combination is selected by the motion properties of quadruped robot, and the details will be given in later section. Then, the above-proposed real-time filtering method (consists of Eqs. 111.8, 111.9, and 111.17) will be substantially testified.

### 111.3 Experiments

To validate the developed real-time filtering method, we compare it with the Kalman filtering algorithm by using speed signals of quadruped robot feet which are sampled by accelerometers. And then, we test the computation time of both methods in LPC2148.



**Fig. 111.2** (a) The sampled quadruped robot feet speed signals. Different color lines represent different robot feet. Figures on the *right* show local enlarge parts of the *left* figure. (b) The comparisons between the developed real-time filtering method and the Kalman filtering algorithm. The *black*, the *blue*, and the *red* represent the measured speed signal, the filtering result of Kalman filtering algorithm, and the filtering result of developed real-time filtering method, respectively. Figures on the *right* show local enlarge parts of the *left* figure

The measured quadruped robot feet speed signals are shown by different color lines in Fig. 111.2a. The sampling points are 20,000, and the sampling frequency is 200 Hz. The interference noises caused by robot natural vibration and signal mensuration process can be seen clearly from local enlarge figures, and the frequencies of interference noises are uncertain. Towards the sampled quadruped robot feet speed signals, the useful signals generated by the motion of quadruped robot feet are quite small, and the frequencies are less than 10 Hz due to the dynamic balance and self-weight of quadruped robot. It is an extreme narrow band-pass width of frequency. Hence, in order to acquire the useful signals and track accurately, h is demanded to select a small value. Here we choose  $h = 0.1$ . The corresponding parameter combination obtain by Eqs. 111.8 and 111.17 is listed in Table 111.1, and the amplitude-frequency figure is shown in Fig. 111.1b.

For the motor behaviors of robot feet are similar, we take one foot as an example. The filtering comparisons are shown in Fig. 111.2b. The black, the red, and the blue represent the measured robot foot speed signal, the filtering result of developed real-time filtering method, and the filtering result of Kalman filtering algorithm, respectively. Obviously, the filtering and tracking performance of developed real-time filtering method is better than that of Kalman filtering algorithm, and the wave form of filtered signal is much smoother. It indicates that the developed real-time filtering method implements filtering and tracking functions at the same time, while the Kalman filtering algorithm tends to track.

We also compare the computation time of two filtering methods in LPC2148 whose main frequency is 60 MHz. For 5,000 sampling points, the comparison result is shown in Table 111.2. It indicates that the computation time of developed real-time filtering method is about 3 times than that of Kalman algorithm under the same condition.

**Table 111.2** The computation time in LPC2148 (milliseconds)

Developed real-time filtering method	Kalman filtering algorithm
2.75	7.75

From Fig. 111.2 and Table 111.2, the real-time filtering method based on neuron filtering mechanism, which shows better performance than Kalman filtering algorithm towards robot signals, meets higher real-time and better tracking and filtering requirements.

## 111.4 Conclusion

A real-time filtering method based on neuron filtering mechanism was proposed and applied to quadruped robot speed signals. Filtering function and preferences under minimum error variance condition are given, and the filtering and tracking performance are discussed compared with Kalman filtering algorithm. Besides, the computation time of the two methods are measured in LPC2148, and the results show that the filtering speed of developed real-time filtering method is several times faster than that of Kalman filtering algorithm.

**Acknowledgements** This work is partially supported by the National Natural Science Foundation of China (Nos. 60901074, 51075092, 61175107, 61005076), the National High Technology Research and Development Program ("863" Program) (No. 2007AA042105) of China, the Natural Science Foundation of Heilongjiang Province in China (No. E200903), and the State Key Laboratory of Robotics and System (HIT) (No. SKLRS200801B).

## References

1. Hubel, D. H., Henson, C. O., Rupert, R., & Galambos, R. (1959). Attention units in the auditory cortex. *Science*, *129*, 1279–1280.
2. Schnupp, J. (2006). Auditory filters, features, and redundant representations. *Neuron*, *51*, 278–280.
3. Romanyshyn, Y., & Pukish, S. (2011). Filtering signals in models of neurons and neurthods in MEMS Design (MEMSTECH), Proal networks. Perspective Technologies and Meceedings of VIIth International Conference on. IEEE, (p.191).
4. Hrocholle-Bossavit, B., & Quenet, B. (2009). Neural model of frog ventilatory rhythmogenesis. *Biosystems.*, *97*, 35–43.
5. Xie, J. L., Wang, Z. J., & Shi, H. B. (2010). Effect of Synaptic Plasticity on Correlation between Neural Spike Trains. Information Engineering and Computer Science (ICIECS), 2nd International Conference on. IEEE, (pp.1–4).
6. Izhikevich, E. M., Desai, N. S., Walcott, E. C., & Hoppensteadt, F. C. (2003). Bursts as a unit of neural information: selective communication via resonance. *Trends in Neurosciences*, *26*, 161–167.
7. Daniel, I. J. (2001). *Engineering vibration* (pp. 28–60). Upper Saddle River, NJ: Prentice Hall.



8. Thompson, W. T. (1996). *Theory of vibrations* (pp. 35–55). Cheltenham, England: Nelson Thornes Ltd.
9. Zhang, J. J., & Jin, Y. F. (2012). Stochastic resonance in FHN neural system driven by non-Gaussian noise. *Acta Physica Sinica*, *61*, 13 (In Chinese).
10. Edwards, C. J., Alder, T. B., & Rose, G. J. (2005). Pulse rise time but not duty cycle affects the temporal selectivity of neurons in the anuran midbrain that prefer slow AM rates. *Journal of Neurophysiology*, *93*, 1336–1341.
11. Xie, X. P., Song, D., Wang, Z., Marmarelis, V. Z., & Berger, T. W. (2006). Interaction of short-term neuronal plasticity and synaptic plasticity revealed by nonlinear systems analysis in dentate granule cells. Engineering in Medicine and Biology Society, 28th Annual International Conference of the IEEE, (pp.5543–5546).
12. Fortune, E. S., & Rose, G. J. (2001). Short-term synaptic plasticity as a temporal filter. *Trends in Neurosciences*, *24*, 381–385.

# Chapter 112

## Multiple-View Spectral Embedded Clustering Using a Co-training Approach

Hong Tao, Chenping Hou, and Dongyun Yi

**Abstract** It is a challenging task to integrate multi-view representations, each of which is of high dimension to improve the clustering performance. In this paper, we aim to improve the clustering performance of spectral clustering method when the manifold for high-dimensional data is not well defined in the multiple-view setting. We abstract the discriminative information on each view by spectral embedded clustering which performs well on high-dimensional data without a clear low-dimensional manifold structure. We bootstrap the clusterings of different views using discriminative information from one another. We derive a co-training algorithm to obtain a most informative clustering by iteratively modifying the affinity graph used for one view using the discriminative information from the other views. The approach is based on the assumption that the clustering from one view should agree with the clustering from another view. Comprehensive experiments on four real-world multiple-view high-dimensional datasets are presented to demonstrate the effectiveness of the proposed approach.

### 112.1 Introduction

In many important data mining applications, an instance may have multiple representations (views) from different feature spaces [1]. For example, a document can be translated to multiple languages, an Internet webpage can be represented as page-text as well as the hyperlinks pointing to it [2]. The phenomenal impact of multiple-view data in many applications has raised interest in the so-called multiple-view learning. Although each single view of the data might be sufficient for a given learning task, the complementary information of different views is ignored by learning from each view separately. So the main challenge of multiple-view

---

H. Tao (✉) • C. Hou • D. Yi  
Department of Mathematics and Systems Science, National University of Defense  
Technology, Changsha 410073, China  
e-mail: [taohong08@sina.com](mailto:taohong08@sina.com)

learning is to develop algorithms that can integrate complementary information of different views to improve the learning performance [3].

Co-training is the first algorithm to deal with multi-view data. It assumes each view of the data is sufficient for learning and the two views are conditionally independent. In the original co-training algorithm, a separate classifier for each view is learned using any labeled examples. The most confident predictions of each classifier on the unlabeled data are then used to iteratively construct additional labeled training data. This process should slowly drive the two classifiers to agree with each other on labels [1].

In this paper, we focus on multi-view clustering particularly in multi-view spectral clustering (SC). The available literature for this topic is growing with encouraging results [4–7]. However, the success of traditional SC methods is largely dependent on the manifold assumption, and this assumption does not always hold on high-dimensional data. When the data do not exhibit a clear low-dimensional manifold structure (e.g., high-dimensional and sparse data), the clustering performance of SC degrades. The spectral embedded clustering (SEC) [8] solves this problem in the single-view setting, but little work takes this issue into consideration in the multiple-view setting.

We propose a new approach named co-trained multi-view spectral embedded clustering (CoSEC) to solve the problem mentioned above. We assume that the clustering from one view should agree with the clustering from the other view and bootstrap the clusterings of different views using information from one another by co-training. In particular, we use the cluster assignment matrix obtained by the SEC algorithm on one view to modify the affinity graph used for the other view. By iteratively applying this approach, the clusterings of the two views tend to each other. And then we extend the proposed co-training framework for more than two views. The cluster assignment matrix on each view got by SEC algorithm can reflect the local and global structure information of the data. Then the co-training iterations retain the information needed for clustering and throw away the within-cluster details which might be confusing. These two factors result in a better clustering performance.

The rest of this paper is organized as follows. In Sect. 112.2, we will briefly review the SEC approach. Our proposed CoSEC framework is then presented in Sect. 112.3. Experiments on four real-world datasets are displayed in Sect. 112.4, and the conclusion remarks are given in Sect. 112.5.

## 112.2 SEC Revisited

SEC is a variant of SC methods to enhance the clustering performance on single-view high-dimensional data, motivated by the observation that the true cluster assignment matrix for high-dimensional data can be always embedded in a linear space spanned by the data. We assume  $X = [x_1, x_2, \dots, x_n] \in \mathbb{R}^{d \times n}$  is the high-dimensional data

of only one view and  $Y$  is the corresponding cluster assignment matrix. Thus, there exist  $W \in \mathbb{R}^{d \times c}$  and  $b \in \mathbb{R}^{c \times 1}$  such that  $Y = X^T W + \mathbf{1}_n b^T$ . This equation usually holds for the high-dimensional and small-sample-size problem, which is usually the case in many real-world applications [9].

Suppose  $A$  is a symmetric matrix with each entry  $A_{ij}$  representing the affinity of a pair of data points; the normalized Laplacian graph  $L$  is then defined by  $L = I_n - D^{-1/2} A D^{-1/2}$ , where  $D$  is a diagonal matrix with the diagonal elements as  $D_{ii} = \sum_j A_{ij}$ ,  $\forall i$ . Denote the relaxed cluster assignment matrix as  $U \in \mathbb{R}^{n \times c}$  and the trace operator of a matrix  $A$  as  $tr(A)$ . SEC expects the learned  $U$  is close to a linear space spanned by the data  $X$ ; thus, the optimization problem is [8, 10]

$$\min_{U^T U = I_c, W, b} tr(U^T L U) + \mu \left( \|X^T W + \mathbf{1}_n b^T - U\|^2 + \gamma tr W^T W \right) \quad (112.1)$$

where  $\mu$  and  $\gamma$  are two trade-off parameters to balance three terms. In Eq. 112.1, the first term reflects the smoothness of data manifold, while the second term characterizes the mismatch between the relaxed cluster assignment matrix  $U$  and the low-dimensional representation of the data.

For simplicity, we assume the data is centered, i.e.,  $X \mathbf{1}_n = \mathbf{0}$ . Set the derivatives of the objective function with respect to  $b$  and  $W$  to zeros; we have

$$b = \frac{1}{n} U^T \mathbf{1}_n \text{ and } W = (X X^T + \gamma I_d)^{-1} X U \quad (112.2)$$

Substitute  $W$  and  $b$  in Eq. 112.1 by Eq. 112.2; the optimization problem Eq. 112.1 becomes

$$\min_{U^T U = I_c} tr \left( U^T \left( L + \mu \tilde{L} \right) U \right) \quad (112.3)$$

where  $\tilde{L} = H_n - X^T (X X^T + \gamma I_d)^{-1} X$ , and  $H_n = I - \frac{1}{n} \mathbf{1}_n \mathbf{1}_n^T$  is the centering matrix. The global optimal solution  $U^*$  to Eq. 112.3 can be relaxed as the eigenvector of  $L + \mu \tilde{L}$  corresponding to the  $c$  smallest eigenvalues. Based on  $Y \in \mathbb{B}^{n \times c}$ , the discrete-valued cluster assignment matrix  $Y$  can be obtained by K-means or spectral rotation [11].

### 112.3 Co-training for Spectral Embedded Clustering

In this section, we will present our CoSEC algorithm, which aims to improve the clustering performance on the high-dimensional data without a clear manifold structure.

We cannot make use of the semi-supervised co-training directly because there is no labeled data in unsupervised learning problems. However, the motivation still remains: to restrict the classification (in our problem, clustering) in one view to be consistent with those in other views [6].

For a Laplacian matrix with exactly  $k$  number of connected components, the first  $k$  eigenvectors of it are the cluster assignment vectors. That is to say, these  $k$  eigenvectors only contain discriminative information about the different clusters, leaving out the details within the same cluster. In the case that the Laplacian matrix is fully connected, the eigenvectors are no longer the cluster assignment vectors, yet they still contain discriminative information which can be used for clustering [12]. In addition, the eigenvectors obtained by SEC capture local and global discriminative information because SEC balances the smoothness of data manifold (local) and the mismatch between the relaxed clustering assignment matrix  $U$  and the low-dimensional representation of the data (global). In multi-view setting, we exploit the eigenvectors obtained by SEC on one view to modify the graph structure in the other view, and vice versa.

How to modify the graph structure in one view using discriminative information from the other view? We consider each column  $\mathbf{a}_i$  of the affinity matrix  $A \in \mathbb{R}^{n \times n}$  as an  $n$ -dimensional vector that indicates the affinities of the  $i$ th point with all the points in the graph. The eigenvectors of the Laplacian matrix are vectors in the space spanned by these  $n$  affinity vectors. As above analyzed, the first  $k$  eigenvectors have the discriminative information for clustering, so projecting the affinity vectors along these directions (eigenvectors) can reserve the information needed for clustering and drop the within-cluster details that might confuse us. Then project them back to the original  $n$ -dimensional space, we get the modified graph. The inverse projection is easily finished by multiplying the transpose of the projection matrix thanks to its orthogonality [6].

Let us denote the data matrix, the affinity matrix, and the relaxed clustering assignment matrix on the  $v$ th view as  $X^{(v)}$ ,  $A^{(v)}$ , and  $U^{(v)}$  ( $v = 1, 2$ ), respectively. Then the modified affinity matrixes on both views are  $S^{(1)} = \text{sym}(U^{(2)}U^{(2)T}A^{(1)})$  and  $S^{(2)} = \text{sym}(U^{(1)}U^{(1)T}A^{(2)})$ , where  $\text{sym}(S)$  is the symmetrization operator on a matrix  $S$ . We symmetrize  $S^{(v)}$  because the projection of affinity matrix  $A^{(v)}$  on the eigenvectors does not yield a symmetric matrix. And we add a rank-1 matrix to  $\text{sym}(S)$  that has all its entries equal to the absolute value of the minimum negative entry of  $\text{sym}(S)$ . This makes sure that the corresponding Laplacian matrix is positive semi-definite at each iteration. We iteratively repeat this process by using  $S^{(1)}$  and  $S^{(2)}$  as new affinity matrixes to conduct SEC on each view to get new relaxed cluster assignment matrixes.

For the data have more than two views, we take the affinity matrix  $A^{(v)}$  of a view and project it onto the union of subspaces spanned by top  $k$  discriminative eigenvectors of the other views, i.e.,  $S^{(v)} = \text{sym}((\sum_{i \neq v} U^{(i)}U^{(i)T})A^{(v)})$  for all the views [6]. Algorithm 1 gives a detailed description of the CoSEC algorithm.

**Algorithm 1:** CoSEC Algorithm**Input:**

Data matrix  $X^{(v)}$  and affinity matrix  $A^{(v)}$  for each view,  $v = 1, 2, \dots, l$

The number of iteration  $iter$ , the number of clusters  $c$

**Output:** Assignment to  $c$  clusters

**Initialize:** Compute the matrixes  $L_0^{(v)} + \mu^{(v)}\tilde{L}_0^{(v)}$ , solve Eq. 112.3 by SEC, and obtain the optimal

$$U_0^{(v)}, v = 1, 2, \dots, l$$

**for**  $i = 1$  to  $iter$  **do**

1:  $S^{(v)} = \text{sym}((\sum_{j \neq v} U_{i-1}^{(j)} U_{i-1}^{(j)T}) A_{i-1}^{(v)})$ ,  $v = 1, 2, \dots, l$ .

2: Use  $S^{(v)}$  as the new affinity matrix, i.e.,  $A_i^{(v)} = S^{(v)}$ , and compute the matrix  $L_i^{(v)} + \mu^{(v)}\tilde{L}_i^{(v)}$  and conduct SEC to obtain  $U_i^{(v)}$ ,  $v = 1, 2, \dots, l$ .

**end for**

3: Row-normalize  $U_{iter}^{(v)}$ .

4: Let  $V = U_{iter}^{(v^*)}$ , where  $v^*$  is believed to be the most informative view. If there is no prior knowledge on the view informativeness, matrix  $V$  can also be set to column-wise concatenation of all  $U_{iter}^{(v)}$ s.

5: Based on  $V$ , compute the discrete cluster assignment matrix  $Y$  by using K-means clustering or spectral clustering.

## 112.4 Experiments

In this section, several experiments are conducted on four real-world datasets to compare our CoSEC approach with a number of baselines. The baselines are:

- Single view: The view that achieves the best SEC performance using a single view of the data, which is called the most informative view.
- Feature concatenation (FC): Concatenating the features of each view, and then running SEC using the joint view representation of the data.
- CCA-based feature extraction (CCA): Applying CCA for feature fusion from multiple views of the data [13] and then running SEC using these extracted features.
- Multi-view partitioning via tensor methods (TensorSC) [7]: Building up an affinity tensor from multiple affinity matrices and obtaining a joint optimal subspace by tensor decomposition, partitioning the subspace to obtain the cluster labels.
- Co-trained multi-view spectral clustering [6] (CoSC): This is most closely related to our approach. The only difference is that CoSC uses SC to obtain the relaxed cluster assignment matrix.

The four real-world datasets used in our experiments are the 3sources dataset, the UCI Handwritten digits dataset (Digits), the Internet advertisement dataset (AD), and the WebKB dataset. Some datasets are resized, and Table 112.1 summarizes the details of the datasets used in the experiments.

**Table 112.1** Dataset description

Dataset	Size	Views	Classes
3sources	169	3	6
Digits	2,000	2	10
AD	3,264	2	2
WebKB	1,051	2	2

We employ the clustering accuracy ( $Acc$ ) to evaluate the performance for all the clustering algorithm.  $Acc$  discovers one-to-one relationship between clusters and the true classes. It measures the extent to which cluster contains examples from the corresponding category.

We adopt the Gaussian similarity with cosine distance to construct the graph for spectral partition. That is to say, the entry of the affinity matrix is defined as

$$A_{ij} = \exp(-(1 - \cos \theta)/(2\delta^2)) \quad (112.4)$$

Usually this local similarity is not sensitive to  $\delta$ , so we put  $\delta = 0.5$ .

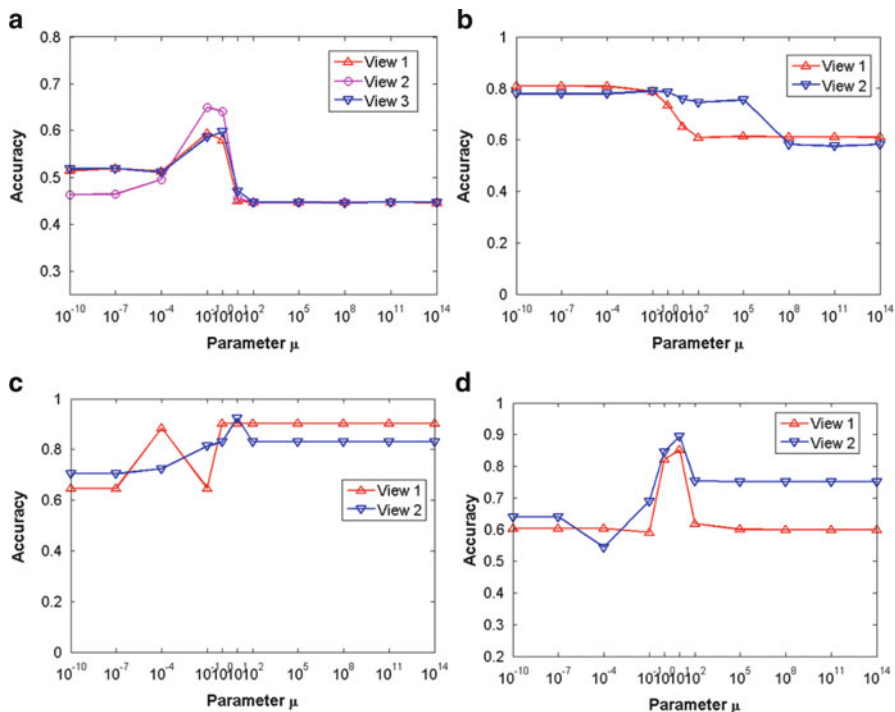
For fair comparison, when conducting SEC, we set the parameter  $\gamma$  as 1, and the parameter  $\mu$  as  $\{10^{-10}, 10^{-7}, 10^{-4}, 10^{-1}, 10^0, 10, 10^2, 10^5, 10^8, 10^{11}, 10^{14}\}$  in all approaches except TensorSC and CoSC. For CoSEC, we set the parameter  $\mu$  the same value on each view though it may be more practical to set different values according to the essence of data on different views.

We use the spectral rotation to calculate the discrete assignment matrix  $Y$  based on the relaxed continuous solution. To reduce statistical variety, we independently repeat spectral rotation for 50 times with random initialization for all methods, and then we report the mean clustering accuracy and standard deviation corresponding to the best parameters.

We first study the sensitivity of the clustering performances of CoSEC with respect to parameter  $\mu$  in Fig. 112.1. CoSEC favors a small value on 3sources data and Handwritten digits data, and its performances on these datasets are relatively stable when a small value is set for  $\mu$ , while it prefers an intermediate value of  $\mu$  for WebKB dataset and a larger one for the Internet advertisement data. These indicate that the 3sources data and Handwritten digits data have a clear manifold structure. Setting a large value for  $\mu$  implies that the linearity regularization is more important for SEC, so the Internet advertisement dataset probably has a strong linearity relationship between the data matrix  $X$  and the cluster assignment matrix  $Y$ . For WebKB data, the balanced combination of data manifold and linearity regularization helps to bring about a better clustering performance.

For the comprehensive study of performances of various clustering methods, the results for the four datasets are shown in Table 112.2. The numbers in the brackets are the standard deviations of the clustering accuracy obtained with 50 different runs of spectral rotation with random initializations.

As it can be seen, our proposed CoSEC outperforms all the baselines on all datasets. For 3sources data, all methods are run first using any two views and then using all three views, and the best results are reported. On 3sources data, all the



**Fig. 112.1** Acc in different views of CoSEC with different  $\mu$  on the four datasets. (a) 3 sources, (b) Handwritten digits, (c) AD, (d) WebKB

**Table 112.2** Clustering accuracy (*Acc*) results on four datasets. Numbers in parentheses are the std. deviations

Method	Single view	FC	CCA	TensorSC	CoSC	CoSEC
3source	0.6523 (0.0058)	0.6256 (0.0046)	0.3275 (0.0129)	0.2522 (0.0139)	0.5339 (0.0189)	0.6749 (0.0131)
Digits	0.7418 (0.0371)	0.6983 (0.0319)	0.7558 (0.0065)	0.7543 (0.0056)	0.7133 (0.0300)	0.8093 (0.0020)
AD	0.9033 (0)	0.8227 (0.0061)	0.8400 (0)	0.8401 (0.0005)	0.6496 (0.0392)	0.9233 (0.0012)
WebKB	0.8173 (0)	0.6100 (0.0043)	0.7175 (0.0053)	0.7187 (0.0059)	0.7093 (0.0005)	0.8943 (0)

baselines except CoSEC perform worse than single-view SEC. On Handwritten digits dataset and WebKB dataset, CoSEC outperforms all the baselines by a significant margin. FC performs worst on both datasets. For the Internet advertisement data, the clustering accuracy of the best single view is slightly lower than that of CoSEC and higher than those of the other baselines. CoSC performs worst on this dataset.



## 112.5 Conclusion

In this paper, a CoSEC was proposed. It aims to improve the clustering performance of spectral clustering on multi-view high-dimensional datasets when the data manifold is not well defined. Based on the key assumption that the true underlying clustering is same for all views, we apply the idea of co-training to the unsupervised clustering. The central idea of our approach is to use the discriminative information on one view to modify the graph structure used for the other view. The discriminative information obtained by SEC on each view is a balance of the global and the local discriminative information. And the co-training iterations retain the information needed for clustering and throw away the within-cluster details. These two factors result in a better clustering performance. Empirical evaluation on the four different kinds of high-dimensional data shows the effectiveness of our CoSEC approach.

## References

1. Blum, A., & Mitchell, T. (1998). Combining labeled data with co-training [C]. *Proceedings of the Workshop on Computational Learning Theory* (pp. 92–100). San Francisco: Morgan Kaufmann.
2. Long, B., Yu, P.S., & Zhang, Z. (2008). A general model for multiple view unsupervised learning [C]. *Proceedings of the SIAM International Conference on Data Mining* (pp. 822–833). Atlanta, GA: SIAM.
3. Tzortzis, G. F., & Likas, C. L. (2010). Multiple view clustering using a weighted combination of exemplar-based mixture models. *IEEE Transaction on Neural Networks*, 21(12), 1925–1938.
4. de Sa, V. R. (2005). Spectral clustering with two views [C]. *Proceedings of the Workshop on Learning with Multiple Views, 22th International Conference on Machine Learning* (pp. 20–27). Bonn, Germany: ACM.
5. Zhou, D., & Burges, C.J.C. (2007). Spectral clustering and transductive learning with multiple views [C]. *Proceedings of the 24th International Conference on Machine Learning* (pp. 1159–1166). Corvallis, OR: ACM.
6. Kumar, A., & Daumé, H. (2011). A co-training approach for multiview spectral clustering [C]. *Proceedings of the 28th International Conference on Machine Learning* (pp. 393–400). Bellevue, WA: Omnipress.
7. Liu, X., Ji, S., Glänzel, W., & De Moor, B. (2013). Multiview partitioning via tensor methods. *IEEE Transactions on Knowledge and Data Engineering*, 25(5), 1056–1069.
8. Nie, F., Xu, D., Tsang, I.W., & Zhang, C. (2009). Spectral embedded clustering [C]. *Proceedings of the International Joint Conference on Artificial Intelligence* (pp. 1181–1186). Westerville, OH: Odyssey Press.
9. Ye, J. (2007). Least squares linear discriminant analysis [C]. *Proceedings of the 24th International Conference on Machine Learning* (pp. 1087–1093). Corvallis, OR: ACM.
10. Shi, J., & Malik, J. (2000). Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(8), 888–905.
11. Yu, S. X., & Shi, J. (2003). Multiclass spectral clustering [C]. *Proceedings of the 9th IEEE International Conference on Computer Vision* (pp. 313–319). Nice, France: IEEE Computer Society.

12. Luxburg, U. (2007). A tutorial on spectral clustering. *Statistics and Computing*, 17(4), 395–426.
13. Chaudhuri, K., Kakade, S.M., Livescu, K., & Sridharan, K. (2009). Multi-view clustering via canonical correlation analysis [C]. *Proceedings of the 26th International Conference on Machine Learning* (pp. 129–136). Montreal, Canada: ACM.

# Chapter 113

## Feedback Earliest Deadline First Exploiting Hardware Assisted Voltage Scaling

Chuansheng Wu

**Abstract** In this paper, we examine the merits of hardware/software co-design of a feedback dynamic voltage scaling algorithm and a new processor are capable of executing instructions in the frequency and voltage conversion. We study several energy-aware feedback schemes based on earliest-deadline-first scheduling, dynamic adjustment of the behavior of the system, for different workload characteristics. An infrastructure for investigating several hard real-time dynamic voltage scaling schemes, including our feedback dynamic voltage scaling algorithm, is implemented on an NEC 530 embedded board. System structure and algorithm overhead is evaluated for different dynamic voltage scaling schemes. Feedback dynamic voltage scaling algorithm saves at least more energy frequently than the previous dynamic voltage scaling algorithm, with an additional 18 % energy reduction peak savings.

### 113.1 Introduction

In order to reduce the power dissipation of CPU, the dynamic voltage scaling technique has received widespread support in recent years to extend battery life. Dynamic Voltage Scaling dynamically scales the processor core voltage up or down depending on the computational demand of the system. Switch to reduce the speed of the transistor with low power supply voltage also allows a lower clock frequency. Assuming that the voltage and frequency are linearly related, lower voltage and frequency result in reduced cubic power consumption [1].

We developed several energy-aware feedback schemes for dynamic voltage scaling algorithm feedback earliest deadline based on priority scheduling, dynamic adjustment according to different load characteristics of real-time system [2]. Feedback dynamic voltage scaling algorithms have been proposed in our previous work,

---

C. Wu (✉)

Software College, University of Science and Technology Liaoning, Anshan 114051,  
Liaoning, China  
e-mail: [gykwcs@163.com](mailto:gykwcs@163.com)

the simulation evaluation. Several of our improved algorithms are developed in this paper feedback scheme considering practical design and implementation issues of the actual embedded system structure. We are interested in studying the performance of the dynamic voltage scaling algorithm in an embedded environment where the overhead and the actual energy consumption can be measured quantitatively. The real-time scheduler itself, when integrated with a dynamic voltage scaling algorithm, may execute at several different CPU frequencies, which also requires accurate modeling of the system overhead.

### 113.2 Earliest-Deadline-First Scheduling with Dynamic Voltage Scaling Support

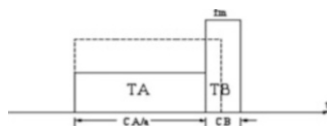
In order to evaluate energy-saving performance of dynamic voltage scaling algorithm in the embedded environment, we consider the scheduling problem with the earliest-deadline-first policy in hard real-time system. The scheduling framework is composed of two parts: (1) the earliest-deadline-first and (2) dynamic voltage scaling scheduling. These two components are independent of each other; therefore, earliest-deadline-first scheduling with dynamic voltage scaling algorithm is different. The earliest deadline is particularly attractive for dynamic voltage scaling algorithm because of its dynamic characteristics, which allows the dynamic voltage scaling scheme, using relaxation. Our dynamic voltage scaling scheduling is based on the feedback control incremental adjustments to the behavior of the system to reduce energy consumption.

A cycle uses the framework of fully preemptive and independent task model. Each task  $T_i$  is defined by a triple  $(C_i, P_i, c_i)$ , where  $C_i$  is the period of  $T_i$ ,  $P_i$  is the measured worst-case execution time of  $T_i$ , and  $c_i$  is the actual execution time of  $T_i$ . Each task's relative deadline,  $d_i$ , is equal to its period, and all tasks are released at time zero. The periodically released instances of a task are called jobs.  $T_j$  is used to denote the  $j$ th job of task  $T_i$ . Its release time is  $C_i * (j-1)$  and its deadline is  $C_i * j$ .  $P_{ij}$  is used to represent the actual execution time of job  $T_j$ . The hyperperiod  $H$  of the task set is defined as the least common multiplier (LCM) among all the tasks' periods. The schedule is repeated in each extended end.

### 113.3 Feedback Dynamic Voltage Scaling Algorithm

Dynamic voltage scaling algorithm for feedback on each task calls for a real execution time (work) based on feedback from a previous call to execution time. Then, a task execution budget is divided into two parts, as shown in Fig. 113.1. Frequency of the real-time CA is proportional to the minimum of the expected. On the contrary, the maximum frequency of the rest of the execution time is scalable,  $C_a + C_B = WCET$  (worst-case execution time).

Fig. 113.1 Task splitting



All future tasks are deferred as far as possible with the use of the maximum (worst) plan and schedule to achieve the task  $K$ . Therefore, currently available relaxation of SK shows the scaling factor and the corresponding minimum frequency. Through feedback schemes, the algorithm is able to capture changes in actual execution time. The current task preemption is expected in the future through the scheduling time slot allocation. This is the implementation of a backward scanning to fill idle times and finish the task ahead of the scheduled time slot (algorithms in detail see). Due to the even more greedy approach than any of the previous schemes, the algorithm was reported to exhibit additional energy savings in simulation experiments, particularly for medium utilization systems, which are quite common [3]. More substantial savings have been observed for execution time in pulsating PID feedback that provides new opportunity for positive scaling. NEC 530 of the embedded development board is in the process of implementation. We propose a feedback scheme by refining the following two feedback mechanisms.

### 113.3.1 Simple Feedback

Some periodic real-time workload is in a certain time interval of relatively stable behavior. The actual execution time of their different works remained almost unchanged or changes only in a very small range interval. For such workloads, we use a very simple feedback mechanism by computing the moving average of previous jobs' actual execution times and feed it back to the dynamic voltage scaling scheduler. We try to avoid the overhead of more complex feedback mechanisms, like the PID feedback controller as described in the next section, because a simple feedback usually in these situations provides a performance that is good enough. Quantitative comparison of the overhead of dynamic voltage scaling algorithm for our PID feedback and several other dynamic voltage scaling algorithms also makes us believe that a complex feedback dynamic voltage scaling scheme can reduce the number of expansion of its energy saving potential [4].

The simple feedback mechanism chooses the value of CA as the controlled variable. Each job  $T_{ij}$ 's actual execution time  $P_{ij}$  is chosen as the set point. CA is assigned to be 50 % WCET for the first job of each task, which means half of the job's execution is budgeted at a low frequency and half of it is reserved at the maximum frequency. The maximum frequency portion guarantees the deadline requirements, even if the worst-case execution time is used in full. Each time a job completes, its actual execution time is fed back and aggregated to anticipate the

next job’s CA. Let CA<sub>ij</sub> denote the CA value for T<sub>ij</sub>. The (j + 1)th job of the task is assigned a CA value according to:

$$CA_i(j + 1) = (CA_{ij} + ci(j + 1) - ci(j - N + 1))/N \tag{113.1}$$

where N is a constant that represents the number of items used in the moving average calculation.

### 113.3.2 PID Feedback

Multiple input-multiple output function is very difficult not only to accurately control the behavior of the system but also to increase the complexity of the algorithm. Therefore, we improve the original PID feedback dynamic voltage scaling mechanism with the use of the following simplified design.

Instead of using CA<sub>i</sub> (i = 1...n) as the controlled variable for each task T<sub>i</sub> and creating n different feedback controller for n different tasks, we now define a single variable r as the controlled variable for the entire system as:

$$r = \frac{1}{n} \sum_{i=1}^n \frac{CA_{ij} - c_{ij}}{c_{ij}} \tag{113.2}$$

$$e(t) = r - o \tag{113.3}$$

where J is the TI’s new job index in the sampling point. Our goal is to make R approximately 0 (i.e., the set point). System error is E (t PID) feedback to control the controlled variable R. PID feedback controller is now defined as:

$$\Delta r_j = K_p e_i(t) + \frac{1}{k_i} \sum IW e_i(t) + K_d \frac{e_i(t) - e_i(t - DW)}{DW} \tag{113.4}$$

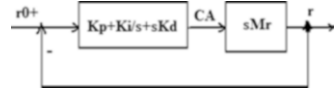
$$r_{j+1} = r_j + \Delta r_j$$

For each r<sub>j</sub>, we adjust the CA value for task T<sub>i</sub> by CA<sub>i</sub>(j + 1) = r<sub>j</sub> c<sub>ij</sub> + c<sub>ij</sub>. The transfer function Gr between r and CA can be derived by taking derivative of both sides of Eq. (113.2):

$$Gr(s) = Mrs \tag{113.5}$$

where  $M_r = \frac{1}{n} \sum_{i=1}^n \frac{1}{c_i}$ . The block diagram of the model is shown in Fig. 113.2. Its transfer function is:

**Fig. 113.2** Control loop model



$$\frac{G_p(s)G_r(s)}{1 + G_p(s)G_r(s)} = \frac{MK_{ps} + MK_i + MK_{ds}^2}{1 + MK_{ps} + MK_i + MK_{ds}^2} \tag{113.6}$$

According to control theory, a system is stable if and only if all the poles of its transfer function are in the negative half-plane of the s-domain. From Eq. (113.6), we can infer the pole of our system as follows:

$$\frac{-MK_p \pm \sqrt{MK_p^2 - 4MK_d(MK_i + 1)}}{2MK_d} \tag{113.7}$$

Note that  $-MK_p + \sqrt{MK_p^2 - 4MK_d(MK_i + 1)}$  is still less than 0 when  $MK_p^2 - 4MK_d(MK_i + 1) > 0$ . Hence, all the poles in the s domain are in the negative half plane, in order to ensure the stability of the system.

## 113.4 Experimental Evaluation

Through dynamic voltage scaling algorithm for our feedback in embedded architecture practical evaluation, we evaluate our algorithm’s potential for energy savings in a real system and simulation environment.

### 113.4.1 Platform and Methodology

The embedded platform used in our experiment is a 495LP embedded board running on a diskless MontaVista Embedded Linux variant, which is based on the 2.4. 21 stock kernel but has been patched to support dynamic voltage scaling on the PPC 405LP. This board provides the hardware support required for dynamic voltage scaling and allows software to scale voltage and frequency via user-defined operation points ranging from a high end of 266 MHz at 1.8V to a low end of 33 MHz at 1V. The board has been modified to 50 % reduced capacitor, which allows the dynamic voltage regulating switch to occur more rapidly, i.e., switch is composed of a maximum of 0 μs duration from 1 to 1.8 V.

This set of pairs was constrained by a need to have a common PLL multiplier of 16 relative to the 33MHz base clock and a divider of two or any multiple of 4. To change the multiplier brings the extra spending for switching, which we hope to eliminate in the research [5]. Dynamic power management (DPM) facilities are

**Table 113.1** Dynamic voltage scaling switching overhead

Activity	Sync. DVS	Sync. DVS	Signal handler
Overhead ( $\mu$ s)	117–162	8–20	0.07–0.6

used as an enhanced Linux kernel support for the dynamic voltage scaling function. The DPM operating point defines the stable frequency/voltage (and the related system parameters) that we experimentally determined.

### 113.4.2 Synchronous vs. Asynchronous Switch

We first evaluate the overhead of different dynamic voltage scaling techniques, through the expansion of DPM test plate support and operating system. A unique dynamic voltage scaling function supported by the NEC 530 embedded motherboard is that the frequency switching can be synchronous or asynchronous. The traditional method of processor frequency and voltage conversion is synchronous switching; the application must stop at the transition interval [6]. On the contrary, asynchronous switching of frequency and voltage allows the process to continue. Figure 113.2 describes the core of the voltage and current changes in the asynchronous switch of the PPC 530 processor [7].

Table 113.1 reports the overhead for synchronous and asynchronous switching in a time range bounded by two extremes: (a) exchange switch adjacent frequency/voltage level between and (b) between the lowest and highest frequency/voltage grade. In addition, the overhead of the following signal processing is also measured for a series of minimum and maximum processor frequencies with each asynchronous switch. The results show that the synchronous dynamic voltage regulating switch has about an order of magnitude greater overhead than asynchronous switching. The timer interrupt handler that is triggered only increases the overall cost by a little amount during each asynchronous switch.

### 113.4.3 Dynamic Voltage Scaling Scheduler Overhead

The overhead of our feedback dynamic voltage scaling algorithm is compared with several other dynamic voltage scaling algorithms. On the embedded development board, we first measured the execution time of the scheduling algorithm of the dynamic voltage scaling in different frequencies, as shown in Table 113.3. The overhead was obtained by measuring the amount of time when a task issues a yield system call till another task was dispatched by the scheduler. The table shows that the static dynamic voltage scaling has the lowest overhead among the four and our PID feedback dynamic voltage scaling has the highest one [8]. This is not surprising since static dynamic voltage scaling uses a very simple strategy to select the



**Table 113.2** Overhead of dynamic voltage scaling-earliest-deadline-first scheduler

CPU freq. (MHz)	Static	DVS	Idling over look-ahead	Egad [spec] CiD-feedback
33	217	487	2,296	3,652
44	170	366	1,714	2,943
66	100	232	1,112	1,728
133	52	120	546	801
266	36	76	229	472

**Table 113.3** Task set, times in msec

Task	Task set 1		Task set 2		Task set 3	
	Period (Ci)	WCET (Pi)	Period (Ci)	WCET (Pi)	Period (Ci)	WCET (Pi)
1	2,400	400	600	80	90	12
2	2,400	600	320	120	48	18
3	1,200	200	400	40	60	6

frequency and voltage falling short in finding the best energy saving opportunities. Cycle-conserving dynamic voltage scaling, look-ahead dynamic voltage scaling and our PID feedback dynamic voltage scaling use more sophisticated and aggressive algorithms for lower energy consumption, albeit at higher overheads. The trade-off between cost and performance needs to be carefully checked.

Next, we evaluate if feedback dynamic voltage scaling algorithm, although having the biggest cost among the four, provides the best energy-saving effect in the actual embedded environment. We measured the actual energy consumption of these dynamic voltage scaling algorithms when executing three medium utilization task sets depicted in Table 113.2 using both synchronous and asynchronous dynamic voltage scaling switchings. As a baseline for comparison, we also implemented a naive dynamic voltage scaling scheme, in which the maximum frequency is selected when a task is scheduled and the minimum frequency is selected when the system is idle.

The first task set in Table 113.3 is harmonic, i.e., the time is an integer multiple of the shortest cycle that facilitates scheduling. This will allow the scheduling algorithm to demonstrate the extreme behavior, usually outperforming any other choice of period. The second and third of the task sets are nonharmonic with longer and short periods, respectively. The actual execution time is half that of the WCET for each of the experimental tasks.

The naive dynamic voltage scaling algorithm is used as a base of comparison for each dynamic voltage scaling algorithm subsequently. In the task set, a static dynamic voltage scaling can reduce energy consumption by about 29 % over the naive scheme. Dynamic voltage scaling cycle can save 47 % energy. The look-ahead real-time dynamic voltage scaling can save more than 50 %, and our feedback method saves 54 % energy compared to the naive dynamic voltage scaling. This clearly shows great potential in energy saving for real-time scheduling.

## 113.5 Conclusion

We evaluated it as well as several other real-time dynamic voltage scaling algorithms on an NEC 530 embedded platform. We compared the energy consumption and scheduling overhead between different dynamic voltage scaling schemes. The experimental results show that the positive feedback dynamic voltage scaling algorithm in our energy consumption reached additional savings 24 % over the look-ahead dynamic voltage scaling algorithm and AGR-2 algorithm and up to 64 % energy savings over the naive dynamic voltage scaling scheme when considering the scheduling overhead.

## References

1. Jejurikar, R., & Gupta, R. (2012). Procrastination scheduling in fixed priority real-time systems. *Proceedings of the Language Compilers and Tools for Embedded Systems*, 9(2), 101–111.
2. Kang, D., Crago, S., & Suh, J. (2011). A fast resource synthesis technique for energy-efficient real-time systems. *IEEE Real-Time Systems Symposium*, 35(5), 204–207.
3. Lu, C., Stankovic, J. A., Abdelzaher, T. F., Tao, G., Son, S. H., & Marley, M. (2009). Performance specifications and metrics for adaptive real-time systems. *Proceedings of the IEEE Real-Time Systems Symposium*, 23(4), 1011–1013
4. Brock, B., & Rajamani, K. (2011). Dynamic power management for embedded systems. In *IEEE international SOC conference* (pp. 23). London: WET publishing.
5. Chandrakasan, A., Sheng, S., & Brodersen, R. W. (2012). Low-power CMOS digital design. *IEEE Journal of Solid-State Circuits*, 27(3), 473–484
6. Govil, K., Chan, E., & Wasserman, H. (2012). Comparing algorithms for dynamic speed-setting of a low-power CPU. *First International Conference on Mobile Computing and Networking*, 63(4), 483–494
7. Gruian, F. (2012). Hard real-time scheduling for low energy using stochastic data and DVS processors. In *Proceedings of the international, symposium on low-power electronics and design ISLPED'02* (pp. 77–86). New York: Addison-wesley Publishing.
8. Dirk Grunwald, Philip Levis, Charles B. Morrey, Michael Neufeld & Keith I. Farkas (2012). Policies for dynamic clock scheduling. In *Symposium on operating systems design and implementation* (pp. 235–247). Paris: Springer.

# Chapter 114

## Design and Realization of General Interface Based on Object Linking and Embedding for Process Control

Jiguang Liu, Jianbing Wu, and Zhiguo He

**Abstract** Based on the analysis of existing problems of interface software development process of industrial control, the importance of building the general interface system based on OPC (Object Linking and Embedding for Process Control) was proposed. The data model was given with database technology. On the basis of the data model, the configurable general interface system based on OPC was implemented. Versatility and configurability is the most important feature of the interface system. By simple modification of configuration information, the interface system will meet the needs of different projects. The application results show that the interface system greatly reduces the development cycle of the related software, improves the reliability and stability of the application system, and reduces costs of system operation and maintenance.

### 114.1 Introduction

In conventional control system, information interchange between hardware device and software of control system was done through device driver. Different control equipment manufacturers often use different communication protocol. To access production data on field device, users have to write particular communication interfaces to talk with those peripherals. Besides, industrial field devices are abundant in variety, and hardware products are updated constantly. Therefore, it is difficult to develop a suit of industrial control interface software for production information interchange which is universally applicable to all industrial devices. Consequently, “Information islands” are formed in industry field, which constrained further development and application in production fields.

---

J. Liu (✉) • J. Wu • Z. He  
School of Mathematics and Computer Science, Panzhuhua University,  
Panzhuhua 617000, China  
e-mail: [liujig@gmail.com](mailto:liujig@gmail.com)

To solve the existing problems in conventional control system, the industry launched OPC technology; the technology as the newest, the most widely used soft bus standards in the industry control domain, has been supported by basic automation manufacturers in general and was widely used in industrial field. Currently, although there are many interface applications based on OPC technology, but these interface systems had a common characteristic of high coupling between itself and application, it caused excessive customizability, poor flexibility, extensibility and independence of the interface system. The OPC interface given in [1–4] belonged to such applications. For these interface systems mentioned above, due to the lack of abstraction of the common characteristics of the interface systems, therefore, it was impossible to develop a reusable generic interface system. In these applications, the application developers wrote some important parameters (such as information of basic automation device, data acquisition point information) into interface program in the form of hard coding; the slight changes of external environment parameters will lead to the entire program to recompile and release, resulting in poor interface system adaptability to changes in the external environment. To buy mature industrial control software can eliminate the coupling between the interface system and application system, but industrial control software is generally not low cost, and different process control project must repeat to purchase the same software, resulting in project cost go straight up. In OPC technology maturing today, therefore, building a general interface system based on OPC is of great significance to solve the problem of “Information islands” of automation, improve the level of factory process control, reduce the cost of implementation of the process control system, and have rapid implementation of the production process control system.

## 114.2 System Design Objective and Essential Function

In order to meet various needs of different external system and automation system, the interface system must be configurable. For various external system and automation system, it can meet different application requirements with a simple modify configuration information. This interface system’s most outstanding characteristic is universality and configurability, which is the biggest difference between this interface system and other system based on OPC. Guided by the principle of universality and configurability, a design objective of the interface system was given as follows:

- The system can simultaneously connect to multiple OPC servers. (Generally, hardware from different manufacturers exist simultaneously in industry field.)
- The system is capable of dynamic configuration for data collecting point.
- The system has the general purpose interface to external system, which facilitates data exchange between basic automation system and external system and which also triggers process tracking logic of external system.
- The system must have producibility.

Interface system is a basis for solving “islands of automation” problem, achieving production data sharing, further expanding and applying of production fields. Therefore, interface system must realize the following two basic functions:

- It implements bidirectional process information exchange between any automation system based on OPC technology and external system.
- To establish triggering mechanism for tracking and monitoring production process, provide the conditions for tracking and monitoring production process.

### 114.3 Data Model Design

Data model is the core part of interface system design and is a basis of implementing universalization and configurability of interface system. Excellent data model can add good extensibility, flexibility, and adaptability to the interface system. On the analysis of domestic and foreign application experience of OPC, a data model as shown in Fig. 114.1 was obtained. Follow-up discussion centers around the data model which is a basis of implementation of the interface system. The following paragraphs introduce the key entities of the model.

#### 114.3.1 OPC\_SERVER

Usually, OPC servers from several control equipment manufacturers exist simultaneously in industry field. The total number of the servers may be changed at any time. To accomplish production data acquisition from control equipment in use, the

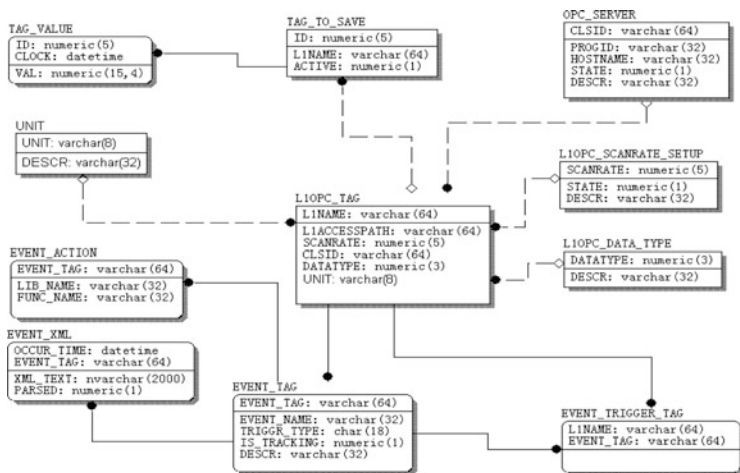


Fig. 114.1 Structure diagram of system data model

interface system must have the capability of dynamic connecting to OPC servers. This entity stores information of all OPC servers in the industry field, and the administrative staff can modify, add, or reduce the data in the table according to actual situations when increasing or decreasing or changing OPC server in the field. As the interface system starts up, it dynamically accesses the information of all OPC servers and connects to each OPC server. Consequently, the system has good extensibility and adaptability.

### ***114.3.2 LIOPC\_SCANRATE\_SETUP***

Every production data has a definite occurrence cycle. User can define scanning period according to characteristic of production data. This entity is used to hold user-defined rate of scanning on field data point. Users can increase or adjust scan rate at any time based on actual conditions. The latest scan rate can be acquired from configuration table when system starts. To facilitate setting up a mechanism of tracking and triggering of the production process, the interface system categorized the production data into four lists according to the characteristics of the production data. The categories are cycle data, event data, trigger data, and download data which are described, respectively, as follows.

#### **114.3.2.1 Cycle Data**

Cycle data is a large amount of analog signal acquired successively in the industry field, such as temperature, pressure, and flow. Data points whose scan rate is above 0 in configuration table all belong to cycle data point. Cycle data is characterized by volatility and continuity. Its value typically fluctuates frequently in a certain range. On-spot operators can quickly find whether the device or production process stays normal through monitoring this type of data. The asynchronous data access method was adopted to acquire cycle data. In consideration of the frequent fluctuation of the cycle data, in order to improve performance but also saving memory, the definition of opc-dead band was provided for interface system. OPC server of automation system will send varying data to the interface system only when automation system variation range of cycle data is above definition of the dead band.

#### **114.3.2.2 Event Data**

Event data is a signal triggered by change of state, position, behavior, or features of field device or materials. For example, continuous cast slab cutting signal and billet entering the furnace signal all belong to digital signal. In configuration table, it belongs to event data if its SCANRATE is 0. Event data is a kind of important production data which is a key to establish triggering mechanism for tracking and

monitoring production process. When an event occurs in the production site, the system will accomplish an invocation of interface of external system according to interface information of external system stored in `EVENT_ACTION`. Meanwhile, it transmits the production process information to external system. That completes the tracking and monitoring of the production process. Asynchronous mode is adopted to acquire the event data.

### **114.3.2.3 Trigger Data**

Trigger data is a set of production process data closely bound to event data, such as pouring weight of big casting ladle, pouring weight of middle casting ladle, and pouring time. In configuration table, it belongs to trigger data if its `SCANRATE` is `-1`. A single event signal can be bound to multiple trigger data. The configuration information of trigger data bound to specific event is stored in `EVENT_TRIGGER_TAG`. When the event occurs, the system gets the trigger data according to the relevant configuration information in `EVENT_TRIGGER_TAG` and transmits it with event data to an external system in a form of package. The external system can track and monitor production process after parsing the received data. It should be pointed out that trigger data must bind to event and it may have no practical meaning to get this type of data alone. Therefore, the system reads trigger data synchronously only when an event occurs.

### **114.3.2.4 Download Data**

Download data refers to the data transmitted to automation system by external system through the interface system. This type of data generally is the control information of production. In configuration table, it belongs to download data if its `SCANRATE` is `-2`. Interface system do not read but write this type of data into automation system if necessary. To avoid industrial accidents or plant accident, automation system must verify the legitimacy of this type of data when used, and valid data can be used.

## ***114.3.3 LIOPC\_TAG***

This entity is used to save data acquisition point information of all OPC servers which include access path and data type. Configuration information is obtained from this entity and is added to the OPC Group object. Since the regular change of field data point, management staff can increase or decrease or modify data point in the light of actual conditions without leading to modification of system source code. This makes the system more universal, flexible, and easily extensible.

### ***114.3.4 EVENT\_ACTION***

This entity is used to save logical program interfaces of external system corresponding to an event. When some event of L1 system happens, interface system will invoke program interface of external system on the basis of configuration of this event to perform business logic and realize the tracking and monitoring of production process.

## **114.4 Implementation of System**

### ***114.4.1 EVENT\_ACTION***

Data access method of OPC is either synchronous access or asynchronous access [3]. The synchronous access means OPC client is in a wait state after sent a request to an automation OPC server. The invocation of OPC client can be returned, and other process will be executed after a requested data from the OPC server was returned to the client. The asynchronous access means the client will return immediately and process other task after sending a data request to OPC server. OPC server will call the callback function of OPC client immediately and send the requested data to OPC client when a request is received. The implementation of synchronization is simpler than that of asynchronous. Synchronization can be adopted when data quantity are less in interaction between OPC server and client. But a large amount of data exchange will necessarily lead to performance and efficiency of system degradation. It may even influence the proper use of basic automation system. The implementation process of asynchronous mode is more complex. Its advantage is high performance of system so that a block caused by a large volume of data request from multiple clients can be avoided as well as CPU and net resources can be saved at maximum.

There are different formulations about which is better of using synchronization for data-gain or asynchronous for data-gain. The characteristic of production data decides which mode should be used. Asynchronous should be adopted for access of cycle data and event data and synchronization for trigger data.

### ***114.4.2 System Implementation***

System implementation is actually converting data model to program. According to data model and knowledge of object-oriented programming, data model is converted to OPC server object management class, OPC server object class, OPC Group object class, OPC Item object class, EventManager object class, and connection point object class. Each of the abovementioned classes is an encapsulation



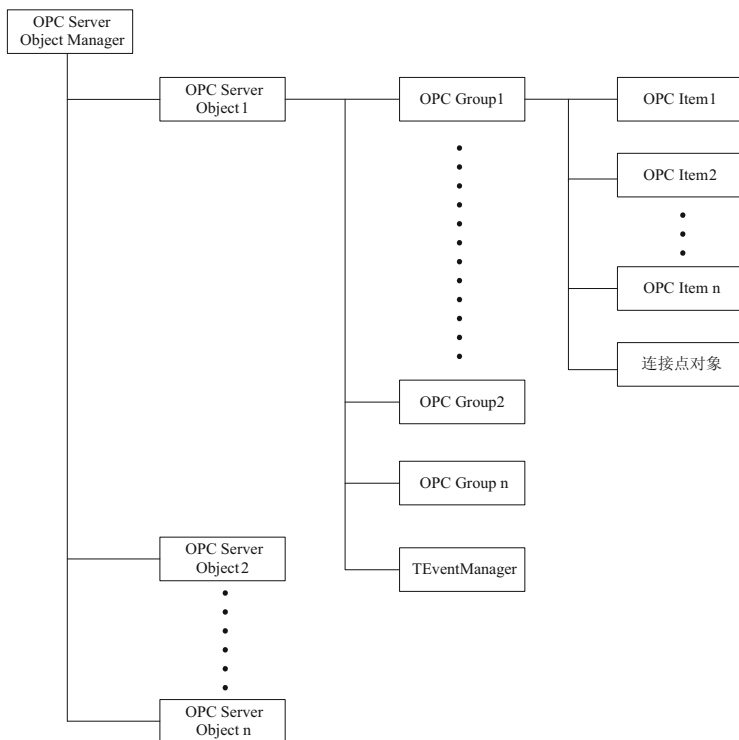


Fig. 114.2 Hierarchy diagram of system object

of OPC interface. To improve the efficiency of program developing and debugging, each class has corresponding management function, for example, connection and disconnection of server and activating and deactivating of OPC Group. Object hierarchy of system is shown as Fig. 114.2. The functions and implementations of each class are described as follows.

### 114.4.2.1 OPC Server Management Object

This class is a container class of OPC server object which is responsible for creating all OPC server objects. The class of OPC server management object traverses the configuration information of OPC\_SERVER data table and creates each connection to OPC server object when the system starts up. If creating OPC server object failed, the log will be recorded by the system so that relevant staff can examine the failure. The operators can conduct the management of a specific OPC server such as start-up and shutdown.

#### **114.4.2.2 OPC Server Object**

This is an encapsulation of OPC interface server in the system which is responsible for executing connection to field OPC server. This object is created by OPC server management object and performs the connection to remote OPC server. It will traverse the information of configuration table of L1OPC\_SCANRATE\_SETUP and create OPC Group object if connecting successfully. OPC server object is a container class of OPC Group object which can conduct management of creating, deleting, activating, deactivating, etc., on OPC Group.

#### **114.4.2.3 OPC Group Object**

The data transmission of OPC is conducted with the group as the unit [4]. This is an encapsulation of OPC interface group and is created by OPC server object. This object will get all data acquisition point information belonging to this group from L1OPC\_TAG and add it to the group item by item. OPC Group is organized by scan rate in interface system and the threads are arranged by OPC Group in OPC server. Therefore, it is not suitable to set excessive scan rate in L1OPC\_SCANRATE\_SETUP, otherwise might result in performance decline of OPC server. The dead band parameter setting is also supported by OPC Group object.

#### **114.4.2.4 OPC Item Object**

This object is an encapsulation of OPC interface Item which is responsible for management and maintenance information about OPC Item. Each OPC Item object is subordinate to the OPC Group object which has the same scan rate as OPC Item. OPC Item represents data connection with OPC server and general corresponding to registers on device. There is a correspondence between the OPC Item object and the records in L1OPC\_TAG. The management staff can increase or decrease or modify L1OPC\_TAG information to adapt to data acquisition requirement in industry field and the latest configuration information will be used when system starts up. The management functions such as activating, deactivating, and value writing can be conducted by this object.

#### **114.4.2.5 Event Management Object**

This is essentially an encapsulation of OPC Group which belongs to OPC server object. This object is a container of event object and facilitates flexible managing event object for the management. The capacity of retrieving event object through index or event name is provided by the system.

#### 114.4.2.6 Event Object

Event object is also an encapsulation of OPC interface Item. Due to special features of event, an individual encapsulation is needed for event object. Event calls external interface function according to a configuration in EVENT\_ACTION and sends process data of automation system to the external system. Thus completing the tracking and monitoring of the production process. It is the important function of event object. Event corresponds to event data point in L1OPC\_TAG. Therefore, increase or decrease of event object can be realized by modification of configuration. Through this, the dynamic expansion of system function can be realized.

#### 114.4.2.7 OPC Connection Point Object

There are two different mechanism of OPC asynchronous access: advisory connection mechanism and connection object mechanism [4] which respectively corresponds to OPC DA 1.0 [5] specification and OPC DA 2.0 [6] specification. The interface system uses connection object mechanism to realize asynchronous mechanism. OPC connection object is an encapsulation of OPC connection object mechanism. This object which belongs to each OPC Group object is a key to realize asynchronous communication. OPC Group adopted asynchronous communication mechanism will register onDataChange function with OPC server of automation system. The onDataChange function registered by client will be invoked automatically by OPC server of automation system when it detects data change of a certain group. This realized pushing change to client. Asynchronous communication is adopted for cycle data and event data of interface system. To avoid frequently invoked onDataChange function by server caused by subtle change of cycle data, dead band is defined for each OPC group. Only the range of data change is above the dead band; a notification of data change will be received by client.

### 114.5 Conclusion

This universal interface system has been applied in many projects of process control in Panzhihua Iron and Steel Co. There are some achievements of this successfully developed system:

- Facilitating the design and integration of industry control software, shortening the product development period.
- Standardize interface specification of automation system and external system, facilitating debugging of process tracing system, and improve the efficiency of software development.
- Configurability of system enhances extensibility and adaptability of interface system. Meanwhile, producibility interface system also improves stability and reliability of application system.

## References

1. Zhou, J., & Zhou, Y. (2004). Application of middleware OPC technology in industrial control system. *Computer Engineering*, 30(23), 165–167 (in Chinese).
2. Chen, J., & Yuan, N. (2003). Research and application of OPC data access specifications. *Techniques of Automation and Applications*, 22(8), 61–64 (in Chinese).
3. Su, M., & Wang, Z. (2006). Research and realization of OPC data access sever. *MicroComputer*, 22(3-1), 11–13 (in Chinese).
4. Tan, J., Jiang, S., Wu, Z., & Shao, H. (2011). The design of remote monitoring system of gas drainage based on OPC. *Coal Engineering*, 7, 128–130 (in Chinese).
5. OPC common definitions and interfaces version 1.0. OPC Foundation, 1998.
6. OPC data access custom interface specification 2.05A. OPC Foundation, 2002.

# Chapter 115

## A Stateful and Stateless IPv4/IPv6 Translator Based on Embedded System

Yanlin Yin and Dalin Jiang

**Abstract** In order to solve intercommunication problem between IPv4 network and IPv6 network more flexibly, this paper has proposed an improved IPv4-IPv6 translator based on embedded system. By using an optimized address mapping regulation, it can support both stateful and stateless translation method. In addition, a lightweight SIP-ALG and Modbus-ALG have been designed to assist the translator to process the datagram, which may take address and domain information at the seven layers of OSI model. The results show the translator can work well between sensor network and Internet, and the mixed use of stateful and stateless method has much less memory usage than stateful method and nearly the same process delay as stateless method.

### 115.1 Introduction

Since the last five IPv4 address spaces had been allocated completely in 2011 by ICANN/IANA and a growing number of users have the requirement for Internet, IPv4 addresses will be exhausted very soon [1]. At the same time, with the rise of the Internet of Things (IOT), which is considered as the third wave of development of global information industry that comes after computer and Internet, a great many of sensors and equipments need billions of IP addresses to accomplish communications. IPv6, as the core of Next Generation Internet (NGN), with exhaustless addresses, more secure and flexible, will take over IPv4 step by step eventually and undoubtedly.

However, the transition from IPv4 to IPv6 is considered as a long-term strategy; therefore, the intercommunication between these two networks becomes the main issue. Typical transition technologies include dual stack technology, which is

---

Y. Yin (✉) • D. Jiang  
College of Electronics Information and Control Engineering, Beijing University  
of Technology, Beijing 100124, China  
e-mail: [alin4187@163.com](mailto:alin4187@163.com)

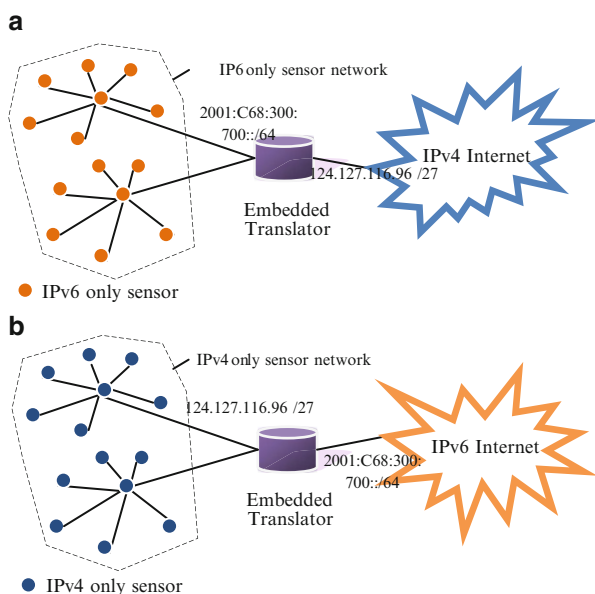
implemented at the early stage and still has to consume IPv4 address when a new node adds in; tunnel technology, which can only support IPv4 to IPv4 communication over IPv6 or IPv6 to IPv6 communication over IPv4; and translation technology, which is considered as the most feasible solution so far and will be discussed throughout this paper.

There are four types of IPv4/IPv6 translation cases which can be divided into eight scenarios depending on whether the IPv6 side or the IPv4 side initiates communication [2]. Translation technology is based on stateful method which refers to the translator that keeps every mapping information in its memory or stateless method which refers to the translator that only uses predetermined rules for address mapping [3, 4]. RFC6145 suggests that stateless translation supports end-to-end address transparency and has better scalability compared with stateful translation. In the early year, Aoun and Davies proposed to move the Network Address Translator-Protocol Translator (NAT-PT), which is a stateful translation mechanism illustrated in RFC2766, to historic status because of its complicated implementation and insecurity [5]. Yet RFC4996 also mentioned that in some circumstances, an IPv6-IPv4 protocol translation solution may be a useful transitional solution. In recent years, some stateless translation methods based on IVI [6], of which IV represents IPv4 and VI represents IPv6, have been proposed one after the other such as IVI/MAP-T/MAP-E [7]. But all these researches and experiments of stateless translation methods were under personal computer environment, and they may not be appropriate for sensor network which considered as the key to the IOT. Considering the scalability and utilization of IPv4 addresses, this paper proposed and designed an improved NAT-PT translator which supported both stateful and stateless translation method based on embedded system in order to fit the particularity of the sensor network.

## 115.2 Translation Model

Translation model includes two parts, address translation and protocol translation, which sometimes probably go along with the functional application layer gateway (ALG) [8]. In fact, the main difference between stateful and stateless translation method is the address translation. For the stateful method, like original NAT-PT, it keeps every entry of address mapping in the memory. Therefore, it is wasted storage and time when looking up corresponding entry in mapping table. For the stateless method, like original IVI, it uses a number of certain rules for address mapping. However, the multiplexing number of its each IPv4 address is suggested up to 256, and it consumes at least  $2^{64}$  available IPv6 addresses for stateless usage [9]. Hence, the improved NAT-PT translator should balance both pros and cons of these two methods. In this paper, we will only discuss the top half IPv4-IPv6 translation scenarios mentioned in RFC6144 [2] because of the particularity of sensor network, as is shown in Fig. 115.1.

**Fig. 115.1** Panel (a) shows the interoperation between an IPv6-only sensor network and the IPv4 Internet in bidirection, panel (b) shows the interoperation between an IPv4-only sensor network and the IPv6 Internet in bidirection



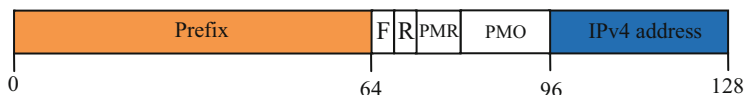
To support both the stateful and stateless translation is a tough thing, thus we should do some work on the IPv4-IPv6 address mapping and describe how it works, and then illustrate the protocol translation algorithm and application level gateway (ALG) solution.

## 115.2.1 Net Address Translation

### 115.2.1.1 Address and Mapping Regulation

The main idea to support both stateful and stateless translation is to utilize IPv6 address mapping regulation to separate them. Figure 115.2 described the address mapping regulation for the improved NAT-PT translator, which is optimized from IVI. Notice that the IPv6 prefix must be smaller than /64 and has to use 1/16 of its address space for stateless implement. Each segment defined in the address format is explained as follows:

- *Prefix*: a 64-bit IPv6 prefix assigned by the Internet service provider (ISP) for global use or defined by the network itself for local use.
- *Flag (F)*: a 4-bit field to decide which kind of translation method to be used, stateful or stateless. 4-bit all zero stands for stateful and 4-bit all one stands for stateless; other combinations are reserved.
- *Reserve (R)*: a 4-bit field reserved for future use, default is all zero.



**Fig. 115.2** Address mapping regulation

- *Port multiplexing ratio (PMR)*: an 8-bit field, in the power of two, is to reveal the multiplex number of each IPv4 address. The theoretical maximum multiplex ratio is  $2^{16}:1$  (PMR value is  $0 \times 0f$ ), while a sensor can only use one port to communicate, and the minimum is 1:1.
- *Port multiplexing offset (PMO)*: a 16-bit field to determine the start position of a port range, which is decided by the PMR.
- *IPv4 address*: a 32-bit IPv4 address for sensors in IPv6-only network to map.

In order to illustrate conveniently, we use the combination of colon hexadecimal and dotted decimal to represent an IPv6 address (e.g., use  $2001::255.255.255.255$  instead of  $2001::FFFF:FFFF$ ). For example,  $2001:C68:300:700:F00A:0100:124.127.116.125$  is an IPv6 address in which  $2001:C68:300:700::/64$  is assigned by the ISP and the prefix  $2001:c68:300:700:F000::/68$  means it uses the stateless address translation. The PMR,  $0 \times 0A$ , reveals that the address multiplex ratio is 1:1m024 while the available port number for a single sensor is 64. The PMO,  $0 \times 0210$ , shows that these 64 port numbers start from  $0 \times 0210$  to  $0 \times 0250$  (512–592 in decimal).

## 1. How It Works

### (a) Interoperation between IPv6 sensor network and IPv4 Internet

Firstly, we will discuss about the case that interoperation happened between an IPv6-only sensor network and the IPv4 Internet shown in Fig. 115.1a. Indeed, this is the situation at the beginning of IPv4-IPv6 transition when the IPv4 network is still the backbone network. In this case, both stateful method and stateless method can be used. We assume that the translator has been assigned an IPv4 prefix  $124.127.116.96/27$ ; one-half is for stateless use and the other half is for stateful use. When a data packet passes through the translator in the direction of IPv6 to IPv4, it will go through the following steps:

- *Step 1*: Check the 4-bit flag field in the source address of the IP header. If it is set to  $0 \times f$ , then go to step 2; else if it is set to  $0 \times 0$ , then go to step 5; else go to step 7.
- *Step 2*: Check the source port field of UDP or TCP header, if it is in the region of  $[PMO, PMO+2^{16-PMR}]$ , then go to step 3; else go to step 7.
- *Step 3*: Extract the 96–127 bits from the IPv6 destination address as the new IPv4 destination address; go to step 4.
- *Step 4*: Extract the 96–127 bits from the IPv6 source address as the new IPv4 source address; go to step 6.



- *Step 5*: Do stateful translation by checking source address in mapping table, if you find the entry, then use the corresponding address and port as the new address and port; else create a new mapping entry, using the mapped IPv4 address and port. Go to step 6.
- *Step 6*: Do protocol translation with ALGs and send the translated packet.
- *Step 7*: Drop the packet.

When a data packet passes through the translator in the direction of IPv4 to IPv6, it will go through the following steps:

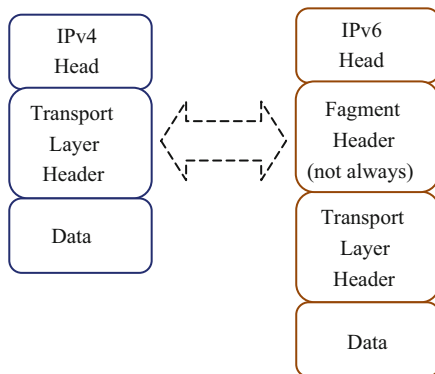
- *Step 1*: Read the translator's configuration and check the IPv4 destination address; if it is for stateless use, then go to step 2; else if it is for stateful use, go to step 5; else go to step 7.
  - *Step 2*: Get the PMR from translator's configuration, and get the source UDP or TCP port as P from the IPv4 packet, calculate PMO,  $PMO = INT(P/PMR) * PMR$ , and go to step 3.
  - *Step 3*: Fill the new IPv6 source address with prefix, F, R, PMR, PMO, and IPv4 source address extracted from the old IPv4 packet; go to step 4.
  - *Step 4*: Fill the new IPv6 destination address with prefix, F, R, PMR, PMO, and IPv4 destination address extracted from the old IPv4 packet; go to step 6.
  - *Step 5*: Do stateful translation by checking destination address in mapping table; if you find the entry, then use the corresponding address and port as the new address and port; else create a new mapping entry, using the mapped IPv6 address and port. Go step 6.
  - *Step 6*: Do protocol translation with ALGs and send the translated packet.
  - *Step 7*: Drop the packet.
- (b) Interoperation between IPv4 sensor network and IPv6 Internet

Next, we will talk about the case that interoperation happened between an IPv4-only sensor network and the IPv6 Internet shown in Fig. 115.1b. This is the situation that happens in the late stage of IPv4-IPv6 transition when legacy sensors or equipments want to access the IPv6 Internet. Unfortunately, stateless transition cannot be used due to the philosophical logic: IPv4 network needs to communicate with all of the IPv6 Internet, not just a small subset, and stateless can only support a subset of the IPv6 addresses. But we can use stateful method to accomplish address translation; the process is similar as previous illustration.

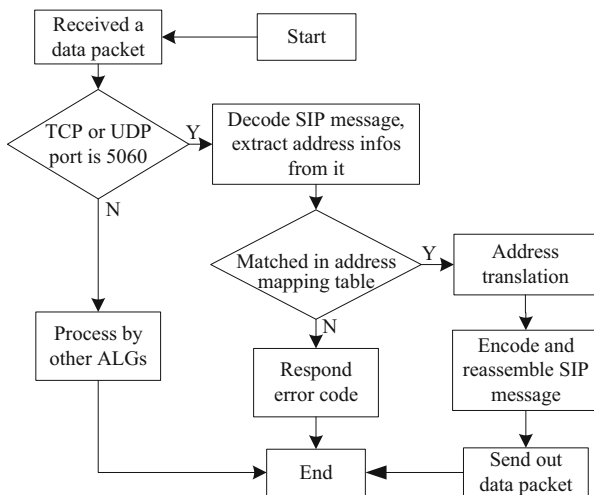
### 115.2.2 Protocol Translation

Protocol translation between IPv4 and IPv6 is a critical part; this is because besides the wide difference in address space and protocol header, the architecture of these two protocols has an extreme variation. In this paper, the protocol translation algorithm is based on Stateless IP/ICMP Translation Algorithm (SIIT) [10], as is shown in Fig. 115.3.

**Fig. 115.3** Protocol translation diagram



**Fig. 115.4** Lightweight SIP-ALG workflow



### 115.2.3 Application Level Gateway

It is inevitable that sensors or equipments are running Layer7 network applications that take address information, leading to communication problem. Thus, ALG should be set to solve the problem. Typical ALG is DNS-ALG and FTP-ALG which have much discussed. Considering that a number of sensors, maybe advanced and intelligent, are using Modbus protocol over TCP/IP stack to signaling [11], or utilizing Session Initiation Protocol (SIP) to control voice and video transportation, this paper has proposed a lightweight SIP-ALG and a Modbus-ALG solution for the translator. Figure 115.4 shows the workflow of lightweight SIP-ALG in stateful method, and Modbus-ALG is similar.

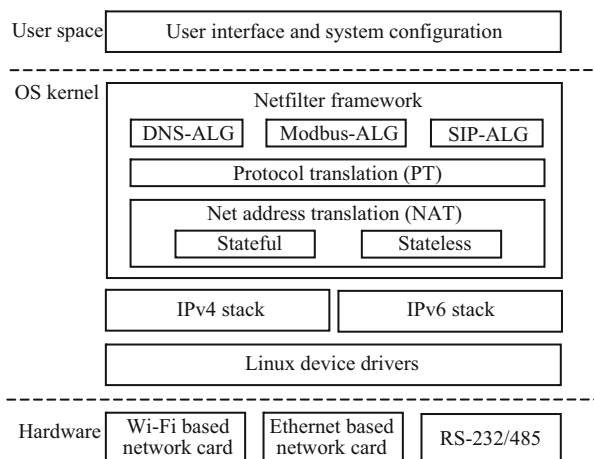
## 115.3 Architecture and Design

The translator is based on the embedded system, which uses ARM Cortex-A8 development board (1 GHz CPU, 256M RAM) as the hardware, uses Linux as the operating system, and takes advantage of netfilter framework for developing. Figure 115.5 shows the system architecture. The Linux device drivers provide some methods to access raw data packet from the hardware devices. The netfilter framework gets the packet from the device drivers and deals with it. Firstly, the NAT module generates new IP source and destination addresses by using stateful or stateless method. Then, the PT module generates a new IP or ICMP header. At last, ALG modules do some further processing when the packet takes address information at the application layer. All the OS kernel-level developments, including device driver development and translation modules development, are using C language. However, the user space development mostly is based on shell or PHP script.

## 115.4 Testing

### 115.4.1 Testing of Connectivity

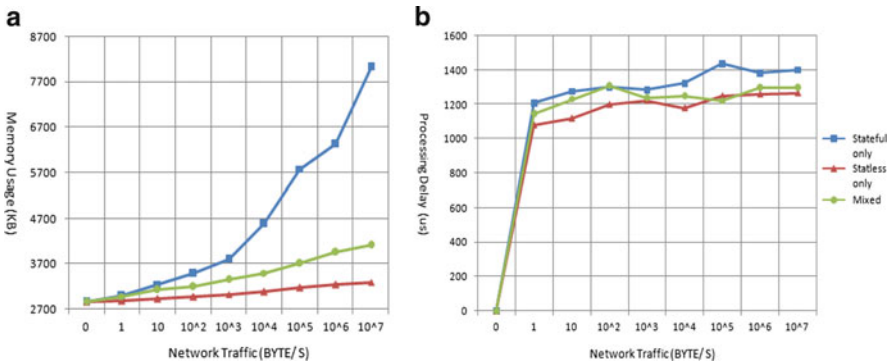
We wrote a shell script (using ping command) running on Linux-based sensors and hosts to test the connectivity between nodes in different networks, including eight cases, and the statistics is shown in Table 115.1. Notice that the success rates of cases 1, 2, 4, 5, 6 are under 100 %, probably because the compatibility of IPv6 stack in some sensors is not very good. Generally, the translator works well.



**Fig. 115.5** Architecture of the translator

**Table 115.1** Results of connectivity testing

Method	Case	Total	Success	Rate (%)
Stateful	v6 sensor Pings v4 host	10,000	9,989	99.89
	v4 host Pings v6 sensor	10,000	9,973	99.73
	v4 sensor Pings v6 host	10,000	10,000	100
	v6 host Pings v4 sensor	10,000	9,999	99.99
Stateless	v6 sensor Pings v4 host	10,000	9,980	99.80
	v4 host Pings v6 sensor	10,000	9,987	99.87
	v4 sensor Pings v6 host	10,000	10,000	100
	v6 host Pings v4 sensor	10,000	10,000	100



**Fig. 115.6** Panels (a) and (b) show the relationship between network traffic and memory usage or processing delay, respectively, in three modes (stateful, stateless, and mixed)

### 115.4.2 Testing of Pressure

In this test, we wrote a shell script running on the translator to monitor memory usage, processing delay while network traffic is changing under different translation methods (MTU is set to 1,460). The statistical results are shown in Fig. 115.6, ignoring the flow direction. We can infer that the mixed use of stateful and stateless method has a lot less memory usage than stateful method and nearly the same process delay as stateless method.

## 115.5 Conclusion

This paper has proposed an improved IPv4-IPv6 translator that can support both stateful and stateless translation for sensor network accessing Internet. Then we implemented it on the ARM Cortex-A8 based on embedded system and did some

system testings. Although stateful method is resource wasting and time-consuming, it is still the only solution under some certain circumstances. Surely, this translator is useful but maybe not perfect in architecture, translation algorithms, connectivity of software modules, table search algorithm, C code efficiency, etc. And these will be optimized gradually in our future researches.

## References

1. IPv4 address report (online). (2013). Retrieved from <http://www.potaroo.net/tools/ipv4/>.
2. Baker, F., Li, X., Bao, C., & Yin, K. (2011). Framework for IPv4/IPv6 translation, RFC 6144.
3. Bagnulo, M., Matthews, P., & van Beijnum, I. (2011). Stateful NAT64: Network address and protocol translation from IPv6 clients to IPv4 servers, RFC6146.
4. Li, X., Bao, C., & Baker, F. (2011). IP/ICMP translation algorithm, RFC6145.
5. Aoun, C., & Davies, E. (2007). Reasons to move the network address translator—protocol translator (NAT-PT) to historic status, RFC4966.
6. Li, X., Bao, C., Chen, M., Zhang, H., & Wu, J. (2011). The China Education and Research Network (CERNET) IVI translation design and deployment for the IPv4/IPv6 coexistence and transition, RFC6219.
7. Li, X., & Bao, C. (2013). IVI/MAP-T/MAP-E: Unified IPv4/IPv6 stateless translation and encapsulation technologies (online). Retrieved from <http://www.cnki.net/kcms/detail/34.1228.TN.20130228.1703.003.html>.
8. Holdrege, M., & Srisuresh, P. (2001). Protocol complications with the IP network address translator, RFC3027.
9. Zhu, Y. C. (2008). Stateless mapping and multiplexing of IPv4 addresses in migration to IPv6 Internet (pp. 2248–2252). In: *2008 I.E. global telecommunications conference*. New York: Institute of Electrical and Electronics Engineers Inc.
10. Nordmark, E. (2000). Stateless IP/ICMP translation algorithm (SIIT), RFC2765.
11. Modbus messaging on TCP/IP implementation guide V1.0b (online). (2006). Retrieved from <http://www.modbus.org/specs.php>.

# Chapter 116

## A Novel Collaborative Filtering Approach by Using Tags and Field Authorities

Zhi Xue, Yaoxue Zhang, Yuezhi Zhou, and Wei Hu

**Abstract** Traditional collaborative filtering is widely used in social media and e-business, but data sparsity and noise problems have not been solved effectively yet. In this chapter, we propose a novel approach of collaborative filtering based on field authorities, which achieves genre tendency of items by mapping tags to genres and simulates a fine-grained word-of-mouth recommendation mode. We select the nearest neighbors from sets of experienced users as field authorities in different genres and assign weights to genres according to genre tendency. Our method can solve sparsity and noise problems efficiently and has much higher prediction accuracy. Experiments on MovieLens datasets show that the accuracy of our approach is significantly higher than traditional user-based kNN CF approach in both MAE and precision tests.

### 116.1 Introduction

Collaborative filtering (CF) is the most popular technology in current recommender systems. The basic idea is to recommend items to active users based on the opinions of other users who have similar tastes. CF approach achieves a great success in research and practice, such as Google News, Netflix, and Amazon [1].

A typical collaborative filtering approach, k-Nearest Neighbor (kNN) [2, 3], for example, is a way to find the “nearest neighbors” of the active user. The items will be recommended to the active user only if they are most liked by the neighbors. However, because of the limitation of data, this approach has data sparsity and noise problems, which cause failures in neighbor searching. In order to solve these problems, a CF approach based on expert opinions has been proposed [4] (ECF). Rather than applying a nearest neighbor algorithm to the user-rating data, predictions

---

Z. Xue (✉) • Y. Zhang • Y. Zhou • W. Hu  
National Laboratory of Information Science and Technology, Department of Computer  
Science and Technology, Tsinghua University, Beijing 100084, China  
e-mail: [raphaelxue@gmail.com](mailto:raphaelxue@gmail.com)

are computed by using a set of expert neighbors from an independent dataset. Compared to traditional kNN method, this approach can effectively solve the problems of data sparsity and noise but leads to several other problems, such as prediction accuracy declination, and more than that, it is hard to find external experts normally.

In this chapter, we present a novel collaborative filtering algorithm based on field authorities. We define field authorities as experienced user sets in different areas. The genre information of items is from an experience-decision dataset, and genre weights are calculated by user-generated tags. We use MovieLens dataset to experiment and find the following results: (1) the prediction accuracy of our approach is significantly higher than traditional user-based CF and ECF; (2) there is no need to use external data, so our approach has good expansibility.

### 116.2 Genre Resolution Model Based on Tags

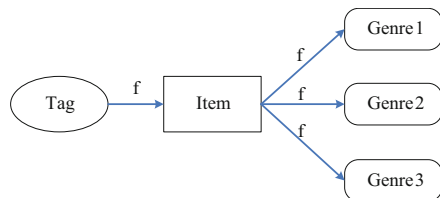
In order to describe the reality more fine-grained, we model the genres of items by using tags marked on the content. Using tags to model genres can solve the credibility problem of items for us, because the nominal and the real are always not consistent and undoubtedly the most authoritative judgments are from the choice of users.

The performance of a single tag can be very unstable, but if we observe the statistical significance from the macrostructure, we can find obvious corresponding relations between tags and genres. We use the associated characteristics between tags and genres and can judge the tendency of each item on different genres.

As shown in Fig. 116.1, we use this relationship to associate the tag and the genres, by mapping the number of times to the genres. It is reasonable because we can treat the tagging behavior as a vote to the item. From statistical sense, a few individual, local views will be erased by the majority of user mainstream view, and individual differences in statistical sense are irrelevant. By focusing on the vote results, we can achieve maps from tags to genres and get the frequency of it, which is shown in Fig. 116.2.

For tag<sub>i</sub>, the number of labeled times to genre<sub>j</sub> is F<sub>ij</sub>, and the number of labeled times to item<sub>m</sub> is f<sub>im</sub>:

$$F_{ij} = \sum_{genre_j \in item_m} f_{im}$$



**Fig. 116.1** The relationship of tag, item, and genre

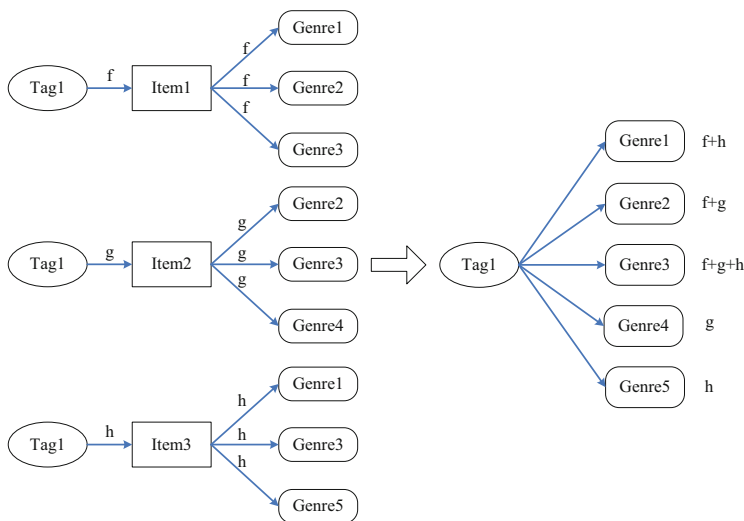
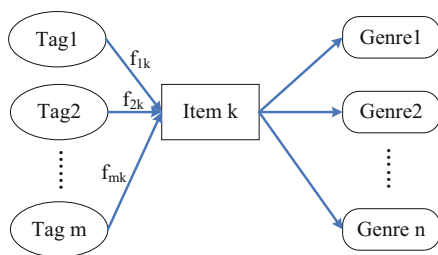


Fig. 116.2 The vote from the tag to genres

Fig. 116.3 Calculate item tendency via tag tendency



Then we can calculate the ratio of tag<sub>i</sub> on genre<sub>j</sub>:

$$ratio_{ij} = \frac{F_{ij}}{\sum_j F_{ij}}$$

Now we have the tendency of each tag on different genres. We calculate the tendency of each item on different genres by using the tag tendency as a bridge, shown in Fig. 116.3.

We calculate the frequency of item<sub>k</sub> on different genres using the matrix shown in Fig. 116.4.

For genre<sub>j</sub>, the equivalent frequency FG<sub>jk</sub> on item<sub>k</sub> is:

$$FG_{jk} = \sum_{i=1}^m f_{ik} \cdot ratio_{ij}$$



**Fig. 116.4** The matrix for calculating frequency

	<i>genre<sub>1</sub></i>	...	<i>genre<sub>j</sub></i>	...	<i>genre<sub>n</sub></i>
<i>tag<sub>1</sub></i>	<i>f<sub>1k</sub> · ratio<sub>11</sub></i>	...	<i>f<sub>1k</sub> · ratio<sub>1j</sub></i>	...	<i>f<sub>1k</sub> · ratio<sub>1n</sub></i>
⋮	⋮		⋮		⋮
<i>tag<sub>i</sub></i>	<i>f<sub>ik</sub> · ratio<sub>i1</sub></i>	...	<i>f<sub>ik</sub> · ratio<sub>ij</sub></i>	...	<i>f<sub>ik</sub> · ratio<sub>in</sub></i>
⋮	⋮		⋮		⋮
<i>tag<sub>m</sub></i>	<i>f<sub>mk</sub> · ratio<sub>m1</sub></i>	...	<i>f<sub>mk</sub> · ratio<sub>mj</sub></i>	...	<i>f<sub>mk</sub> · ratio<sub>mn</sub></i>

Finally we get  $Tendency_{jk}$  between  $genre_j$  and  $item_k$ :

$$Tendency_{jk} = \frac{FG_{jk}}{\sum_j FG_{jk}}$$

In order to facilitate the comparison between different items, we have to normalize the tendency of different genres on the same item.

$$Tendency'_{jk} = \frac{Tendency_{jk}}{\max_j Tendency_{jk}}$$

### 116.3 Field Authority Nearest Neighbors

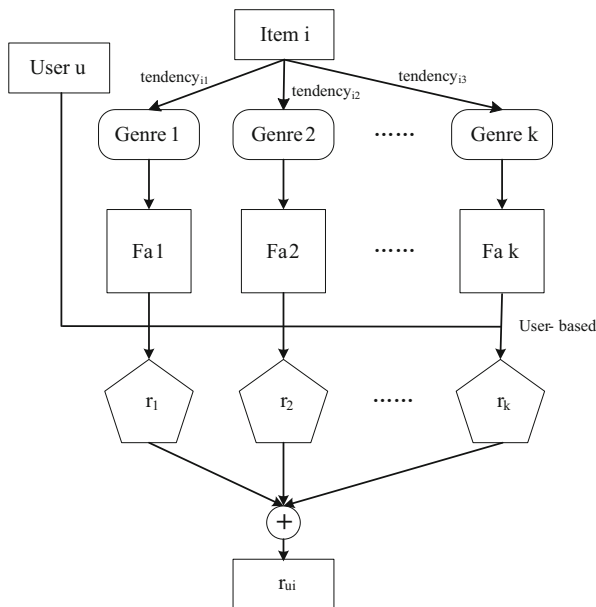
On the basis of genre resolution model, we design our collaborative filtering approach based on field authority (FACF), and our core tasks are to predict ratings for blank “user-item” pairs. Figure 116.5 shows the algorithm process. Traditional CF usually predicts rating using user-based [5–7] or item-based [1, 8] mode. Since our algorithm is based on field authorities, we use the user-based mode in the calculating process.

For user  $u$  and item  $i$ , we predict  $r_{ui}$  step by step as follows. First we have got the genre tendency of item  $i$ . In each genre involved, we use FACF to calculate a prediction rating and use genre tendency as weight to sum the ratings of different genres to get the final result  $r_{ui}$ .

The most important aspect of FACF is to find field authority nearest neighbors. Here we require that field authorities are in the top 100 sorted by rating numbers and at least rate 10 movies in a certain genre. For user  $u$  and  $Fa_k$ , which stands for field authorities on genre  $k$ , we calculate the similarities between  $u$  and member  $v$  of  $Fa_k$ , and there are  $K$  field authorities who are the most similar with user  $u$  as his nearest neighbors. We use the cosine similarity here:

$$w_{ui} = \frac{\sum_{i \in I} (r_{ui} - \bar{r}_u) \cdot (r_{vi} - \bar{r}_v)}{\sqrt{\sum_{i \in I} (r_{ui} - \bar{r}_u)^2 \cdot \sum_{i \in I} (r_{vi} - \bar{r}_v)^2}}$$

**Fig. 116.5** FACF algorithm process



where  $I$  stands for the set that items are co-rated by  $u$  and  $v$ ;  $\bar{r}_u$  and  $\bar{r}_v$  are the average ratings in current genre for  $u$  and  $v$ .

We can predict rating by using the similarity as a weight. This is done by means of a similarity-weighted average of the ratings input from each field authority [9]:

$$\hat{r}_{ui} = \bar{r}_u + \frac{\sum_{v \in S(u,K) \cap N(i)} w_{uv} \cdot (r_{vi} - \bar{r}_v)}{\sum_{v \in S(u,K) \cap N(i)} |w_{uv}|}$$

where  $\hat{r}_{ui}$  is the prediction rating for genre  $k$ .  $\bar{r}_v$  is the average rating for  $Fa_k$  member  $v$  on item  $i$ .  $\bar{r}_u$  is the average rating of user  $u$ . After we get the prediction rating of each genre, we use weighted summation of predicted ratings to get the final prediction result:

$$r_{ui} = \sum_{e=1}^k r_e \cdot tendency_{ie}$$

In summary, the core idea of FACF is to predict ratings for user-item pairs from different genre perspective and then use user’s preferences of each genre as weights to get the final result. The results are able to take the preferences between the user and field authorities in common into account and reflect the item tendency on different genres quite suitable.

## 116.4 Experiments

We test our algorithm on MovieLens dataset. We can verify FACP performance in prediction accuracy obviously by contrasting with kNN CF and localized ECF. The MovieLens dataset is a subset of MovieLens community, which contains 544 thousand ratings on 4,988 movies rated by 1,000 users. Each user rates 20 ratings and each movie has 1 rating at least.

In order to evaluate the accuracy of predictions, we divide our data set into 80 % training – 20 % testing sets and calculate the average results of a 5-fold cross-validation.

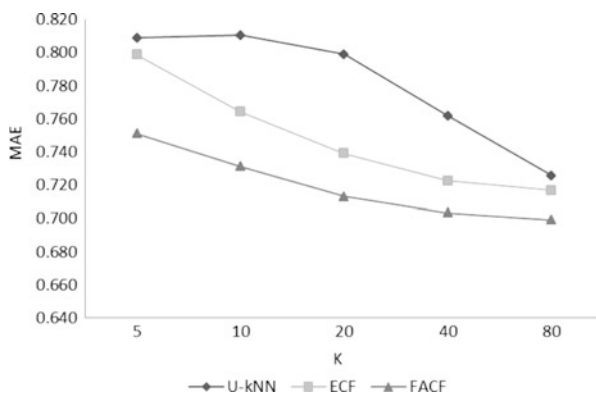
### 116.4.1 Mean Absolute Error

For each user-item pair in the test set,  $r_{ui}$  is its actual rating, and  $\hat{r}_{ui}$  is its prediction rating. MAE is defined as follow:

$$MAE = \frac{\sum_{u,i \in T} |r_{ui} - \hat{r}_{ui}|}{|T|}$$

To test the performance of our algorithm under different conditions, we take the number of nearest neighbors  $K$  as a variable. We also run the same experiment on user-based kNN CF (U-kNN) and ECF as a comparison.

Figure 116.6 shows the MAE results of three approaches. FACP has the best performance on MAE test. We can see details in Table 116.1. FACP has a reduction of MAE from 3.6 to 10.8 % compared to U-kNN and from 2.4 to 6.0 % compared to ECF. It is strong evidence that FACP is significantly more accurate than U-kNN and ECF.

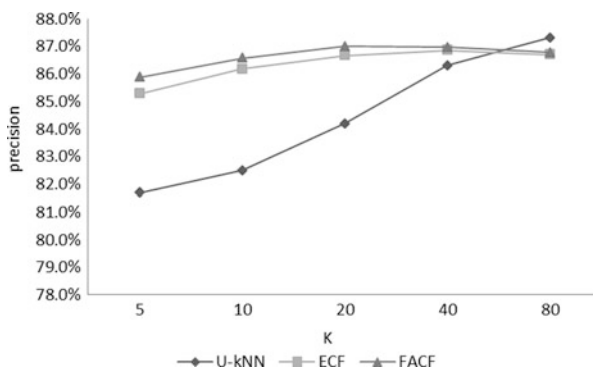


**Fig. 116.6** MAE of U-kNN, ECF, and FACP

**Table 116.1** MAE and reductions of U-kNN, ECF, and FACF

	U-kNN	ECF	FACF	Reduction (1-FACF/U-kNN)	Reduction (1-FACF/ECF)
K=5	0.809	0.799	0.751	7.1 %	6.0 %
K=10	0.811	0.764	0.731	9.8 %	4.3 %
K=20	0.799	0.739	0.713	10.8 %	3.5 %
K=40	0.762	0.723	0.703	7.7 %	2.6 %
K=80	0.726	0.717	0.699	3.6 %	2.4 %

**Fig. 116.7** Precision when threshold = 3



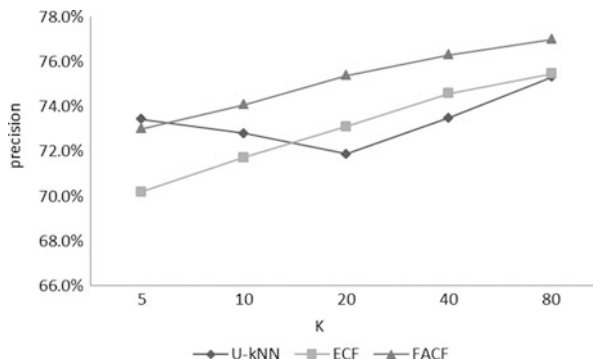
### 116.4.2 Precision of Recommendation Lists

Even though MAE is a good indicator, users do not consider it the way as MAE does in real life. We are more concerned with whether users feel good to what we recommend, like a top-N recommendation list [10]. Here we observe the prediction accuracy by setting a threshold to distinguish a movie between “recommended” and “not recommended” [4].

We set a threshold  $\tau$  as a criterion. When our predicted rating is higher than  $\tau$ , we think this movie is recommendable, otherwise is not recommendable. So if the actual rating on this movie is also higher than  $\tau$ , we consider it as a successful recommendation. By calculating the success rate, we can compare the performance of user-based kNN CF, ECF, and FACF.

Figures 116.7 and 116.8 show the precision under different threshold conditions. When threshold  $\tau=3$ , the precision of FACF has been around 87 %, which is much higher than that of U-kNN and a little higher than that of ECF. When threshold  $\tau=4$ , the precision of FACF is 3 % higher than both U-kNN and ECF. We can see that when we need a high standard recommendation by using a tough condition, FACF can give us a much better result.

**Fig. 116.8** Precision when threshold = 4



## 116.5 Conclusion

We propose a novel CF approach by using tags and field authorities. By mapping tags to genres of each item and introducing field authorities based on genre information, we simulate the word-of-mouth recommendation mode properly. We perform experiments on MovieLens dataset. Compared to kNN CF and localized ECF, (1) we have a much higher accuracy than user-based kNN CF and ECF; (2) there is no need to use external data, so our algorithm has a larger scope than ECF. Besides, because of using the concept of field authorities, the inherent characteristics, such as high degree of participation and consistent tendency, can ameliorate the data sparsity and noise problems existing in traditional collaborative filtering naturally.

## References

- Linden, G., Smith, B., & York, J. (2003). Amazon.com recommendations: Item-to-item collaborative filtering. *IEEE Internet Computing*, 7(1), 76–80.
- Beyer, K., Goldstein, J., Ramakrishnan, R., & Shaft, U. (1998). When is “Nearest Neighbor” meaningful? In *ICDT’99* (pp. 217–235). LNCS 1540. Berlin: Springer.
- Hall, P., Park, B., & Samworth, R. (2008). Choice of neighbor order in nearest-neighbor classification. *Annals of Statistics*, 36(5), 2135–2152.
- Amatriain, X., Lathia, N., Pujol, J., Kwak, H., & Oliver, N. (2009). The wisdom of the few: A collaborative filtering approach based on expert opinions from the web. In *Proceedings of the 32nd international ACM SIGIR conference on research and development in information retrieval* (pp. 532–539). New York, NY: ACM.
- Schafer, B., Frankowski, D., Herlocker, J., & Sen, S. (2007). Collaborative filtering recommender systems. In *The adaptive web* (pp. 291–324). LNCS 4321. Berlin: Springer.
- Koren, Y. (2008). Tutorial on recent progress in collaborative filtering. In *Proceedings of the 2008 ACM conference on recommender systems* (pp. 333–334). New York, NY: ACM
- Su, X., & Khoshgoftaar, T. (2009). A survey of collaborative filtering techniques. *Advances in Artificial Intelligence, 2009*, 1–19.

8. Sarwar, B., Karypis, G., Konstan, J., & Riedl, J. (2001). Item-based collaborative filtering recommendation algorithms. In *Proceedings of the 10th international conference on world wide web* (pp. 285–295). New York, NY: ACM.
9. Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P., & Riedl, J. (1994). GroupLens: An open architecture for collaborative filtering of netnews. In *Proceedings of the 1994 ACM conference on computer supported cooperative work* (pp. 175–186). New York, NY: ACM.
10. Deshpande, M., & Karypis, G. (2004). Item-based top-N recommendation algorithms. *ACM Transactions on Information Systems*, 22(1), 143–177.

# Chapter 117

## Characteristics of Impedance for Plasma Antenna

Bo Yin and Feng Yang

**Abstract** Impedance analysis is very important for antenna design. In this chapter, the internal impedance of the plasma antenna is analyzed by building the model of high-frequency electromagnetic waves acting with plasma. At the same time, a model of surface current for plasma antenna is developed in accordance with the eigenvalue equation of guided mode, and the radiation resistance of plasma antenna is analyzed according to the method of Poynting vector. From the results, we find that the internal impedance and the radiation resistance of the plasma antenna are affected distinctly by the plasma density and electron-neutral collision frequency. The internal resistance could be reduced, and the radiation resistance would be added efficiently by increasing the plasma density and decreasing the collision frequency.

### 117.1 Introduction

More and more people are interested in plasma antenna which is based on plasma elements instead of metal conductors in recent years. Plasma antenna's behavior is determined by a circular plasma column in which a surface wave is propagating along it. A plasma antenna may work immediately once it is energized through using of an RF source, and will stop instantly with the source removed. When de-energized, the plasma antenna becomes a dielectric tube filled with inert gas, and

---

B. Yin (✉)

School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China

College of Electronic Engineering, Chongqing University of Posts and Telecommunications, Chongqing 400065, China

e-mail: [byin0520@163.com](mailto:byin0520@163.com)

F. Yang

School of Electronic Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China

it reflects little return wave signals to radar [1]. Reconfiguration is another advantage of plasma antenna in which its radiation characters can be changed conveniently by electrical rather than mechanical control [2].

In the past years, many scientific or technical documents concerning plasma antenna have been reported. Kumar [3] investigated the radiation properties of a plasma column as a reconfigurable plasma antenna by controlling the operating parameters, such as drive frequency, input power, and argon gas. Zhu [4] presented the characteristics of the AC-biased plasma antenna by experimental observations. Wu [5] analyzed radiation pattern of plasma antenna through a model for a plasma antenna of beam-forming, and results indicated a good performance in the aspects of beam-forming, beam-scanning, and radiation efficiency. However, from these studies it is very difficult to infer the information of impedance of plasma antenna. As we all know, the impedance of a plasma antenna is important for impedance matching and efficiency analysis in the antenna design. In this chapter, the internal impedance and radiation resistance of plasma antenna are analyzed through the theory of impedance and the method of Poynting vector, respectively. Since physics property of plasma varies as the signal frequency due to a disperse material, the plasma density and the collision frequency have obviously influence on the impedance of plasma antenna, and results show that the internal resistance could be reduced and the radiation resistance be increased efficiently by adjusting the plasma density and the collision frequency properly.

## 117.2 The Internal Impedance of a Plasma Antenna

Plasma is a collection of free charged particles moving in random directions which consists of electrons, ions, and neutrons; in general, it is neutral in the steady state. When the density of ionized gas is very high enough, the electromagnetic wave cannot go deep into the plasma; the skin depth is quite small, and then plasma exhibits properties of a conductor [6]. The electric conductivity determines the ohmic power dissipation, which is an important mechanism for electron heating in discharges.

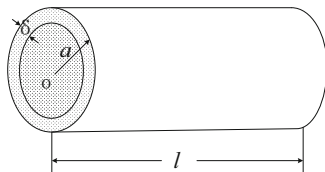
A plasma column antenna is modeled in Fig. 117.1, assume that its density is homogeneous in axial and radial directions, and the length and radius of the plasma column are 1.2 and 0.0125 m, respectively. Then the wave vector propagating in the plasma column is a complex quantity, and it can be expressed as

$$k = \omega\sqrt{\mu\epsilon_p} = k_0\sqrt{\epsilon_r} = \beta - j\alpha \quad (117.1)$$

where  $\epsilon_r = 1 - \omega_{pe}^2/\omega(\omega - j\nu)$  is the plasma relative dielectric constant [7],  $\omega$  is the frequency of propagation signal, and  $\nu_m$  is the electron-neutral collision frequency. The quantity  $\omega_{pe}^2 = n_e e^2/m_e \epsilon_0$  is the electron plasma frequency, where  $n_e$  and  $m_e$  are the plasma density and quality, respectively [7]. As the relative



**Fig. 117.1** Structure of plasma column antenna



permittivity of plasma is complex, so it can be expressed as  $\epsilon_r = \epsilon_1 - j\epsilon_2$ . Then the attenuation constant and phase constant are

$$\beta = \omega \sqrt{\frac{\mu_0 \epsilon_0}{2}} \sqrt{\epsilon_1 + \sqrt{\epsilon_1^2 + \epsilon_2^2}} \tag{117.2}$$

$$\alpha = \omega \sqrt{\frac{\mu_0 \epsilon_0}{2}} \sqrt{-\epsilon_1 + \sqrt{\epsilon_1^2 + \epsilon_2^2}} \tag{117.3}$$

So we can obtain skin depth of the plasma column  $\delta = 1/\alpha$ .

For a plasma column with certain density, it shows conductive properties like a conducting column. So the internal impedance of plasma antenna can be counted as the following:

When  $\delta < a$ ,

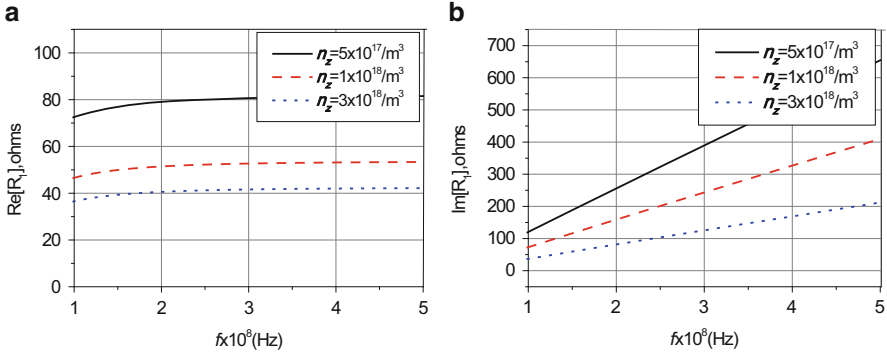
$$R_1 = R_s = \frac{l}{\sigma S} = \frac{l}{\sigma \pi a^2} = \frac{l}{2\pi a \delta \sigma} \tag{117.4}$$

When  $\delta > a$ ,

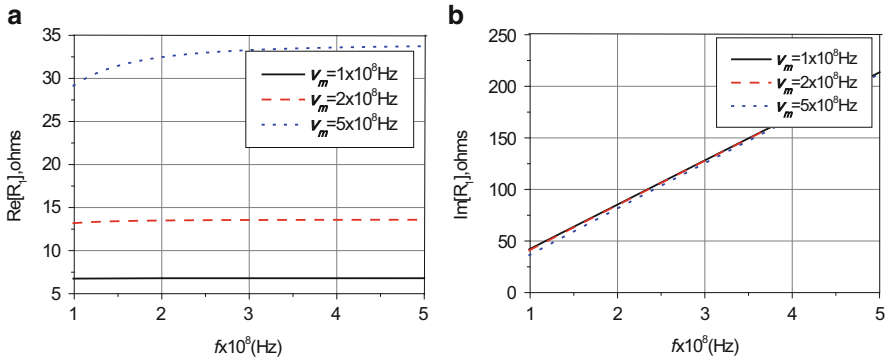
$$R_1 = R_s = \frac{l}{\sigma S} = \frac{l}{\sigma \pi a^2} = \frac{l}{\pi a^2 \sigma} \tag{117.5}$$

The internal impedance of the plasma antenna is plotted in Fig. 117.2a, b at different plasma densities, when the electron-neutral collision frequency is  $5 \times 10^8$  Hz. From Fig. 117.2a, the internal resistance of the plasma antenna becomes bigger slowly as the signal frequency rises, and increases rapidly in the low-frequency band. And if we enhance the plasma density, the internal resistance can be decreased observably, which is the results of the plasma conductivity increasing with the increase of the plasma density. In Fig. 117.2b, the internal reactance of plasma antenna increases linearly with the increase of signal frequency. And for a higher plasma density, the internal reactance becomes smaller, and the slope is smaller too.

It is the collision frequency that is relevant to plasma’s conductive and dielectric properties, which impact antennas’ characteristics with using plasma as a conducting column. The internal impedance of the plasma antenna is plotted in Fig. 117.3a, b at electron-neutral collision frequencies, when the plasma density is  $n_e = 1 \times 10^{18} \text{ m}^{-3}$ . As shown in Fig. 117.3a, the internal resistance of plasma



**Fig. 117.2** (a) Internal resistance and (b) internal reactance of the plasma antenna at different plasma densities



**Fig. 117.3** (a) Internal resistance and (b) internal reactance of the plasma antenna at different collision frequencies

antenna becomes bigger as the signal frequency increases, and arises rapidly in the low-frequency band. At same time, if we enhance the collision frequency of plasma, the internal resistance can be increased evidently. This is easily understood on the basis of the collision model for plasma. Electromagnetic (EM) waves propagating in plasma will cause oscillation of electrons. If no collision, the plasma will show characteristics of a lossless medium. However, electrons will collide with each other and with neutral particles and their kinetic energy is generally exchanged between particles. And the energy transfer is critically dependent on the electron's oscillation cycle when the collision occurs [8]. So the higher collision frequency is, the greater plasma resistance will be. In Fig. 117.3b, the results indicate that the internal reactance of the plasma antenna increases linearly with the increase of signal frequency, and the internal reactance is almost not affected by collision frequency of plasma.

### 117.3 The Radiation Resistance of a Plasma Antenna

When frequencies of incident EM waves are far lower than the plasma frequency, the surface wave will appear on the interface between the plasma and the surrounding dielectric just like traveling on the surface of a metal column. However, the wave vector of surface wave should vary with the plasma density. For a uniform density, we may obtain the plasma surface wave dispersion relation by Helmholtz's wave equation and the boundary condition [9, 10], that is,

$$\varepsilon_r T_0 I_1(T_p a) K_0(T_0 a) + T_p K_1(T_0 a) I_0(T_p a) = 0 \quad (117.6)$$

where  $T_p^2 = k^2 - \varepsilon_r k_0^2$  and  $T_0^2 = k^2 - k_0^2$ ,  $a$  is the radius of the plasma column,  $k_0 = \omega/c$  is the wave number in free space,  $I_i(\cdot)$  and  $K_i(\cdot)$  denote modified Bessel function of the first and second kind, respectively, and  $\varepsilon_r$  is the plasma relative dielectric constant.

It is through the mechanism of the radiation resistance that power is transferred from the guided wave of the plasma antenna to the free-space wave. The greater the radiation resistance is, the higher will be the power radiated for a given electric current. In order to find the radiation resistance of plasma antenna, the Poynting vector is formed in terms of the electric field and magnetic field radiated by the antenna. By integrating the Poynting vector over a closed surface (usually a sphere of very large radius), the total power radiated by the source is found. As approximation, the current distribution of monopole for plasma column is

$$I(z) = I_0 \left( e^{jk_p z} - e^{jk_p(2l-z)} \right) \quad (117.7)$$

and the current in the fed end of plasma column is

$$I_1 = I_0 (1 - e^{j2k_p l}) \quad (117.8)$$

where  $k_p$  is the propagation constant of surface wave and can be got from Eq. 117.6. The far-zone field of electric current element is given by [11]

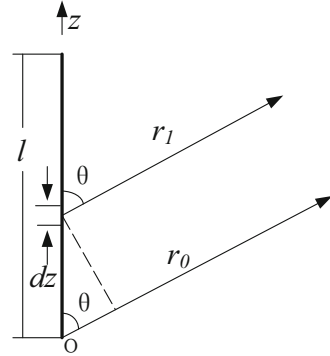
$$dE_\theta = j \frac{60\pi I dz}{\lambda R} \sin \theta e^{-jk_0 R} \quad (117.9)$$

where  $\lambda$  and  $k_0$  are the wave length and the wave number in free space, respectively, and  $R$  is the distance between field point and source point.

So the far-zone electric field of the plasma antenna with a length  $l$  is

$$E_\theta = \int_0^l j \frac{60\pi I_z}{\lambda r_1} \sin \theta e^{-jk_0 r_1} dz \quad (117.10)$$

**Fig. 117.4** Plasma antenna placed along the  $z$  axis



As shown in Fig. 117.4, a plasma column antenna,  $l = 1.2$  m and  $a = 0.0125$  m, is placed along the  $z$  axis.  $r_0$  is the distance between the origin to the field point, and  $r_l$  is the distance between the electric current element ( $Idz$ ) to the field point in far region. Consider  $l \ll r_0$ , so that we take  $1/r_0 \approx 1/r_l$ . Due to  $r_0 \parallel r_l$  in the far field, as the first approximation, we can take  $r_l = r_0 \cos \theta$ , then

$$E_\theta = j \frac{60\pi}{\lambda r_0} e^{-jk_0 r_0} \sin \theta \int_0^l I_z e^{jk_0 z \cos \theta} dz \tag{117.11}$$

The power radiated by the plasma antenna is

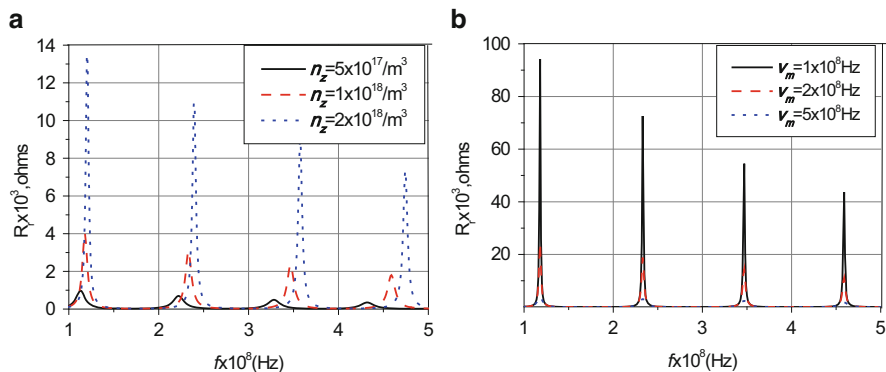
$$P_r = \oint_\Omega \vec{S} \cdot d\vec{A} \tag{117.12}$$

where  $\Omega$  is a sphere surface of its radius in the far-field region. Poynting vector  $\vec{S} = \frac{1}{2} \text{Re} \left[ \vec{E} \times \vec{H}^* \right]$ .

Finally, the radiation resistance can be obtained by the principle of the equivalent circuit.

$$R_r = P_r / I_1^2 \tag{117.13}$$

When the electron-neutral collision frequency is  $5 \times 10^8$  Hz, the radiation resistance of the plasma antenna is plotted in Fig. 117.5a at different plasma densities. In the figure, we get that the plasma antenna displays resonance characteristic similar to that of the metal one, the curve is moved to the right, and its peak values become bigger with the increase of concentration for plasma. For a plasma column, the increasing plasma density will cause a good conductivity, and this leads to a stronger ability to radiate EM wave. On the other hand, the peak value of radiation resistance is approximately periodic, which will help us select a frequency band to realize a high-efficiency plasma antenna. Figure 117.5b shows that the radiation resistance of the plasma antenna varies with different electron-neutral



**Fig. 117.5** Radiation resistance for the plasma antenna (a) at different plasma densities and (b) at different collision frequencies

collision frequencies, when  $n_e = 1 \times 10^{18} \text{ m}^{-3}$ . The results tell us that the plasma antenna displays resonance characteristics too and the peak values of the curve with the increase of collision frequencies of plasma. As mentioned previously, the power is consumed in the plasma column, which reduces the EM power radiation, so the radiation resistance becomes smaller.

### 117.4 Conclusion

Since high-frequency electromagnetic waves act with plasma and charged particles collide with each other, a plasma column not only shows the ability of EM power radiation but also demonstrates ohmic loss characteristics. In this chapter, the internal impedance and radiation resistance of plasma antenna are analyzed through the theory of impedance and the method of Poynting vector respectively. The results indicate that the internal resistance could be reduced and the radiation resistance be increased efficiently by adjusting the plasma density and the collision frequency properly; accordingly, the efficiency of the antenna is improved too. It is very useful for the analysis and design of plasma antenna.

### References

1. Cerri, G., De Leo, R., Mariani Primiani, V., & Russo, P. (2008). Measurement of the properties of a plasma column used as a radiated element. *IEEE Transactions on Instrumentation and Measurement*, 57(2), 242–247.
2. Alexe, I., Anderson, T., Farshi, E., Karnam, N., & Pulasani, N. R. (2008). Recent results for plasma antennas. *Physics of Plasmas*, 15(5), 057104.

3. Kumar, R., & Bora, D. (2010). A reconfigurable plasma antenna. *Journal of Applied Physics*, 107(5), 053303.
4. Zhu, A., Chen, Z., Lv, J., & Liu, J. (2012). Characteristics of AC-biased plasma antenna and plasma antenna excited by surface wave. *Journal of Electromagnetic Analysis and Applications*, 4(7), 279–284.
5. Wu, X. P., Shi, J. M., Chen, Z. S., & Xu, B. (2012). A new plasma of beam-forming. *Progress in Electromagnetics Research*, 126, 539–553.
6. Yin, B., Yang, F., Wang, B., & Hao, H. G. (2011). Mutual impedance of plasma antennas, 2011. In *7th International Conference on Wireless Communications, Networking and Mobile Computing* (pp. 1–4). Wuhan, China: IEEE Press.
7. Lieberman, M. A., & Lichtenberg, A. J. (2005). *Principles of plasma discharges and materials processing* (pp. 46–88). New York, NY: Wiley.
8. Yuan, C. X., Zhou, Z. X., & Sun, H. G. (2010). Reflection properties of electromagnetic wave in a bounded plasma slab. *IEEE Transaction on Plasma Science*, 38(12), 3348–3355.
9. Rayner, J. P., Whichello, A. P., & Cheetham, A. D. (2004). Physical characteristics of plasma antennas. *IEEE Transactions on Plasma Science*, 32(1), 269–281.
10. Rayner, J. P., & Cheetham, A. D. (2010). Travelling modes in wave-heated plasma sources. *IEEE Transactions on Plasma Science*, 38(2), 62–72.
11. Balanis, C. (2005). *Antenna theory-analysis and design* (3rd ed., pp. 31–69). New York, NY: Wiley.

# Chapter 118

## A Low-Voltage 5.8-GHz Complementary Metal Oxide Semiconductor Transceiver Front-End Chip Design for Dedicated Short-Range Communication Application

Jhin-Fang Huang, Jiun-Yu Wen, and Yong-Jhen Jiagn

**Abstract** A 5.8-GHz transceiver front-end applied in dedicated short-range communication (DSRC) systems which is developed in public traffic transportation to improve the safety is fabricated on a chip using TSMC 0.18- $\mu\text{m}$  CMOS process. The proposed prototype includes an asymmetric T/R switch, a current-reused LNA, and a class A power amplifier (PA) on the low-voltage operation in order to minimize the power consumption. Measured results achieve the power gain of 11 dB, the NF of 4.9 dB, the third-order intercept point (IIP3) of  $-5.4$  dBm, and the power consumption of 3.9 mW in the receiving (Rx) mode. On the other hand, the power gain of 12.4 dB, the output 1 dB compression point ( $\text{OP}_{-1\text{dB}}$ ) of 11.4 dBm, the PAE of 14.7 % at  $\text{P}_{-1\text{dB}}$ , the IMD3 of  $-15.8$  dBc at 1 dB compression level, the output power of 2.6 dBm with a 50  $\Omega$  load, and power consumption of 116.3 mW are obtained in the transmitting (Tx) mode. The overall chip area is 1.5 (1.32  $\times$  1.14)  $\text{mm}^2$ . This RF CMOS transceiver front-end includes all matching circuits and biasing circuits, and no external components are required.

### 118.1 Introduction

For mobile wireless communication, handheld sets of small size and light weight are more attractive, and the battery endurance plays an important role in the market. In order to minimize the required power consumption, operating the circuit at a reduced supply voltage is apparently an effective approach. ITS communication system can effectively improve the mobile safety and traffic efficiency in vehicle transportation. The DSRC protocol is defined in the physical layer of ITS and

---

J.-F. Huang (✉) • Y.-J. Jiagn  
Department of Electronic Engineering, National Taiwan University of Science and Technology, Taipei 106, Taiwan  
e-mail: [jfhuang@mail.ntust.edu.tw](mailto:jfhuang@mail.ntust.edu.tw)

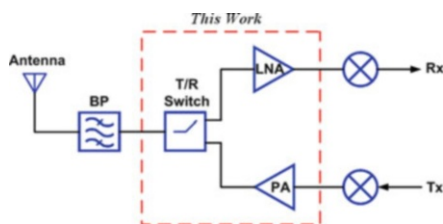
J.-Y. Wen  
National Communications Commission, Taipei 106, Taiwan

supports both public safety and private operations in vehicle to roadside communications and provides a high-speed radio link between the roadside unit (RSU) and onboard unit (OBU). Meanwhile, the DSRC protocol has been developed worldwide and practically applied for electronic toll collection (ETC) system. Furthermore, the 5.8-GHz band is located in the unlicensed industrial, scientific, and medical (ISM) bands and is widely applied for medium distant communication applications.

A fully integrated 0.25- $\mu\text{m}$  SiGe-BiCMOS transceiver for DSRC applications is presented, but this chip is more expensive than CMOS one [1]. In addition, larger power consumption is unpractical for OBU's application under the supply voltage of 3.3 V. A series-shunt T/R switch is integrated with transceiver front-end amplifier, but its power-handling capability is too low for Tx path [2]. A fully integrated transceiver front-end is proposed with the T/R switch, LNA, and PA devices, yet this chip is unpractical due to larger power consumption [3]. A fully integrated high-efficiency linear CMOS class E PA for 5.8-GHz ETC applications is presented [4]. Class E PA has high efficiency, but it is only suitable for nonlinear modulation communication systems. For DSRC systems, it is not suitable. Hence, with those considerations, a low-voltage 5.8-GHz CMOS transceiver front-end chip design for DSRC applications is presented in this paper.

## 118.2 Transceiver Front-End Circuit Design

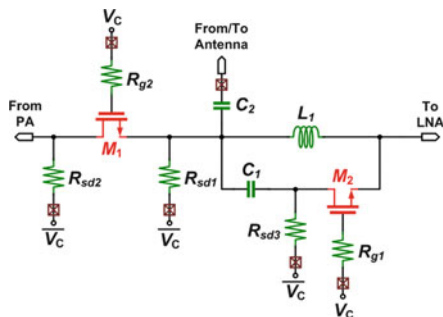
Figure 118.1 shows the structure of the proposed prototype which includes a T/R switch, an LNA, and a PA. The T/R switch is a key block for a time division duplex (TDD)-based radio system, and it connects radio transmitter and receiver alternatively to a shared antenna. It receives the radio signal from antenna to the LNA or transmits the radio signal from the PA to antenna. In Rx mode, the LNA appropriately amplifies the weak radio signal from the antenna through the T/R switch but not adding too much noise to it. In Tx mode, the PA amplifies the radio signal from the up-converter to appropriate signal power level through the T/R switch to the antenna.



**Fig. 118.1** The structure of the proposed transceiver front-end



**Fig. 118.2** The proposed asymmetric T/R switch circuit



### 118.2.1 An Asymmetrical T/R Switch Circuit

The proposed T/R switch is an asymmetrical architecture shown in Fig. 118.2. In the Tx path, a series switch,  $M_1$ , is used, whereas a shunt switch,  $M_2$ , is employed in the Rx path. A digital control signal,  $V_C$ , is applied to the gates of the switches to select Tx or Rx mode operation. Series gate bias resistors,  $R_{g1}$  and  $R_{g2}$ , are used, so the gate potential is bootstrapped to the source and drain. The size of  $M_1$  is determined based on the trade-off of the on-resistance,  $R_{on}$  which affects insertion loss, and  $C_{sd}$  in cutoff region, which affects isolation. The on-resistance can be evaluated as

$$R_{on} = \frac{1}{\mu_n C_{ox} W (V_{gs} - V_{TH}) / L}. \tag{118.1}$$

It is desired to keep  $R_{on}$  small to reduce the insertion loss. This can be achieved by choosing large mobility,  $\mu$ ; increasing transistor aspect ratio,  $W/L$ ; and keeping  $V_{gs} - V_{TH}$  large where  $V_{TH}$  is the threshold voltage.

One criterion uses NMOS transistors rather than PMOS transistors in the design. The other rule designs transistors with minimum allowable channel length,  $L$ . Because the minimum value of  $L$  is limited by the process, low  $R_{on}$  eventually requires large  $W$ . However, broadening a transistor definitely increases its junction and parasitic capacitances proportionally. The source and drain voltages ( $V_{S/D}$ ) of  $M_1$  and  $M_2$  are biased by the inverted  $V_C$  through large bias resistors ( $R_{sd1} - R_{sd3}$ ). In Tx mode,  $V_C$  is set at 1.8 V, so  $V_{SD}$  of  $M_1$  and  $M_2$  is at 0 V. The Rx path presents a parallel resonant tank with  $L_1$  and  $C_1$  shorted through  $M_2$ . The large bias resistor  $R_{sd3}$  causes the source terminal of  $M_2$  to be floating. This keeps the impedance between  $M_2$ 's source and drain small in spite of large signal swings at the antenna node. As a result, the quality factor ( $Q$ ) of the  $L_1 C_1$  tank remains sufficiently high to effectively block out leakage power from the Tx branch.

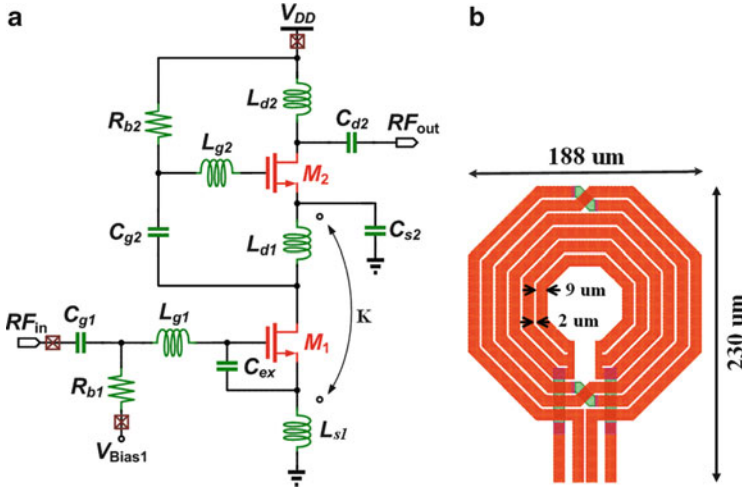


Fig. 118.3 The proposed current-reused LNA: (a) schematic and (b) transformer layout

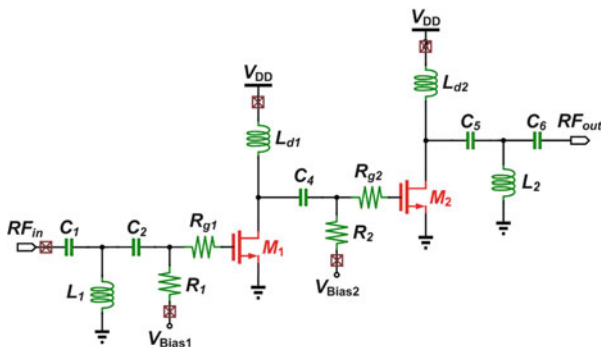
### 118.2.2 A Current-Reused Low-Noise Amplifier

To reduce power consumption, a two-stage cascade amplifier is folded into a single-stage cascode amplifier. Applying this idea obtains the proposed current-reused LNA circuit which comprises an input matching network matched to 50, and an on-chip transformer-degenerated cascode amplifier shown in Fig. 118.3a. Figure 118.3b illustrates this transformer layout. The two inductors  $L_{d1}$  and  $L_{s1}$  forming this transformer are connected at the source terminals of  $M_1$  and  $M_2$  for reducing chip area. The transformer has turn ratio of 1:3, 9- $\mu\text{m}$  metal width, 2- $\mu\text{m}$  space, 30- $\mu\text{m}$  inner length, 188- $\mu\text{m}$  outer length, and a chip area of  $188 \times 230 \mu\text{m}^2$ . This LNA provides higher gain than the common cascode amplifier since when operating in a 5.8-GHz frequency band,  $C_{g2}$  is shorted and the LNA acts as a cascode amplifier. On the contrary, when operating in lower frequency band,  $C_{g2}$  is open and the LNA acts as a cascode amplifier. Obviously, the power consumption of this circuit will keep the same as a cascode amplifier. The input series gate matching inductor actually comes from the  $L_1$  used in Fig. 118.2. The chip area can then be reduced. The capacitive coupling  $C_{g2}$  is needed to achieve DC isolation between the active devices of both transistors  $M_1$  and  $M_2$ .  $L_{g2}$  is added to optimize interstage matching with  $C_{g2}$ , the effective parasitic capacitance of the active device  $M_2$  and the on-chip transformer.

Neglecting the effect of the bias resistor,  $R_{b1}$ , the LNA input impedance  $Z_{in}$  is solved by writing Kirchoff’s voltage law in its phasor form across its input loop:

$$Z_{in} = \frac{g_{m1}L_{s1}}{C_{ex}} + j \left[ \omega(L_{s1} + L_{g1}) + -\frac{1}{\omega C_{g1}} - \frac{1}{\omega C_{ex}} \right], \quad (118.2)$$

**Fig. 118.4** Schematic of the proposed power amplifier circuit



where  $g_{m1}$  is the transconductance of the transistor  $M_1$  and the added  $C_{ex}$  contains  $C_{gs1}$ , the parasitic capacitance between the gate and source nodes.

For input matching,  $Z_{in} = R_s = 50 \Omega$ , from Eq. 118.2, the input matching occurs at the frequency  $\omega_c$  as

$$\omega_c = \sqrt{\frac{1}{L_{s1} + L_{g1}} \cdot \frac{C_{g1} + C_{ex}}{C_{g1} + C_{ex}}}. \tag{118.3}$$

At this frequency, the input impedance becomes

$$R_s = \text{Re}(Z_{in}) = \frac{g_{m1}L_{s1}}{C_{ex}}, \tag{118.4}$$

where  $\text{Re}(Z_{in})$  means the real part of  $Z_{in}$ .

### 118.2.3 Class A Power Amplifier

The proposed PA circuit is shown in Fig. 118.4. It is a two-stage cascade common source amplifier. Both the drive stage and the power stage are biased on class A operation to achieve high linearity. To evaluate the PA function, power-added efficiency,  $PAE$  includes information on the driving power for the PA and is more commonly used than power conversion efficiency.  $PAE$  is defined as

$$PAE = \frac{P_{out} - P_{in}}{P_{DC}}, \tag{118.5}$$

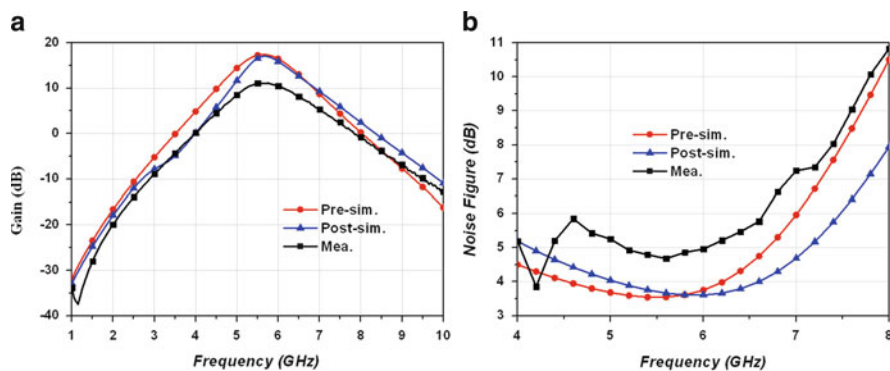
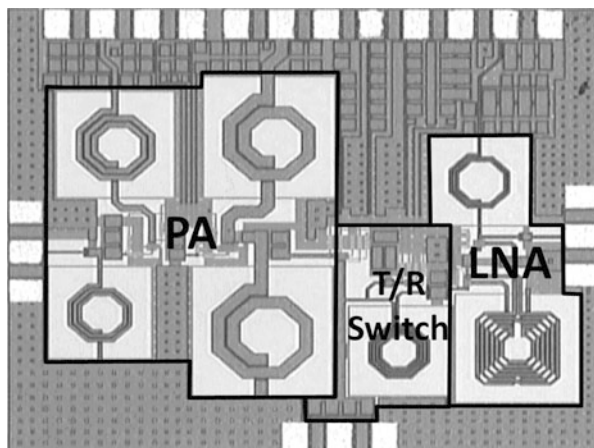
where  $P_{in}$  and  $P_{out}$  are input power and output power, respectively, at the frequency of interest.  $P_{DC}$  is the DC supply power.

### 118.3 Measured Results

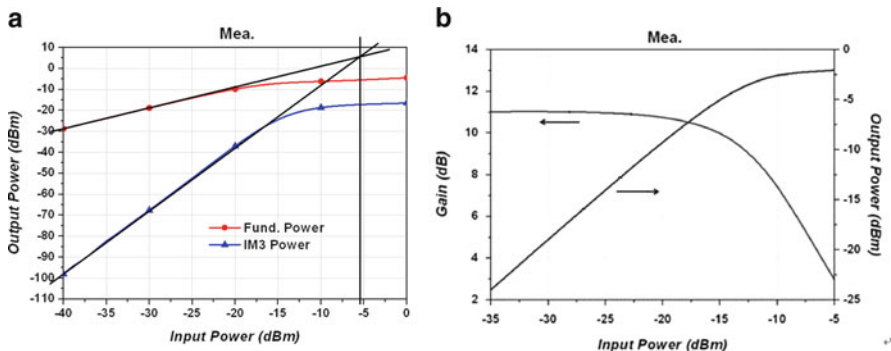
The proposed transceiver front-end is fabricated in TSMC 0.18- $\mu\text{m}$  CMOS process. The die photograph is shown in Fig. 118.5, and the chip area including pads is 1.5 ( $1.32 \times 1.14$ )  $\text{mm}^2$ . Measurements have been performed with a GSG probe bench with an HP 8510C network analyzer, an Agilent 8975A NF analyzer, and an Agilent E4407B spectrum analyzer. The power consumption is 3.9 mW with 1-V supply voltage in the Rx mode and 116.3 mW with 1.8-V supply voltage in the Tx mode.

Figure 118.6a, b shows the measured power gain and noise figure, respectively, in the Rx mode. The measured power gain is 11 dB somewhat below the simulated value about 6 dB at 5.8 GHz, but their data curves are pretty matched. The measured NF is 4.9 dB at 5.8 GHz. Figure 118.7a, b shows the measured  $P_{-1\text{dB}}$  in the Rx mode. The measured  $P_{-1\text{dB}}$  is  $-15$  dBm. Figure 118.8 shows the measured IIP3 in the Rx mode. The measured IIP3 is  $-5.4$  dBm. The  $P_{-1\text{dB}}$  of

**Fig. 118.5** Die photomicrograph of the transceiver front-end with a chip area of  $1.32 \times 1.14$  ( $1.50$ )  $\text{mm}^2$

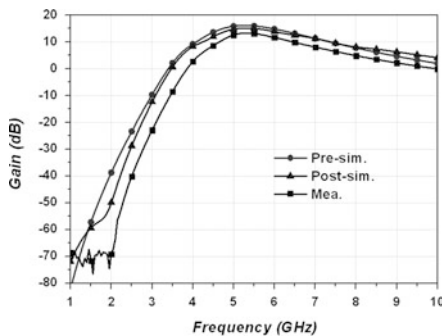


**Fig. 118.6** Measured power gain and noise figure vs. frequency in the Rx mode



**Fig. 118.7** (a) Measured power gain and output power vs. input power in the Rx mode with  $P_{-1dB} = -15$  dBm and (b) measured IIP3 in the Rx mode with an IIP3 value of  $-5.4$  dBm

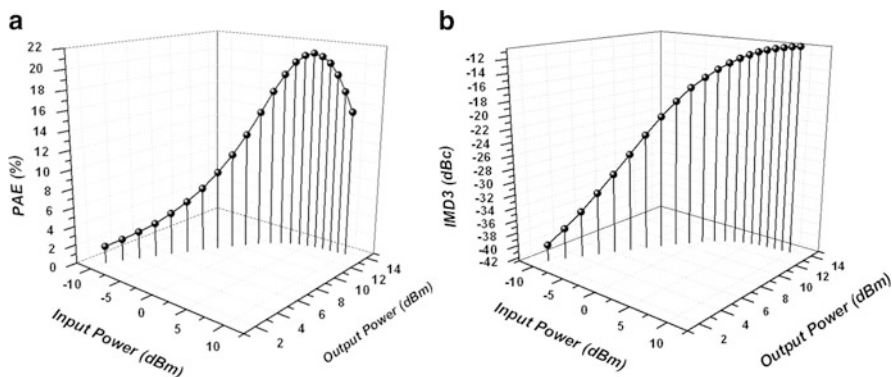
**Fig. 118.8** Measured power gain vs. frequency in the Tx mode with a power gain of 12.4 dB at 5.8 GHz



output power at the linear region of operation can also be verified by  $P_{-1dB} = IIP3 - 9$  (dBm).

Figure 118.8 shows the measured power gain in the Tx mode with a value of 12.4 dB at 5.8 GHz. Figure 118.9a shows that the measured result exhibits an  $OP_{-1dB}$  of 11.4 dBm with a PAE of 14.7 % in the Tx mode. The maximum PAE of 21 % is achieved at an output power of 13.7 dBm, where the power gain is greater than 12 dB.

Figure 118.9b depicts the measured IMD3 and output power as a function of input power in the Tx mode, and the measured IMD3 is  $-15.8$  dBc at  $P_{-1dB} = -15$  dBm. The measured output power at single port with no buffer circuits is 2.6 dBm at the frequency  $f_o$  of 5.8 GHz with input power of  $-10$  dBm. The measured performance of the proposed transceiver front-end is summarized and compared to recently other reported works in Table 118.1. The proposed transceiver front-end achieves the smallest chip area, the lowest NF, and the lowest power consumption while attaining very good performances compared to other features.



**Fig. 118.9** (a) Measured PAE and output power vs. input power in the Tx mode with a measured PAE of 14.7 % at  $P_{-1\text{dB}}$  of 11.4 dBm and (b) measured IMD3 and output power vs. input power in the Tx mode with a measured IMD3 of  $-15.8$  dBc at  $P_{-1\text{dB}} = -15$  dBm

**Table 118.1** Comparison of some transceiver front-ends in recent publications

	[5] 2005	[6] 2010	[7] 2006	This work
Process ( $\mu\text{m}$ )	BiCMOS 0.35	CMOS 0.13	CMOS 0.18	CMOS 0.18
$f_{\text{RF}}$ (GHz)	5.2	5.8	5.5	5.8
Chip area ( $\text{mm}^2$ )	2.86	6.24	NA	1.5
<i>Rx mode</i>				
$V_{\text{DD}}$ (V)	N/A	1.2	3.0	1
DC current (mA)	9.8	52 <sup>a</sup>	40 <sup>b</sup>	3.9
Conver. gain (dB)	21.9	48	NA	11
NF (dB)	N/A	5	NA	4.9
IIP3 (dBm)	N/A	1.1	N/A	$-5.4$
<i>Tx mode</i>				
$V_{\text{DD}}$ (V)	N/A	3.3	3.0	1.8
Current (mA)	59.1	150	40 <sup>b</sup>	64.6
Power gain (dB)	12.1	NA	NA	12.4
$OP_{-1\text{dB}}$ (dBm)	N/A	$-41$	NA	11.4
PAE @ $P_{-1\text{dB}}$ (%)	N/A	9.2	N/A	14.7

<sup>a</sup>52 mA includes Rx RF front-end, Rx baseband analog circuits, and Rx LO buffers. 150 mA includes Tx RF front-end, Tx baseband analog circuits, and Tx LO buffers

<sup>b</sup>Both Tx/Rx modes

## 118.4 Conclusion

A 0.18- $\mu\text{m}$  5.8-GHz CMOS fully integrated transceiver front-end for DSRC applications was presented. A T/R switch of asymmetric topology is advantageous for handling high power since it takes into account the asymmetrical power level in the Tx and Rx branches of a typical transceiver. The T/R switch in Rx side is merged with the LNA architecture consuming very small chip area and is almost with no

power consumption. On-chip T/R switch is built in the transceiver front-end chip, allowing for reduced signal loss in the whole architecture. This integrated CMOS transceiver front-end includes all matching circuits and biasing circuits, and no external components are required. Our work can provide a compact, low-power, and low-cost solution to DSRC payloads.

**Acknowledgements** The authors would like to acknowledge the fabrication support and chip fabrication provided by the National Chip Implementation Center (CIC). Thanks are also given to Dr. Ron-Yi Liu for his layout guidance and Taiwan Mobile-Phone Inc. for the financial support.

## References

1. Sasho, N., Minami, K., Fujita, H., Takahashi, T., Iimura, K., Abe, M., et al. (2008). Single chip 5.8GHz DSRC transceiver with dual-mode of ASK and Pi/4-QPSK[C]. *Proceedings of IEEE Radio and Wireless Symposium* (pp. 799–802). Orlando, FL.
2. Yamamoto, K., Heima, T., Furukawa, A., Ono, M., Hashizume, Y., Komurasaki, H., et al. (2001). A 2.4-GHz-band 1.8-V operation single-chip Si-CMOS T/R-MMIC front-end with a low insertion loss switch. *IEEE Journal of Solid-State Circuits*, 36(8), 1186–1197.
3. Jou, C.-F., Huang, P.-R., & Cheng, K.-H. (2003). Design of a 0.25- $\mu\text{m}$  transceiver front-end[C]. *Proceedings of IEEE International Symposium on Electronics, Circuits, and Systems* (Vol. 3, pp. 1090–1093). Sharjah, United Arab Emirates.
4. Suh, Y., Sun, J., Horie, K., Itoh, N., & Yoshimasu, T. (2009). Fully-integrated novel high efficiency linear CMOS power amplifier for 5.8 GHz ETC applications[C]. *Proceedings of Asia Pacific Microwave Conference* (pp. 365–368). Singapore.
5. Kanaya, H., Koga, F., Seki, K., & Yoshida, K. (2005). Impedance matching circuit for wireless transceiver amplifier based on transmission line theory[C]. *Proceedings of Asia Pacific Microwave Conference* (pp. 19–22). Suzhou, China.
6. Kwon, K., Choi, J., Choi, J., Hwang, Y., Lee, K., & Ko, J. (2010). A 5.8 GHz integrated CMOS dedicated short range communication transceiver for the Korea/Japan electronic toll collection system. *IEEE Transactions on Microwave Theory and Techniques*, 58(11), 2751–2763.
7. Nagata, M., Masuoka, H., Fukase, S.-I., Kikuta, M., Morita, M., & Itoh, N. (2006). 5.8 GHz RF transceiver LSI including on-chip matching circuits[C]. *Proceedings of IEEE Bipolar Circuits and Technology Meeting* (pp. 263–266). Maastricht Dutch.

# Chapter 119

## A 5.8-GHz Frequency Synthesizer with Dynamic Current-Matching Charge Pump Linearization Technique and an Average Varactor Circuit

Jhin-Fang Huang, Jia-Lun Yang, and Kuo-Lung Chen

**Abstract** A 5.8-GHz frequency synthesizer is implemented in TSMC 0.18- $\mu\text{m}$  CMOS process. This paper proposes a dynamic current-matching charge pump linearization technique and uses a current-switching differential Colpitts VCO to lower the phase noise and an averaged varactor circuit to increase the linearity of the VCO tuning range. At the supply voltage of 1.8 V, measured results achieve the locked tuning frequency from 5.55 to 5.94 GHz, corresponding to 6.8 % and the phase noise of  $-105.83$  dBc/Hz at 1 MHz offset frequency from 5.8 GHz. The overall power consumption is 21.6 mW. Including pads, the chip area is 0.729 ( $0.961 \times 0.761$ )  $\text{mm}^2$ .

### 119.1 Introduction

Frequency synthesizer is an important component used in wireless transceiver front-end to perform signal up- and down-conversion. The integer-N frequency synthesizer is considered to be well understood and less complicated to design. Several multiband frequency synthesizers have been published [1–3]. A 5-GHz frequency synthesizer utilizes injection-locked frequency divider (ILFD) in the first divider stage to save power found, but it may cause the frequency synthesizer becoming unlocked due to the narrow locked range of ILFD and consuming more chip area due to the inductor of ILFD [1]. A 1-V frequency synthesizer for

---

J.-F. Huang (✉) • J.-L. Yang  
Department of Electronic Engineering, National Taiwan University of Science and Technology, Taipei 106, Taiwan, China  
e-mail: [jfhuang@mail.ntust.edu.tw](mailto:jfhuang@mail.ntust.edu.tw)

K.-L. Chen  
National Communications Commission, Taipei 106, Taiwan, China



low-voltage applications is presented, but it consumes large chip area of 0.988 mm<sup>2</sup> and much power of 27.5 mW [2]. A locking detector to detect the locking situation of the PLL is adopted, but it is more complex as illustrated in [3]. In the phase/frequency detector (PFD) and charge pump (CP) circuits, the nonlinearity is mainly attributed to the up/down current mismatch and the gain (slope) variation around the region of phase error; while in the dividers, the circuit timing jitters modulate the zero-crossing points of a signal and cause the system to exhibit nonlinear behavior. Hence, considering those factors of power consumption, phase noise, tuning range and chip area, a low-phase noise, wide tuning range, and small chip area, frequency synthesizer is proposed and fabricated in TSMC 0.18-μm CMOS process.

### 119.2 Architecture of Frequency Synthesizer

The proposed frequency synthesizer consists of a PFD, a CP, an off-chip 3rd-order passive loop filter and a VCO in the feed-forward path and a programmable frequency divider in the feedback path as shown in Fig. 119.1. An accurate VCO with low noise is essential in designing a quality frequency synthesizer. The PFD detects the phase error between the reference signal  $F_{REF}$  and the feedback signal  $F_{DIV}$ . The digital output signals of PFD control the VCO through the CP and filter circuits. The locked frequency synthesizer output frequency  $f_{VCO}$  is expressed as follows:

$$f_{VCO} = 4 \times N_{MMFD} \times F_{REF}, \tag{119.1}$$

where  $N_{MMFD}$  is the programmable divider ratio and  $F_{REF}$  is the reference frequency.

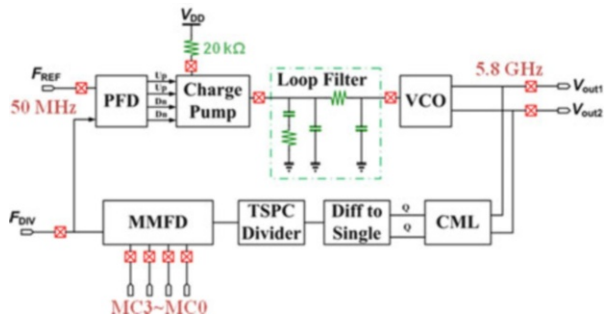


Fig. 119.1 Architecture of the proposed frequency synthesizer

### 119.3 Frequency Synthesizer Functions

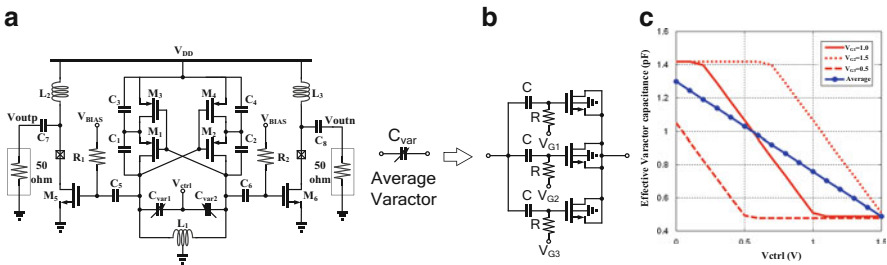
#### 119.3.1 $G_m$ -Boosting Voltage-Controlled Oscillator

The cross-coupled VCO in CMOS has attracted considerable interest due to its easy start-up and good phase noise characteristics. Colpitts VCO features superior phase noise because noise current from active devices is injected into the tank during minima of the tank voltage when the impulse sensitivity is low. Unfortunately, the conventional Colpitts VCOs suffer from poor start-up characteristics; i.e., higher power consumption is needed to ensure reliable start-up.

In order to resolve the poor start-up characteristic, and improve the phase noise, the current-switching differential Colpitts VCO shown in Fig. 119.2 where the differential oscillator is built with symmetry around the resonator tank is modified from [4]. The phases of gate voltages of  $M_1$  and  $M_3$  and  $M_2$  and  $M_4$  are the same as the phases of drain voltages of  $M_2$  and  $M_1$ , respectively. Connecting them together will have the effect of  $G_m$ -boosting scheme. The negative resistance  $-2/g_m$  where  $g_m$  denotes the transconductance of each transistor generated by the cross-coupled pMOS transistors is to compensate for the loss with the LC-tank. The differential outputs of the VCO connect to the common source amplifiers which function as analog buffers. This proposed balanced VCO consists of two single-ended LC-tanks and two pairs of pMOSFETs so that their gate-source voltages become small. To enhance the start-up oscillation condition of the balanced Colpitts VCO, the pMOSFET core is chosen to reuse the dc current. Therefore, the proposed balanced VCO uses four MOSFETs, and the consumed current still remains very small; therefore, low power dissipation can be achieved.

The capacitance  $C_{var1, 2}$ , realized from a high-Q MOSFET capacitor, in parallel with the inductors forms the LC-tank resonator which determines the oscillating frequency. The oscillating frequency is given by

$$f_o = \frac{1}{2\pi\sqrt{L_1 C_{vnet}}} \tag{119.2}$$



**Fig. 119.2** The gain-boosting Colpitts VCO: (a) the overall circuit; (b) an averaged varactor circuit by using distributed bias voltages of 1.0, 1.5, and 0.5 V; and (c) linear property of the averaged capacitance

where  $C_{vnet}$  is the effective capacitance of the VCO. Three MOS varactors are connected in parallel with dc bias voltages  $V_{G1}$ ,  $V_{G2}$ , and  $V_{G3}$ . Rather than using a fixed dc bias voltage (i.e.,  $V_{G1} = V_{G2} = V_{G3}$ ) in the traditional varactor design, the averaged varactor circuit uses distributed voltage values for  $V_{G1}$ ,  $V_{G2}$ , and  $V_{G3}$  which are 1.0, 1.5, and 0.5 V, respectively, as shown in Fig. 119.2b [5]. The tuning capacitance of the combination of the three varactors is approaching linear shown in Fig. 119.2c and is therefore insensitive to the varactor errors. The resonant frequency of the LC-tank will become more stable and the nonlinearities of the varactors are averaged, but at the sacrifice of the chip area of  $0.15 \text{ mm}^2$ . The averaged varactor increases the linearity of VCO tuning range. Therefore, the phase noise is lowered and this feature is applied to these VCO varactors.

### 119.3.2 Phase Frequency Detector

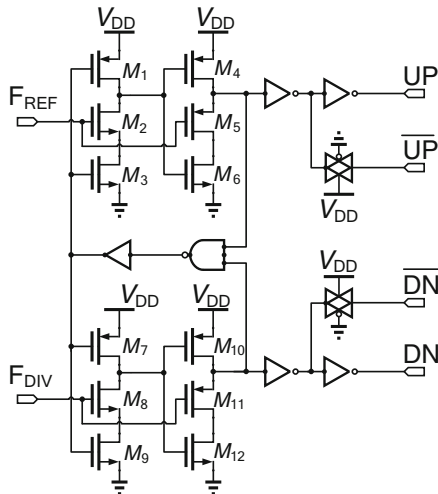
Important techniques to design PFD operating at high frequency with minimum dead zone are adopted to reach the minimum phase offset and to reduce the dead zone in the circuit. With these considerations, a domino-logic PFD shown in Fig. 119.3 is presented [6]. On the reset path, the reset delay cell is inserted to reduce the minimum UP and DN pulse widths and then to improve the dead zone in the PFD. When the input and output frequencies are sufficiently close, the PFD operates as a phase detector, performing phase lock. The loop locks when the phase difference drops to zero and the charge remains relatively idle.

### 119.3.3 Dynamic Current-Matching Charge Pump

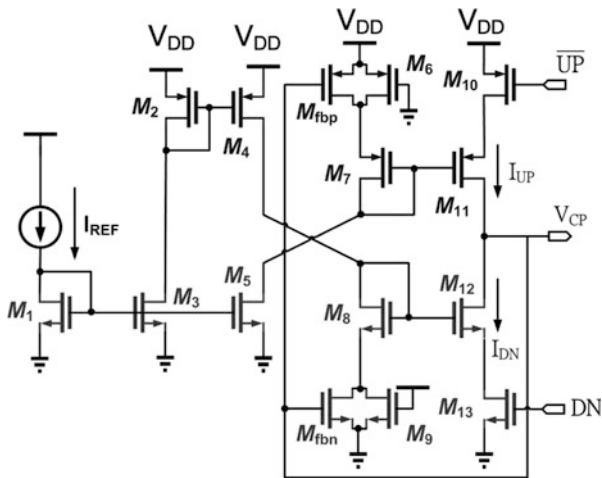
A charge pump consists of two switched current sources that pump charge into or out of the loop filter according to logic inputs. The main objective is to design a reasonably sized CP circuit to achieve the performance of a large-sized one for area efficiency and minimizing the unwanted transient corruptions. Figure 119.4 illustrates the proposed CP PLL driven by a PFD, such an implementation senses the transitions at the input and output, detects phase or frequency differences, and activates the charge pump accordingly [7].

This CP is based on a switches-in-source architecture. The feature is that two extra feedback transistors,  $M_{fbN}$  and  $M_{fbP}$ , are added to compensate for the channel-length modulation effect of the up/down current mirrors via negative feedback. The amount of compensation is dynamically adjusted according to the CP output voltage ( $V_{CP}$ ). The technique reduces the current error between charge up and charge down due to the mismatch of charge-sharing effects when the switches are turned ON and hence improves the in-band phase noise.

**Fig. 119.3** The proposed PFD circuit with minimum dead zone



**Fig. 119.4** Dynamic current-matching charge pump circuit



### 119.3.4 Low-Pass Filter

The low-pass filter eliminates noise from the tuning voltage. The VCO output signal is altered by the control data inputted into the filter which follows some parameter specs in the PLL, including reference frequency ( $F_{REF}$ ), divide ratio ( $N$ ), loop bandwidth ( $K$ ), charge pump current ( $I_{CP}$ ), slope of VCO ( $K_{VCO}$ ), and phase margin ( $\phi_P$ ). The filter suppresses spurs introduced by the reference frequency.

**Fig. 119.5** The third-order low-pass loop filter circuit

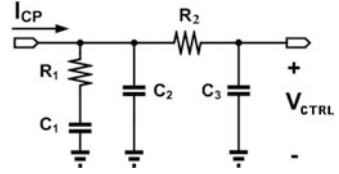


Figure 119.5 shows the schematic of the third-order low-pass loop filter which provides more attenuation of spurs by placing a series resistor  $R_2$  and a shunt capacitor  $C_3$ .  $V_{CTRL}(s)/I_{cp}(s)$  will become

$$Z(s) = \frac{sR_1C_1 + 1}{s^3R_1R_3C_1C_2C_3 + s^2(R_1C_1C_2 + R_1C_1C_3 + R_3C_2C_3 + R_3C_1C_3) + s(C_1 + C_2 + C_3)}. \tag{119.3}$$

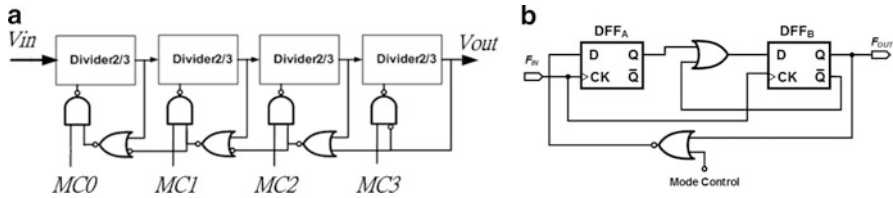
Resistor  $R_1$  and capacitor  $C_1$  generate a pole at the origin and a zero at  $1/(R_1C_1)$ .  $C_2$  and the combination of  $R_2$  and  $C_3$  generate extra poles at frequencies higher than frequency synthesizer bandwidth of interest to reduce the feedthrough at reference frequency and decrease spurious harmonics of the reference frequency. The component parameters used in this 3rd low-pass loop filter with consideration of CP currents, VCO tuning range gain, frequency divider ratio, loop filter bandwidth, phase noise, etc., are listed as follows:  $F_{REF} = 50$  MHz,  $K = 350$  kHz,  $I_{CP} = 100$   $\mu$ A,  $K_{VCO} = 216$  MHz/V, phase margin =  $62^\circ$ ,  $C_1 = 271$  pF,  $C_2 = 18$  pF,  $C_3 = 1$  pF,  $R_1 = 6.8$  k $\Omega$ , and  $R_2 = 13.6$  k $\Omega$ .

### 119.3.5 Multi-Modulus Frequency Divider

Figure 119.6a shows the programmable MMFD circuit which contains 2/3 true-single-phase-clock (TSPC) divider and traditional logic gate circuits, and Fig. 119.6b shows the 2/3 TSPC divider architecture which contains two traditional DFFs [8]. The MMFD has to treat frequency division over large continuous range and can be programmable.

In this design, the 16-modulus divider is chosen to deal with all of integer divide ratio values from 16 to 31. The divide ratio of MMFD is defined in Eq. 119.4. When the frequency synthesizer oscillates at 5.8 GHz, the divide ratio must be 29, and then the control code, MC0-MC3, is set to be (1, 1, 0, 1):

$$N_{Divide\_Ratio} = 16 + (2^3 \times MC3) + (2^2 \times MC2) + (2^1 \times MC1) + (2^0 \times MC0). \tag{119.4}$$



**Fig. 119.6** (a) The MMFD circuit containing 2/3 dividers and logic gates, and (b) the schematic of the proposed 2/3 divider circuit

## 119.4 Measured Results

The proposed frequency synthesizer was implemented in TSMC 0.18- $\mu\text{m}$  CMOS process. Figure 119.7 shows the die microphotograph including the wire-bonded pads. The overall chip area is 0.729 (0.961  $\times$  0.761)  $\text{mm}^2$  including measured pads. Under the supply voltage of 1.8 V, the power consumption is 21.6 mW.

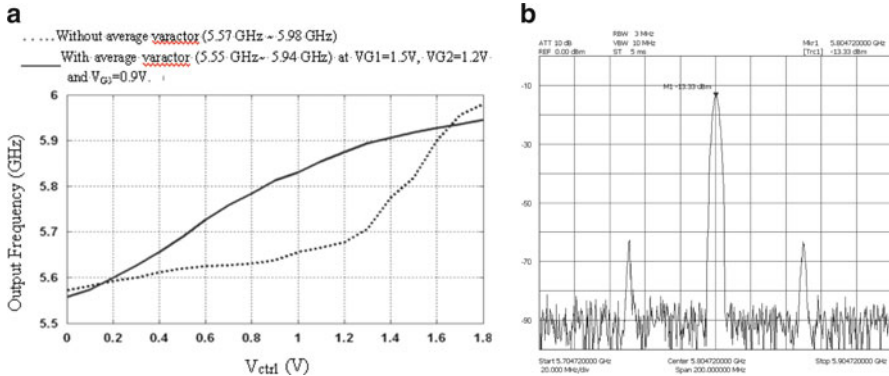
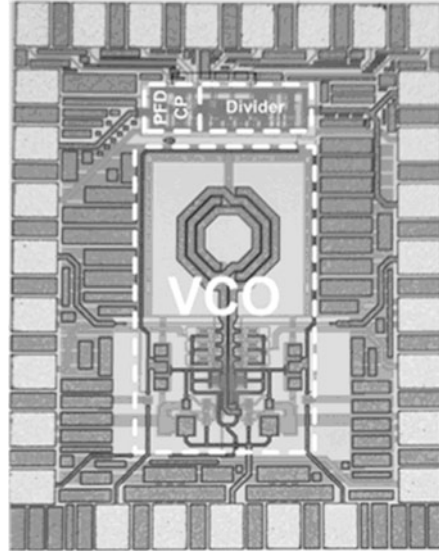
Measurements have been performed with an Agilent E4446A spectrum analyzer and an HP 8110A 150-MHz function generator which provides 50 MHz for reference frequency used to perform the measurement. The plot of the tuning characteristic of the locked frequency synthesizer versus the controlled voltage varying from 0 to 1.8 V is shown in Fig. 119.8a. From it we can find the much linearity of using average varactor circuit. The VCO output frequency is tunable from 5.5 to 5.94 GHz by varying the controlled voltage  $V_{ctrl}$ . Figure 119.8b shows the measured output power of  $-13.33$  dBm of the prototype after locking at 5.805 GHz. The reference spur is about 50 dBc and appears exactly at the reference frequency of 50 MHz.

Figure 119.9 shows both the measured phase noises of free-running VCO and the phase-locked VCO. According to the measured results, the phase noise of the locked VCO is  $-105.83$  dBc/Hz at 1 MHz offset frequency from 5.805 GHz. The measured performances of the proposed frequency synthesizer are summarized in Table 119.1 in comparison with other recently published papers. The proposed prototype achieves the highest output frequency, the widest tuning range, the least chip area, and lower phase noise, comparing to the other three references. The phase noise is  $-111$  dBc/Hz @ 1 MHz, but with a power consumption of 27.5 mW and a larger chip area of 0.988  $\text{mm}^2$  shown in [2]. Our prototype only needs power of 21.6 mW and chip area of 0.729  $\text{mm}^2$ .

## 119.5 Conclusion

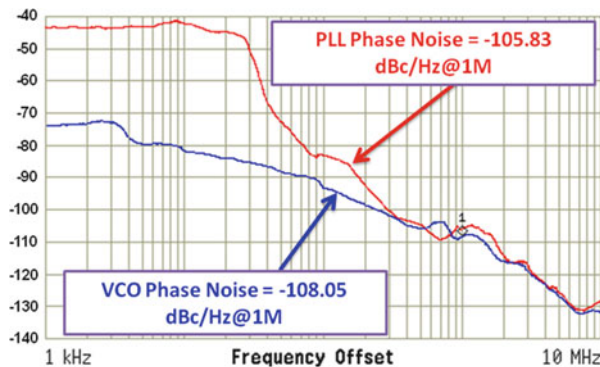
In this paper, a 5.8-GHz frequency synthesizer was fabricated in TSMC 0.18- $\mu\text{m}$  CMOS process. To improve phase noise, a cross-coupled Colpitts VCO with an average varactor circuit was adopted. The CP employed dynamic current-matching

**Fig. 119.7** Die micrograph of the proposed frequency synthesizer with a chip area of  $0.761 \times 0.961 \text{ mm}^2$



**Fig. 119.8** (a) Measured tuning range versus varying the control voltage, with average varactor circuit; the tuning frequency range is much linear, and (b) output spectrum of the locked frequency synthesizer with  $V_{DD} = 1.8 \text{ V}$  and  $V_{ctrl} = 0.9 \text{ V}$

**Fig. 119.9** Measured phase noises of VCO and locked PLL



**Table 119.1** Performance comparison of proposed PLL with previously published papers

	[1] (2009)	[2] (2004)	[3] (2009)	This work
Technology ( $\mu\text{m}$ )	0.18	0.18	0.18	0.18
Supply voltage (V)	1.8	1	1.8	1.8
Tuning range (GHz) (%)	5.15~5.35 (3.8)	5.45~5.65 (3.6)	5.1~5.35 (3.8)	5.55~5.94 (6.8)
Phase noise @ 1 MHz offset (dBc/Hz)	-104	-111	-104	-105.83
KVCO (MHz/V)	200	75	111	216
Loop bandwidth (kHz)	200	N.A.	200	350
Power consump (mW)	18	27.5	18	21.6
Chip area ( $\text{mm}^2$ )	1.05	0.988	1.045	0.729

circuit to compensate for the channel-length modulation effect. The measured phase noise and power consumption are  $-105.83$  dBc/Hz and 21.6 mW, respectively, and the locked tuning frequency is from 5.55 to 5.94 GHz and the chip area is only  $0.729$   $\text{mm}^2$ .

**Acknowledgements** The authors would like to thank Prof. Ron-Yi Liu for his layout guidance and the National Chip Implementation Center (CIC) for the chip fabrication and technical supports. We also thank the Taiwan Mobile Phone Company for the financial support.

## References

- Deng, P.-Y., & Kiang, J.-F. (2009). A 5-GHz CMOS frequency synthesizer with an injection-locked frequency divider and differential switch capacitors. *IEEE Transactions on Circuits and Systems*, 56(2), 320–326.
- Leung, G.-C., & Luong, H.-C. (2004). A 1-V 5.2-GHz CMOS synthesizer for WLAN applications. *IEEE Journal of Solid-State Circuits*, 36(11), 1873–1882.
- Chiu, W.-H., Huang, Y.-H., & Lin, T.-H. (2009). A 5GHz phase-locked loop using dynamic phase-error compensation technique for fast settling in 0.18- $\mu\text{m}$  CMOS. In *IEEE Symposium on VLSI Circuits* (pp. 128–129).
- Li, X., Shekhar, S., & Allstot, D.-J. (2005). Gm-boosted common-gate LNA and differential Colpitts VCO/QVCO in 0.18- $\mu\text{m}$  CMOS. *IEEE Journal of Solid-State Circuits*, 40(12), 2609–2619.
- Wu, T., Hanumolu, P.-K., Mayaram, K., & Moon, U.-K. (2009). Method for a constant loop bandwidth in LC-VCO PLL frequency synthesizers. *IEEE Journal of Solid-State Circuits*, 44(2), 427–435.
- Huang, J.-F., Mao, C.-C., & Liu, R.-Y. (2011). The 10 GHz wide tuning and low phase-noise PLL chip design. In *IEEE International Security and Identification Conference*, Xiamen, China (pp. 1–4).
- Lin, T.-H., Ti, C.-L., & Liu, Y.-H. (2009). Dynamic current-matching charge pump and gated-offset linearization technique for delta-sigma fractional-N PLLs. *IEEE Transactions on Circuits and Systems I*, 56(6), 877–885.
- Huang, J.-F., Shih, C.-W., & Liu, R.-Y. (2011). A 5.8-GHz frequency synthesizer chip design for worldwide interoperability for microwave access application. *Microwave and Optical Technology Letters*, 53(12), 2931–2935.



# Chapter 120

## Full-Wave Design of Wireless Charging System for Electronic Vehicle

Yongxiang Liu, Yi Ren, and Yi Wang

**Abstract** This chapter studies magnetic resonance based on wireless power transmission (WPT) system for electronic vehicle (EV). In this system, the two resonant coils mounted on the bottom of the vehicle and on the ground were simultaneously analyzed by the method of moments (MoM), an accurate and efficient full-wave electromagnetic analysis method. Then, compared with traditional WPT in ideal circumstance, the different performance of WPT in wireless charging system of EV is studied. Finally, a new design of the WPT integrated with circumstance is proposed, which achieves 90 % energy transmission efficiency at the resonant frequency of 13.56 MHz with the distance between two resonant coils varying within 15–25 cm.

### 120.1 Introduction

Higher power delivery through electromagnetic wave transmission is a fantastic technology after Tesla's hypothesis, which has been a topic of continued interests since the last several decades. This technology is not achieved until the electromagnetic inductance-based wireless power transmission (WPT) occurred [1]. The inductance-based WPT can only work in short range as several centimeters, which is far apart from industry's requirements. Meanwhile, another WPT was proposed as the radio-frequency WPT, which can transmit higher power energy by microwave in long distance. However, the ultra-high-gain antenna array is required which is usually not easy to achieve and very expensive. Besides that, many other

---

Y. Liu

Electric Power Research Institute, Chongqing 404100, China

e-mail: [L\\_yx123@qq.com](mailto:L_yx123@qq.com)

Y. Ren (✉) • Y. Wang

Chongqing University of Posts and Telecommunications, Chongqing 404100, China

e-mail: [renyi@cqupt.edu.cn](mailto:renyi@cqupt.edu.cn); [wangyi@cqupt.edu.cn](mailto:wangyi@cqupt.edu.cn)

shortages hampered this technology to apply and develop [2]. Recently, André Kurs proposed a new WPT based on nonradiative near-field magnetic resonant which is called magnetic resonant WPT [3]. Compared to the former two WPT systems, this new one worked in middle range, typically 2 m with energy transmit efficiency as high as 40 %. The high performance generates great interests in industry and scientific research, which is thought to have broad application prospects.

Nowadays, many potential applications of magnetic resonant WPT are proposed, typically in EV, mobile communication system, implant medical devices, etc. Specially, considering the shortage of fossil energy in the future and the convenience of WPT in EV, many research institute and company pay much more attention on its application research in EV [4]. To author's best knowledge, most of those researches relied on experimental tests and equivalent circuit theory analyses, in which the electromagnetic circumstance is usually simplified. Typically, most of the WPT design in EV system is not considering the affluence of the metallic car chassis and the ground. However, according to the knowledge of electromagnetic theory, the circumstance will strongly affect the performance of WPT. As a result, the design considering electromagnetic environmental integration is required for stability and higher-efficiency performance of magnetic resonant WPT in EV application.

In this work, the method of moments (MoM) [5, 6], an accurate and efficient full-wave electromagnetic analysis method, is applied to simulate the performance of WPT in our design. Furthermore, a new structure of WPT design is proposed, in which the chassis and the surface of the ground is set as perfect electronic conductor (PEC) and the circuits of WPT is mounted in the cavity on PEC ground. Our simulation results show that the power delivery efficiency is higher than 90 % when the distance of transmit and receive ports varied by 15–25 cm, which is the distance between the ground and the majority of car's chassis.

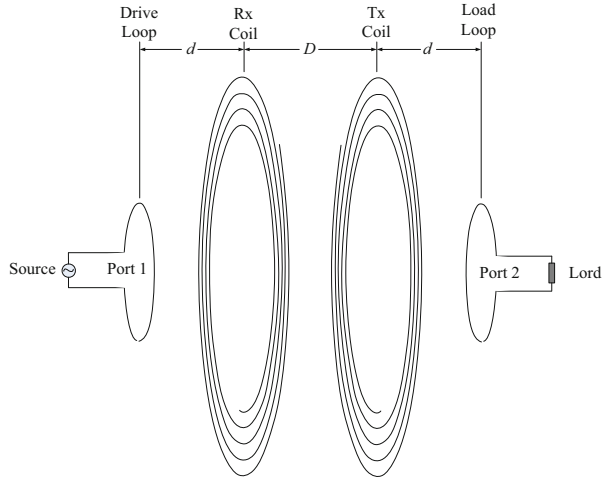
## 120.2 Wireless Power Transmission System

### 120.2.1 *Power Transmission with Nonradiative Near-Field System*

A typical nonradiative near-field WPT system consists of 2 or 4 circles, and the 4 circles system is considered much more efficient than the 2 circuits system. Therefore, in this paper, the 4 circles WPT system is analyzed as shown in Fig. 120.1.

Usually, this structure is analyzed with the equivalent circuit theory, and the details can be found in reference [3]. This method can only work in ideal circumstance where the WPT is applied without considering the influence of electromagnetic circumstances. Namely, this model cannot be applied in the EV circumstances analysis, where the chassis and ground will affect WPT strongly. Therefore, the best

**Fig. 120.1** Circles WPT system



way will be the full-wave electromagnetic simulation method. Finally, in this work, MoM is introduced to simulate the WPT in EV application.

### 120.2.2 Method of Moments

MoM is deduced from the time-harmonic Maxwell’s equations and the boundary conditions, which is thought as the most accurate and efficient method in electromagnetic simulation method. Usually, MoM can be classified as several kinds of forms for the different applied circumstance and integral function core. Given the WPT is usually constructed as PEC, the electric field surface integral equation (EFIE) is applied in this simulation. According to the boundary integral equations, the EFIE is expressed as

$$(j\omega\mathbf{A} + \nabla\Psi)_{\tan} = \mathbf{E}_{\tan}^{\text{sou}}, \quad r \in S \tag{120.1}$$

where  $\mathbf{A}$  is the magnetic vector potential,  $\Psi$  is the electronic scalar potential,  $\omega$  is the angular frequency, and  $\mathbf{E}_{\tan}^{\text{sou}}$  is the imposed voltage source on port 1.  $\mathbf{A}$  and  $\Psi$  can be expressed as

$$\mathbf{A}(\mathbf{r}) = \mu_0 \int_S \mathbf{J}(\mathbf{r}') \frac{e^{-jk|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} ds' \tag{120.2}$$

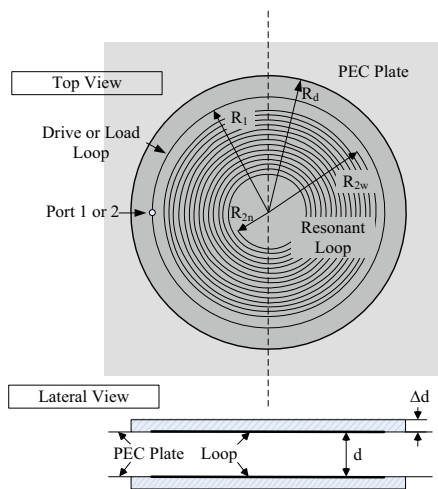
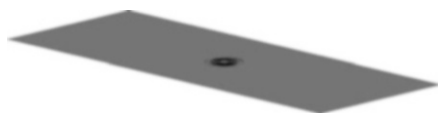
$$\Psi(\mathbf{r}) = \frac{1}{\epsilon_0} \int_S \sigma(\mathbf{r}') \frac{e^{-jk|\mathbf{r}-\mathbf{r}'|}}{4\pi|\mathbf{r}-\mathbf{r}'|} ds' \tag{120.3}$$

where  $\mathbf{J}(\mathbf{r}')$  and  $\sigma(\mathbf{r}')$  are the induced current and charge, respectively. Usually,  $\mathbf{J}(\mathbf{r}')$  and  $\sigma(\mathbf{r}')$  can be contacted with the charge consistent law as  $\mathbf{J} + j\omega\sigma = 0$ . Finally,  $\mathbf{J}(\mathbf{r}')$  is the wanted solution which can be solved by the matrix method. The readers can find more details about MoM in reference [5]. Furthermore, the application property of MOM in the designing of WPT is proved in reference [7].

### 120.3 The Designed New WPT System

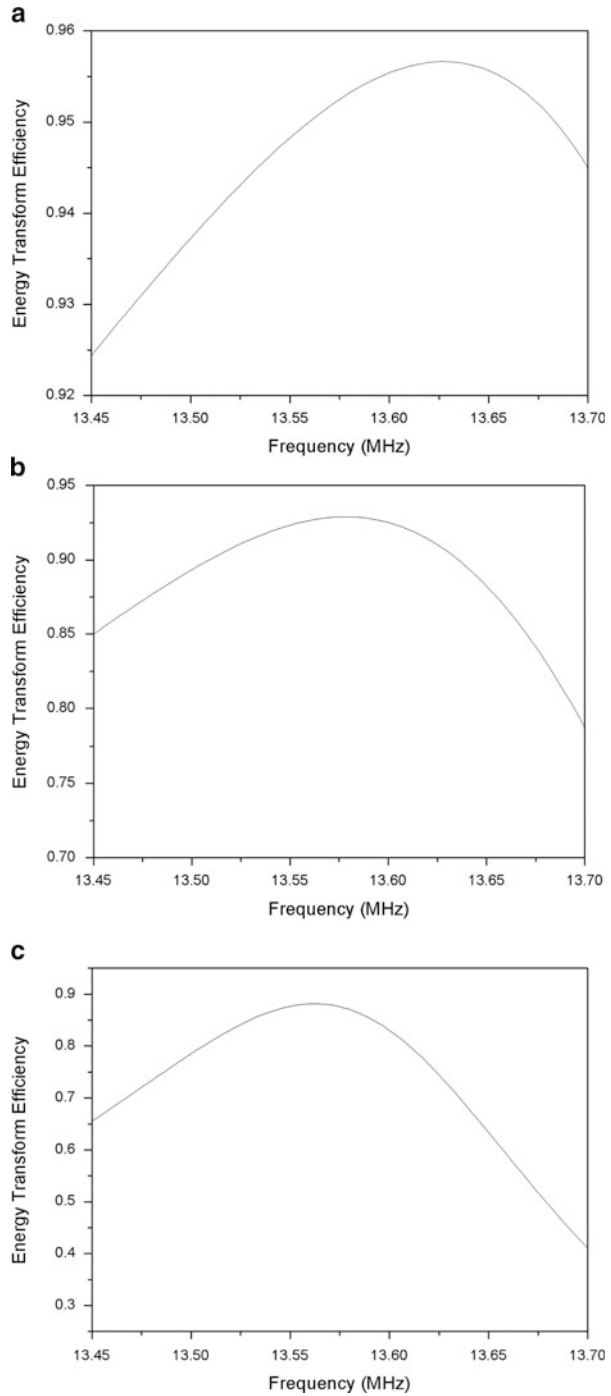
The purpose of this work is introducing MoM to design a new WPT system in EV application, and the key point is that the design should be integrated with circumstance in which the influence of the background should be considered. As this purpose, a typical model of the PWT is shown in Fig. 120.2. Here, in order to express clearly, only the transmit port on the ground is shown in Fig. 120.2, where the ground is modeled as the PEC surface will be easily achieved in the engineering. It is noted that the receive port on car's chassis is symmetrical with the transmit port in our design. In our design, the size of rectangular PEC ground is set to  $4.6 \times 1.8$  m, which is the typical size of a small car's chassis. We proposed that the WPT should be mounted in the backed cavity which is in the middle of the PEC plate. In order to display more clearly, the WPT with backed cavity is enlarged in Fig. 120.3.

**Fig. 120.2** The structure of the WPT in EV application (transmit port)



**Fig. 120.3** The detailed structure of WPT with backed cavity

**Fig. 120.4** The energy transmission efficiency versus frequency and d. (a)  $d = 15$  cm; (b)  $d = 20$  cm; (c)  $d = 25$  cm



This WPT is constructed as drive loop, load loop, and resonant loop. All of the loops are mounted in the backed cavity. The drive loop and resonant loop are mounted in the same plane and the same thing is the load loop and resonant loop in Rx. Given the versatility, the distance  $d$  between Tx and Rx should vary by 15–25 cm, which is the distance of most small car's chassis to the ground. Therefore, the design should make sure that the resonant frequency is fixed at 13.56 MHz and the efficiency is higher than 85 % in the variable range. Therefore, the appropriate parameter should be set to satisfy this target. In this work, the genetic algorithm (GA) [8] combined with MoM is applied in parameter optimization. Given the limited space of this literature, GA will not be detailed here. Finally, the optimized parameter is  $R_1 = 16.2$  cm,  $d_1 = 1$  mm,  $R_{2n} = 5$  cm,  $R_{2w} = 15$  cm,  $d_2 = 3$  mm,  $R_d = 20$  cm,  $\Delta d = 1.5$  cm, and  $n = 14.515$  ( $n$  is the number of turns of the resonant loop). Finally, with the optimized parameters, the simulated efficiency with different  $d$  is shown in Fig. 120.4. It is obvious that the resonant frequency is fixed at 13.56 MHz and the efficiency is higher than 85 %, which achieves the designing requirements.

## 120.4 Conclusion

In this work, the MoM was introduced to design the WPT system in EV application. This new structure consisted of four loops which are mounted in the backed cavity on a large PEC plate, and the PEC plate was used to offset the effect of the ground and chassis. Furthermore, the structure was improved by setting the drive loop around the resonant loop to decrease the sensitivity of the resonant frequency within the distance. Finally, the MoM and GA were combined to optimize the related parameters. The optimized results showed that the designed structure can fix the resonant frequency at 13.56 MHz and the energy transmission efficiency can be kept higher than 85 % when the distance varied within 15–25 cm.

## References

1. PowerMat Inc. (2009). Retrieved from <http://www.powermat.com>.
2. McSpadden, J., & Mankins, J. (2002). Space solar power programs and microwave wireless power transmission technology. *IEEE Microwave Magazine*, 3(4), 46–57.
3. Kurs, A., Karalis, A., Moffatt, R., Joannopoulos, J. D., Fisher, P., & Soljacic, M. (2007). Wireless power transfer via strongly coupled magnetic resonances. *Science*, 317(18), 83–91.
4. Tan, L.-L., Huang, X.-L., Huang, H., Zou, Y.-W., & Li, H. (2011). Transfer efficiency optimal control of magnetic resonance coupled system of wireless power transfer based on frequency control. *SCIENCE CHINA, Technological Sciences*, 54(6), 1428–1434.
5. Harrington, R. F. (1968). *Field computation by moment methods* (pp. 76–124). Malabar, FL: R. E. Krieger.
6. Rao, S. M., Wilton, D. R., & Glisson, A. W. (1982). Electromagnetic scattering by surfaces of arbitrary shape. *IEEE Transactions on Antennas Propagation*, 30(2), 409–418.

7. Moshfegh, J., Shahabadi, M., & Rashed-Mohassel, J. (2011). Conditions of maximum efficiency for wireless power transfer between two helical wires. *IET Microwaves, Antennas & Propagation*, 5(5), 545–550.
8. Dong, Y.-F., Gu, J.-H., Li, N.-N., Hou, X. D., & Yan, W. L. (2007, August 19–22). [Combination of genetic algorithm and ant colony algorithm for distribution network planning](#). *2007 International Conference on Machine Learning and Cybernetics* (Vol. 2, pp. 999–1002), Hong Kong.

# Chapter 121

## A Hierarchical Local-Interconnection Structure for Reconfigurable Processing Unit

Yujia Zou, Leibo Liu, Shouyi Yin, Min Zhu, and Shaojun Wei

**Abstract** Reconfigurable computing is being widely used in Computation-intensive applications. With the rapid development of applications, we have higher requirements for the computational efficiency of reconfigurable computing. In order to improve the computational efficiency, the array size gradually increased for applications that are more complex. With the upgrade of the array size, the hardware overhead of traditional interconnection structure used for reconfigurable processing unit (RPU) increases significantly. This paper proposed a new interconnection structure called hierarchical local interconnection for RPU. Comparing to traditional full-mesh structure used in MorphoSys, the hierarchical local interconnection greatly enhanced the area efficiency while retaining the flexibility of interconnection. When the array scale is  $8 \times 8$ , hardware overhead of new structure is 28.6 % of the traditional structure.

### 121.1 Introduction

In the past 2 decades, the reconfigurable computing technology developed rapidly. Now it is more and more popular in computation-intensive applications, such as media processing [1] and communication [2]. In order to handle the more complex applications, the structure and size of reconfigurable processor is becoming more and more complex. With the improvement of the array size, the proportion of interconnection structure is gradually increasing and the interconnection structure is also changing gradually.

---

Y. Zou • L. Liu (✉) • S. Yin • M. Zhu • S. Wei  
Research Centre for Mobile Computing, Tsinghua University, Beijing 100084, China  
Institute of Microelectronics, Tsinghua University, Beijing 100084, China  
Tsinghua National Laboratory for Information Science and Technology, Beijing 100084, China  
e-mail: [liulb@tsinghua.edu.cn](mailto:liulb@tsinghua.edu.cn)



This paper proposes a method with which we can maintain the area efficiency of the interconnection, increasing the array scale and computing capability. When the array scale is  $8 \times 8$ , the ratio of processing unit hardware overhead and interconnection hardware overhead in our structure is 3.22 compared to 0.536 in MorphoSys. When the array scale increases to  $128 \times 128$ , the ratio in our structure can maintain 0.345 compared to 0.002 in MorphoSys.

## 121.2 Related Work

The RaPiD [3] of Washington University and the Garp [4] of Berkeley used a one-dimensional interconnection. This structure is simple for wiring and widely used for early reconfigurable processor. The MorphoSys [5] of University of California Irvine and the ADRES [6] of IMEC used a traditional full-mesh structure. The full-mesh structure is more flexible than the one-dimensional interconnection. Good flexibility is conducive for mapping algorithm on the array. However, with the expansion of the array scale, traditional full-mesh structure will bring more hardware overhead. With the rapid development of applications, larger array is required to enhance the computing capacity. If the traditional full mesh structure is used in a large reconfigurable architecture, the interconnection hardware overhead will be significantly enhanced. The area efficiency of the array will drop sharply. Therefore, we must develop a new interconnection structure for large-scale reconfigurable processor, which is flexible enough for algorithmic mapping and with less hardware overhead.

REMUS [7] is a reconfigurable processor mainly used for multimedia system, which can be dynamically configured as the H.264, AVS, or MPEG2 video decoder, as well as other multimedia applications.

The architecture of REMUS processor is shown in Fig. 121.1a, which consists of an ARM, two reconfigurable processing units (RPUs), an entropy decoder (EnD), and some assistant modules. The ARM is a RISC CPU with two tightly coupled memories to accelerate the specific loaded codes and is used as the host processor mainly to handle control application and generate the configuration information for RPU.

RPU is a coarse-grained dynamic reconfigurable computing module and is the core element for computation-intensive task in the system (Fig. 121.1b). Each RPU consists of four units of reconfigurable computing array (RCA). Each RCA consists of an  $8 \times 8$  PE array (Fig. 121.2), an internal memory for the temporarily storage, and some control modules. In RCA, reading external data may require multiple cycles, or when the RCA is processing, the input and output data need cache. We design input FIFO and output FIFO for data cache. The depth of FIFO is designed by simulation results of the algorithm mapping.

Fig. 121.1 (a) REMUS architecture, (b) architecture of RPU

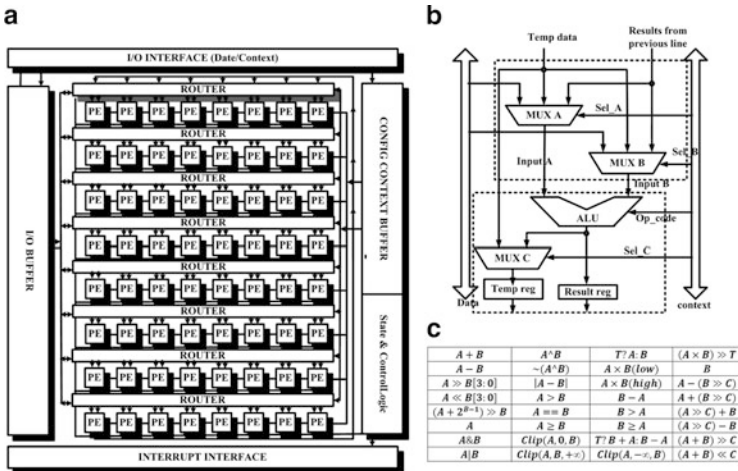
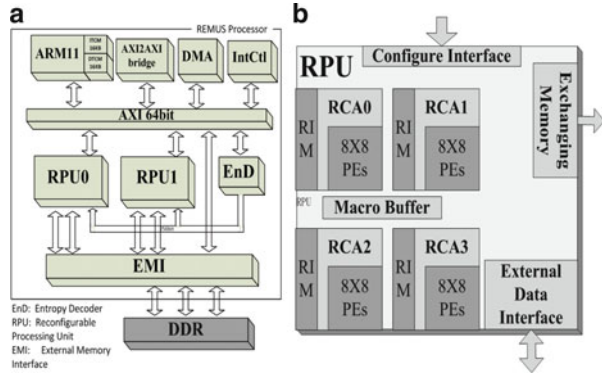


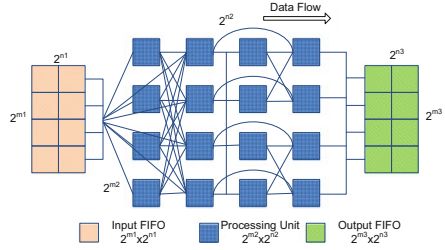
Fig. 121.2 (a) PE 8 × 8 arrays, (b) architecture of PE, and (c) PE functions

### 121.3 Analysis of Interconnect Structure

#### 121.3.1 Traditional Full-Mesh Structure

The traditional full-mesh structure has a great deal of flexibility, which is adopted in MorphoSys [5]. The traditional full-mesh structure contains three parts: (1) the interconnection between input FIFO to processing unit, (2) the interconnection between processing units, and (3) the interconnection between processing unit to output FIFO. The basic traditional full-mesh structure is shown in Fig. 121.3.

**Fig. 121.3** Traditional full-mesh structure



We assume that the width of the input FIFO is  $2^{m1}$  data, the depth of data required for a single operation is  $2^{n1}$  rows. The data required by a task may be larger than the width of input FIFO. So we suppose that the scale of the arrays is  $2^{m2} \times 2^{n2}$ , and the direction of data flow is shown as Fig. 121.2. The scale of output FIFO is  $2^{m3} \times 2^{n3}$ . The processing unit may obtain the data from any layer of the input FIFO. The output data from the processing unit may be transferred to any layer of the output FIFO.

In order to illustrate the interconnection of the hardware scale more clearly, we give a simple example: there are two rows of data interconnected by the full-mesh interconnection structure, the first row is the source with  $2^m$  data and the second row is the target with  $2^n$  data. Firstly, the MUX overhead of transfer one from  $2^m$  data to target address is  $2^{m-1} + 2^{m-2} + \dots + 2^0 = 2^m - 1$ . Then repeat the operation  $(2^n - 1)$  times, making  $2^n$  data transfer from the source to the target. By multiplication principle, we can derive that all the MUX overhead is  $(2^m - 1) * 2^n$ .

Then we calculate the hardware scale of traditional full-mesh structure in three parts:

1. The interconnection between input FIFO and processing unit:

$$(2^{m1} \times 2^{n1} - 1)(2^{m2} \times 2^{n2}) \tag{121.1}$$

2. The interconnection between processing units:

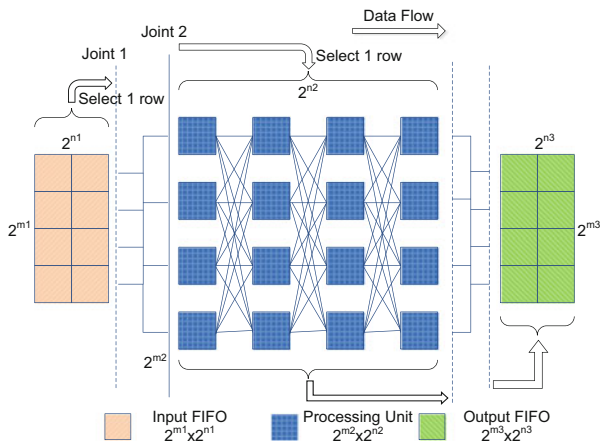
$$\begin{aligned} &(2^{m2} - 1)2^{m2}(2^{n2} - 1 + 1)/2 \times (2^{n2} - 1) \\ &= (2^{m2} - 1)(2^{n2} - 1) \times 2^{m2} \times 2^{n2-1} \end{aligned} \tag{121.2}$$

3. The interconnection between processing unit and output FIFO:

$$(2^{m2} \times 2^{n2} - 1)(2^{m3} \times 2^{n3}) \tag{121.3}$$

We assume that m and n increase in equal proportion, so the hardware scale is in direct proportion to  $2^{m*}2^n$ .

**Fig. 121.4** Hierarchical local-interconnection structure



### 121.3.2 New Local-Connection Structure

Although the traditional full-mesh structure has such a great deal of flexibility, the hardware overhead and configuration information of the traditional full-mesh structure is too large for reconfigurable array. The huge hardware overhead of traditional full-mesh structure leads to the number of RPU's decrease in the same area of a chip. It may not meet the scale of processing units required by the application.

Under the premise of ensuring the flexibility of reconfigurable array as much as possible, we develop a hierarchical local-interconnection structure for reconfigurable array. Generally, the adjacent data in DFG are stored in adjacent memory. The input data position of processing unit array can adjust by passing through the idle processing unit in array. By using the localization of data in the algorithm, hierarchical local interconnection meets the requirement of the algorithm mapping.

The structure of our hierarchical local interconnection is shown in Fig. 121.4. In the hierarchical local interconnection, one row is selected from  $2^{n1}$  rows of the input FIFO connected to Joint 1, the structure between Joint 1 and Joint 2 is full mesh, and one row is selected from  $2^{n2}$  rows of processing unit connected to Joint 2, thus completing the interconnection between input FIFO and processing unit array. In the same manner, we complete the interconnection between the processing unit array and output FIFO. In processing unit array, only the two adjacent rows are full-mesh structure.

The calculation of hardware overhead of the hierarchical local-interconnection structure is a little different from 3-D traditional full-mesh structure. Firstly, selecting one row with  $2^{m1}$  data from  $2^{n1}$  rows needs  $(2^{n1} - 1) * 2^{m1}$  MUX. Secondly, achieving full mesh between  $2^{m1}$  and  $2^{m2}$  needs  $(2^{m1} - 1) * 2^{m2}$  MUX; finally, selecting one row with  $2^{m2}$  from  $2^{n1}$  rows needs  $(2^{n1} - 1) * 2^{m2}$  MUX. So the hardware overhead of hierarchical local-interconnection structure is in three parts:

1. The interconnection between input FIFO and processing unit:

$$(2^{n_1} - 1)2^{m_1} + (2^{m_1} - 1)2^{m_2} + (2^{n_2} - 1)2^{m_2} \quad (121.4)$$

2. The interconnection between processing units:

$$(2^{m_2} - 1)2^{m_2}(2^{n_2} - 1) \quad (121.5)$$

3. The interconnection between processing unit and output FIFO:

$$(2^{n_2} - 1)2^{m_2} + (2^{m_2} - 1)2^{m_3} + (2^{n_3} - 1)2^{m_3} \quad (121.6)$$

Then, we will compare the traditional full-mesh structure to hierarchical local-interconnection structure.

### 121.3.3 Area Optimization

In the section of area optimization, we compare the three parts of interconnection between two structures. Firstly, we compare the interconnection between input FIFO and processing unit in two structures. In order to compare the area overhead of two structures, we use Eq. 121.1/Eq. 121.4  $\cong 2^{m_1+n_1+m_2+n_2}/(2^{m_1+n_1} + 2^{m_1+m_2} + 2^{n_2+m_2})$ ; under normal circumstances,  $n_1$  is smaller (because of the width of the input FIFO is larger, the number of rows in FIFO is less in one circulation); we derived the Eq. 121.1/Eq. 121.4  $\cong 2^{n_1} \times 2^{m_1+n_2}/(2^{m_1} + 2^{n_2})$ . Secondly, we compare the interconnection between processing units in two structures. We derived Eq. 121.2/Eq. 121.5 =  $2^{n_2-1}$ . In the same way, we could derive Eq. 121.3/Eq. 121.6 =  $2^{n_3} \frac{2^{m_3+n_2}}{2^{m_3}+2^{n_2}}$ . When the scale of array is  $16 \times 16$ , that is,  $n_2 = 4$ , the hardware overhead of hierarchical local interconnection is 1/8 of the traditional full-mesh structure. The hierarchical local-interconnection structure greatly reduce the hardware overhead of array.

### 121.3.4 Configuration Optimization

Since each MUX needs to be configured, the configuration of two interconnection structures is different. The configuration of input terminal, output terminal, and processing units is, respectively,  $(m_1 + n_1) \cdot 2^{m_2} \cdot 2^{n_2}$ ,  $(m_3 + n_3) \cdot 2^{m_2} \cdot 2^{n_2}$ , and  $(m_2 + n_2) \cdot 2^{m_2} \cdot 2^{n_2}$  in traditional full-mesh structure. The configuration of input terminal, output terminal, and processing units is, respectively,  $n_1 \cdot 2^{n_2} + m_1 \cdot 2^{m_2}$ ,  $n_3 \cdot 2^{n_2} + m_3 \cdot 2^{m_2}$  and  $m_2 \cdot 2^{m_2} \cdot 2^{n_2}$  in hierarchical local-interconnection structure.

Comparing the configuration of two structures, we could derive the equation as follows:

1. The input terminal:

$$\frac{(m1 + n1)2^{m2+n2}}{n1 \times 2^{n2} + m1 \times 2^{m2}} \approx 2^{n2} \quad (121.7)$$

2. The output terminal:

$$\frac{(m3 + n3)2^{m2+n2}}{n3 \times 2^{n2} + m3 \times 2^{m2}} \approx 2^{n2} \quad (121.8)$$

3. The array:

$$\frac{(m2 + n2)2^{m2+n2}}{m2 \times 2^{m2+n2}} = \frac{m2 + n2}{m2} \quad (121.9)$$

When the array size is  $16 \times 16$ ,  $m2 = n2 = 4$ , the configuration of input terminal and output terminal in hierarchical local interconnection is 1/16 of full-mesh structure, and the configuration of processing units reduces to 50 %.

## 121.4 Experimental Results and Analysis

The relationship between the scale of array and area in traditional full-mesh structure and hierarchical local-interconnection structure is shown, respectively, in Tables 121.1 and 121.2. Based on the architecture of MorphoSys and our structure, we extended the array scale. Then, we synthesize the circuit by Synopsys Design Compiler under TSMC65nmLP processing technique and get the data in Tables 121.1 and 121.2.

Comparing the data in Tables 121.1 and 121.2, the hardware overhead of hierarchical local interconnection is 28.6 % of the traditional structure when the array scale is  $8 \times 8$ . The hardware overhead of interconnection in traditional full-mesh structure is double of the processing unit. When the array scale expanded to  $32 \times 32$ , the hardware overhead of interconnection in traditional full-mesh structure is about 32 times of the processing unit. That area efficiency is too low to be used. On the other hand, even if the array scale expanded to  $128 \times 128$ , the hardware overhead of hierarchical local interconnection is only triple of the processing unit. The area efficiency is still reasonable.

Figures 121.5 and 121.6 show the relationship between area percentage and array size. With the growth of the array size, the area percentage of interconnection in MorphoSys increases rapidly to more than 80 %, which makes the area efficiency

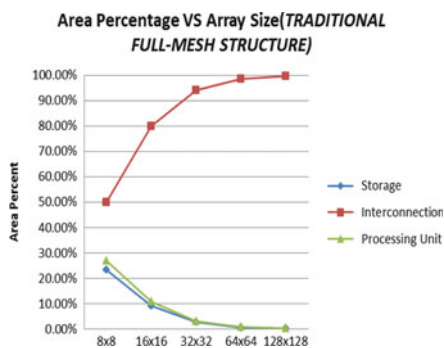
**Table 121.1** Full-mesh area percentage vs. array size

Array scale	Storage	Interconnection	Processing unit
$8 \times 8$	249974 $\mu\text{m}^2$ 23.20 %	540160 $\mu\text{m}^2$ 50.00 %	289456 $\mu\text{m}^2$ 26.80 %
$16 \times 16$	9.30 %	80.00 %	10.70 %
$32 \times 32$	2.70 %	94.10 %	3.20 %
$64 \times 64$	0.70 %	98.50 %	0.80 %
$128 \times 128$	0.20 %	99.60 %	0.20 %

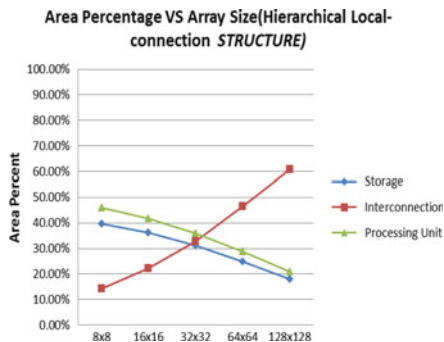
**Table 121.2** Hierarchical local-interconnection area percentage vs. array size

Array Scale	Storage	Interconnection	Processing unit
$8 \times 8$	249974 $\mu\text{m}^2$ 39.70 %	90000 $\mu\text{m}^2$ 14.30 %	289456 $\mu\text{m}^2$ 46.00 %
$16 \times 16$	36.10 %	22.10 %	41.80 %
$32 \times 32$	31.10 %	32.90 %	36.00 %
$64 \times 64$	24.80 %	46.40 %	28.80 %
$128 \times 128$	18.10 %	60.90 %	21.00 %

**Fig. 121.5** Area percentage vs. array size (traditional full-mesh structure)



**Fig. 121.6** Area percentage vs. array size (hierarchical local-interconnection structure)



of array reduce drastically. On the contrary, the area percentage of interconnection increases linearly with the array size; even if the array size expands to  $128 \times 128$ , the area percentage of our interconnection structure is 60 %.

## 121.5 Conclusion

In this paper, we proposed a novel interconnection structure called hierarchical local interconnection for RPU. Comparing to traditional full-mesh structure, which is adopted in MorphoSys, the hierarchical local interconnection greatly enhanced the area efficiency while retaining the flexibility of interconnection. When the array scale is  $8 \times 8$ , hardware overhead of new structure is 28.6 % of the traditional structure. Even if the array size expands to  $128 \times 128$ , hierarchical local-interconnection structure can still make the array maintain reasonable area efficiency.

**Acknowledgements** This work is supported in part by the China National High Technologies Research Program (No. 2012AA012701), the Tsinghua Information S&T National Lab Creative Team Project, the International S&T Cooperation Project of China grant (No. 2012DFA11170), the Tsinghua Indigenous Research Project (No. 20111080997), the Special Scientific Research Funds for Commonweal Section (No. 200903010), the Science and Technology Project of Jiangxi Province (No. 20112BBF60050) and the NNSF of China grant (No. 61274131).

## References

1. Veredas, F. J., Scheppler, M., Moffat, W., & Mei, B. (2005). Custom implementation of the coarse-grained reconfigurable ADRES architecture for multimedia purposes. In *International Conference on Field Programmable Logic and Applications, 2005, IEEE* (pp. 106–111).
2. Ebeling, C., Fisher, C., Xing, G., Shen, M., & Liu, H. (2004). Implementing an OFDM receiver on the RaPiD reconfigurable architecture. *IEEE Transactions on Computers*, 53(11), 1436–1448.
3. Ebeling, C. Cronquist, D. and Franklin P. (1997), Configurable Computing: The Catalyst for High-Performance Architectures, ” Proc. IEEE Int’l Conf. Application-Specific Systems, Architectures, and Processors, pp. 364–372
4. Hauser, J. R., & Wawrzynek, J. (1997). Garp: A MIPS processor with a reconfigurable coprocessor. In *Proceedings of the 5th Annual IEEE Symposium on FPGAs for Custom Computing Machines, IEEE* (pp. 12–21).
5. Singh, H., Lee, M. H., Lu, G., Kurdahi, F. J., Bagherzadeh, N., & Filho, E. M. C. (2000). MorphoSys: An integrated reconfigurable system for data-parallel and computation-intensive applications. *IEEE Transactions on Computers*, 49(5), 465–481.
6. Novo, D., Moffat, W., Derudder, V., & Bougard, B. (2005). Mapping a multiple antenna SDM-OFDM receiver on the ADRES coarse-grained reconfigurable processor. In *IEEE Workshop on Signal Processing Systems Design and Implementation, 2005, IEEE* (pp. 473–478).
7. Zhu, M., Liu, L., Yin, S., & Wang, Y. (2010). A reconfigurable multi-processor SoC for media applications. In *Proceedings of 2010 IEEE International Symposium on Circuits and Systems (ISCAS), IEEE* (pp. 2011–2014).



# Chapter 122

## High Impedance Fault Location in Distribution System Based on Nonlinear Frequency Analysis

Jinqian Zhai, Di Su, Wenjian Li, Feng Li, and Guohong Zhang

**Abstract** A methodology is presented to detect and locate high impedance faults (HIFs) in radial distribution system by means of nonlinear frequency analysis. The proposed technique is based on the analysis of the feeder responses to power line carrier signals, which are periodically injected at the outlet of transformer. The effectiveness of the method has been verified through simulation studies. The results demonstrated that the proposed method has the potential to be applied in practice to resolve HIF real-time monitoring problem.

### 122.1 Introduction

As to power line fault location, most of the research aims at finding the positions of transmission line faults; the locations of faults on distribution systems have started receiving much attention [1]. The detection and location algorithms in power transmission systems are not useful in power distribution systems. Fault location in distribution system is much more difficult than in a transmission network. Due to the complex topology of downstream network, the calculated value is far from being as accurate as needed for a fast and reliable service restoration; so any method that helps to locate faults, as soon as possible, is welcomed. For fault such as short circuit with low impedance fault, several techniques have been proposed by many authors [2–4]. But there is no appropriate solution to high impedance fault (HIF) in distribution system [5–7].

A great number of HIFs can be detected today, but cannot be localized due to lack of communication among distribution feeder sections. HIF are inherently nonlinear and always result in distorted currents. According to nonlinear characteristics of HIF, nonlinear frequency analysis method [8] is proposed to detect the

---

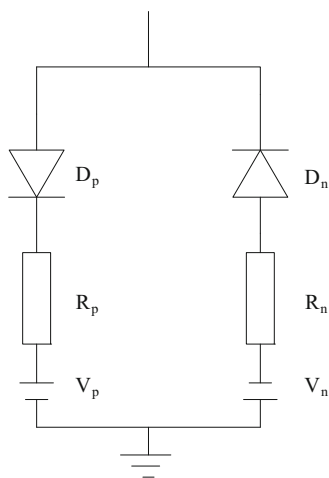
J. Zhai (✉) • D. Su • W. Li • F. Li • G. Zhang  
ZhengZhou Power Supply Company, ZhengZhou 450000, China  
e-mail: [jinqianzhai@163.com](mailto:jinqianzhai@163.com)

position of HIF in distribution system without turning off the electric power. In outlet side of substation, high-frequency signal is injected onto power line through coupling equipment; this high-frequency signal flows through power line, which was received by receiving device installed in tower. By wireless network, the device sends the received signal to control center. In monitoring center, by analyzing all receiving signals, using nonlinear frequency analysis method, the position of HIF in distribution system is found. The purpose of this work is to evaluate the applicability of nonlinear frequency analysis to HIF location in distribution system.

## 122.2 Modeling of HIF

The electric arc has a voltage/current nonlinear relation and might show an asymmetric behavior of the positive half cycle with respect to the negative one. The nonlinear HIF model is shown in Fig. 122.1 [9].

The model includes two dc sources:  $V_p$  and  $V_n$  which present the arcing voltage between the trees and line [10]. During the positive half cycle, the current flows through  $V_p$  and during the negative, through  $V_n$ . When the phase voltage is greater than the positive DC voltage  $V_p$ , the fault current flows toward the ground. The fault current reverses when the line voltage is less than the negative DC voltage  $V_n$ . The values in the phase voltage between  $V_n$  and  $V_p$  have no fault current flows. Two resistances— $R_p$  and  $R_n$ —present the resistance of trees and/or earth resistance.



**Fig. 122.1** Simplified two-diode fault model of HIFs

## 122.3 Nonlinear Frequency Analysis Approach for Locating HIF

### 122.3.1 Derivation of the Approach

For the power transmission system with HIF, the equivalent circuit of power transmission line is shown in Fig. 122.2 [11].

For system without HIF, from the current and voltage laws, the following equations are obtained:

$$\frac{u_{n-1}(s) - u_n(s)}{L_n s + R_n} = \frac{u_n(s)}{L_{n+1} s + R_{n+1} + Z_{load}} + (C_n s + G_n) u_n(s) \tag{122.1}$$

The expansion equation is shown below:

$$\begin{aligned} & C_n L_n L_{n+1} s^3 u_n(s) + [L_n C_n R_{n+1} + R_n C_n L_{n+1} + G_n L_n L_{n+1} + L_n C_n Z_{load}] s^2 u_n(s) \\ & + [R_n R_{n+1} C_n + L_n G_n R_{n+1} + R_n G_n L_{n+1} + L_{n+1} + L_n + (R_n C_n + L_n G_n) Z_{load}] s u_n(s) \\ & + [R_n + (1 + G_n R_n) R_{n+1} + (1 + G_n R_n) Z_{load}] u_n(s) \\ & - L_{n+1} s u_{n-1}(s) - (R_{n+1} + Z_{load}) u_{n-1}(s) = 0 \end{aligned} \tag{122.2}$$

So, the mathematical model of transmission line system can be written in time domain as

$$A\ddot{U}(t) + B\dot{U}(t) + CU(t) + DU(t) = U_s(t) \tag{122.3}$$

where  $A$ ,  $B$ ,  $C$ , and  $D$  are the system parameter matrices, respectively.  $U = (u_1, \dots, u_n)'$  is the voltage vector, and  $U_s = (L_2 \dot{u}_s + R_2 u_s, 0, \dots, 0)'$  is the external input vector acting on the system. Obviously, this system is a linear system.

Assume that HIF is located at  $J$ -th section with  $J \in \{2, \dots, n\}$ . There must be a change to the circuit parameter  $R$  at the  $J$ -th section of practical transmission line, that is,

$$\Delta R_J = R_{J-fault} - R_{J-normal}, \quad \Delta u = U(\Delta R_J), \quad \text{so } NU = \left( \overbrace{0, \dots, 0}^{J-2}, -\Delta u, \Delta u, \overbrace{0, \dots, 0}^{n-J} \right)' \tag{122.4}$$

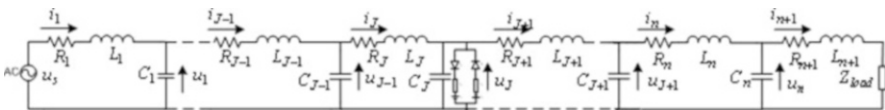


Fig. 122.2 The equivalent circuit of power line system with one high impedance fault

In this case, the power line system can be described as

$$A\ddot{U}(t) + B\dot{U}(t) + CU(t) + DU(t) = NU + U_s(t) \tag{122.5}$$

The system described by eq. (122.5) is a typical locally nonlinear multi-degree of freedom (MDOF) system [12]. By applying the nonlinearity detection approach in [12] to model (122.5), a nonlinear frequency analysis technique for locating HIF of transmission line is proposed. The proposed technique can be described as follows [12]:

1. Excite power line system separately using two different input voltages  $u_s^{(q)}(t)$ ,  $q = 1, 2$ , and measure the corresponding voltage response at each receiving site of power line to obtain  $u_i^{(q)}(t)$ ,  $q = 1, 2$ ,  $i = 1, \dots, n$ .
2. Calculate the FFT spectrum of  $u_s^{(q)}(t)$ ,  $q = 1, 2$  and  $u_i^{(q)}(t)$ ,  $q = 1, 2$ ,  $i = 1, \dots, n$  to produce  $U_s^{(q)}(j\omega)$ ,  $q = 1, 2$  and  $U_i^{(q)}(j\omega)$ ,  $q = 1, 2$ ,  $i = 1, \dots, n$ .
3. Evaluate the functions of  $E^{i,i+1}(j\omega)$ ,  $i = 1, \dots, n - 1$  from the results obtained in Step (2) as follows:

$$E^{i,i+1}(j\omega) = [1 \ 0] \begin{bmatrix} U_s^{(1)}(j\omega), U_{i+1}^{(1)}(j\omega) \\ U_s^{(2)}(j\omega), U_{i+1}^{(2)}(j\omega) \end{bmatrix}^{-1} \begin{bmatrix} U_i^{(1)}(j\omega) \\ U_i^{(2)}(j\omega) \end{bmatrix} \tag{122.6}$$

4. Evaluate  $\bar{E}^{i,i+1}$  for  $i = 1, \dots, n - 1$  as

$$\bar{E}^{i,i+1} = \frac{\int_{\omega_1}^{\omega_2} |E^{i,i+1}(j\omega)| d\omega}{\max_{i \in (1, \dots, n-1)} \left[ \int_{\omega_1}^{\omega_2} |E^{i,i+1}(j\omega)| d\omega \right]} \tag{122.7}$$

where  $[\omega_1, \omega_2]$  is a frequency band within the frequency range of the input spectrum  $U_s^{(q)}(j\omega)$ ,  $q = 1, 2$ .

5. Examine  $\bar{E}^{i,i+1}$  for  $i = 1, \dots, n - 1$ . If  $\bar{E}^{i,i+1}$  is found to change sharply in the J-th section, it can be concluded that HIF is located at the J-th section of power line system.

### 122.3.2 Implementation of the Approach in Distribution System Using Power Line Carrier Method

In order to test the feasibility of the proposed technique in distribution system, a relatively simple distribution system is selected; it is shown in Fig. 122.3. The proposed technique is based on the analysis of the feeder responses to high-frequency signals, which are periodically injected at the feeder inlet. The detection

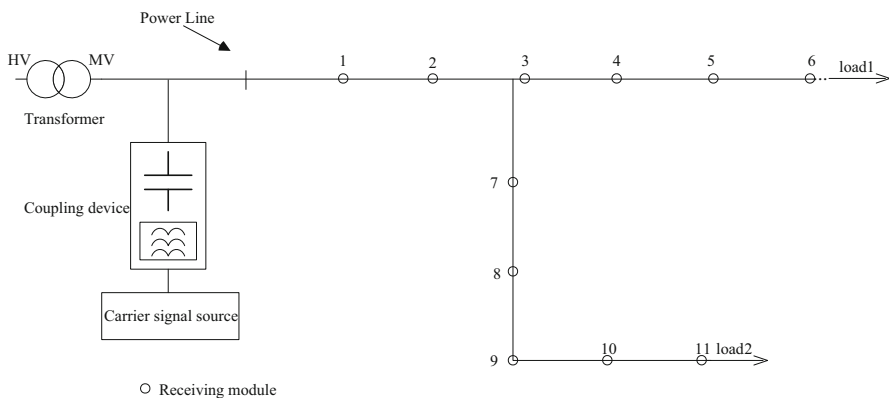


Fig. 122.3 The schematic of distribution line monitoring module

procedure requires exciting the distribution line networks twice using two sinusoidal high-frequency input signals. Several receiving devices are installed along with the feeder, which are used to receive the high-frequency signals and filter out the power frequency signals. The responses to two separate exciting high-frequency signals are used as the information collected in Step (1) of the technique. The procedure in distribution system is as follows:

1. Suppose the receiving module 1–6 is in the main feeder, the receiving module 7–11 is in the branch.
2. Judge the index value of  $\bar{E}^{i,i+1}$  and  $\bar{F}^{i,i+1}$ .

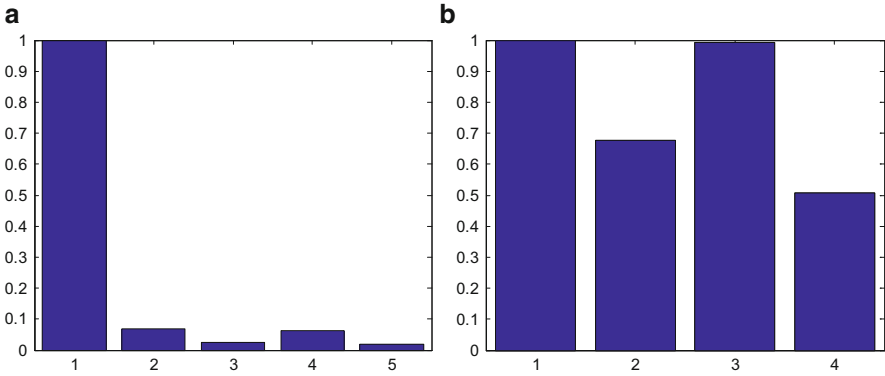
If  $\bar{E}^{i,i+1}$  is found to change sharply in the J-th section, and  $\bar{F}^{i,i+1}$  is found to change randomly, then there is HIF in the main feeder.

If  $\bar{E}^{i,i+1}$  is found to change sharply in the J-th section, and  $\bar{F}^{i,i+1}$  is found to change sharply in the K-th section, then that illustrates the branch is in the J-th section, and there is HIF in the K-th section of branch.

If  $\bar{E}^{i,i+1}$  and  $\bar{F}^{i,i+1}$  are both found to change randomly, there is no HIF in distribution network.

### 122.4 Simulation Studies

In order to verify the effectiveness of the proposed approach, simulation studies using the Matlab/Simulink facilities for distribution networks were conducted. 35 kV power line level was used for the simulation studies. Six inspecting devices were installed in the main feeder of the distribution network, e.g., receiving module 1–6; five inspecting devices were installed in the branch feeder, e.g., receiving module 7–11. To simplify analysis, only one HIF is supposed to exist in the distribution networks. For power line system parameters, each section of line length



**Fig. 122.4** Illustration of  $\bar{E}^{i,i+1}$  and  $\bar{F}^{i,i+1}$  in Table 122.1(a) is illustration of  $\bar{E}^{i,i+1}$ , (b) is that of  $\bar{F}^{i,i+1}$

is 1 km, line type is ACSR LGJ-70. Simulation study is used in the configuration system shown in Fig. 122.4. The used HIF model parameters are as follows:  $R_p = 1250\Omega$ ,  $R_n = 500\Omega$ ,  $V_p = 5000V$ ,  $V_n = 7000V$ . The load parameter is 15,000  $\Omega$ .

### 122.4.1 Case Study 1

When fault is in the main feeder between 1 and 2, from the five-step procedure, the proposed technique was implemented in the main feeder as follows:

1. Two sinusoidal voltage inputs,  $u_s^{(q)}(t) = \alpha_q \sin(\omega_u t)$ ,  $q = 1, 2$ , where  $\omega_u = 2\pi \times 50000$ ,  $\alpha_1 = 12$ ,  $\alpha_2 = 20$ , were applied to excite the system and to generate two sets of output responses on the distribution system.
2. Calculate the FFT spectrum of  $u_s^{(q)}(t)$ ,  $q = 1, 2$  and  $u_i^{(q)}(t)$ ,  $q = 1, 2$ ,  $i = 1, \dots, n$  to produce  $U_s^{(q)}(j\omega)$ ,  $q = 1, 2$  and  $U_i^{(q)}(j\omega)$ ,  $q = 1, 2$ ,  $i = 1, \dots, n$ .
3. Evaluate function of  $E^{i,i+1}(j\omega)$ ,  $i = 1, \dots, n - 1$  from the results obtained in Step (2).
4. Evaluate  $\bar{E}^{i,i+1}$  for  $i = 1, \dots, n - 1$ ; the results obtained are given in Table 122.1 and illustrated in Fig. 122.4a
5. From Table 122.1, it is shown that the index value of  $\bar{E}^{i,i+1}$  changes sharply from  $\bar{E}^{1,2}$  to  $\bar{E}^{2,3}$  from large to small. The index value of  $\bar{E}^{i,i+1}$  is very small after  $\bar{E}^{2,3}$ . So the HIF happened in the second section in the main feeder.

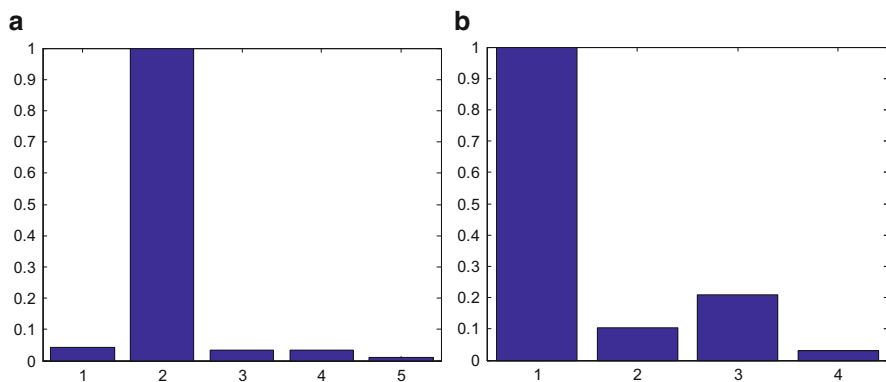
Again the above five-step procedure is used in the branch line. Here, the input voltage has changed, because the branch line is in the third section in the main

**Table 122.1**  $\bar{E}^{i,i+1}$  and  $\bar{F}^{i,i+1}$  evaluated for a case where HIF is located at the Sect. 122.2 of main feeder

$\bar{E}^{1,2}$	$\bar{E}^{2,3}$	$\bar{E}^{3,4}$	$\bar{E}^{4,5}$	$\bar{E}^{5,6}$
1.0000	0.0686	0.0235	0.0623	0.0201
$\bar{F}^{1,2}$	$\bar{F}^{2,3}$	$\bar{F}^{3,4}$	$\bar{F}^{4,5}$	
1.0000	0.6758	0.9930	0.5061	

**Table 122.2**  $\bar{E}^{i,i+1}$  and  $\bar{F}^{i,i+1}$  evaluated for a case where HIF is located at the Sect. 122.2 of branch line

$\bar{E}^{1,2}$	$\bar{E}^{2,3}$	$\bar{E}^{3,4}$	$\bar{E}^{4,5}$	$\bar{E}^{5,6}$
0.0411	1.0000	0.0332	0.0332	0.0091
$\bar{F}^{1,2}$	$\bar{F}^{2,3}$	$\bar{F}^{3,4}$	$\bar{F}^{4,5}$	
1.0000	0.1031	0.2080	0.0304	



**Fig. 122.5** Illustration of  $\bar{E}^{i,i+1}$  and  $\bar{F}^{i,i+1}$  in Table 122.2, (a) is illustration of  $\bar{E}^{i,i+1}$  and (b) is that of  $\bar{F}^{i,i+1}$

feeder, so  $u_3(t)$  is source signal of the branch line. The value of index  $\bar{F}^{i,i+1}$  is shown in Table 122.1 and illustrated in Fig. 122.4b; it is shown that the index value of  $\bar{F}^{i,i+1}$  changes randomly. According to the evaluation standard of Sect. 122.3.2, the HIF is found in the second section of main feeder, which is obviously correct.

### 122.4.2 Case Study 2

In this case, supposed HIF happened in the branch line. First, fault is in the branch line between 7 and 8. The five-step procedure is implemented in the main feeder and branch line, respectively. As the same case study 1, due to the branch line is in the Sect. 122.3 of the main feeder,  $u_3(t)$  is used to the signal source of branch line. The results obtained are given in Table 122.2 and illustrated in Fig. 122.5.

From Table 122.2 or Fig. 122.5a, it is shown that the index value of  $\bar{E}^{i,i+1}$  changes sharply from  $\bar{E}^{2,3}$  to  $\bar{E}^{3,4}$  from large to small; the index value is very small after  $\bar{E}^{3,4}$ , not associated with the index value before  $\bar{E}^{2,3}$ . From Table 122.2 or Fig. 122.5b, it is shown that the index value of  $\bar{F}^{i,i+1}$  change sharply from  $\bar{F}^{1,2}$  to  $\bar{F}^{2,3}$ , the index value is very small after  $\bar{F}^{2,3}$ . According to the topology of distribution line and the evaluation standard of Sect. 122.3.2, the conclusion is that the branch line is in the third section of the main feeder, and the HIF is located in the second section of the branch line, which is correct.

## 122.5 Conclusion

In this paper, the application of a nonlinear frequency analysis-based approach is proposed to detect and locate the HIF in distribution system. The power carrier signal technology has been suggested for the practical implementation of the proposed technique. Numerical simulation studies have been conducted. The results verified the effectiveness of the new technique and demonstrated the potential to apply the technique in practice to resolve the important HIF location problem. Further research will be focused on laboratory tests to make necessary preparations for future experimental studies on real distribution systems.

## References

1. Saha, M.M., Das, R., Verho, P., & Novosel, D. (2002). Review of fault location techniques for distribution systems[C]. *Power systems and communications infrastructures for the future* (pp 1–6), Beijing.
2. Choowong-Wattanasakpubal, & Teratum-Bunyagul. (2010). Algorithm for detecting, indentifying, locating and experience to develop the automate faults location in radial distribution system[J]. *JEET*, 5(1), 36–44.
3. Campoccia, A. Silvestre, M.L.D., Incontrera, I., Sanserverino, E.R., & Spoto, G. (2010). An efficient diagnostic technique for distribution systems based on under fault voltages and currents[J]. *Electric Power Systems Research*, 80(10), 1205–1214.
4. Seung-Jae Lee, Myeon-Song Choi, Sang-Hee Kang, Bo-Gun Jin, Duck-Su Lee, Bok-Shin Ahn, et al. (2004). An intelligent and efficient fault location and diagnosis scheme for radial distribution systems[J]. *IEEE Transactions on Power Delivery*, 19(2), 524–532.
5. Flauzino, R. A., Ziolkowski, V., Silva, I.N., de Souza, & D.M.B.S. (2009). Hybrid intelligent architecture for fault identification in power distribution systems[C]. *Power & Energy Society General Meeting. PES'09* (pp. 1–6). Calgary, AB.
6. Elkalashy, N.I., Lehtonen, M., Darwish, H.A., Taalab, A.M.I., Izzularab, & M.A. (2007). DWT-based extraction of residual currents throughout unearthed MV networks for detecting high-impedance faults due to leaning trees[J]. *ETEP*, 17(6), 597–614.
7. Borghetti, A., Corsi, S., Nucci, C.A., Paolone, M., Peretto, L., & Tinarelli, R. (2006). On the use of continuous-wavelet transform for fault location in distribution power systems[J]. *Electrical Power and Energy Systems*, 28(9), 608–617.



8. Lang, Z. Q., & Billings, S. A. (1996). Output frequency characteristics of non-linear system [J]. *International Journal of Control*, 64(16), 1049–1067.
9. Aboul-Zahab, E.M., Eldin, E.-S.T., Ibrahim, D.K., & Saleh, S.M. (2008). High impedance fault detection in mutually coupled double-ended transmission lines using high frequency disturbances[C]. *12th Middle-East Power System* (pp. 412–419). Aswan.
10. Ibrahim, D.K., El Sayed, T.E., El-Zahab, E.E.-D.A. & Saleh S.M. (2010). Unsynchronized fault-location scheme for nonlinear HIF in transmission lines[J]. *IEEE Transactions on Power Delivery*, 25(2), 631–637.
11. Lonngren, K. E., & Bai, E. W. (1996). Simulink simulation of transmission lines[J]. *IEEE Transactions on Circuits and Device Magazine.*, 12(3), 10–16.
12. Lang, Z.Q. & Peng, Z.K. (2008). A novel approach for nonlinearity detection in vibrating systems[J]. *Journal of Sound and Vibration*, 314(64), 603–615.

# Chapter 123

## Early Fault Detection of Distribution Network Based on High-Frequency Component of Residual Current

Jinqian Zhai, Di Su, Wenjian Li, Feng Li, and Guohong Zhang

**Abstract** A methodology is presented to detect incipient faults in distribution networks by means of DWT and energy detection algorithm. The proposed technique is to extract the characteristic of incipient fault by DWT method, that is, to extract the  $d_5$  coefficient of wavelet decomposition of residual current and residual voltage. Compare energy value with normal situation using an energy detection algorithm; incipient faults are detected. The proposed technique has been investigated by ATP/EMTP. Simulation results show that this technique is effective and robust, and the proposed method has the potential to be applied in practice to resolve incipient fault real-time monitoring problem.

### 123.1 Introduction

Incipient faults in power lines are normally characterized as the faulty phenomena with the relatively low fault currents, such as high impedance faults, insulator leakage current faults, and intermittent/transient faults [1]. Unlike low impedance short circuits, which involve relatively large fault currents and are readily detectable by conventional overcurrent protection, these faults represent little threat of damage to power system equipment [2]. But with time, they may lead to a catastrophic failure (i.e., a permanent damage beyond repair). So, early detection of power line faults would undoubtedly be a great benefit to the utilities, enabling them to avoid catastrophic failures, unscheduled outages, and thus loss of revenues.

Various methods have been proposed by researchers and protection engineers. Among them, harmonic analysis [3], randomness detection [4], artificial neural networks [5], Hilbert-Transform based [6], and wavelet transform [7–11] are used to extract the feature of incipient fault signals in distribution line. But, due to

---

J. Zhai (✉) • D. Su • W. Li • F. Li • G. Zhang  
Zhengzhou Power Supply Company, Zhengzhou 450000, China  
e-mail: [jinqianzhai@163.com](mailto:jinqianzhai@163.com)

high-time resolution and low-frequency resolution for high frequencies and high-frequency resolution and low-time resolution for low frequencies, wavelet transform can achieve a better solution. Recently, a power line condition monitoring system [12] is proposed; according to the project's goal, it is directed to develop a new power line sensornet, along with sensors to be scattered along the line, for prediction of incipient faults and momentary line contact (such as incipient failure of insulators and tree limb contact). However, due to the complexity of fault mechanism and environments, a proven approach for incipient failure detection in distribution system is not yet available and needs high maintenance cost.

The focus of the work reported in this paper is to develop an efficient online system that uses measured voltage and current values over a period to diagnose power line incipient faults. Because it is difficult to collect a large amount of incipient fault data, we propose a deterministic method, rather than a training-based intelligent system method. The proposed algorithm is as follows: The incipient fault feature is extracted by DWT. The energy value and average energy value is computed by using the detail  $d_5$  coefficients of the residual current signal. Comparing the energy value with the set value, incipient fault in distribution network can be detected.

## 123.2 Incipient Faults of Distribution Line

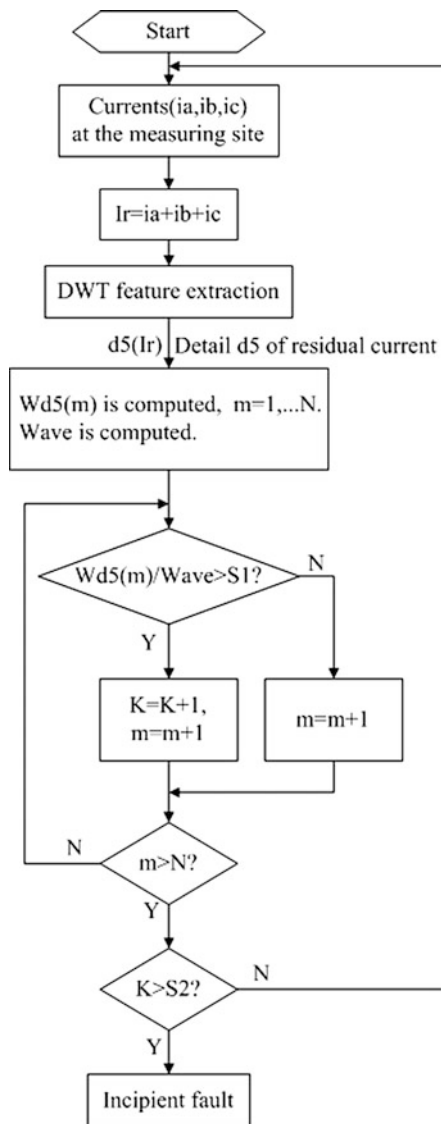
The term “incipient faults” refers to certain pre-fault “symptoms” or electrical activities taking place prior to a power system failure or blackout. Power line incipient faults are the primary causes of catastrophic failures in distribution network. For underground cable, these faults develop in the extruded cables from gradual deterioration of the solid insulation due to the persisting stress factors. For overhead lines, incipient faults are associated with degraded equipment (insulators, arresters, transformer insulation and bushings, etc.) and the gradual intrusion of tree limbs as they grow into the overhead power line.

Unlike short circuit faults, incipient faults in distribution networks do not draw sufficient currents from the line to trigger the protective devices. Incipient faults may present intermittent, asymmetric, and sporadic spikes, which are random in magnitude and could involve sporadic bursts as well, and exhibit complex, nonlinear, and dynamic characteristics [13].

## 123.3 Proposed Technique Principles

The detail coefficients in the high frequency are used to analyze incipient fault signal; the energy value above set value is used to characterize incipient fault signal in distribution network. The flow chart of incipient fault detection in distribution networks is shown in Fig. 123.1.

**Fig. 123.1** Flow chart of incipient fault based on DWT



At the measuring site, three phase currents are measured. The corresponding residual currents are computed and they are extracted using DWT. The sampling frequency of the raw signals is 100 kHz. In this algorithm, only the wavelet transform coefficients on scales 3 to scales 5 will be considered, because the incipient fault which have been detected has a frequency range from 2 to 10 kHz [14], which is verified by a large number of field experiments, and scales 3–5 correspond to a frequency range of 3.125–12.5 kHz,  $d_3$ ,  $d_4$ , and  $d_5$  including the frequency bands 25–12.5, 12.5–6.25, and 6.25–3.125 kHz are investigated, where

the sampling frequency is 100 kHz. It is obvious that detail d5 is included in the frequency band of incipient fault. So the detail d5 is selected as the research frequency band. The residual current detail d5 coefficient over one measuring period is divided into  $N$  section for the fault detection purpose. The sum over absolute value for every section is computed, which is described as  $W_{d5}(m)$ ,  $m = 1, \dots, N$  as

$$W_{d5}(m) = \sum_{n=(m-1)*p+1}^{m*p} |d5\_Ir(n)| \quad (123.1)$$

where  $W_{d5}(m)$  means the detector in the discrete samples according to  $d5\_Ir$ , which is the detail level d5 of the residual current with incipient fault,  $n$  is used for carry out a sliding window covering 5 ms,  $P$  is sample number in a window and  $N$  is a number of window samples.

The average value of the absolute sum over one measuring period is computed, which is shown as follows:

$$W_{ave} = \frac{1}{N * P} \sum_{i=1}^{N * P} |d5\_Ir(n)| \quad (123.2)$$

The ratio value for  $W_{d5}(m)$  to  $W_{ave}$  is obtained as

$$J = \frac{W_{d5}(m)}{W_{ave}} \quad (123.3)$$

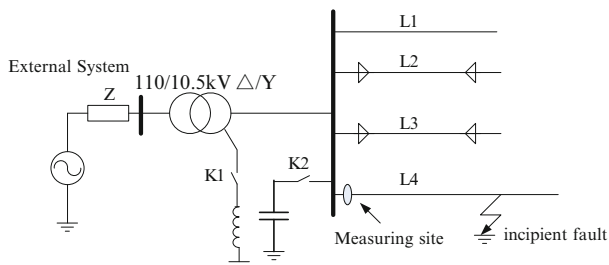
When the ratio  $J$  is above the setting value  $S1$ , the counter  $K$  is triggered once. During one measuring period, when the counter  $K$  is above the setting value  $S2$ , the incipient fault of distribution system is detected.

## 123.4 Simulated System

### 123.4.1 Test Power System

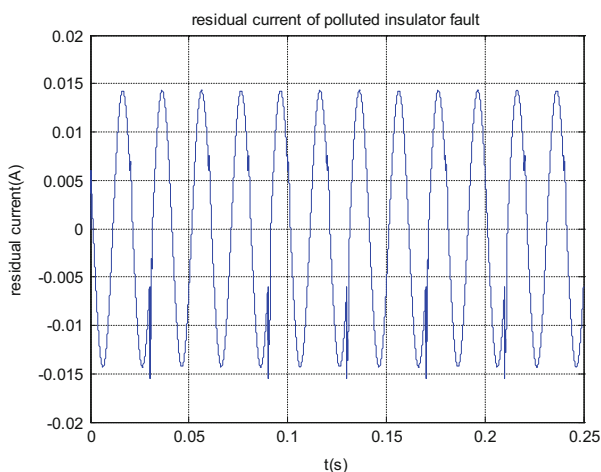
The 10-kV distribution system is supplied with power by a 110-kV grid via 40-MVA transformer as shown in Fig. 123.2; the resonant-earthed 10-kV system consists of several overhead lines and cables as radial feeders, 4-feeder distribution network simulated using ATP/EMTP, in which the processing is created by ATPDraw. L1 and L4 are overhead lines, 15 and 10 km, respectively. L2 and L3 are cables, 5 and 7 km, respectively. The feeder overhead line and cable are represented using the frequency-dependent JMarti model.

The polluted insulator fault is simulated according to the model of reference [15]; transient faults are generated by a fast electronic switch to simulate Peterson arc mechanism [6]. They are supposed at 5 km far away the measuring site.



**Fig. 123.2** Configuration of simulated system

**Fig. 123.3** Residual current of polluted insulator fault

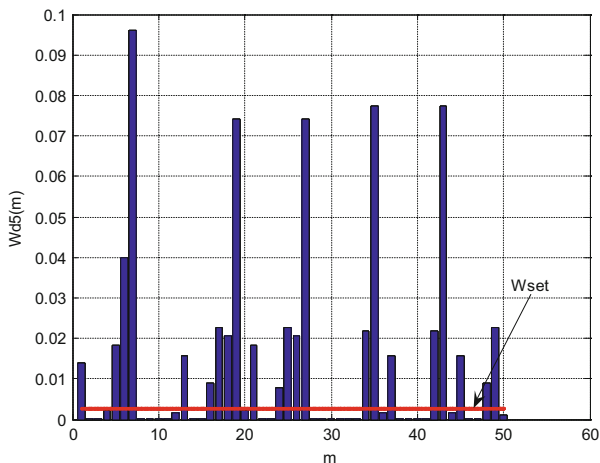


### 123.4.2 Simulation Results

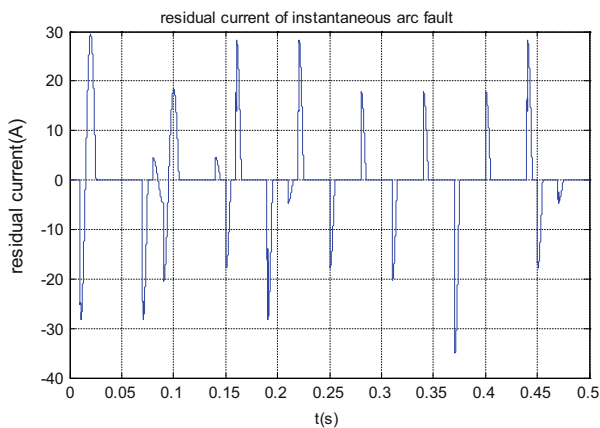
Here, simulated system is an isolated neutral system; polluted insulator fault signals were simulated with a sampling rate of 100 kHz for a duration of 0.25 s; instantaneous arc fault signals were simulated with a sampling rate of 100 kHz for a duration of 0.5 s. Herein, the setting values S1 and S2 are set to 100 and 15, respectively. Each value is set according to a large amount of simulation experiments;  $S1 < 100$  and  $S2 < 15$  account for little incipient fault, and we do not care for them.

The residual current for the fault case, which occurs on polluted insulator far away 5 km from the measuring site, is depicted in Fig. 123.3. After the discrete wavelet analysis, the details  $d_5$  are investigated. According to the proposed algorithm, the simulation time is divided into 50 time segments. The value of  $W_{d_5}(m)$  and  $W_{ave}$  are calculated, the value of  $W_{d_5}(m)$  and  $W_{ave}$  are calculated by Eq. 123.1

**Fig. 123.4** Detector  $W_{d5}(m)$  of residual current details shown in Fig. 123.3



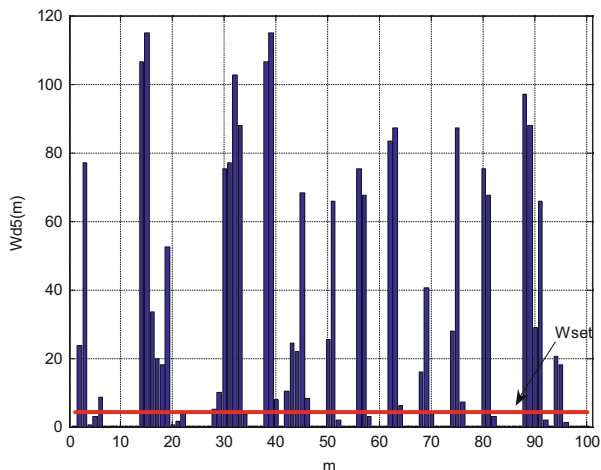
**Fig. 123.5** Residual current of instantaneous arc fault



and Eq. 123.2, respectively,  $W_{set}$  is obtained from  $W_{ave}$  multiply by the setting value S1, which is shown in Fig. 123.4. In Fig. 123.4, it is shown that  $K = 22$ , it accounts for among 50 time segments, 22 time segment  $W_{set}$  is above the setting value S2. So this method can effectively detect the occurrence of polluted insulator in distribution system.

Similar to the above fault case, Fig. 123.5 depicts the residual current of instantaneous arc fault in distribution network. From Fig. 123.5, although the residual current of instantaneous arc fault is very large, it is intermittent signal; intermittence appears as a series of transients. The value K, which is the number of the ratio above the setting value S1, are obtained from Fig. 123.6. It is 43, which are above the setting value S2. Perhaps, instantaneous arc fault will access to the brink of collapse. These results clearly again demonstrate the effectiveness of the proposed technique.

**Fig. 123.6** Detector  $W_{d5}(m)$  of residual current details shown in Fig. 123.5



### 123.4.3 Discussion

From the aforementioned results, the proposed technique has a good performance for incipient fault signals in distribution network, especially for incipient fault signals such as instantaneous arc fault. The proposed technique has a distinctive indicator of incipient fault. But, for polluted insulator fault signal, due to very large resistance, the amplitude of fault feature is small, although it is implemented by this approach, it accounts for the approach is relevant with the fault resistance. Still, the proposed technique is a good approach for implementing these kinds of faults.

### 123.5 Conclusion

Successful detection of abnormalities would be a great benefit to the utilities, enabling them to detect severe faults at an early stage of their development, and consequently preventing unscheduled outages due to failures in distribution line. An ideal incipient fault detection should capture the degrading path of equipment and detect the root causes. The ultimate goal is to improve the overall distribution network reliability and reduce the operational costs strategically.

For incipient fault detection, the key issue is difficult to extract the characteristic of fault information. In this manuscript, a dwt-based method is proposed to extract the characteristic of incipient fault. The basic idea is extract the d5 coefficient of wavelet decomposition of residual current and residual voltage. Together with an energy detection algorithm, a scheme for incipient fault detection has been proposed. Simulation results verify the effectiveness of the proposed approach.



## References

1. Sidhu, T. S., & Zhihan, X. (2010). Detection of incipient faults in distribution underground cables. *IEEE Transactions on Power Delivery*, 25(3), 1363–1371.
2. Wester, C. G. (1998). High impedance fault detection on distribution systems. In *42nd Annual Conference on Rural Electric Power Conference*, St. Louis, MO (pp. c5-1–5).
3. Kim, C. J., Shin, J. H., Yoo, M.-H., & Lee, G. W. (1999). A study on the characterization of the incipient failure behavior of insulators in power distribution line. *IEEE Transactions on Power Delivery*, 14(2), 519–524.
4. Benner, C. L., & Russell, B. D. (1997). Practical high-impedance fault detection on distribution feeders. *IEEE Transactions on Industry Applications*, 33(3), 635–640.
5. Al-Dabbagh, M., & Al-Dabbagh, L. (1999). Neural networks based algorithm for detecting high impedance faults on power distribution lines. In *Proceedings of International Joint Conference on Neural Networks*, Washington, DC (Vol. 5, pp. 3386–3390).
6. Cui, T., Dong, X., Bo, Z., & Juszczyk, A. (2011). Hilbert-transform-based transient/intermittent earth fault detection in noneffectively grounded distribution systems. *IEEE Transactions on Power Delivery*, 26(1), 143–151.
7. Lovisolo, L., Moor Neto, J. A., Figueiredo, K., de Menezes Laporte, L., & dos Santos Rocha, J. C. (2012). Location of faults generating short-duration voltage variations in distribution systems regions from records captured at one point and decomposed into damped sinusoids. *IET Generation, Transmission and Distribution*, 6(12), 1225–1234.
8. Butler, K. L. (1999). An expert system based framework for an incipient failure detection and predictive maintenance system. In *Intelligent System Application to Power Systems Conference*, Orlando, FL (pp. 321–326).
9. Miri, S. M., & Privette, A. (1994). A survey of incipient fault detection and location techniques for extruded shielded power cables. In *The 26th Annual Southeastern Symposium on System Theory*, Athens, OH (pp. 402–405).
10. Kim, C. J., Lee, S.-J., & Kang, S.-H. (2004). Evaluation of feeder monitoring parameters for incipient fault detection using Laplace trend statistic. *IEEE Transactions on Industry Applications*, 40(6), 1718–1724.
11. Apostolos, N. M., Andreou, G. T., & Labridis, D. P. (2012). Enhanced protection scheme for smart grids using power line communications techniques-part II: Location of high impedance fault position. *IEEE Transactions on Power Delivery*, 3(4), 1631–1640.
12. Yang, Y., Divan, D., Harley, R. G., & Habetler, T. G. (2006). Power line sensornet—A new concept for power grid monitoring. In *Power Engineering Society General Meeting, IEEE* (pp. 1–8).
13. Mousavi, J., & Rasoul, M. (2005). *Underground distribution cable incipient fault diagnosis system*. Ph.D. dissertation, TEXAS Digital Library, Texas A&M University, College Station, TX.
14. Ebron, S., Lubkeman, D. L., & White, M. (1990). A neural network approach to the detection of incipient faults on power distribution feeders. *IEEE Transactions on Power Delivery*, 5(2), 905–914.
15. Tsarabaris, P. T., Karagiannopoulos, C. G., & Theodorou, N. J. (2005). A model for high voltage polluted insulators suffering arcs and partial discharges. *Simulation Modelling Practice and Theory*, 13(2), 157–167.

# Chapter 124

## A Complementary Metal Oxide Semiconductor D/A Converter with R-2R Ladder Based on T-Type Weighted Current Network

Junshen Jiao

**Abstract** The mathematical expression and physical implementation are analyzed for a D/A converter and illuminated the T-type network framework of a binary digital-to-analog transform by dividing current means in this paper. Based on it, slice of half-dividing current is suggested by way of the symmetry of the drain and the source terminals in CMOS transistor. The paper puts forward a novel CMOS D/A converter based on T-type weighted current network with R-2R ladder. It has the merits of low power consumption and easy making of integration. Simulation result reveals a monotonic characteristic of the D/A converter.

### 124.1 Introduction

With the digital technology, especially the rapid development of computer technology, modern control, communication, and testing, the signal processing is widely adopted in digital computer technology in order to improve the performance of the system. Digital-to-analog (D/A) converter provides the interface between the analog world and digital signal processing systems. D/A converter is widely used in computer, automatic control, measurement, and many other areas. It is an indispensable device in modern communication.

D/A converter of traditional MOSFET architecture is a binary-weighted current source, which is composed of identical complementary metal oxide semiconductor (CMOS), and it is current steering. It has been used in a wide range of applications for conversion [1–3]. The architecture designing allows high-speed data converter, but a serious drawback is that it has consumed current source due to the high number of units, wherein D/A converter doubles the number of bits and a large

---

J. Jiao (✉)

Department of Electronic Engineering Technology, Tongling University,  
Anhui 244061, China  
e-mail: [jiaotlu@sina.com](mailto:jiaotlu@sina.com)

silicon area. In addition, consumption of large areas of the current source array is difficult to match all the MOS transistors [4–6].

This study evaluates the R-2R ladder D/A converter to present that the current-or inverse-mode framework is best suited for low-power operation. Then, methods to characterize the current-mode converter are put forward. The capability of an 8-bit D/A converter fabricated using CMOS process is finally brought forward to display the characterization techniques.

### 124.2 Mathematical Expression and Physical Implementation of Binary D/A

A  $n$ -bit binary digital quantity  $D(b_{n-1}, b_{n-2} \dots, b_1, b_0)$  can be given as follows:

$$A = b_{n-1} \times 2^{n-1} + b_{n-2} \times 2^{n-2} + \dots + b_1 \times 2^1 + b_0 \tag{124.1}$$

where  $b_i \in \{0,1\}$ ; thus, there are  $2^n$  possible values for  $A$ . As long as the number  $n$  is large enough, the obtained  $A$  is probably regarded as an analog quantity. If signals are represented by voltage, the converted analog voltage signal can be expressed as

$$\begin{aligned} V_A &= K \times V_{REF} \times A = K \times V_{REF} \times (d_{n-1} \times 2^{n-1} + \dots + d_1 \times 2 + d_0 \times 2^0) \\ &= 2^{n-1} \times K \times V_{REF} \times \left( d_{n-1} + d_{n-2} \times \frac{1}{2} + \dots + d_1 \times \frac{1}{2^{n-2}} + d_0 \times \frac{1}{2^{n-1}} \right) \end{aligned} \tag{124.2}$$

where  $K$  is a constant and  $V_{REF}$  is the normal reference voltage. It is known that it is difficult to gain voltage signals, but the current signal facility is to be added by tying wires. Thus, it is guessed that each item in summing is represented by current. These currents are found weighted and switched by  $d_{n-1}d_{n-2} \dots d_1d_0$ , respectively. D/A converter with weighted current signal can be given as

$$I_{n-1} : I_{n-2} : \dots : I_1 : I_0 = 1 : \frac{1}{2} : \dots : \frac{1}{2^{n-2}} : \frac{1}{2^{n-1}} \tag{124.3}$$

It can be replaced as T-type resistor network with R-2R, as shown in Fig. 124.1.

In Fig. 124.1, the current passing through the resistor  $R$  is never changed whether  $S_{n-1}S_{n-2} \dots S_1S_0$  is connected to ground ( $d_i = 0$ ) or virtual ground ( $d_i = 1$ ) because of summing amplifier input  $V^-$  approach 0. Each branch current is always the same in T-type resistor network with R-2R. It should be noted that  $V^-$  is a virtual ground. It can get net equivalent resistance  $R$  from A–A, B–B, C–C, D–D, and E–E port. So,

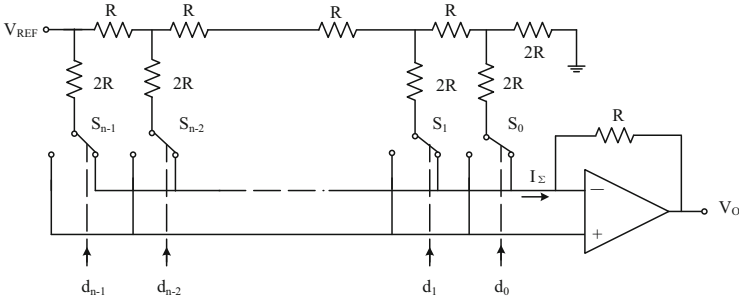


Fig. 124.1 Binary D/A converter T-type resistor network with R-2R

$I = V_{REF}/R$ , each branch current is  $I/2, I/4, I/8, I/16 \dots I/2^{n-1}$  and  $I/2^n$ , respectively.

A main error source of the R-2R ladder D/A converter is the mismatch in switch on resistance, but the advances in CMOS fine-line technology have greatly improved the matching accuracy of switches [7, 8]. Based on these technological aspects, the R-2R ladder is revisited.

### 124.3 Current Division Principle of MOS Transistor

In order to testify MOS transistor division of current, it makes use of N-MOS transistor, which also uses P-MOS. N channel MOS transistor cross-sectional diagram can be given in Fig. 124.2.

If the conductive channel voltage is  $V(x)$ , electron diffusion and drift give birth to arbitrary position current. Assume  $I(x)$  is the current of inversion layer [9, 10]. In this way,

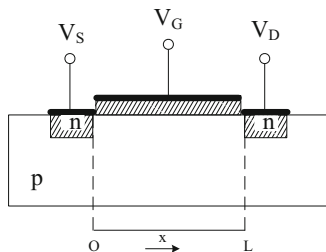
$$I(x) = I_{drift}(x) + I_{diff}(x) \tag{124.4}$$

It is in proportion with  $I_{diff}(x)$ , channel charge density ( $Q_C$ ), electron mobility ( $\mu$ ), channel field strength ( $dV_C/dx$ ), and the channel width ( $W$ ).  $I_{drift}(x)$  can be written as

$$I_{drift}(x) = -W\mu Q_C \frac{dV_C}{dx} \tag{124.5}$$

The diffusion current is in proportion with  $\mu$ , thermal voltage ( $KT/Q$ ), and derivative of  $Q_C$ .  $I_{diff}(x)$  can be given as

**Fig. 124.2** N channel MOS transistor section



$$I_{diff}(x) = W\mu \frac{KT}{q} \frac{dQ_C}{dx} \quad (124.6)$$

Thus,

$$I(x) = W\mu \left( -Q_C \frac{dV_C}{dx} + \frac{KT}{q} \frac{dQ_C}{dx} \right) \quad (124.7)$$

The current along the channel is constant because the channel length is  $L$ . Consequently,

$$I \times L = W \int_0^L -\mu \left( Q_C \frac{dV_C}{dx} - \frac{KT}{q} \frac{dQ_C}{dx} \right) dx \quad (124.8)$$

It is known that  $I_D = -I$ , so

$$I_D = \frac{W}{L} \int_0^L \mu \left( Q_C \frac{dV_C}{dx} - \frac{KT}{q} \frac{dQ_C}{dx} \right) dx \quad (124.9)$$

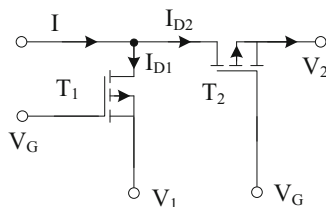
Equation 124.9 clearly reveals the MOS transistor drain and gate symmetry characteristics.  $I_D$  is in proportion with  $W/L$  in the same substrate.

Consequently, the principles of half-dividing current can be used in MOS transistor replacement in Fig. 124.1. In spite of V-I identities of MOS transistor having nonlinear connection, symmetric outcomes can be shown in Eq. 124.10.

$$\frac{I_{D1}}{I_{D2}} = \frac{W_1/L_1}{W_2/L_2} \quad (124.10)$$

Equation 124.10 shows that the ratio  $I_{D1}/I_{D2}$  is constant and independent of the input current and the terminal voltages  $V_G$ ,  $V_1$ , and  $V_2$ . Also, for an equal division of the input current, the transistors  $T_1$  and  $T_2$  should have the same size. The resistors in the R-2R ladder can be replaced by CMOS transistors and still preserve the current division principle, despite the nonlinear current-voltage connection in CMOS transistors. Principle of half-dividing current for MOS is shown in Fig. 124.3.

**Fig. 124.3** Principle of half-dividing current for MOS



## 124.4 Half-Dividing Current Component of MOS with R-2R Ladder

MOS D/A conversion current component is shown in Fig. 124.4. It is composed of four patches PMOS transistors.  $T_2$  assumes the R branch function. 2R branch is connected in series by  $T_1$  and  $T_3$  (or  $T_4$ ). The input current is divided into equal parts.  $T_3$  or  $T_4$  not only takes in hand switching function but also occupies the matching role. Therefore, the entire unit layout is compact and has good match.

Each unit consists of  $T_3$  and  $T_4$  tube of two complementary as the switch transistor. If  $d_i = 1$ , weighted current can be exported by  $I_{OUT}$ ,  $T_4$  transistor is close. If  $d_i = 0$ , weighted current connects virtual ground terminal of the amplifier. Output current accords with the Eq. 124.3.

## 124.5 8-Bit D/A Converter of MOS with R-2R Ladder

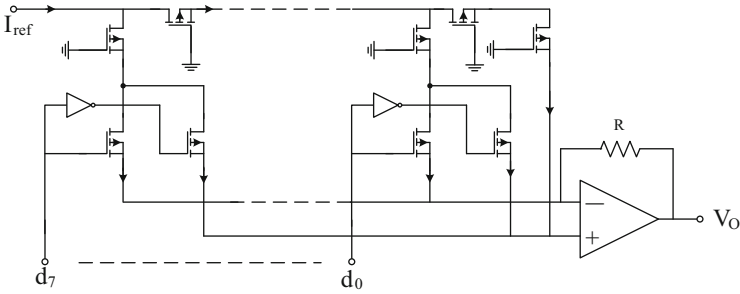
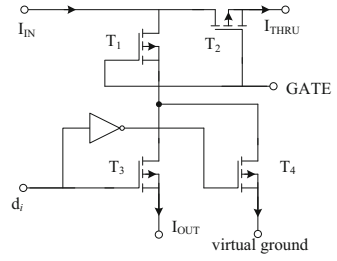
An 8-bit D/A converter fabricated using CMOS process is shown in Fig. 124.5 with R-2R ladder. The resistor R in Fig. 124.5 is replaced by the slice PMOS transistor operating in the linear region and the four slice transistors form the slice cell for 1-bit conversion. Those unit transistors are driven by the digital input  $d_i$  and also operate as switches.

## 124.6 Simulation Result

Based on the 0.35- $\mu\text{m}$  CMOS slice and power voltage of 3 V, it makes use of HSPICE simulation in order to test the proposed architecture linearity.  $T_1$  and  $T_2$  parameters are  $W/L = 2 \mu\text{m}/20 \mu\text{m}$  and  $T_3$  and  $T_4$  parameters are  $W/L = 1 \mu\text{m}/0.7 \mu\text{m}$ .

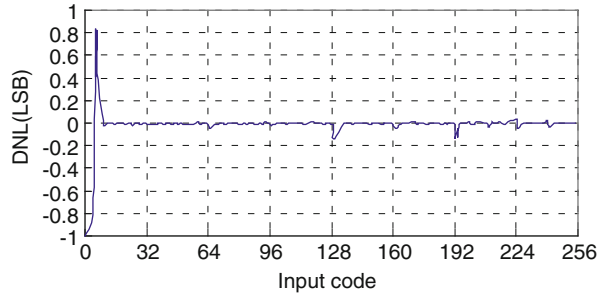
Simulation results are shown in Figs. 124.6 and 124.7. DNL and INL have skip courses because of the output of the amplifier limiting the cause in the initial moment. Despite the initial skips, DNL has no excess of the range of  $\pm 1$  LSB, which reveals a monotonic characteristic of the D/A converter. The INL stays within the limit of  $\pm 0.9$  LSB.

**Fig. 124.4** Half-dividing current component of MOS with equivalent R-2R network

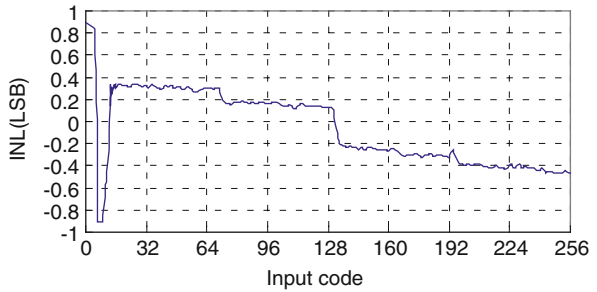


**Fig. 124.5** 8-bit D/A converter of MOS with R-2R ladder

**Fig. 124.6** DNL simulation result



**Fig. 124.7** INL simulation result



## 124.7 Conclusion

By analyzing the mathematical expressions of D/A converter, a T-type resistance network is introduced by using the dividing current. Based on this principle, CMOS transistor as a resistor ladder is proposed in this paper. Compared with the traditional resistance network, the architecture has good linearity. In this network, the CMOS parameter transistors can be precisely controlled in size. It can maintain the ratio between the MOS transistor. So, it has important theoretical and application value in the field of electronic technology.

**Acknowledgements** The work was supported by Educational Commission of Anhui Province of China (No. 2007JY YM443).

## References

1. Yin, G. M., Eynde, F. O., & Sansen, W. (1992). A high-speed CMOS comparator with 8-bit resolution. *IEEE Journal of Solid-State Circuits*, 27(2), 208–211.
2. Lee, S. C., & Cho, M. H. (2002). 10-Bit 200 ms/s CMOS D/A converter employing high-speed limiter. *IET Journal of Electronics Letters*, 38(23), 1407–1408.
3. Borremans, M. A. F., & Steyaert, M. S. J. (2001). A 10-bit 1-Gsample/s Nyquist current-steering CMOS D/A converter. *IEEE Journal of Solid-State Circuits*, 36(1), 315–324.
4. Zhou, Y. J., & Yuan, J. (2003). An 8-bit 100-MHz CMOS linear interpolation DAC. *IEEE Journal of Solid-State Circuits*, 38(10), 1758–1761.
5. Ripley, D., Balteanu, F., & Gheorghe, I. (2004). Quad-band GSM/GPRS/EDGE polar loop transmitter. *IEEE Journal of Solid-State Circuits*, 39(12), 2179–2188.
6. Vleugels, K. (2001). A 2.5-V sigma-delta modulator for wideband communication applications. *IEEE Journal of Solid-State Circuits*, 36(12), 1887–1898.
7. Tseng, W. H., Wu, J. T., & Chu, Y. C. (2011). A CMOS 8-bit 1.6-gs/s DAC with digital random return-to-zero. *IEEE Transactions on Circuits and Systems II: Express Briefs*, 58(1), 1–5.
8. Deveugele, J., & Steyaert, M. S. J. (2006). A 10-bit 250-MS/s binary-weighted current-steering DAC. *IEEE Journal of Solid-State Circuits*, 41(2), 320–329.
9. Woo, J. K., & Shin, D. Y. (2009). High-speed 10-bit LCD column driver with a split DAC and a class-AB output buffer. *IEEE Transactions on Consumer Electronics*, 55(3), 1431–1438.
10. Marche, D., Savaria, Y., & Gagnon, Y. (2008). Laser fine-tuneable deep submicron CMOS 14 bit DAC. *IEEE Transactions on Circuits System I*, 55(8), 2157–2165.



# Chapter 125

## Detecting Repackaged Android Applications

Zhongyuan Qin, Zhongyun Yang, Yuxing Di, Qunfang Zhang,  
Xinshuai Zhang, and Zhiwei Zhang

**Abstract** The rapid development of the smartphone brings immense convenience to people. Recently more and more developers publish their own applications (or apps) on the android markets to make profits. The so-called repackaged apps emerge by embedding malicious codes or injecting ads into the existing apps and then republishing them. In this paper, focusing on the shortcomings of existing detection system, we propose an efficient repackaged apps detection scheme based on context-triggered piecewise hash (CTPH). We also optimize the similarity calculation method (edit distance) and filter unnecessary matching process to make the matching more efficient. Experimental results show that there are about 5 % repackaged apps in pre-collected data. The proposed scheme improves the detection accuracy of the repackaged apps and has positive significance to the ecosystem of android markets.

### 125.1 Introduction

In the past few years, android has developed strikingly which occupies a dominant position in the smartphone markets since its market share exceeds Apple in 2010. The latest data from the research company Strategy Analytics show that android's market share has risen to 70 % to the end of 2012 [1].

---

Z. Qin (✉) • Z. Yang • X. Zhang • Z. Zhang  
School of Information Science and Engineering, Southeast University, Nanjing 210096, China  
Information Security Research Center, Southeast University, Nanjing 210096, China  
e-mail: [zyqin@seu.edu.cn](mailto:zyqin@seu.edu.cn)

Y. Di  
Communication Department, Nanjing Institute of Artillery Corps, Nanjing 210000, China

Q. Zhang  
Computer Department, Nanjing Institute of Artillery Corps, Nanjing 210000, China

Because android is free and open, app developers can publish their own apps in the [1] android markets. Repackaged app developers first download original apps, disassemble, modify configuration file, inject malicious code, insert ads, re-sign with a private key, and then release to android markets again. Further analysis indicates that these repackaged apps are typically used to steal ad revenues and obtain user location, phone number, and other personal information, even to control user's phone remotely.

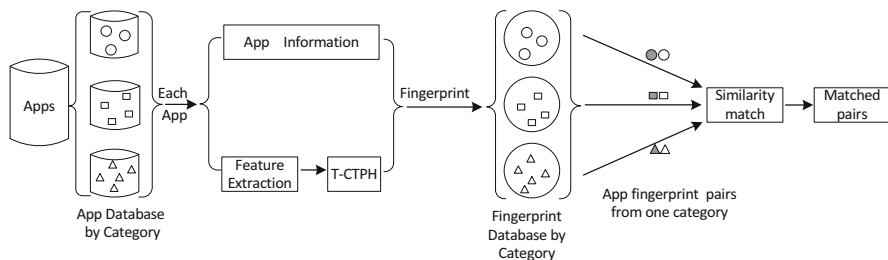
Some schemes have been proposed to detect the repackaged apps. DEXCD [2], developed by Ian Davis, extracts the opcodes from Java class in Dex file and tries to find a steam match of opcodes to detect the cloned apps. Clint et al. propose DNADroid to detect apps copying [3], which utilizes WALA to construct program dependency graphs (PDGs) [4], applies lossless and lossy filters to discard the method pairs, and computes the similarity by subgraph isomorphism. However, the robust techniques in DNADroid are largely expensive. DroidMOSS [5], presented by Wu et al., computes the fuzzy hashes of apps and compares similarity of all fingerprints; the similarity scoring algorithm is memory-consuming, thus slowing down the process.

In this paper, we proposed an improved repackaged android apps detection system based on CTPH (i.e., fuzzy hash) [6], which uses two small primes to do twice CTPH (T-CTPH) process and generates two fingerprints for each app. The main contributions of this paper are the following: (1) We propose an improved fingerprint-generating algorithm, two small primes are selected as the trigger values for T-CTPH to increase the randomness against possible attacks and improve the accuracy, which can be further used to filter unnecessary matching processes. (2) An efficient edit distance calculation method is proposed to speed the calculation of the similarity between fingerprints, which greatly reduce the memory usage. (3) We have realized our system to detect repackaged apps and found that about 5 % repackaged apps in pre-collected 6,438 samples of four app types.

The remainder of this paper is organized as follows: Sect. 125.2 introduces our approach, including feature extraction, fingerprint generation, and similarity matching. Section 125.3 illustrates the evaluation results based on 6,438 real applications from several android markets. Finally, we make our conclusion in Sect. 125.4.

## 125.2 System Design

Usually the repackaged apps have the following features: (1) They always have the same app type with the official ones. (2) The size will not have a big difference from the original one. Furthermore, we assume that the signing keys are not leaked. Therefore, the matched pairs (with high similarity) with the same author information are ignored because they are often the different versions of the same application.



**Fig. 125.1** System architecture

### 125.2.1 System Overview

To detect the repackaged apps, we have designed the following system shown in Fig. 125.1, which contains three parts: (1) sample collection and classification, (2) feature extraction and fingerprint generation, and (3) similarity matching. We first download apps from android markets by app category; for each app we extract the signature information and instruction sequence for fingerprint generation. Then we store fingerprints in their own databases and select any two fingerprints from each database for similarity calculation. Given a comparison threshold, suspicious repackaged app pairs will be fixed.

### 125.2.2 App Feature Extraction

To extract the app feature, we first uncompress it and use keytool to extract certificate information in *META-INF* directory [7]. As shown in Fig. 125.2, we get the MD5 as the unique information of one app, for it will be different between the original and repackaged apps. Then we leverage existing Dalvik disassembler *baksmali* to disassemble *classes.dex* file [8] and extract instructions by the following rules: (1) depth traversal with the alphabetical order of generated smali files and folders; (2) before releasing, some names of class are modified, we ignore the confused names of classes to reduce the error of instruction extraction; and (3) extracting methods of different classes.

### 125.2.3 Fingerprint Generation

For an actual APK file, the extracted instruction sequences may be very long. To generate the fingerprints, one common way presented in DroidMOSS is to use fuzzy hash directly, which does not consider the nature that the size of repackaged app will not have a big difference from the original one. Figure 125.3 shows the once

```

administrator@ubuntu:~$ keytool -printcert -v -file \
> /renren/META-INF/CERT.RSA
Owner: CN=renren, OU=renren.com, O=opi, L=beijing, ST=beijing, C=CN
Issuer: CN=renren, OU=renren.com, O=opi, L=beijing, ST=beijing, C=CN
Serial number: 4b85da7d
Valid from: Thu Feb 25 10:03:41 CST 2010 until: Fri Nov 28 10:03:41 CST 2064
Certificate fingerprints:
    MD5: FB:5C:BF:1E:21:6D:40:74:54:5C:72:17:84:DB:18:48
    SHA1: CE:4A:B1:BB:60:4F:74:67:3A:1B:8B:3B:C7:F7:D4:10:71:F6:E7:93
Signature algorithm name: MD5withRSA
Version: 1
    
```

Fig. 125.2 Certificate information of one application

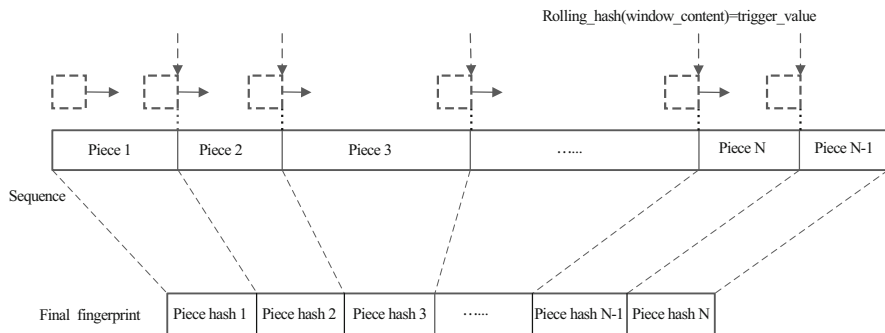
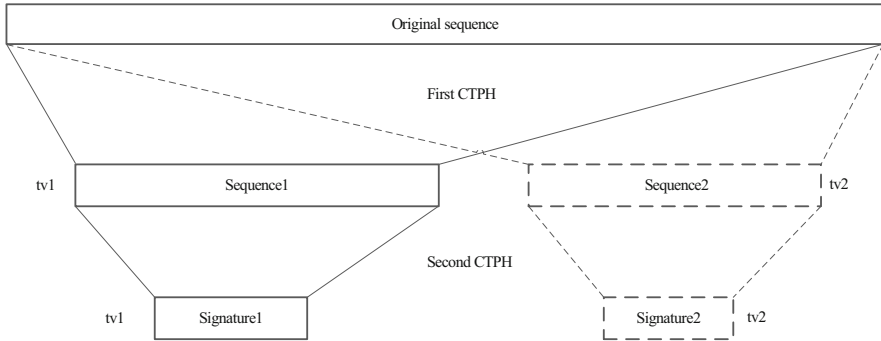


Fig. 125.3 Context-triggered piecewise hash

CTPH process, the sequence is the input and a trigger value is for dividing pieces, then all the piece hashes are calculated and concatenated directly as the final fingerprint.

In this paper, we present an improved approach. Specifically, we remove the compression mapping process in Spamsun [9]; and for sequences of different lengths, we use different primes to calculate the fingerprints; for an instruction sequence, we use two small primes to do twice CTPH processes, respectively. As visually shown in Fig. 125.4, the *original sequence* is the instructions extracted by the rules in Sect. 125.2.2. Suppose the length of the original sequence is  $N$  and  $S1$  pieces generate after the first CTPH process (e.g., the left in solid line in Fig. 125.4), then the average size of the trigger value  $tv$  is about  $\lfloor N/S1 \rfloor$  ( $\lfloor \cdot \rfloor$  means round downwards). We represent each 32-bit binary piece hash as an 8-digit hexadecimal number in this paper, so the length of the sequence generated by the first CTPH process is about  $8*S1$ . If the piece number is  $S$  after the second CTPH process, then the trigger  $tv'$  of the second CTPH is about  $\lfloor 8*S1/S \rfloor$ . So we have  $tv \approx \frac{N}{S1}$  and  $tv' \approx \frac{8*S1}{S}$ . Further,  $tv * tv' \approx \frac{8*N}{S}$ . In order to make the trigger values in similar size, we have  $tv = tv'$ . That is,  $tv = tv' \approx \sqrt{\frac{8*N}{S}}$ . And we use primes to trigger pieces to increase the randomness against possible attacks. Given  $tv$  (or  $tv'$ ), we have two adjacent primes  $r1$  and  $r2$ , that is,  $r1 \leq tv = tv' \leq r2$ ; thus we take  $r1$  and  $r2$  as the trigger values to generate two fingerprints. In order to balance efficiency and



**Fig. 125.4** T-CTPH for fingerprint generation

similarity between fingerprints,  $S$  is taken as 128, that is, *signature1* (or *signature2*) is made up about 1,024 characters in formula Eq. 125.1. Here,  $r_1$  is  $tv_1$  and  $r_2$  is  $tv_2$  in Fig. 125.4. Therefore, the final fingerprint of an app is

$$signature = (tv_1, signature_1) || (tv_2, signature_2) \tag{125.1}$$

Next the fingerprint of each app is calculated and stored in its corresponding type database. If the lengths of two sequences are far away, it is little possible they are repackaged pairs. For this nature, not all the fingerprints are compared, and we just concern the fingerprints triggered by the same prime. For two different apps, the fingerprints are  $x$  and  $y$ :  $x = (x_1, xsign1) || (x_2, xsign2)$  and  $y = (y_1, ysign1) || (y_2, ysign2)$ . If  $x_1 = y_1$ , it will definitely have  $x_2 = y_2$ ; we compare  $(xsign1, ysign1)$  and  $(xsign2, ysign2)$  with the method in Sect. 125.2.4, respectively, and put the larger similarity score as the final result; if  $x_1 = y_2$  or  $x_2 = y_1$ , we compare  $(xsign1, ysign2)$  or  $(xsign2, ysign1)$ . In other cases, we do not treat them as repackaged apps. With such filtering method, it will be more efficient to detect the repackaged apps in the android market.

### 125.2.4 Similarity Matching

Our above steps are applied for each app generating the fingerprint. Then we compare the similarity between the fingerprints by adopting an optimized edit distance method. The edit distance is the minimum edit operations to turn one fingerprint into another, including insertion, deletion, and substitution. The conventional way is to use a two-dimensional array to calculate the edit distance of two strings (with lengths of  $len_1$  and  $len_2$ , respectively) by filling the array circularly. So the final result is  $array[ len_1-1, len_2-1 ]$ .

However, if two strings are very long,  $len_1 * len_2$  size of memory will be needed for calculating, slowing the matching speed. In order to speed up the calculation for

long strings, we have optimized the method by using three one-dimensional arrays, *array1*, *array2*, and *array3* (with sizes of *len1*, *len2*, *len2*, respectively), to calculate the edit distance. Here, *array1* denotes the first column of the conventional two-dimensional array; *array2* and *array3* denote two adjacent rows of that two-dimensional array. We fill *array2* and *array3* circularly with an iterative method and exchange *array2* and *array3* continuously to denote the two adjacent rows. At last, if *len1* is odd, the edit distance is *array2*[*len2*-1]. Otherwise, the edit distance is *array3*[*len2*-1]. This process is described in Algorithm 1, and it just needs about *len1* + *len2*\*2 memory space.

---

**Algorithm 1** Calculate the edit distance between two apps

---

**Input:** Two fingerprints fp1 and fp2  
**Output:** Edit distance between fp1 and fp2  
1:  $len1 \leftarrow strlen(fp1)$ ,  $len2 \leftarrow strlen(fp2)$   
2:  $array1[0] = (fp1[0] == fp2[0] ? 0 : 1)$  //initialize array1  
3: **for**  $i = 1 \rightarrow len1 - 1$  **do**  
4:    $cost = (fp1[i] == fp2[0] ? 0 : 1)$   
5:    $array1[i] = \min(array1[i-1] + 1, i + cost)$   
6: **end for**  
7:  $array2[0] = array1[0]$  //initialize array2  
8: **for**  $i = 1 \rightarrow len2 - 1$  **do**  
9:    $cost = (fp2[i] == fp1[0] ? 0 : 1)$   
10:  $array2[i] = \min(array2[i-1] + 1, i + cost)$   
11: **end for**  
13: **for**  $i = 1 \rightarrow len1 - 1$  **do**  
14:    $(i \bmod 2 == 0) ? array2[0] = array1[i] : array3[0] = array1[i]$   
15: **for**  $j = 1 \rightarrow len2 - 1$  **do**  
16:    $cost = (fp1[i] == fp2[j] ? 0 : 1)$   
17:   **if**  $(i \bmod 2 == 0)$  **then**  
18:      $array2[j] = \min(array2[j-1] + 1, array3[j] + 1, array3[j-1] + cost)$   
19:   **else**  
20:      $array3[j] = \min(array3[j-1] + 1, array2[j] + 1, array2[j-1] + cost)$   
21:   **end if**  
22: **end for**  
23: **end for**  
24:  $edit\_dist = ((len1 \bmod 2 == 0) ? array3[len2 - 1] : array2[len2 - 1])$   
25: **return**  $edit\_dist$

---

After the edit distance between two fingerprints is calculated, we use the following formula to measure the similarity between the two fingerprints [5]:

$$Sim\_Score = \left[ 1 - \frac{edit\_dist}{\max(len1, len2)} \right] * 100 \quad (125.2)$$

If two apps are signed with different keys and the similarity score exceeds a certain threshold, we treat them as repackaged matching pairs. Note that the choice of threshold largely affects the false-positive and false-negative rate, thus influencing the accuracy of our test results. In our experiments, we choose 70 as the threshold, and it shows a good balance between the false-positive and false-negative rate.

### 125.3 Evaluation

In our experiment, we collect 6,438 apps from several android markets and store them with different categories. There are four types: social networking, game, system tool, and shopping, which are shown in column 2 of Table 125.1.

From Table 125.1, we find that our scheme (T-CTPH) improves the accuracy of detecting repackaged apps, and the results are closer to the results of manual analysis. In addition, the apps about social networking have the highest rate. This is because the many repackaged apps are used for stealing user's Internet traffic and phone bill by injecting malicious code to control user's phone remotely, which will inevitably require Internet service. Also, games related are rather high. It is the reason that developers can reroute or steal ad revenues by replacing or embedding ads to the games.

In order to detect so many applications, the complexity of the algorithm is very important. In our scheme, the time and space complexity of fingerprint generation are both  $O(n)$ , which has little influence on the overall efficiency. For the edit distance of two fingerprints, the time and space complexity are  $O(n^2)$  and  $O(n)$ , respectively. Note that before optimizing, they are both  $O(n^2)$ . Here we reduce the space complexity from  $O(n^2)$  to  $O(n)$ , which largely improve the efficiency. Then we test the optimized similarity algorithm on a computer with a Linux system (Ubuntu 10.04). The CPU is Intel (R) Pentium 4 running at 2.93 GHz, 512 MB RAM. Table 125.2 shows the consumed time of sequences in different length ranges. We find that the optimized algorithm largely speeds up the calculation and is more significant to reduce the time when the sequences are longer.

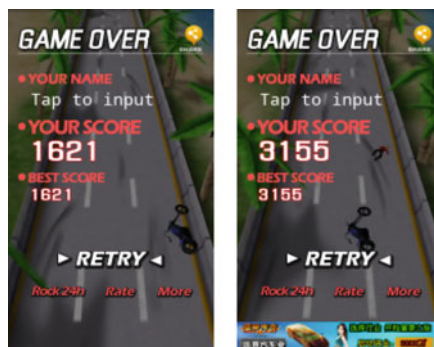
To perform a concrete study of the repackaged apps and reveal how one app is repackaged, we show the analysis of repackaged apps detected by our scheme. Usually, advertising SDK needs to add a publisher identifier to `AndroidManifest.xml` and modify the layout description and the program bytecodes to show ads. As shown in Fig. 125.5, it is a detected example that repackage a normal app (*com.racingstudio.racingmoto*) by including AdMob SDK [10]. We find that the signed keys are different, and they are similarity matching pairs. Further, with manual analysis, the repackaged app (right) always pops up ads in the bottom when running the two games and functions like *setVisibility*, *findViewById*, and *loadAd* are inserted into *onCreate* to display ads in the disassembled files, by which developers can steal the ads revenues.

**Table 125.1** The results of different detecting repackaged apps methods and manual analysis

Category	Number	DroidMOSS	T-CTPH	Manual analysis	Percentage
Social	2,557	157	156	155	6.1
Game	2,396	140	140	138	5.8
Tool	838	39	38	38	4.5
Shopping	647	34	34	33	5.1

**Table 125.2** Consumed time of before and after optimization

Lengths	0.5k	2k	4k	6k	8k	10k
Before (ms)	10.27	127.41	468.56	1,028.37	1,818.52	2,992.54
After (ms)	4.04	68.24	272.91	604.49	1,032.50	1,443.57

**Fig. 125.5** Screenshots of repackaging

## 125.4 Conclusion

In this paper, we propose an efficient method to detect repackaged apps based on CTPH. We remove the compression mapping step, do twice CTPH process with two small primes to improve the fingerprint accuracy, optimize the similarity algorithm, and filter unnecessary matching processes to make the matching more efficient. Our experimental results show there are about 5 % repackaged apps in pre-collected samples, and it has a positive and practical significance for the ecological system of the android markets.

**Acknowledgements** This paper is funded by the Information Security Special Projects of National Development and Reform Commission. The authors would like to thank the anonymous reviewers for their insightful comments that helped improve the presentation of this paper.

## References

1. Bicheno, S. (2013). *Global smartphone OS market share by region: Q4 2012*. <https://www.strategyanalytics.com/default.aspx?mod=reportabstractviewer&a0=8222>
2. Davis, I. (2012). *Dex clone detector*. <http://www.swag.uwaterloo.ca/dexcd/index.html>
3. Crussell, J., Gibler, C., & Chen, H. (2012). Attack of the clones: Detecting cloned applications on Android markets. In *Computer Security—ESORICS 2012* (pp. 37–54). Heidelberg: Springer.
4. IBM T.J. Watson Research Center. (2012). *Watson libraries for analysis (WALA)*. [http://wala.sourceforge.net/wiki/index.php/Main\\_Page](http://wala.sourceforge.net/wiki/index.php/Main_Page)



5. Zhou, W., Zhou, Y., Jiang, X., & Ning, P. (2012). Detecting repackaged smartphone applications in third-party android marketplaces. *Proceedings of the Second ACM Conference on Data and Application Security and Privacy* (pp. 317–326). New York, NY: ACM.
6. Kornblum, J. (2006). Identifying almost identical files using context triggered piecewise hashing. *Digital Investigation*, 3S, S91–S97.
7. *Android development guide: Signing your applications*. <http://developer.android.com/tools/publishing/app-signing.html>
8. *Smali-An assembler/disassembler for Android's dex format*. <https://code.google.com/p/smali/>
9. Andrew, T. *Spamsum README*. <http://www.samba.org/ftp/unpacked/junkcode/spamsum/>
10. Google Inc. *Admob for android developers*. [http://support.google.com/admob/topic/1307236?hl=zh-Hans&ref\\_topic=1307209](http://support.google.com/admob/topic/1307236?hl=zh-Hans&ref_topic=1307209)

# Chapter 126

## Design of Wireless Local Area Network Security Program Based on Near Field Communication Technology

Pengfei Hu and Leizhen Wang

**Abstract** In order to solve wireless local area network (WLAN) security problem due to the open-wide nature of wireless radio and the improvement of computing power, a design of WLAN security program based on near field communication (NFC) is presented in this paper. In this paper, the importance of having access to handshake for WPA2 brute force is explained. The proposed design protects the four-way handshake by taking advantage of NFC short-range character to eliminate the risk of intercept. For implementation, Android system is selected as a mobile device development platform. The design is compatible with the IEEE 802.11i which ensures the massive expansion in the future. Furthermore, the design simplifies operations to improve users' experience without much extra hardware cost and offers an option to the owner of WLAN to control the access physically, which benefits commercialization of NFC. From one perspective, this design can solve the wireless network security problem effectively.

### 126.1 Introduction

With the popularization of mobile terminals and Internet of Things, wireless communication technology has gained great progress. The near field communication (NFC) has got considerable concern due to its security, low power consumption characteristics. Rapidly growing NFC is expected to be one of the most important trends and continues to gain popularity in the business and IT industry.

At the same time, wireless networking has been experiencing an explosive growth and offers attractive mobility and flexibility to both network users and operators [1] in the age of mobile Internet. The wireless local area network (WLAN) systems like IEEE 802.11 networks become common access networks in

---

P. Hu (✉) • L. Wang  
Northeastern University at Qinhuangdao, Qinhuangdao 066004, China  
e-mail: [hupengfei1993@gmail.com](mailto:hupengfei1993@gmail.com)

public and private environments. Due to the wide-open nature of wireless radio, security over a wireless environment is more complicated than in a wired environment. Many attacks make the wireless network insecure [2]. The security problem has already become the Achilles' heel of the further development of the wireless network. To overcome the security challenges, IEEE 802.11i (also called WPA2) has been developed to enhance the security. However, current wireless technologies in use allow hackers to monitor and even change the integrity of transmitted data [3]. The improvement of computing power makes it possible for the effective WPA2 brute-force cracking.

To solve these security challenges, we propose a program to set up a WLAN that combine Wi-Fi with NFC. The proposed program provides a simple but safe mean to set up a WLAN among mobile devices. We realize the importance of capturing handshake packet to hack WPA/WPA2-psk, so we introduce the NFC by taking advantage of NFC short-range character to protect the 4-way handshake to guarantee WLAN security.

The rest of this chapter is organized as follows. In Sect. 2 we introduce some related works about NFC and IEEE 802.11i security threats. After that, our design of WLAN security program based on NFC technology is explained in Sect. 3. In Sect. 4, we give some discussion. This chapter is concluded with a summary and future works in Sect. 5.

## 126.2 Related Work

### 126.2.1 Introduction of NFC

NFC is a short-range wireless connectivity technology which is heavily based on *radio frequency identification* (RFID). Extending the ability of the contactless card technology, NFC also enables devices to share information over a distance of a few centimeters with a maximum communication speed of 424 kbps operating in the 13.56 MHz frequency band. NFC has three operating modes: read/write mode, card emulation mode, and peer-to-peer mode. The peer-to-peer mode is an operating mode specific to NFC and allows two NFC devices to communicate directly with each other [4] with the lightweight, binary message data format called *NFC Data Exchange Format* (NDEF). NFC bidirectional communication is an ideal ability for establishing connections with other technologies. This technology can be integrated into an existing system to simplify and speed up the process of monitoring and control of the system [5]. To support the peer-to-peer communication, the *Logical Link Control Protocol* (LLCP) was designed to make peer-to-peer transactions smoother as it enables NFC devices to be equal in communication [6].

Because the communication distance is short, it is almost impossible to intercept. NFC and WLAN do not share the same frequency band [7]. The difference of frequency obstructs the radio interception among the transmission of NFC data.

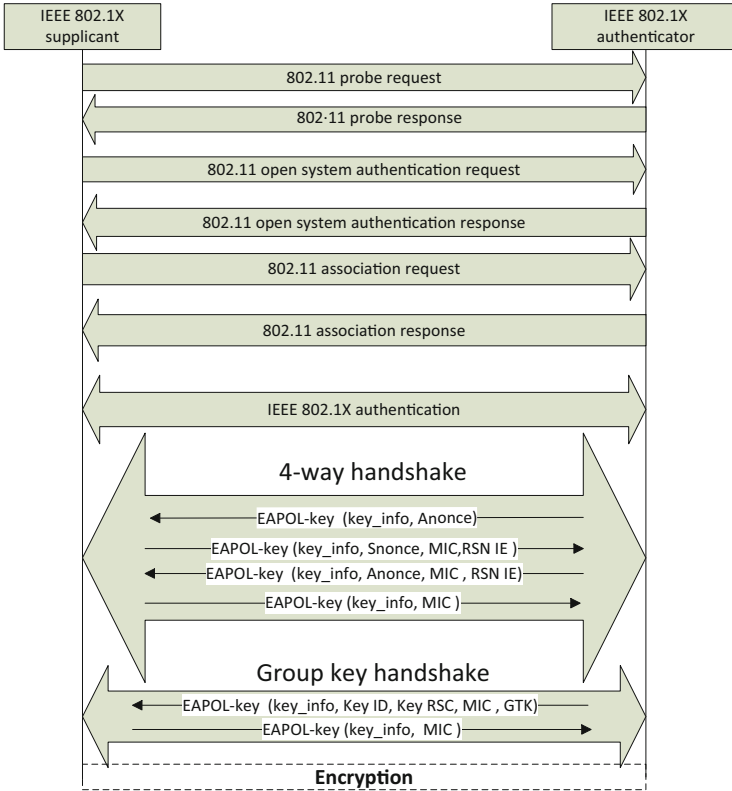


Fig. 126.1 RSNA establishment flow

### 126.2.2 IEEE 802.11i Security Threats

The IEEE 802.11 committee established the Task Group i (TGi) and ratified IEEE 802.11i on July 2004 [8]. 802.11i incorporates IEEE802.11X as its authentication enhancement and adopts a concept called *robust security network associations* (RSNAs) which is established as shown in Fig. 126.1. The success of authentication is witnessed by the fact that the access point (AP) and the client station (STA) own the same *Pairwise Master Key* (PMK) at the end. The four-way handshake provides a new temporary key called the *Pairwise Transient Key* (PTK) for confidentiality, data authentication, and anti-replay attack [8]. All the messages during the four-way authentication handshake are sent as EAPOL-Key frames.

Considering the fact that WPA2 has its own advanced system, there is no specific effective attack method till now except for brute-force attack or dictionary attacks. But the security risk cannot be ignored with computing power booming. From theory to practice, the high-speed brute force for WPA/WPA2-psk based on

distributed multi-core CPU and GPU get a performance optimization in recent years. The development of cloud computing provides the optimal balance between cost and benefit as well. Both of them let WPA/WPA2-psk cracking available.

To make the matter worse, people will be prone to choose easy key for convenience in home or cafe. In such environment, privacy information and sensitive data are in transmission every day without enough protection. The high-speed brute force for WPA/WPA2-psk based on distributed multi-core CPU and GPU can crack 8 bits digital key in tens of seconds [9]. As for the required data to generate and verify, the key is broadcast with normal traffic and is really obtainable for the attacker. And the weakness is that the PMK was derived from the concatenation of the passphrase, SSID, length of the SSID, and nonces [3]. Once the four-way authentication handshake has been captured, the attacker has enough information required to do a dictionary attack to subject the passphrase. While PTK is a keyed-HMAC function based on the PMK, the challenges still exist in WPA2.

Dictionary and brute-force attacks are typically done automatically with tools. We must find a way to jump out the infinite loop that computing power's gain means our wireless network security loss. We have realized that capturing the four-way authentication handshake is the basement of the cracking. So what we need to do is to find a way of avoiding the four-way handshake being captured.

### 126.3 Design of the Program

Through the explanations of the weakness of WPA/WPA2-psk, we conclude that protecting the four-way handshake will be a good way to make sure our WLAN is secured. Meanwhile WLAN must transmit its authentication information in air to meet the flexibility requirement. So our program must be able to prevent the attacker from obtaining the four-way handshake and meet the flexibility requirement. NFC as a short-range wireless connectivity technology exactly fixes these requirements. Consequently, we hold the belief that the security of NFC should be introduced to the handshaking to connect the WLANs.

Considering that most mobile NFC devices available in the market are for Android system, the most popular open-source operating system, we select Android system as the mobile device development platform. The function of NFC was introduced by Google into Android2.3 device. The Android 4.0 system provides API support for NFC tag read/write mode and NFC P2P mode (Android Beam).

The Android SDK provides an NFC API to develop NFC applications which conduct peer-to-peer data exchange [10]. As the Wi-Fi functions are hiding behind the SDK interface, we cannot call the bottom relevant class. So our design pays more attention on the NFC part. The *.nfc* android package provides access to NFC function. On *android.nfc*, several classes could be used to running NFC function [11]:

- *Android.app.Activity*
- *android.content.Intent*
- *android.nfc.NdefMessage*
- *android.nfc.Record*
- *android.nfc.NfcAdapter*
- *android.nfc.NfcEvent*
- *android.nfc.NfcAdapter.CreatNdefMessageCallback*

We can use *NdefMessage* (byte[ ]) to construct an NDEF message from binary data or *NdefMessage (NdefRord [ ])* to construct from one or more *NdefRecords* and then transform the information from EAPOL-Key frames into NDEF frames. So we construct *NdefRecord* (short, byte [ ], byte [ ], byte [ ]) according to the EAPOL-Key which contains the *key\_info*, *Annoce*, *MIC*, and *RSN IE*. The program is designed as follows [12] (shown in the Fig. 126.2):

#### 1. Initialization and Anti-collision

The mobile devices should initialize first. The initiator periodically probes the presence of a target by scanning the surrounding Wi-Fi radio signals and then verifies the signature. If the signature does not match, continue scanning.

#### 2. Activation and Parameters Selection

Once a target has been detected by two devices in range to communicate, the initiator gets into activation state to let the IEEE 802.1X authentication process as normal. Once the authentication phase is completed, the initiator creates an *NdefMessage* that contains the *NdefRecords* which are constructed according to the EAPOL-Key and a set of parameters like *key\_info*, *Annoce*, *MIC*, and *RSN IE* are notified or negotiated. LLCP services will be selected after those abovementioned are finished.

#### 3. Data Exchange

As the old EAPOL-Key frames have already exchanged into NDEF, the initiator calls the *setNdefPushMessage()* with the *NdefMessage* created before. Data exchange over the LLCP is in good protection. Once the message is received by the initiator, the initiator checks the replay counter to avoid replay attack. The initiator will load the PTK if the MIC is valid.

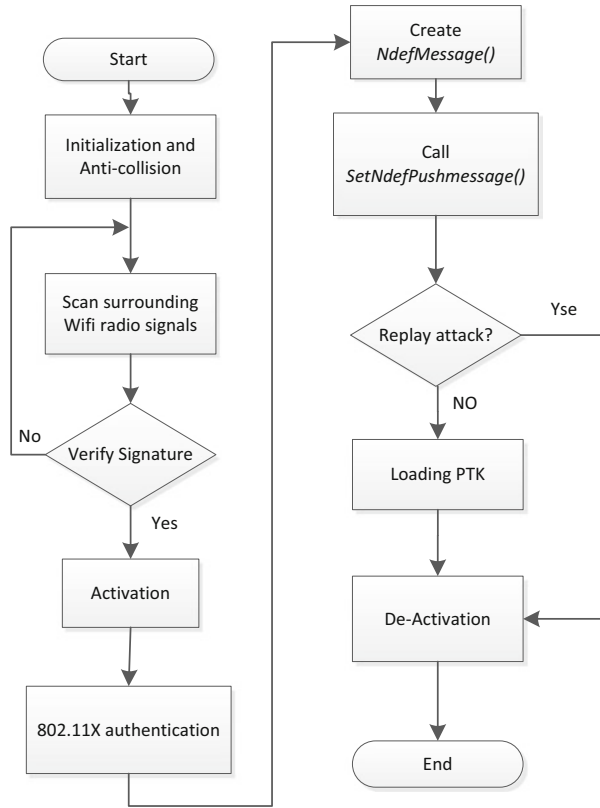
#### 4. Deactivation

The initiator can release the NFC session after the success of authentication, via Release-Request/Response messages.

## 126.4 Discussion

The factor that our supplicant/initiator and authenticator are both on the Android system makes our program similar with the Android Beam or Wi-Fi Direct. The reason/cause is that there is no NFC-enable wireless router availed. In order to implement this program, we have to take the NEC-enable Ultrabook or tabletPC on

**Fig. 126.2** Approach schema



the Android platform as access point to be the authenticator. In our ideal condition (shown in Fig. 126.3), there is a machine like point-of-sale (POS) or wireless router ingrate NFC, so we can finish our payment conveniently and join the WLANs in safety by simple touch.

Base on the description of the program, we can get several points as what are written below:

1. We found that the importance of hacking WPA/WPA2-psk is capturing handshake via the wide-open nature of wireless radio. So we take advantage of the short-range character from NFC to protect the four-way handshake and use NFC to share WPA/WPA2-psk link setup parameters to eliminate the risk of intercept.
2. Due to the compatibility with IEEE 802.11i without changing its framework and protocol, our program could be massively expanded easily if the NFC-enable router appears. Along with the Internet of Things walking into our life, wireless telecommunication equipment that supports NFC function will be mass produced.

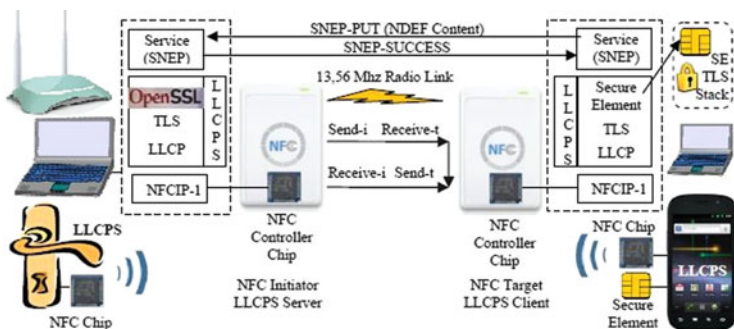


Fig. 126.3 Program overview

3. Considering the convenience and economy, people usually choose WPA/WPA2-psk (key pre-shared) model in their office and home. This key management cannot provide high-level security. In our program, the balance point between cost and security has been found. Our program provides the management of dynamic key to guarantee the WLAN security. People could say no to the trouble of remembering passwords without much extra hardware cost.
4. As the joining WPA2 via application calling is not supported by the Android system, Wi-Fi connection is sort of complicated. In the era of mobile Internet, these obstacles will disappear along with mobile device system upgrade [13]. In our program, the user can join the WLAN just by simple touch. The users' experience gain is improved, and making the power consumption becomes less.
5. Because exchanging the pairing data is allowed, the user's device does not need to store the password before. And the unsafety of temporary authority is disappeared. The owner could control the access to the WLAN physically on condition that the access methods except NFC are restricted. This point will benefit the commercialization of NFC. Businesses could use this way to let their own customers join the WLAN in an open field while closing the door to someone who is not their customer. This option could decrease network cost, improve quality of service, and especially offer a safe network environment to customers.

## 126.5 Conclusion

This paper introduced NFC, a promising technology, to the wireless network access and tried to solve the problem of wireless network security completely from one perspective. This program made full use of the short-range characteristics of NFC to eliminate the possibility of intercept of handshaking package. The main benefit of our program is compatibility with IEEE 802.11i without changing the network



framework. So we can make full use of existing resource to reduce upgrade cost. To compare with the similar technology such as Android Beam, our program used IEEE 802.11i for mutual authentication, key generation, key exchange, encryption, and integrity. Since NFC can establish a connection faster, the cost of time decreases. The user can join the WLAN just by a simple touch to improve user experience. As for future research, we will study the implementation of our program in other systems and further optimize the program.

**Acknowledgement** This work was supported by the National Natural Science Foundation of China (Grant No.61273203).

## References

1. Samiah, A., Aziz, A., & Ikram, N. (2007). An efficient software implementation of AES-CCM for IEEE 802.11i Wireless St. In *31st Annual International Computer Software and Applications Conference (COMPSAC 2007)* (pp. 689–694). IEEE.
2. Chen, J. C., Jiang, M. C., & Liu, Y. W. (2005). Wireless LAN security and IEEE 802.11i. *Wireless Communications*, 12(1), 27–36.
3. Lashkari, A. H., Danesh, M. M. S., & Samadi, B. (2009). A survey on wireless security protocols (WEP, WPA and WPA2/802.11i). In *2nd IEEE International Conference on Computer Science and Information Technology* (pp. 48–52). IEEE.
4. Monteiro, D. M., Rodrigues, J. J., & Lloret, J. (2012). A secure NFC application for credit transfer among mobile phones. In *International Conference on Computer Information and Telecommunication Systems* (pp. 1–5). IEEE.
5. Opperman, C. A., & Hancke, G. P. (2011). A generic NFC-enabled measurement system for remote monitoring and control of client-side equipment. In *3rd International Workshop on Near Field Communication (NFC)* (pp. 44–49). IEEE.
6. Lotito, A., Mazzocchi, D. (2012). OPEN-NPP: an open source library to enable P2P over NFC. In *4th International Workshop on Near Field Communication (NFC)* (pp. 57–62). IEEE.
7. Jie, MA, & Jin-long, E. (2013) Program of establishing connection of WiFi transmission rapidly based on NFC technology. *Computer Engineering*, 39(6), 1–6 (In Chinese).
8. Hori, Y., & Sakurai, K. (2006) Security analysis of MIS protocol on wireless LAN comparison with IEEE802. 11i. In *3rd International Conference on Mobile Technology, Applications & Systems* (p. 11). ACM.
9. Liu, Y. L., Jin, Z. G., Chen, Z., & Liu, J. W. (2007). Design and implement of high-speed brute forcer for wpa/wpa2-psk. *Computer Engineering*, 37(10), 125–127 (In Chinese).
10. Serfass, D., & Yoshigoe, K. (2012). Wireless sensor networks using android virtual devices and near field communication peer-to-peer emulation. *Proceedings of IEEE Southeastcon* (pp. 1–6). IEEE.
11. Android SDK Developer Guide. Available: <http://developer.android.com/>
12. Urien, P. (2013). LLCPS: A new security framework based on TLS for NFC P2P applications in the Internet of Things. In *Consumer Communications and Networking Conference* (pp. 845–846). IEEE.
13. Arakawa, Y., Sonoda, Y., Tagashira, S., & Fukuda, A. (2012). WiFiTag: Direct link from the real world to online digital contents. In *Seventh International Conference on P2P, Parallel, Grid, Cloud and Internet Computing* (pp. 339–344). IEEE.

# Chapter 127

## A Mechanism of Transforming Architecture Analysis and Design Language into Modelica

Shuguang Feng and Lichen Zhang

**Abstract** One of the fundamental challenges in research related to cyber-physical system is accurate modeling and representation of these systems. The main difficulty lies in developing an integrated model that represents both cyber and physical aspects with high fidelity. Among existing techniques, an approach to integrate Modelica with AADL is a suitable choice, as it can encapsulate diverse attributes of cyber-physical systems. AADL modeling language provides a comprehensive set of diagrams and constructs for modeling many common aspects of systems engineering problems, such as system requirements, architectures, components, and behaviors. Complementing these AADL constructs, the Modelica language has emerged as a standard for modeling the continuous dynamics of cyber-physical systems in terms of hybrid discrete event and differential algebraic equation systems. Integrating the descriptive power of AADL models with the analytic and computational power of Modelica models provides a capability that is significantly greater than provided by AADL or Modelica individually. A transformation of AADL into Modelica is developed that will support implementations to transfer efficiently the modeling information between AADL and Modelica models without ambiguity. This chapter proposes an approach to transform the models of AADL into the models of Modelica, to clarify the transformation principles, and to illustrate the important synergies resulting from the integration between these two languages.

### 127.1 Introduction

Cyber-physical systems (CPS) are becoming an integral part of modern societies [1]. As an application domain, CPS is not new. For example, early automotive-embedded systems in the 1970s already combined closed-loop control of the brake

---

S. Feng • L. Zhang (✉)  
Shanghai Key Laboratory of Trustworthy Computing, East China Normal University,  
Shanghai 200062, China  
e-mail: [zhanglichen1962@163.com](mailto:zhanglichen1962@163.com)

and engine subsystems (physical parts) with the embedded computer systems (cyber parts) [2]. Since then, new requirements, functionalities, and networking have dramatically increased the scope, capabilities, and complexities of CPS. This has created needs to bridge the gaps between the separate CPS subdisciplines (computer science, automatic control, mechanical engineering, etc.) and to establish CPS as an intellectual discipline in its own right [3]. The development of a CPS involves many stakeholders who are interested in different aspects of the system [4]. CPS require more advanced modeling techniques to capture physicality including time and space, reliability in terms of probabilistic models, and connectivity in terms of communication links, adaptivity, context awareness, interoperability, and autonomy. This requires a comprehensive integrated modeling framework for specification, modeling of architecture, and tracing their relationships [5].

In order to meet the challenge of cyber-physical system design, we need to realign abstraction layers in design flows and develop semantic foundations for composing heterogeneous models and modeling languages describing different physics and logics. We need to develop new understanding of compositionality in heterogeneous systems that allows us to take into account both physical and computational properties. One of the fundamental challenges in research related to cyber-physical system is accurate modeling and representation of these systems. The main difficulty lies in developing an integrated model that represents both cyber and physical aspects with high fidelity. Among existing techniques, an approach to integrate Modelica [6] with AADL [7] is a suitable choice, as it can encapsulate diverse attributes of CPS. AADL is designed for modeling system architecture. The Society of Automotive Engineers (SAE) released the AADL in November 2004 [8]. AADL can design system architecture and analyzes the time property, reliability, efficiency, and some other properties. According to the principle of MDA (model-driven architecture), a complete model can reduce the risk of consistency and security of a system. Modelica [9] is a multi-domain modeling language. This language is put forward in 1997 with a group of international efforts. Modelica is object-oriented, a causal-modeling, and equation-based language. With an object-oriented property, model reusability can be improved. Modelica describes the physical world in a direct way. In the period of model checking, Modelica can analyze the dynamic change of a variable in continuous time. AADL cannot model the continuous time properties for a system, while Modelica can do this. AADL focuses on design system architecture. In a system architecture, there are many components. The consistency among components can be checked. But in a component, changes with continuous time cannot be visualized. For this reason, we can use Modelica to model these components.

Integrating the descriptive power of AADL models with the analytic and computational power of Modelica models provides a capability that is significantly greater than that provided by AADL or Modelica individually. A transformation of AADL into Modelica is developed that will support implementations to transfer efficiently the modeling information between AADL and Modelica models without ambiguity. AADL and Modelica are two complementary languages supported by two active communities. By integrating AADL and Modelica, we combine the very

expressive, formal language for differential algebraic equations and discrete events of Modelica with the very expressive AADL constructs for requirements, structural decomposition, logical behavior, and corresponding crosscutting constructs. In addition, the two communities are expected to benefit from the exchange of multi-domain model libraries and the potential for improved and expanded commercial and open-source tool support.

In this chapter we propose an approach to transform the models of AADL into the models of Modelica, to clarify the transformation principles, and to illustrate the important synergies resulting from the integration between these two languages. Based on this transformation mechanism, the properties of AADL components can be checked with the Modelica tool, which validates the security and consistency of architecture.

## **127.2 Models of AADL and Modelica**

Many CPS applications are systems-of-systems, integrating various mechanical, electronic, and information technology systems. The design of these systems depends more and more on effective solutions that can address heterogeneity and interplay of physical and software elements. In particular, design languages used for specifying CPS should incorporate, in a consistent manner, essential concepts from multiple disciplines, such as mechanical, electronic, and software engineering. Model-driven engineering (MDE) approaches to system development have been adopted in diverse domains, in particular, CPS. This is because the use of models has shown to be promising in addressing the above issues, as well as in handling the increasing complexity of CPSs, reducing their cost of construction, and supporting efficient maintenance and evolution. AADL and Modelica are two complementary languages supported by two active communities. By integrating AADL and Modelica, we combine the very expressive, formal language for differential algebraic equations and discrete events of Modelica with the very expressive AADL constructs for requirements, structural decomposition, logical behavior, and corresponding crosscutting constructs. Although both Modelica and AADL have extension mechanisms, in unifying the two languages, we can better use AADL to model the cyber part of the cyber-physical system. This is because its dynamic analysis mechanism comes from Modelica, which is much more widespread and better supported by tools and which is also more powerful compared to that of AADL (annexes).

### ***127.2.1 AADL Introduction***

AADL is a design for architecture. The fundamental element of AADL is component. AADL provides standard components for modeling system architecture.

**Table 127.1** AADL components and layers

AADL component layer	Label	Component	Description
Application software	1	Thread	An active component, can be initialized in process
	2	Thread group	An abstraction for thread, data and thread group
	3	Process	Represents the protected address
	4	Data	Represents statics data and data types within a system
	5	Subprogram	Represents executable source text-a callable component with or without parameters
Execution platform	6	Processor	Responsible for scheduling and executing threads
	7	Memory	Storage components for data and executable code
	8	Device	Entities that interface with the external environment of an application system
	9	Bus	Represents hardware and associated communication protocols that enable interactions among other execution platform components
Composite	10	System	Represents a composite of software, execution platform, or system components

These components are divided into three layers: application software, execution platform, and composite. Each layer has its focus. The application software layer is used to build the software architecture. The execution platform is used to build the hardware architecture. The composite is used to integrate the software and hardware architecture. Each layer has its components.

The detailed statements about layers and components are in Table 127.1. AADL components are divided into three categories: application software, execution platform, and composite. Each component has its description for different utilizations.

### 127.2.2 Modelica Introduction

Modelica is a multi-domain modeling language for the physical world. Modelica has three features: object-oriented modeling, a casual modeling, and equation-based modeling [3]. With these features, Modelica can reduce the complexities of modeling. Models can be built for inheriting from its father. The behaviors in a model can be described with equations and physical principles.

Modelica model is formed with classes. Modelica classes contains: class, model, record, block, function, connector, type, and package. The class “class” is a non-special class while other classes are special classes.

## 127.3 Transforming Mechanism

The mechanism is based on projecting AADL components into Modelica classes. Not all classes in Modelica can be used. Firstly, analyze the features of each AADL component. Then find a proper class in Modelica that can represent this AADL component. Check the relationship of this component. If the type of component is Type, it can be directly transformed into Modelica class. If the type of component is Implementation, transform this component into Modelica class and replace the relation between Type component and Implementation component with extents in Modelica.

The project of AADL package can be transformed into Modelica project package for the hierarchy similarity of the two kinds of packages.

### 127.3.1 Project Structure Transformation

The AADL project structure is in a package. The components of a system are contained in a package. Packages are contained in a project.

The Modelica project structure is in a package. In the Modelica project, there is a package. But a package can use classes from other packages with importing these packages.

The project projection between AADL and Modelica can be carried out in a package level, which means transforming an AADL package into a Modelica package. In this level, the private components in AADL package may be public classes in Modelica.




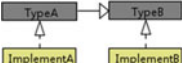


### 127.3.2 Inheritance Transformation

Modelica is object oriented, so the class has inheritance. In AADL, there is some sense of inheritance between components. The transformation of inheritance relationship is listed in Table 127.2. In AADL, the same kind of components can have an inheritance relationship. In Modelica, the same kind of classes can have an inheritance relationship.

### 127.3.3 AADL Components Projected into Modelica Classes

AADL components can be projected into Modelica classes. After the analysis of AADL components and Modelica classes, a transformation table is listed in Table 127.3. While mapping components of AADL into Modelica classes, there may be some missed information. According to Table 127.3, AADL components can be transformed into Modelica classes.

**Table 127.2** The mapping between AADL and Modelica

Label	Modelica class relationship	AADL component relationship	Description
1			In AADL, the <i>Type</i> component define features and implement complement implements <i>Type</i> component
2			The relationship of component extension can be represented with Modelica class extension
3			AADL component instantiation can be represented by Modelica class instantiation

**Table 127.3** The projection from AADL components into Modelica classes

Label	AADL components	Modelica classes	Projection pair
1	Thread	Model	(1, 5)
2	Thread group	Class	(2, 2)
3	Process	Model	(3, 1)
4	Data	Record and type	(4, [3, 7])
5	Sub program	Model	(5, 1)
6	Processor	Model	(6, 1)
7	Memory	Model	(7, 1)
8	Device	Model	(8, 1)
9	Bus	Connector	(9, 1)
10	System	Model	(10, 1)

**Table 127.4** Keywords transformation from AADL into Modelica

Label	AADL keywords	Modelica keywords
1	Extends	Extends
2	Features	Parameter
3	Flows	Connect
4	Properties	Equation
5	Packages	Packages
6	Implementation	Extends
7	Port	Connector
8	Connections	Connect
9	Subcomponents	Declared as variable
10	Modes	Transformed as comments in Modelica

### 127.3.4 Keywords Transformation

The keywords in AADL are different from those in Modelica. So, we need to transform the keywords in AADL into Modelica. The transformation table is in Table 127.4.

The port in AADL is transformed as connector class in Modelica.

## 127.4 Conclusion

By integrating AADL and Modelica, we combine the very expressive, formal language for differential algebraic equations and discrete events of Modelica with the very expressive AADL constructs for requirements, structural decomposition, logical behavior, and corresponding crosscutting constructs. In this chapter we propose an approach to transform the models of AADL into the models of Modelica, to clarify the transformation principles, and to illustrate the important synergies resulting from the integration between these two languages. Based on this transformation mechanism, the properties of AADL components can be checked with a Modelica tool, which validates the security and consistency of architecture.

In the future work, we will work on implementing the transformation tool from AADL to Modelica, which makes the transformation from AADL models to Modelica automatic.

**Acknowledgments** This work is supported by the Shanghai Knowledge Service Platform Project (No. ZF1213). This work is supported by the National High Technology Research and Development Program of China (No.2011AA010101); National Basic Research Program of China (No.2011CB302904); the National Science Foundation of China under grant No.61173046, No.61021004, and No.61061130541; Doctoral Program Foundation of Institutions of Higher Education of China (No. 200802690018); and National Science Foundation of Guangdong Province under grant No. S2011010004905.

## References

1. Broy, M. (2012). Cyber physical systems (Part 1). *it—Information Technology*, 54(6), 255–256.
2. Broy, M. (2013). Cyber physical systems (Part 2). *it—Information Technology*, 55(1), 3–4.
3. Lee, E. A. (2008). Cyber physical systems: Design challenges. In *Proceedings of the 11th IEEE Symposium on Object/Component/Service-Oriented Real-Time Distributed Computing (ISORC '08)* (Vol. 100, Part 1, pp. 363–369).
4. Broman, D., Lee, E. A., Tripakis, S., & Törngren, M. (2012). Viewpoints, formalisms, languages, and tools for cyber-physical systems. *Proceedings of the 6th international workshop on multi-paradigm modeling (MPM'12)*, ACM SIG (pp. 56–63).
5. Eidson, J., Lee, E. A., Matic, S., Seshia, S. A., & Zou, J. (2012). Distributed real-time software for cyber-physical systems. *Proceedings of the IEEE (Special Issue on CPS)*, 100(1), 45–59.
6. Junjie, T., et al. (2012). Cyber-physical systems modeling method based on Modelica. In *2012 I. E. sixth international conference on software security and reliability companion (SERE-C)* (pp. 188–191).
7. Feiler, P. H., Gluch, D. P., & Hudak, J. H. (2006). *The architecture analysis & design language (AADL): An introduction*. Pittsburgh, PA: Software Engineering Institute, Carnegie-Mellon University.
8. Feiler, P. H., Lewis, B. A., & Vestal, S. (2006). The SAE Architecture Analysis & Design Language (AADL) a standard for engineering performance critical systems. In *Computer Aided Control System Design, 2006 I.E. International Conference on Control Applications, 2006 I.E. International Symposium on Intelligent Control* (pp. 1206–1211).
9. Fritzson P, & Modelica, E. V. (1998). A unified object-oriented language for system modeling and simulation. In *ECOOP'98-Object-Oriented Programming* (pp. 67–90). Heidelberg: Springer.



# Chapter 128

## Aspect-Oriented QoS Modeling of Cyber-Physical Systems by the Extension of Architecture Analysis and Design Language

Lichen Zhang and Shuguang Feng

**Abstract** Cyber-physical systems have varying quality-of-service (QoS) requirements driven by the dynamics of the physical environment in which they operate. Developing cyber-physical systems is hard because of their end-to-end QoS requirements. Aspect-oriented development method can decrease the complexity of models by separating their different concerns. We can model QoS as a crosscutting concern of cyber-physical systems to reduce the complexity of cyber-physical system development. In this paper, we propose an aspect-oriented QoS modeling method based on AADL. We present our current effort to extend AADL to include new features for separation of concerns, and we make an AADL extension for QoS by aspect-oriented method. Finally, we illustrate QoS aspect-oriented modeling via an example of transportation cyber-physical system.

### 128.1 Introduction

The very recent development of cyber-physical systems (CPS) provides a smart infrastructure for connecting abstract computational artifacts with the physical world. As new CPS applications start to interact with the physical world using sensors and actuators, there is a great need for ensuring that the actions initiated by the CPS are timely. This will require new quality-of-service (QoS) functionality and mechanisms for CPS [1]. Cyber-physical systems are characterized by their stringent requirements for QoS, such as predictable end-to-end latencies, timeliness, and scalability. Delivering the QoS needs of cyber-physical systems entails the need to specify and analyze QoS requirements correctly.

Cyber-physical systems share characteristics giving rise to tangled concerns in their development and maintenance lifecycle [2]. The characteristics must

---

L. Zhang (✉) • S. Feng  
Shanghai Key Laboratory of Trustworthy Computing, East China Normal University,  
Shanghai 200062, China  
e-mail: [zhanglichen1962@163.com](mailto:zhanglichen1962@163.com)

simultaneously support. Distributed, Real-Time, and Embedded “software controllers are increasingly replacing mechanical and human control of critical systems. These controllers must simultaneously support many challenging QoS constraints, including (1) real-time requirements, such as low latency and bounded jitter, (2) availability requirements, such as fault propagation/recovery across boundaries, (3) security requirements, such as appropriate authentication and authorization, and (4) physical requirements, such as limited weight, power consumption, and memory footprint. For example, a distributed patient monitoring system requires predictable, reliable, and secure monitoring of patient health data that can be distributed in a timely manner to healthcare providers.” [2]

Fundamental limitations for CPS include [3]:

- Lack of good formal representations and tools capable of expressing and integrating multiple viewpoints and multiple aspects
- Lack of strategies to cleanly separate safety-critical and non-safety-critical functionality, as well as for safe composition of their functionality during human in-the-loop operation
- Ability to reason about, and trade off between, physical constraints and QoS of the CPS [4]

Aspect-oriented programming (AOP) [5] is a new software development technique, which is based on the separation of concerns. Systems could be separated into different crosscutting concerns and designed independently by using AOP techniques. Every concern is called an “aspect.”

As the QoS concern needs to be considered as the most parts of the system, it is a crosscutting concern. Crosscutting concerns [6] are concerns that span multiple objects or components. Crosscutting concerns need to be separated and modularized to enable the components to work in different configurations without having to rewrite the code. By using AOP, concerns can be modularized in an aspect and later weaved into the code. The QoS of CPS [7] is very complex; currently QoS research still does not have a completely technical system, and there is no solution meeting all the QoS requirements.

This paper proposes an aspect-oriented QoS modeling method based on Architecture Analysis and Design Language (AADL) [8]. In this paper, we present our current effort to extend AADL to include new features for separation of concerns. We make an in-depth study of AADL extension for QoS. Finally, we illustrate QoS aspect-oriented modeling via an example of transportation cyber-physical system.

## 128.2 Related Works

Developing cyber-physical systems is hard since it requires a coordinated, physics-aware allocation of CPU and network resources to satisfy their end-to-end QoS requirements. Jaiganesh Balasubramanian et al. make two contributions to address these challenges.

The development of Distributed, Real-Time, and Embedded (DRE) systems is often a challenging task due to conflicting QoS constraints that must be explored as trade-offs among a series of alternative design decisions. Jeff Gray et al. present a model-driven approach for generating QoS adaptation in DRE systems.

Carsten Köllman, Lea Kutvonen, Peter Linington, and Arnor Solberg present an approach for managing several dependability dimensions [9]. They use aspect-oriented and model-driven development techniques to separate and construct QoS independent models and graph-based transformation techniques to derive the corresponding QoS specific models.

QoS-UniFrame classifies quantifiable QoS requirements into static and dynamic. *Static QoS* is design-related, whereas *dynamic QoS* is substantially influenced by the deployment environment.

Bikash Sabata et al. specify QoS as a combination of metrics and policies. QoS metrics are used to specify performance parameters, security requirements, and the relative importance of the work in the system. They define three types of QoS performance parameters: timeliness, precision, and accuracy.

Mohammad Mousavi et al. present an extension to the GAMMA formalism, which they name AspectGAMMA [10], and they show how non-computational aspects can be expressed separately from the computation in this framework.

Dionisio de Niz and Peter H. Feiler discuss their effort to extend the AADL to include new features for separation of concerns [11]. These features include not only constructs to describe design choices but also routines to verify the proper combination of constructs from different concerns.

Lydia Michotte, Thomas Vergnaud, Peter H. Feiler, and Robert B. France extended the AOM approach to support the separation of crosscutting concerns in component architectures using AADL components [12].

AO4AADL is an aspect-oriented extension for AADL [13]. This language considers aspects as an extension concept of AADL components called aspect annex. Instead of defining a new aspect-oriented ADL, they extend AADL, a well-known ADL, with an aspect annex. Therefore, they consider, in their work, that aspects can be specified in a language other than AADL and then integrated in AADL models as annexes.

Ana-Elena Rugina, Karama Kanoun, and Mohamed Kaâniche proposed a four-step modeling dependability method based on AADL [14].

### 128.3 The Extension of AADL by Aspect-Oriented Method

In its conformity to the ADL definition, AADL provides support for various kinds of nonfunctional analyses along with conventional modeling:

*Flow latency analysis*: Understands the amount of time consumed for information flows within a system, particularly the end-to-end time consumed from a starting point to a destination.

**Fig. 128.1** Property sets of AADL

```

property set Clemson is
MbitPerSec : type units (MPS, GPS => MPS*1000);
Band_width: type aadlinteger units Clemson::MbitPerSec;
Radio_band_width: Clemson::Band_width applies to (all);
Band_width_802_11g: constant Clemson::Band_width => 54 MPS;
Band_width_802_11n: constant Clemson::Band_width => 300 MPS;
Band_width_fast_ethernet: constant Clemson::Band_width => 100 MPS;
end Clemson;

```

*Resource consumption analysis:* Allows system architects to perform resource allocation for processors, memory, and network bandwidth and analyze the requirements against the available resources.

*Real-time schedulability analysis:* AADL models bind software elements such as threads to hardware elements like processors. Schedulability analysis helps in examining such bindings and scheduling policies.

*Safety analysis:* Checks the safety criticality level of system components and highlights potential safety hazards that may occur because of communication among components with different safety levels.

*Security analysis:* Like safety levels, AADL components can be assigned various security levels.

AADL defines two main extension mechanisms: property sets, as shown in Fig. 128.1, and sublanguages (known as annexes). Annexes and properties allow the addition of complex annotations to AADL models that accommodate the needs of multiple concerns.

In this paper, we extend AADL by aspect-oriented method in the following aspect:

*Physical world aspect:* Cyber-physical systems are often complex and span multiple physical domains, whereas mostly these systems are computer-controlled.

*Dynamic continuous aspect:* Cyber-physical systems are mixtures of continuous dynamic and discrete events. These continuous and discrete dynamics not only coexist but also interact, and changes occur both in response to discrete, instantaneous events and in response to dynamics as described differential or difference equations in time.

*Formal specification aspect of data:* A formal specification aspect of data captures the static relation between the object and data. Formal data aspect emphasizes the static structure of the system using objects, attributes, operations, and relationships based on formal techniques.

*Formal specification aspect of information flow and control flow:* Formal specification aspect of information flow and control flow aims at facilitating the description and evaluation of various flow properties measures.

*Spatial aspect:* The analysis and understanding of railway cyber-physical systems' spatial behavior—such as guiding, approaching, departing, or coordinating movements—is very important.

### 128.4 Case Study: Aspect-Oriented Specification of QoS of Lunar Rover

A lunar rover or moon rover is a space exploration vehicle designed to move across the surface of the moon, as shown in Fig. 128.2.

We use the AADL to model the software part of the lunar rover. The whole system is split into measurement system, control system, and perform system. The data stream of the control system of the lunar rover is shown in Fig. 128.3.

The specification of the delays of control system stream is

```
f1:flow source AttitudeData{Latency => 20 Ms; };
f2:flow source HeadingData{Latency => 20 Ms; };
f3:flow source LocationData{Latency => 20 Ms; };
f1:flow sink SurveyData{Latency => 20 Ms; };
f2:flow sink SurveyData{Latency => 20 Ms; };
f3:flow sink SurveyData{Latency => 20 Ms; };
```



Fig. 128.2 Lunar rover

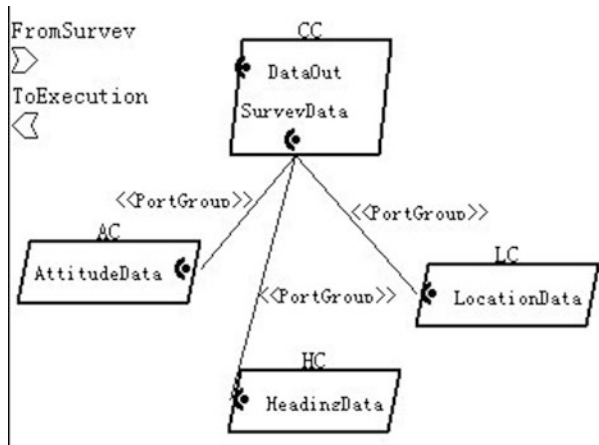


Fig. 128.3 Control system stream

Infos (6 items)	
i	End-to-end flow AC2CC_E2E calculated latency (ASynchronous)40.0 ms is less than expected latency 100.0 ms
i	End-to-end flow AC2CC_E2E calculated latency (Synchronous)40.0 ms is less than expected latency 100.0 ms
i	End-to-end flow HC2CC_E2E calculated latency (ASynchronous)40.0 ms is less than expected latency 100.0 ms
i	End-to-end flow HC2CC_E2E calculated latency (Synchronous)40.0 ms is less than expected latency 100.0 ms
i	End-to-end flow LC2CC_E2E calculated latency (ASynchronous)40.0 ms is less than expected latency 100.0 ms
i	End-to-end flow LC2CC_E2E calculated latency (Synchronous)40.0 ms is less than expected latency 100.0 ms

**Fig. 128.4** Control system stream analysis

In the flow implementation, we define the end-to-end delays as follows:

```

AC2CC_E2E: end to end flow AC . f1->AC2CC->CC . f1 {
  Latency => 100 Ms ;
} ;
HC2CC_E2E: end to end flow HC . f2->HC2CC->CC . f2 {
  Latency => 100 Ms ;
} ;
LC2CC_E2E: end to end flow LC . f3->LC2CC->CC . f3 {
  Latency => 100 Ms ;
} ;

```

We use the OSATE tool of AADL to make flow analysis; we obtain the analysis results, as shown in Fig. 128.4.

## 128.5 Conclusion

This paper proposed an aspect-oriented QoS modeling method based on AADL. Aspect-oriented development method can decrease the complexity of models by separating their different concerns. In model-based development of cyber-physical systems, this separation of concerns is more important given the QoS concerns addressed by cyber-physical systems. These concerns can include timeliness, fault tolerance, and security. Architecture Analysis and Design Language (AADL) is a standard architecture description language used to design and evaluate software architectures for embedded systems already in use by a number of organizations around the world. In this paper, we presented our current effort to extend AADL to include new features for separation of concerns and made an in-depth study of AADL extension for QoS. Finally, we illustrated QoS aspect-oriented modeling via an example of the specification of VANET based on AADL.

The future work focuses on the integration of AADL and formal techniques to specify and verification of QoS of cyber-physical systems.

**Acknowledgements** This work is supported by Shanghai Knowledge Service Platform Project (No. ZF1213), national high-technology research and development program of China (No. 2011AA010101), national basic research program of China (No. 2011CB302904), the national science foundation of China under grant (Nos. 61173046, 61021004, 61061130541, 91118008), doctoral program foundation of institutions of higher education of China (No. 20120076130003), and national science foundation of Guangdong province under grant (No. S2011010004905).

## References

1. Dillon, T., Potdar, V., Singh, J., & Talevski, A. (2011). Cyber-physical systems -providing quality of service (QoS) in a heterogeneous systems-of-systems environment. In *IEEE international conference on digital ecosystems and technologies* (pp. 330–335), Daejeon.
2. Gokhale, A., & Gray, J. (2005). (2005). An integrated aspect-oriented model-driven development toolsuite for distributed real-time and embedded systems. In *Workshop on aspect-oriented modeling workshop, held at AOSD 2005* (pp. 20–26). Chicago, IL: IEEE Computer Society.
3. Wolf, W. (2009). Cyber-physical systems. *Computer.*, 42(3), 88–89.
4. Lee, E. A. (2008). Cyber physical systems design challenges. In *11th IEEE international symposium on object oriented real-time distributed computing (ISORC), 2008* (pp. 363–369), Orlando, FL.
5. Kiczales, G., et al. (1997). Aspect-oriented programming. In *Proceedings of ECOOP, LNCS* (Vol. 1241, pp. 220–242). Heidelberg: Springer.
6. Wehrmeister, M. A., Freitas, E. P., Pereira, C. E., et al. (2007) An aspect-oriented approach for dealing with non-functional requirements in a model-driven development of distributed embedded real-time systems. In *Tenth IEEE international symposium on object and component-oriented real-time distributed computing* (pp. 428–432). Santorini Island, Greece: IEEE Computer Society.
7. Frolund, S., & Koistinen, J. (1998). Quality of service specification in distributed object systems. *IEE/BCS Distributed Systems Engineering Journal.*, 5(4), 179–202.
8. AE Aerospace. (2009). SAE AS5506A: Architecture analysis and design language V2.0.
9. Köllman, C., Kutvonen, L., Linington, P., & Solberg, A. (2007). An aspect-oriented approach to manage QoS dependability dimensions in model driven development. In *International workshop on model-driven enterprise information systems* (pp. 85–94), Vienna.
10. Mousavi, M. R., Russello, G., Chaudron, M., Reniers, M., Basten, T., Corsaro, A., et al. (2002). Using aspect-GAMMA in the design of embedded systems. In *Proceedings of the seventh IEEE international workshop on high level design, verification and test (HLDVT'02)* (pp. 69–75), Cannes, France.
11. de Niz, D., & Feiler, P. H. (2007). Aspects in the industry standard AADL. In *AOM '07 proceedings of the tenth international workshop on Aspect-oriented modeling* (pp. 15–20), Nashville.
12. Michotte, L., Vergnaud, T., Feiler, P., & France, R. (2008). Aspect oriented modeling of component architectures using AADL. In *Proceedings of the second international conference on new technologies, mobility and security* (pp. 5–7), Auckland.
13. Loukil, S., Kallel, S., Zaliila, B., & Jmaiel, M. (2010). Toward an aspect oriented ADL for embedded systems. In *The fourth European conference on software architecture (ECSA 2010), LNCS 6285* (pp. 489–492). Copenhagen: Springer.
14. Rugina, A.-E., Kanoun, K., & Kaaniche, M. (2007). An architecture-based dependability modeling framework using AADL. In *Proceedings of tenth IASTED international conference on software engineering and applications* (pp. 222–227), Dallas, USA.

# Chapter 129

## Using RC4-BHF to Construct One-way Hash Chains

Qian Yu and Chang N. Zhang

**Abstract** Cryptographic hash functions play a fundamental role in today's security applications. In general terms, the principal applications of a cryptographic hash function are to verify the integrity of the data, which refers to data authentication or data integrity. The one-way hash chain is an important topic in key management and is also an important cryptographic primitive in many security applications. As one-way chains are very efficient to verify, they are also the primitives to design security protocols for ultra-low-power devices. In this chapter, an RC4-based hash function RC4-BHF is introduced and how to use RC4-BHF to construct efficient one-way hash chains is proposed. The proposed construction for one-way hash chains is efficient and is designed for ultra-low-power devices.

### 129.1 Introduction

Cryptographic hash functions play a fundamental role in today's security applications. Hash functions take a variable-sized message as input and produce a small fixed-sized string as output [1]. Hash functions are indispensable for a variety of security applications, and the principal application is to verify the integrity of a message, which refers to data authentication or data integrity. In addition, cryptographic hash functions can be used to one-way hash chain, password file generation, intrusion and virus detection, pseudorandom function, and pseudorandom number generator.

Examples of well-known cryptographic hash functions are SHA-family [2, 3], MD4 [4], MD5 [5], and RIPEMD [6]. Analysis has demonstrated that weaknesses are found in some well-known hash functions [2]. Reported a collision in SHA-0 and [7, 8] reported collisions in MD4, MD5, SHA-1, HAVAL-128, and RIPEMD.

---

Q. Yu (✉) • C.N. Zhang  
Department of Computer Science, University of Regina, 3737 Wascana Parkway,  
Regina, SK, Canada S4S 0A2  
e-mail: [yu209@cs.uregina.ca](mailto:yu209@cs.uregina.ca)



The weaknesses may compromise the security of the applications in which those functions are being used.

Therefore, it is really needed to design some hash functions with totally different internal structures from the broken classes. Furthermore, the emerging ultra-low-power technology sets new challenges for cryptographic algorithms and applications because their resources are limited. The traditional cryptographic hash functions are not well suited to the resource-limited environment, and therefore a hash function designed for ultra-low-power devices is also really needed. RC4-BHF [9] is such a new hash function, which is light-weight, structurally different from the broken hash function classes, and is able to reuse existing RC4 hardware to implement. RC4-BHF is very simple and efficient, designed to run on an eight-bit processor, and rules out most major generic attacks of hash functions, which enables it to run on ultra-low-power devices.

The one-way chain is an important cryptographic primitive in many security applications. In key management, a hash chain is a method to produce many one-time keys or session keys from a single key. For non-repudiation, a hash function can be applied successively to additional pieces of data in order to record the chronology of data's existence [10]. Furthermore, as one-way chains are very efficient to verify, they became the primitive to design security protocols for ultra-low-power devices, such as sensor nodes, as their low-powered processors can compute a one-way function within milliseconds, but would require tens of seconds or up to minutes to generate or verify a traditional digital signature [11], or even impossible to verify the traditional digital signature. Based on RC4-BHF, the proposed construction for one-way hash chains is very efficient and is designed for ultra-low-power devices.

The following is the outline of the rest of the chapter. Sections 129.2 and 129.3 cover the basic knowledge of cryptographic hash function and RC4 stream cipher, respectively. Section 129.4 introduces RC4-BHF and Sect. 129.5 proposes the construction for one-way hash chains. Section 129.6 provides the analysis and Sect. 129.7 provides the conclusion of this chapter.

## 129.2 Hash Function

A cryptographic hash function  $H$  is a transformation that accepts a variable-length block of data  $M$  as input and produces a fixed-size hash value  $h = H(M)$ . The following characteristics are required in a cryptographic hash function:

- The input of  $H$  can be of any length.
- The output of  $H$  has a fixed length.
- $H(x)$  is relatively easy to compute for any given input, making both hardware and software implementations practical.
- $H(x)$  is a one-way mapping, which means that for any given value  $h$ , it is computationally infeasible to find  $x$  such that  $H(x) = h$ .

- $H(x)$  is weak collision resistance, which means that for any given block  $x$ , it is computationally infeasible to find  $y \neq x$  with  $H(y) = H(x)$ .
- $H(x)$  is strong collision resistance, which means it is computationally infeasible to find any pair  $(x,y)$  such that  $H(x) = H(y)$ .

Cryptographic hash function is used in a wide variety of applications. In general terms, the principal application of a cryptographic hash function is to verify the integrity of a message, which refers to data authentication or data integrity. Message authentication protects from unauthorized data alteration to assure that the data received are exactly as sent. Two common applications of the cryptographic hash function are message authentication code and digital signature. Following are the security attacks for hash function [12]. We say the hash function is resistant to this attack if it is hard to find one of the attacks, or we say the hash function is resistant to these attacks if it is hard to find all of the attacks [13].

- Collision Attack: can find  $M_1 \neq M_2$ , but  $H(M_1) = H(M_2)$
- Preimage Attack: Given a random  $y$ , can find  $M$  that  $H(M) = y$
- Second Preimage Attack: Given  $M_1$ , can find  $M_2$  that  $H(M_1) = H(M_2)$

### 129.3 RC4 Stream Cipher

The stream cipher is an important class of cryptographic algorithms and they encrypt each bit or byte of plaintext one at a time, using a simple time-dependent encryption transformation. Stream ciphers are almost always faster and use far less code than block ciphers, and RC4 is the most widely used stream cipher nowadays because of its high efficiency and simplicity [14, 15]. RC4 has been selected as the encryption algorithm in some security protocols [16, 17] for ultra-low-power devices.

RC4 is a variable key-size stream cipher. It is based on a 256-byte internal state  $S$  and two 1-byte indexes  $i$  and  $j$ . RC4 consists of two algorithms that are key-scheduling algorithm (KSA) and pseudo-random generation algorithm (PRGA) which are described in Algorithm 1 and Algorithm 2, respectively. For a given RC4 base key, KSA generates an initial 256 bytes permutation state. This permutation state is the input to PRGA. PRGA is a repeated loop procedure and each loop generates a 1-byte output as the stream key which to XOR with 1-byte of the plaintext, in the meantime a new 256-byte permutation state  $S$  and two 1-byte indexes  $i$  and  $j$  are getting updated. We call  $(S, i, j)$  is an RC4 state.

```

for i from 0 to 255
    S[i] := i
    T[i] := K[i mod keylength]
endfor
j := 0
for i from 0 to 255
    j := (j + S[i] + T[i]) mod 256
    swap values of S[i] and S[j]
endfor

```

Algorithm 1: KSA

```

i := 0
j := 0
while GeneratingStreamOutput:
    i := (i + 1) mod 256
    j := (j + S[i]) mod 256
    swap values of S[i] and S[j]
    Output k := S[(S[i] + S[j]) mod 256]
endwhile

```

Algorithm 2: PRGA

There are many papers to analyze the security strength of RC4, but none is practical against RC4 with a reasonable key length, such as 128 bytes [15]. So far, the practical attacks (e.g., [17–20]) against RC4 applications remain with WEP attacks, which aim to a key derivation problem in WEP standard [20]. Essentially, this attack is not in RC4 itself, it is in the scheme how to generate the secure keys. This particular problem does not appear to be applicable to other applications using RC4 and can be avoided in WEP by changing the scheme on how to generate the secure keys.

## 129.4 RC4-BHF

In this section we introduce RC4-BHF, which is an RC4-based hash function. The proposed construction for one-way hash chains is based on it. RC4-BHF includes three steps: padding and dividing step, compression step, and truncation step. Two algorithms  $KSA^*$  and  $PRGA^*$  are given below in Algorithm 3 and Algorithm 4.  $KSA$  and  $PRGA$  are already introduced in Algorithm 1 and Algorithm 2. The input of RC4-BHF can be of any length, but equal or less than  $2^{64}$  bits long, and the output is 128-bit or 256-bit long. The internal state of RC4-BHF is 256-byte long.

```

Input:  $M_k$  and  $STATE_k$ 
Output:  $STATE_{Mk}$ 
for i from 0 to 255
    j := (j + S[i] +  $M_k[i \text{ mod } 64]$ ) mod 256
    swap values of S[i] and S[j]
endfor

```

Algorithm 3:  $KSA^*$ 

```

Input: len and  $STATE_{Mk}$ 
Output:  $STATE_k$ 
for i from 0 to len
    i := (i + 1) mod 256
    j := (j + S[i]) mod 256
    swap values of S[i] and S[j]
endfor

```

Algorithm 4:  $PRGA^*$ 

### Step 1: Padding and Dividing Step

The input of the padding and dividing process is the plain message, and the outputs are 512-bit data blocks. The input message is padded as: *Padded Message* := M

**Fig. 129.1** The padding process



$\| 1 \| 00 \dots 0 \| L$  where  $M$  is the input message of  $N$  bits long,  $\|$  is the concatenation symbol, the number of zero  $v$  is the least non-negative integer in which  $N + 65 + v \equiv 0 \pmod{512}$ , and  $L$  is a 64-bit data segment indicating the length of message  $M$ . The padding process makes the padded message a multiple of 512 bits in length. Figure 129.1 illustrates the padding process. For dividing process, the padded message is divided into 512-bit data blocks, notated by  $m_1, m_2, \dots, m_n$ .

**Step 2: Compression Step**

The first 512-bit data block  $m_1$  and offset integer are input to initialize the internal state  $S$ :  $State_{m_1} = PRGA^*(offset, KSA(m_1))$ , then the function  $PRGA^*$  modifies the internal state depending on the length  $len_1$  where

$$len_1 = \begin{cases} m_1 \bmod 2^5 & \text{if } (m_1 \bmod 2^5) \neq 0 \\ offset & \text{if } (m_1 \bmod 2^5) = 0 \end{cases}$$

$State_1 = PRGA^*(len_1, State_{m_1})$ .

For  $k$  ( $k = 2, 3 \dots n$ ),  $S$  are updated as below:

$State_{m_k} = KSA^*(m_k)$ ,  $State_k = PRGA^*(len_k, State_{m_k})$

where  $len_k = \begin{cases} m_k \bmod 2^5 & \text{if } (m_k \bmod 2^5) \neq 0 \\ offset & \text{if } (m_k \bmod 2^5) = 0 \end{cases}$ .

The compression process is illustrated in Fig. 129.2. The number of rounds of  $PRGA^*$  is controlled by  $len_k$ .

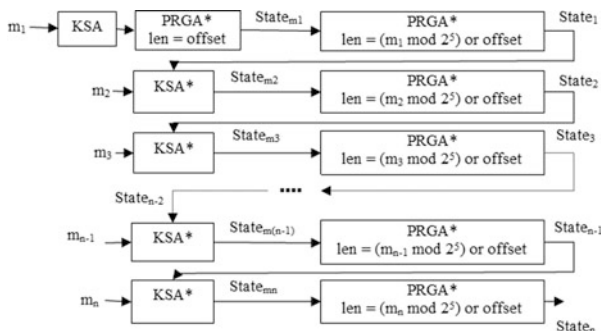
**Step 3: Truncation Step**

The input of the truncation process is 256-byte  $State_n$ , which is the output of the compression process, and the output of the truncation process, which is the final hash value of the hash function, is 128 or 256 bits long.

The detail of the truncation process is that it applies  $PRGA$  to generate 512 bytes output in total, and discards the first 256 bytes and only uses the last 256 bytes in calculation.  $State_n$  XOR the last 256 bytes of the  $PRGA$  output to get a 256 bytes data  $H$ .  $RC4\text{-BHF}$  does not define how to input to applying  $PRGA$  so for different applications it can define different input, such as a new random state, or a defined internal state in the compression process.

In order to reduce the overhead, many schemes can be used to reduce the size of the final hash value. Two options have been adopted in  $RC4\text{-BHF}$  in the truncation process. The first option is to select the least significant bit of each byte of  $H$ , and the final hash value is 256-bit value long. The second option is to select the least significant bit of each even or odd number byte of  $H$ , and the 128-bit value is the final hash value.

**Fig. 129.2** The compression process of RC4-BHF



### 129.5 Proposed Construction for One-way Hash Chains

The one-way chain is an important cryptographic primitive in many security applications. In key management, some applications require key change frequently. A hash chain is a method to produce many one-time keys or session keys from a single key. For non-repudiation, a hash function can be applied successively to additional pieces of data in order to record the chronology of data's existence [10]. Furthermore, as one-way chains are very efficient to verify, they became the primitive to design security protocols for ultra-low-power devices, such as sensor nodes, as their low-powered processors can compute a one-way function within milliseconds, but would require tens of seconds or up to minutes to generate or verify a traditional digital signature [11], or even impossible to verify due to the ability of the device and the resource required by the security algorithms.

A one-way function  $F : K_j = F(K_{j+1})$  can be used to generate a one-way key chain  $(K_n, K_{n-1}, \dots, K_1, K_0)$ . Through this function, anybody can compute forward (e.g., computing  $K_0, \dots, K_j$  from a given  $K_{j+1}, K_0, \dots, K_j$ ) but nobody can compute backward (e.g., computing  $K_{j+1}$  for only given  $K_0, \dots, K_j$ ) in the key chain. RC4-BHF is such a function and can be used to generate the key chain.

The detailed processes of the key chain generation and distribution, as well as of the key self-authentication, are illustrated in the following. Each key generated here is 256-byte long and the key chain is generated by a sender or trusted third party, and a new base key picking up in sequence from the one-way key chain is distributed to receiver(s) when a new key is needed. The sender or trusted third party first generate a random 256-byte permutation state as the last key  $K_n$  of the one-way key chain, and generates the rest of the keys by successively applying RC4-BHF  $F : K_j = F(K_{j+1})$ . Please note that the truncation process is not needed here. By this approach, the key chain is generated in the order of  $K_n \rightarrow K_{n-1} \rightarrow \dots \rightarrow K_1 \rightarrow K_0$ . Here we assume n is large enough that the key chain is sufficiently long in relation to the duration of the data transmission. The distribution order is from  $K_0, K_1, K_2, \dots$  to  $K_n$ .

In the one-way key chain generated as above, the keys are self-authenticated. The receiver can easily and efficiently authenticate subsequent keys of the one-way

key chain using an authenticated key. For example, as soon as a receiver receives a new key  $K_i$ , the receiver can authenticate the new received key by its current key  $K_{i-1}$ . The new received key is verified once the two keys matches by applying the function  $F : K_{i-1} = F(K_i)$ . That is, a receiver can use its current key or any previous keys which are from the key chain to verify the new disclosed key.

In many security applications, the integer number is also needed and will need to be updated when requested. The sizes of the key and integer number are different, so we cannot use RC4-BHF to generate the integer number chain directly. In the following, we illustrate the detailed processes of the integer number chain generation and distribution, as well as of the integer number self-authentication.

The integer number chain is generated by a sender or trusted third party, and a new integer number picked up in sequence from the one-way integer number chain is distributed to receiver(s) when a new integer number is needed. We assume the biggest of the integer number is  $(256^4 - 1)$ . The sender or trusted third party first generates two random integer numbers  $O_n$  and  $O_{n-1}$  as the last two integer numbers of the one-way integer number chain, and generates the rest of the integer numbers by successively applying RC4-BHF  $F : O_j = F[(O_{j+1} + O_{j+2}) \bmod 256^4 \times [(O_{j+1} - O_{j+2}) \bmod 256^4 \times O_{j+1} \times O_{j+2}]$ . Please note that the truncation process will only select the least significant bit of every 8 bytes of  $H$  as the output, which is 32-bit long.

By this approach, the integer number chain is generated in the order of  $O_n \rightarrow O_{n-1} \rightarrow \dots \rightarrow O_1 \rightarrow O_0$ . Here we assume  $n$  is large enough that the integer number chain is sufficiently long in relation to the duration of the data transmission. The distribution order is from  $O_0, O_1, O_2, \dots$  to  $O_n$ .

In the integer number chain generated as above, the integer numbers are self-authenticated. The receiver can easily and efficiently authenticate subsequent integer numbers of the one-way integer number chain using two recent integer numbers. For example, as soon as the receiver receives a new integer number  $O_i$ , the receiver can authenticate the new received integer number by its current integer number  $O_{i-1}$  and most recent integer number  $O_{i-2}$ . The new received integer number is verified once the three numbers match by applying RC4-BHF  $F : O_{i-2} = F[(O_{i-1} + O_i) \bmod 256^4 \times [(O_{i-1} - O_i) \bmod 256^4 \times O_{i-1} \times O_i]$ , that is, the receiver can use any two continuous integer numbers no matter they are current or previous from the integer number chain to verify the new disclosed integer number.

## 129.6 Analysis

RC4-BHF is based on RC4. The security analysis can be made in the view of the security analysis of RC4 which is well studied as well as the resistance to the major hash function attacks: preimage attack, second preimage attack, and collision attack.

The simulation to analyze the randomness of the RC4 state generation on KSA\* and PRGA\* shows that the generation of a new RC4 state by KSA\* and PRGA\* maintains the same randomness of KSA and PRGA. In addition, the truncation process which to generate the final hash value does not reduce the randomness level. Therefore, the generated hash value of RC4-BHF is close to uniform.

The maximum input length of RC4-BHF is  $2^{64}$  bits and the output is fixed 128 or 256 bits. RC4-BHF is relatively easy to compute for any given message. Since the generated hash value of RC4-BHF is close to uniform, it is impossible to find the input through output and it is also computationally infeasible to find any two messages  $x$  and  $y$  such that  $H(x) = H(y)$ . Therefore, RC4-BHF is the one-way mapping and strongly collision-free. RC4-BHF satisfies the requirements of a hash function which listed in Sect. 129.2. Since the generated hash value of RC4-BHF is close to uniform, it is hard to find  $M_1$  and  $M_2$  that  $M_1 \neq M_2$ , but  $H(M_1) = H(M_2)$  and to find  $M_2$  when given a message  $M_1$  to make  $H(M_1) = H(M_2)$ . In addition, RC4-BHF is not reversible, so for a random  $y$  we cannot find an  $M$  that  $H(M) = y$ . In conclusion from the above, RC4-BHF is collision-resistant, preimage-resistant, and second preimage attack-resistant. RC4-BHF rules out the major security attacks of hash function. The compression process of RC4-BHF has output size of about 2,048 bits which is much larger than two times of the size of the hash output. Thus, generic attacks such as Kelsey-Schneier second-preimage attack does not work.

For the performance, RC4-BHF is based on the RC4 structure which is very efficient. We conducted a benchmark simulation to compare the relative speeds among RC4-BHF, MD4, MD5, SHA-1, RIPEMD-128, and RIPEMD-160 on a 32-bit processor. Assuming the required time for one time of memory access is two times of the required time for a logic operation or a simple arithmetic operation. The comparison result shows that RC4-BHF is much faster than the other algorithms.

Because the security of RC4-BHF is sound, the proposed construction for one-way hash chains is secure as it is based on RC4-BHF. We have implemented the one-way key chain and one-way integer number chain and both of them work as expected.

## 129.7 Conclusion

This chapter introduced RC4-BHF, which is an RC4-based hash function, and proposed how to use RC4-BHF to construct efficient one-way hash chains, such as one-way key chain and one-way integer number chain. RC4-BHF is an attempt to use RC4 algorithm to design a cryptographic hash function. The design structure of RC4-BHF is totally different from the broken hash function classes. Moreover, RC4-BHF can be used to ultra-low-power devices, to which most other hash functions cannot be applied. RC4-BHF is very efficient compared to other hash functions, is collision-resistant, preimage-resistant, and second preimage-resistant, and rules out the important generic attacks for hash functions. Based on RC4-BHF,

an efficient construction for one-way hash chains is proposed. One-way hash chains are important cryptographic primitives in many security applications and it is an important topic in key management. As one-way chains are very efficient to verify, they are also the primitives to design security protocols for ultra-low-power devices. We believe that RC4-BHF and the proposed construction for one-way hash chains could be applied to many applications, especially to the ultra-low-power devices which have limited resource and capability.

## References

1. Stallings, W. (2011). *Cryptography and network security, principles and practice* (5th ed.). Upper Saddle River, NJ: Prentice Hall.
2. SHA-0: A federal standard by NIST. NIST. (1993).
3. FIPS 180–1: Secure hash standard. US Department of Commerce, Washington, DC, Springer. (1996).
4. Ronald, L. (1991). Rivest. The MD4 message-digest algorithm. In *Proceedings of the Crypto'1990, LNCS 537* (pp. 303–311), Springer.
5. Ronald, L. (1992). Rivest. The MD5 message-digest algorithm. RFC 1320, Internet Activities Board, Internet Privacy Task Force.
6. RIPE, Integrity Primitives for secure Information systems, Final report of RACE Integrity Primitive Evaluation. LNCS 1040, Springer. (1995).
7. Wang, X., Yu, H., & Yin, Y. L. (2005). Efficient collision search attacks on SHA-0. In *Proceedings of the Crypto'2005, LNCS 3621* (pp. 1–16), Springer.
8. Wang, X., Yin, Y. L., & Yu, H. (2005). Finding collisions in the full SHA-1. In *Proceedings of the Crypto'2005, LNCS 3621* (pp. 17–36), Springer.
9. Yu, Q., Zhang, C. N., Orumiehchiha, M. A., & Li, H. (2012). RC4-BHF: An improved RC4-based hash function. In *Proceedings of the CIT 2012* (pp. 322–326), IEEE CS Press.
10. Hash chain. [http://en.wikipedia.org/wiki/Hash\\_chain](http://en.wikipedia.org/wiki/Hash_chain).
11. Hu, Y. C., Jakobsson, M., & Perrig, A. (2005). Efficient constructions for one-way hash chains. In *Proceedings of the ACNS 2005, LNCS 3531* (pp. 423–441), Springer.
12. Stinson, D. R. (2002). *Cryptography, theory and practice* (2nd ed.). Boca Raton: CRC Press.
13. Chang, D., Gupta, K. C., & Nandi, M. (2006). RC4-Hash: A new hash function based on RC4. In *Proceedings of the INDOCRYPT, LNCS 4329* (pp. 80–94), Springer.
14. Karlof, C., Sastry, N., & Wagner, D. (2004). TinySec: A link layer security architecture for wireless sensor networks. In *Proceedings of the 2nd international conference on Embedded networked sensor systems* (pp. 162–175), ACM.
15. Mantin, I. (2001). *Analysis of the stream cipher RC4*. Master's thesis, The Weizmann Institute of Science, Israel.
16. Yu, Q., & Zhang, C. N. (2010). *A lightweight secure data transmission protocol for resource constrained devices. Security and communication networks*. Wiley Press, 3(5), 362–370.
17. Mitchell, S., & Srinivasan, K. (2004). State based key hop protocol: A lightweight security protocol for wireless networks. In *Proceedings of the MSWiM'04* (pp. 112–118), ACM.
18. Stubblefield, A., Loannidis, J., & Rubin, A. D. (2001). Using the Fluhrer, Mantin, and Shamir attack to break WEP. *AT&T Labs Technical Report*.
19. Fluhrer, S., Mantin, I., & Shamir, A. (2001). Weakness in the key scheduling algorithm of RC4. In *Proceedings of the Workshop in Selected Areas of Cryptography*.
20. IEEE 802.11-1999: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications. (1999).



# Chapter 130

## Leakage Power Reduction of Instruction Cache Based on Tag Prediction and Drowsy Cache

Wei Li and Jianqing Xiao

**Abstract** Tag prediction is proposed to reduce the leakage power consumption of instruction cache and the power consumption of branch prediction that represent a sizeable fraction of the total power consumption of embedded processors in this chapter. By extending the architectural control mechanism of the drowsy cache to predict the cache line read in the next access, the tag prediction wakes up the necessary cache line in advance, while the rest of cache line is in the drowsy mode. Empirical results show that the tag prediction reduces the 77 % power consumption compared to the policy adopting branch prediction, and the accuracy of tag prediction is roughly same with the accuracy of BTB prediction. By removing the BTB and adopting the technique of drowsy cache, the tag prediction effectively reduces the power consumption without significant impact on performance of processors.

### 130.1 Introduction

In modern embedded processor, cache occupies a large portion of the power consumption of the processor chip. On one hand, performance improvements need to use the larger and faster cache that uses the faster transistors consuming the more leakage energy when they are turned off. On the other hand, new applications favor low power and high performance. Figure 130.1 [1] shows the trend of static and dynamic power consumption. As the figure shows, the static power takes up more and more portion of the total power consumption. Because of the popularity of portable device, the power problem is being more serious.

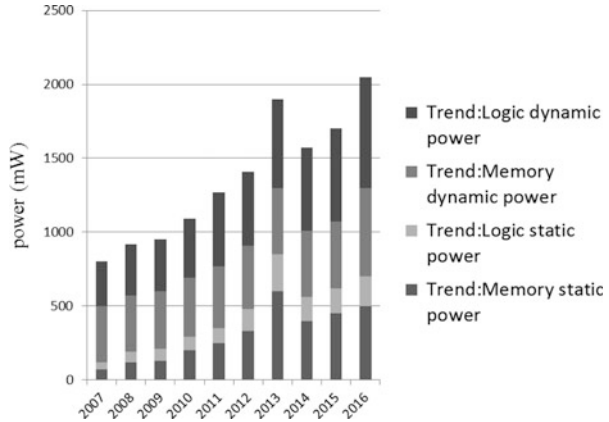
Static power is problem for all the transistors. But it is more important for the cache, because that always need using, especially the instruction cache, cannot be simply turned off. Based on this situation, the drowsy cache [2] is an effective

---

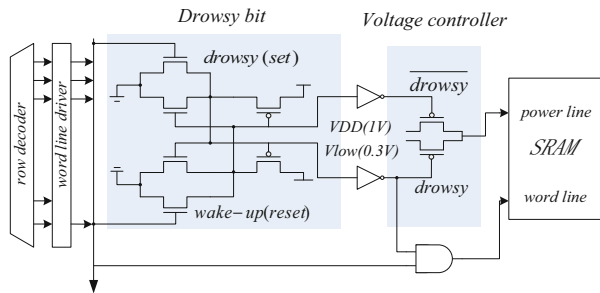
W. Li (✉) • J. Xiao

Xi'an Microelectronic Technology Institute, Xi'an 710054, Shanxi, China  
e-mail: [lw@stu.xjtu.edu.cn](mailto:lw@stu.xjtu.edu.cn)

**Fig. 130.1** SOC consumer portable power consumption trends



**Fig. 130.2** Implementation of the drowsy cache line



method to reduce the leakage power consumption of instruction cache. As shown in Fig. 130.2, the main idea of the drowsy cache is that cache line can be in one of the two modes: in a low-leakage drowsy mode using the lower supply voltage to preserve the data and in a normal mode using the normal supply can be a normal access. To reduce the leakage energy consumption, an algorithm is used to decide which line will be turned on or off in the near future. Then the rest of the cache lines can reduce the power consumption. Therefore, the quality of algorithms decides the power consumption and performance of processors.

There are several wake-up polices including wake-up next set of cache [3], wake-up next block of cache [4, 5], and Just-in-Time Activation [6]. The main idea of these methods that predict which block or line of cache will be read in the next access are based on the principle of spatial locality of programs. The predictor of these methods uses the BTB (branch target buffer) and BHT (branch history table) that can give the target address and the direction of the branch. By using this information, the cache can ahead make the drowsy line or block of cache active. These methods focus on reducing the power consumption of cache and ignore the power consumption of BTB and BHT.

In addition, there is another opposite strategy of no-access [7] based on the principle of cache decay [8] that is used to decide which block or line of cache to switch to drowsy mode when the block or line of cache is not accessed after a period of time called decay time. LRU (Least Recently Used)-assist strategy [9] is also a statistics strategy based on cache decay. This method also uses temporal locality of programs. In executing the program, the cache line accessed is woken up and holds the active mode. When the counter of no-access is overflow, the cache line with no access is switched to drowsy mode. So there are many cache lines in normal mode to consume power. And the method of no-access is more effective for data cache rather than for instruction cache.

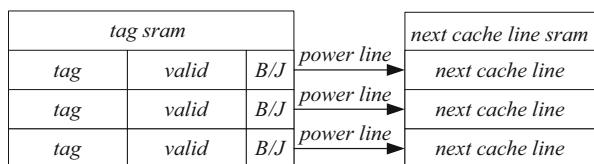
In the techniques mentioned previously, some only focus on reducing the power consumption of instruction cache and ignore the extra power consumption of BHT and BTB. Other methods are more appropriate for the data cache rather than for the instruction cache. There are no methods to reduce the power consumption of branch prediction and the leakage power consumption of instruction cache at the same time. In order to reduce the leakage power consumption of instruction cache and reduce the power consumption of branch prediction, this study proposes a method of tag prediction. As Sect. 130.2 presents, the tag prediction extends the structure of drowsy cache to predict the branch target, while the branch prediction is removed. By this way, the extra power dissipation consumed by BTB and BHT can be reduced. As Sect. 130.3 presents, the results show that the tag prediction has almost no loss of performance compared to branch prediction.

## 130.2 Tag Prediction

### 130.2.1 Tag Prediction Technique

In order to use the tag prediction to replace the branch prediction, an NCL SRAM (next cache line SRAM) is added to record a number of cache lines. As shown in Fig. 130.3, each cache line needs more than one bit of B/J (Branch or Jump) in the tag word. This bit is used to indicate whether there is a branch or jump instruction in the current cache line. If there is a branch or jump instruction, the target number of next cache line that will be accessed is stored in the NCL SRAM.

The traditional target address buffer modified also can predict the next cache line that will be accessed in the near future. But the BTB uses the full-associated structure or set-associated structure. That means it will consume more dynamic



**Fig. 130.3** The relationship diagram between tag SRAM and NCL SRAM

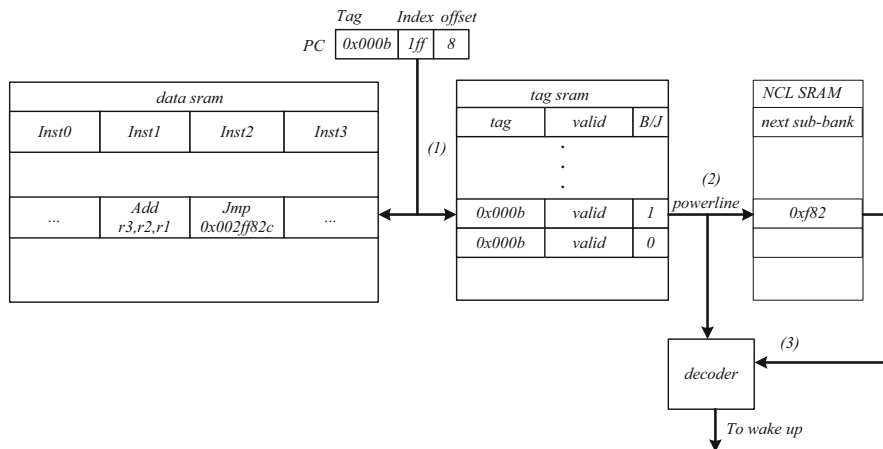


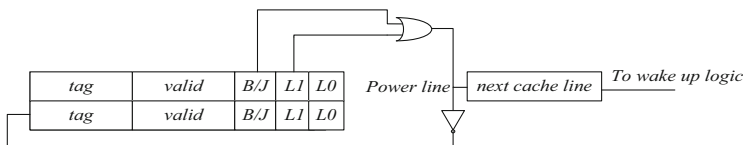
Fig. 130.4 The working scheme of tag prediction in cache line

power. The tag prediction technique can reduce the leakage power and dynamic power consumed by using the BTB technique. The disadvantage is that the information of prediction would be lost when the current line would be replaced.

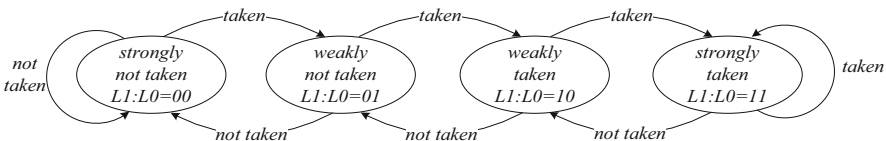
Figure 130.4 shows how the tag predictor works. Whenever the processor accesses the cache line, the instruction cache checks whether the valid bit and B/J bit are valid. If these bits are valid, the instruction cache sends the address of the next cache line accessed by the processor to the wake-up logic that can early make the corresponding cache line active.

### 130.2.2 Tag Predictor Design

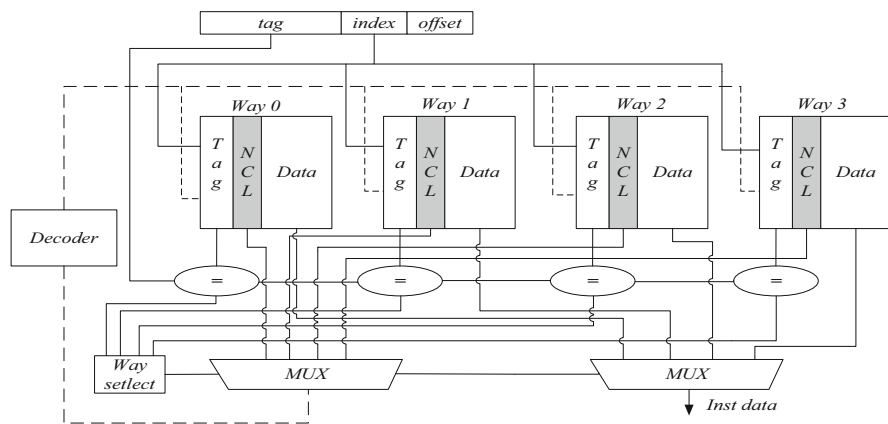
Figure 130.5 shows the implementation of tag predictor; there are two more bits including L1 (latest access bit 1) and L0 (latest access bit 0) that are used to predict the direction of branch instruction by 2-bit saturating method. B/J is still used to indicate whether there is a branch instruction or a jump instruction in the current cache line. When there is a jump instruction, tag predictor will wake up the NCL SRAM and send the address of cache line to wake-up logic that will make the corresponding cache line active. When there is a branch instruction and the L1 bit is valid, it will make the corresponding cache line also active. The difference between jump instruction and branch instruction is that the direction of branch instruction is decided by 2-bit saturating counter. The state chart of saturating counter is shown in Fig. 130.6. When the direction of prediction is correct, called taken, the state is



**Fig. 130.5** The implementation of tag prediction used to wake up the next cache or the predicted cache line



**Fig. 130.6** 2-Bit saturating counter used to predict the direction of branch



**Fig. 130.7** Implementation of 4-way set associative cache

switched to the corresponding taken state. When the direction is not correct, called no-taken, the state is switched to the corresponding no-taken state.

Figure 130.7 shows the implementation of 4-way set associative cache. Each way contains tag SRAM, NCL SRAM, and data SRAM. When the cache line matches, tag prediction checks B/J, L1, and L0 bits to decide which cache line will be accessed by the processor as described above. The solid line indicates the instruction data flow as the common cache dose. The dotted line indicates the wake-up signal flow that is used to wake up predicted cache line.

## 130.3 Experiments and Analysis

### 130.3.1 Experimental Setup

In the study, hot-leakage toolsets [10] are used to model an out-of-order speculative processor with a two-level cache hierarchy that is used to evaluate the power consumption and performance. The simulation parameters that roughly correspond to Alpha 21264 is listed in Table 130.1.

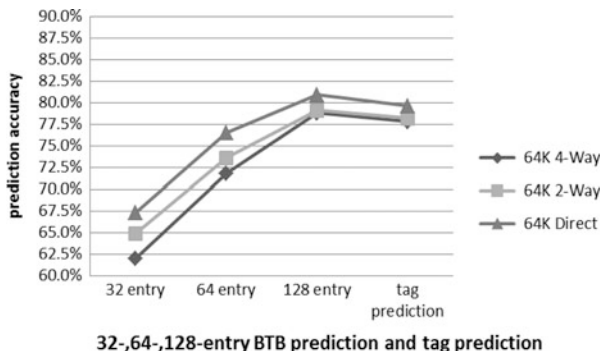
There are three different BTB entries: 32-entry, 64-entry, and 128-entry. Instruction cache size is 64 KB and various degrees of associativity are as follows: direct-map, 2-way, and 4-way.

In this study, benchmarks are from SPEC2000 suites that were run on a modified hot-leakage simulator.

**Table 130.1** Parameters of out-of-order speculative processor with BTB prediction or tag prediction

Parameter	BTB prediction	Tag prediction
Fetch/issue/ decode/ commit	4 Instructions	Same with left
Fetch queen/ speed	4 Instructions/1x	Same with left
Branch prediction	Bimodal, 2k	No branch prediction
BTB	32-/64-/128-entry	No BTB
RAS	8-entry	Same with left
RUU size	64-entry	Same with left
LSQ size	32-entry	Same with left
Integer ALUs/ multi-divs	4/1	Same with left
Floating point ALUs/multi- divs	1/1	Same with left
Memory bus width/latency	8 Bytes/80 and 8 cycles for the first and inter chunks	Same with left
Inst./data TLBs	16-entry/32-entry in each way, 4 KB page size, 4-way, LRU, 30-cycle latency	Same with left
L1 cache	64 KB, 4-way, 32-byte blocks, LRU, 1-cycle latency, write-back	Parameter is the same with left, add the NCL SRAM
L2 unified cache	256 KB, 4-way, 64-byte blocks, LRU, 8-cycle latency	Same with left

**Fig. 130.8** Accuracy comparison between BTB prediction and tag prediction



### 130.3.2 Results and Analysis

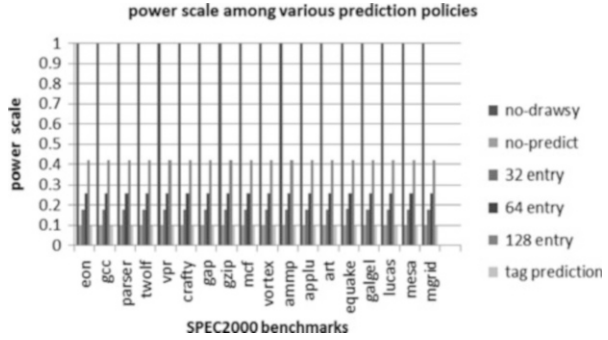
For more effective evaluation of the quality of the two different wake-up policies, the accuracy of drowsy cache policy is defined as

$$\text{accuracy} = \frac{\text{total\_correct\_predictions}}{\text{total\_wakeups}} \tag{130.1}$$

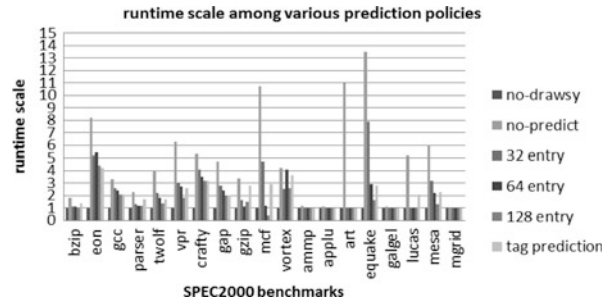
In the tag prediction, the prediction information is stored in NCL SRAM. In the BTB prediction, the BTB structure is modified to send two addresses. One is the target address of branch. The other is the predicted address used to wake up the cache line in the next access as the most of wake-up methods of BTB prediction dose. The result is shown in Fig. 130.8. The prediction accuracy of tag prediction is quite near the accuracy of BTB using 128-entry. By analyzing the results, the accuracy of tag prediction is little less than the accuracy of BTB prediction; the reasons are that there may be more than one branch instruction in the same cache line and the information of tag prediction may be lost when the replacement of instruction cache line occurs. Nevertheless dynamic and static power is reduced without using the BTB and PHT.

In the following figures, the policy of no-drowsy is the basic reference. The scales of rest policies are given in these figures. Figures 130.9 and 130.10 show the power comparison among different policies considering the power consumption of BTB and PHT and run time of different policies separately. The policies include no-drowsy without using the drowsy technique, no-predict with waking up the drowsy cache on demand, and 32-, 64-, and 128-entry BTB prediction and tag prediction used to wake up drowsy cache ahead. As shown in Fig. 130.9, the leakage power consumption of instruction cache using tag prediction is significantly less than the leakage power consumption of instruction plus the dynamic power of BTB and PHT. The run time using different policies is quite different. As shown in Fig. 130.10, the run time using tag prediction is basically the same with the run time using the 128-entry BTB. And the run time using the rest policies significantly increases because of the low accuracy of prediction.

**Fig. 130.9** Power comparison among various policies



**Fig. 130.10** Run-time comparison among various policies



The experimental results show that the tag prediction has about 78 % prediction accuracy just like the branch prediction does. So there is no loss of performance compared to branch prediction. While removing the branch prediction, the tag prediction effectively reduces the power consumption of processors. While the tag prediction reduces the negative performance impact by 76 % compared to the no-prediction policy, the tag prediction reduces 90 % leakage power consumption compared to the no-drowsy policy.

### 130.4 Conclusion

By using the drowsy policy of tag prediction, the cache line accessed by processor was active and the rest of cache lines were in the drowsy mode. The extra dynamic power consumption of BTB and BHT was eliminated. Compared to the 32-, 64-, and 128-entry BTB prediction, 46, 63, and 77 % power consumption was reduced. The accuracy of tag prediction was near the accuracy of 128-entry BTB prediction at the same time. Though the area of SRAM increased, the area occupied by BTB and BHT can be removed. In the design of embedded processors, the cache is generally small, so increased area used for NCL SRAM is quite small. So this



technique is very appropriate for embedded processors using the drowsy technique. During the investigation of drowsy instruction caches in the study, tag prediction of instruction can significantly reduce the dynamic and static power consumption of processors.

## References

1. ITRS Organization. (2008). International technology roadmap for semiconductors 2008 updates. Retrieved from <http://public.itrs.net/>.
2. Flaunter, K., Kim, N. S., et al. (2002). Drowsy caches: simple techniques for reducing leakage power. In *SIGARCH, proceedings of the 29th annual international symposium on computer architecture* (pp. 148–157). Washington: IEEE Computer Society.
3. Zhang, C., Zhou, H. W., et al. (2006). Architectural leakage power reduction method for instruction cache in ultra deep submicron microprocessors. In *The 11th Asia-Pacific computer systems architecture conference* (pp. 588–594). Berlin: Springer.
4. Hu, J., et al. (2003). Exploiting program hotspots and code sequentiality for instruction cache leakage management. In *International symposium on low power electronics and design (ISLPED'03)* (pp. 25–27). Berlin: Springer.
5. Chung, S. W., & Skadron, K. (2006). Using branch prediction information for near-optimal I-Cache leakage. In *The 11th Asia-Pacific computer systems architecture conference* (pp. 24–37). Berlin: Springer.
6. Kim, N. S., Flautner, K., et al. (2004). Single-VDD and single-VT super-drowsy techniques for low-leakage high performance instruction caches. In *International symposium on low power electronics and design (ISLPED'04)* (pp. 54–57). Berlin: Springer.
7. Kim, N. S., Flautner, K., et al. (2004). Circuit and microarchitectural techniques for reducing cache leakage power. *IEEE Transaction on VLSI Systems.*, 12(2), 167–184.
8. Kaxiras, S., Hu, Z., & Martonosi, M. (2001). Cache decay: Exploiting generation behavior to reduce cache leakage power. In *ISCA 2001* (pp. 240–251). Goteborg, Sweden: IEEE Computer Society.
9. Zhang, C. Y., Zhang, M. X., et al. (2006). LRU-assist: An efficient algorithm for cache leakage power controlling. *Acta Electronica Sinica.*, 34(9), 1626–1630 (In Chinese).
10. Zhang, Y., Parikh, D., et al. (2003). *Hotleakage: An architectural, temperature-aware model of subthreshold and gate leakage*. Virginia, USA: Department of computer sciences, University of Virginia.

**Part VI**  
**Network Optimization**

# Chapter 131

## The Human Role Model of Cyber Counterwork

Fang Zhou

**Abstract** The essence of cyber counterwork mainly reflects as the counterwork process between people, in order to solve effectively the cyber counterwork test problem about the cognitive level and decision level. In this paper, firstly, the dynamic adaptive cyber attack and defense “observation, orient, decision, act” (OODA) loop process models are established, which are based on the traditional military command and control operational process model. Secondly, establishing the cyber attacker and defender role models during the cyber counterwork process, which are mainly from role’s identities, role’s function, role’s capability, and role’s relationship utilizing the multi-attribute group description method. At last, establishing capability evaluation index system for each role and evaluating the capability through Delphi method. The human role model can provide theoretical basis for configuring attacker and defender role during the cyber cognitive and decision level test process.

### 131.1 Introduction

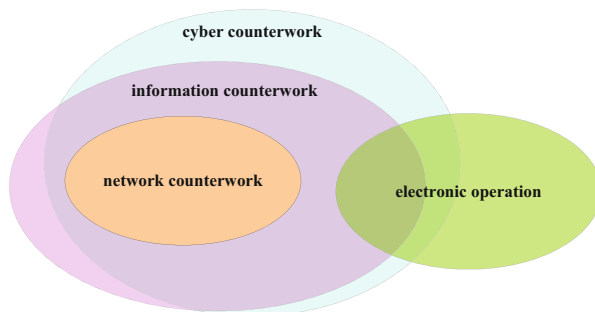
With the development of information technology, cyber counterwork will be the primary information warfare form and become an important part of the concept of full spectrum operation [1–3]. Cyberspace operation is closely related to the land, sea, air, and aerospace operation domain and provides guarantee for military operations within above domains. Presently, the warfare form is translated from physical and transport layer to decision and cognitive layer. And the human function is becoming more and more important, expressed as the man-centered and man–machine combination.

---

F. Zhou (✉)

The Key Laboratory of Information System Engineering, The 28th Research Institute of China Electronics Technology Group Corporation, Nanjing 210007, China  
e-mail: [326zhoufang@163.com](mailto:326zhoufang@163.com)

**Fig. 131.1** The relation between cyber, network, electronic, and information counterwork



In this paper, we firstly research the concept and characteristic of cyber counterwork. Then the cyber attack and cyber defense process models of dynamic adaptive are established, and several human role models are presented. Those models can provide theoretical basis and guidance for cyber counterwork test.

## 131.2 The Concept of Cyber Counterwork

Cyber counterwork is defined as the operation action that exploits cyberspace, adopts the reconnaissance, attack, defense and control confrontation, and applies to the information domain; it will also have impact on the physical domain and cognitive domain [4, 5]. The relation between the cyber and network counterwork is inclusive; the relation between electronic and cyber counterwork is intersectant but not inclusive; the relation between cyber and information counterwork is intersectant and convergent. Figure 131.1 shows the above relations.

Similar to the traditional counterwork forms, cyber counterwork mainly includes reconnaissance, attack, protection, and control forms and involves physical domain, information domain, cognitive domain, and social domain [6].

## 131.3 The Human Role of Cyber Attacking

### 131.3.1 The OODA Loop Pocess Model of Attacking

Considering the attacking side of cyber counterwork, we build the OODA loop process model with dynamic adaptive OODA loop theory, as shown in Fig. 131.2.

During the observation stage, the main task is to gather objective system information with electronic reconnaissance or network detection tools, such as network topology structure, open ports, OS type and version, system service, and other information.

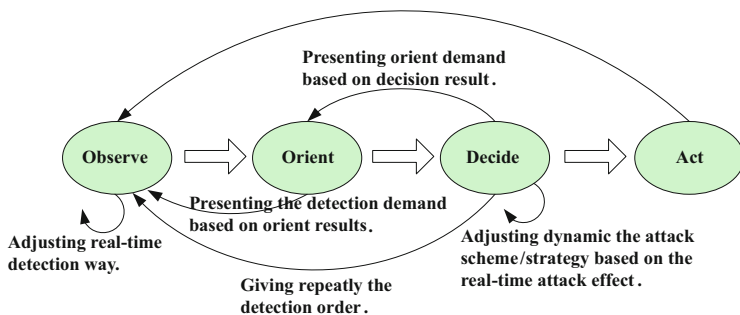


Fig. 131.2 The OODA loop process model of attacking

During the orient stage, the main task is to analyze objective system type, mine system vulnerability, analyze the captured packets, and restructure network or system.

During the decision-making stage, the main task is to make the attack plan that includes determining the attack targets, setting attack way, and choosing an attack weapon.

During the action stage, the main task is to implement the attack and provide real-time feedback information about attack result.

Because of the complexity and uncertainty of cyber counterwork, there is one or more OODA embedded loops, which will be generated during the attack process and which are manifested in three aspects: (1) There are also the observation, orient, decision-making, and action stages; (2) during the orient stage, the detection demand will be repeatedly presented based on orient results; and (3) during the decision-making stage, the detection and orient demand will be presented once the detected information or orient results cannot effectively support the decision making, and the attack plan shall be adjusted based on the attack effect. So the OODA loop of attacking is the iterative, embedded, and repeatedly rising process.

### 131.3.2 The Human Role Model of Attacking

#### 1. The Definition and Function of Human Role

Based on the analysis about attacking process [7, 8], we propose four roles: detecting controller, attack situation analyzer, attack decision maker, and attack performer. These human roles are defined as follows with multi-property group description method.

$$\text{Role} = \{\text{name, obligation, relation, capability}\}$$

The term “name” is the identifier of human role, such as detecting controller and the attack situation analyzer. The term “obligation” denotes the role’s function,

**Table 131.1** The human role’s function of attacking

Detecting controller	<ul style="list-style-type: none"> <li>✓ Detecting objective system and gathering network topology, the port, protocol, service, OS type and version, vulnerabilities, and defense measures information using reconnaissance and detection tool</li> <li>✓ Adjusting detection way based on information gathered</li> </ul>
Attack situation analyzer	<ul style="list-style-type: none"> <li>✓ Analyzing and evaluating the vulnerabilities of objective system based on information gathered</li> <li>✓ Mining the key node of objective system and business relations among nodes</li> </ul>
Attack decision maker	<ul style="list-style-type: none"> <li>✓ Making the attack plan or strategy, which includes determining attack targets, setting attack way, and choosing attack weapon</li> <li>✓ Adjusting dynamically the attack plan or strategy based on attack effect</li> </ul>
Attack performer	<ul style="list-style-type: none"> <li>✓ Manipulating attack tool to implement attack activities and providing feedback information about attack effect</li> </ul>

**Table 131.2** The capability evaluation index system of detecting controller

The capability of detecting controller ( $C_{obs}$ )	The capability of sensitive information gathering ( $C_{11}$ )	<ul style="list-style-type: none"> <li>The quantity of sensitive information gathering (<math>C_{111}</math>)</li> <li>The efficiency of sensitive information gathering (<math>C_{112}</math>)</li> </ul>
	The capability of vulnerability scanning ( $C_{12}$ )	<ul style="list-style-type: none"> <li>The quantity of vulnerability scanning (<math>C_{121}</math>)</li> <li>The efficiency of vulnerability scanning (<math>C_{122}</math>)</li> </ul>

describing the assignment undertaken by each role. The term “capability” denotes the skill needed in accomplishing the task, such as scanning system capability, breaking code capability, exploiting vulnerabilities capability, raising privilege capability, and avoiding detection capability. The term “relation” denotes the associated relationship among roles. According to the above definition, we define the role’s function, as shown in Table 131.1.

2. *The Capability Level of Human Role*

According to the role’s function category, firstly, we propose the measure factor set of each kind of role and establish the evaluation index system and measure criterion of role’s ability. Secondly, the Delphi evaluation method is used to determine the weight of each evaluation index in this paper. The basic idea of this method is making full use of expert experience and knowledge to review the evaluation index, and index weight is determined by using the method of geometric mean.

The capability will be evaluated by the capability of sensitive information gathering and the capability of vulnerability scanning. The evaluation indexes of these two capabilities are shown in Table 131.2.

The capability of attack situation analyzer can be evaluated by the time and the success rate of mining the key node of objective system and business relations among nodes. Table 131.3 shows the evaluation index system.

**Table 131.3** The human role’s function of defense

Role’s type	Role’s function
Defense supervisor	✔ Utilizing system security tools to monitor system running state and collect system or system security equipment log file information
Defense situation analyzer	✔ Generating system security situation, conducting threat estimation and prediction, mining the potential attack or abnormal behavior
Defense decision maker	✔ Making defense plan according to the analysis results ✔ Adjusting dynamically security defense strategy or plan
Defense performer	✔ Preventing illegal operation and intrusion behavior, scanning and remedying system vulnerability, data backup, and fault recovery

**Table 131.4** The capability evaluation index system of defense performer

The capability of system recovering ( $C_{Da1}$ )	The time of system recovering ( $T_{Ds}$ ) The degree of system recovering ( $P_{Dd}$ )
The capability of attack blocking ( $C_{Da2}$ )	The time of attack blocking ( $T_{Da}$ ) The number of attack threat ( $N_{Dn}$ )
The capability of honeypot technology ( $C_{Da3}$ )	The time of responding attack ( $T_{Dr}$ ) The time of tracking attack threat ( $T_{Dt}$ ) The number of decoying attack threat source ( $N_{Dd}$ )

**Table 131.5** The capability evaluation index system of attack performer

The capability of password decoding ( $C_{a1}$ )	The time of decoding ( $T_{ad}$ )
The capability of system penetrating ( $C_{a2}$ )	The access privilege obtained ( $P_a$ ) The time of system penetrating ( $T_{ap}$ ) The concealment of system penetrating ( $P_{ac}$ )
The capability of system palsyng ( $C_{a3}$ )	The duration time of palsyng ( $T_{as}$ ) The difficulty of system recoving ( $D_a$ )

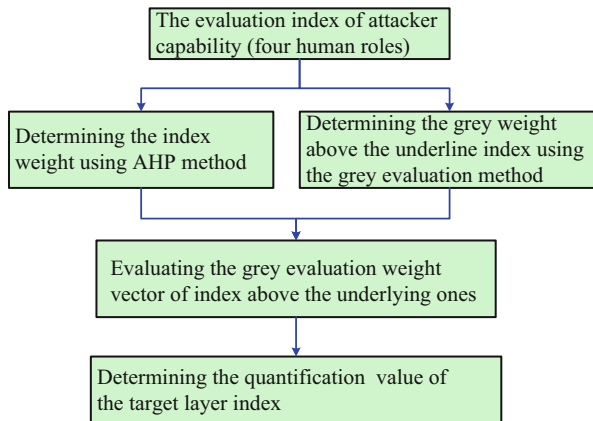
The capability evaluation index system of attack decision maker can be evaluated by the time and the effectiveness of making attack plan. Table 131.4 shows the detail index system.

The capability of attack performer is evaluated by three indexes: the access privilege obtained, the duration time of attack, and the concealment of attack. Table 131.5 shows the detail index system.

### 3. The Evaluation Model of Attacker’s Capability

In this paper, AHP (analytical hierarchy process) and the gray evaluation method are used to build the above human roles’ capability evaluation model. The basic idea is that the weights of each level index are obtained using the AHP method, to improve the effectiveness, reliability, and viability of index computed. Then the weights of underlying indexes are quantified using the gray evaluation method. Under this basis, the gray weight vector above each index and the gray class are determined. The specific implementation process of the evaluation method is shown in Fig. 131.3.

**Fig. 131.3** The flow chart of ability evaluation



Firstly, denote the first level index as  $C_{1i}(i = 1,2)$ , and the second level index as  $C_{1ij}$ . The computing process with AHP method includes constructing the judgment matrix, calculating the weight vector, and testing the consistency.

The weight vector of the first level index is denoted as  $W = (w_1, w_2)$ , and the weight vector of underline level index is denoted as  $W_1 = (w_{11}, w_{12})$ ,  $W_2 = (w_{21}, w_{22})$ .

Then the underlying index weights are determined with gray evaluation method. The essence of this method is through partial information known to generate and extract valuable information.

The central triangle whitening weight function is superior to the traditional whitening weight function on the crossover phenomenon, clustering coefficient, and endpoint selection aspects. Therefore, the gray evaluation method of central triangle whitening weight function is used to solve the problem of quantization based on the attacker role of factors. The specific steps are as follows:

(a). Determining the gray class

The value of each index is divided into 5 gray classes, such as “very low,” “low,” “medium,” “high,” and “very high.” Then selecting the closest center point as  $K_1, K_2, K_3, K_4$ , and  $K_5$ , these point values are set to 2, 4, 6, 8, and 10. And the index values are divided in the 4 gray intervals  $[K_1, K_2], [K_2, K_3], [K_3, K_4]$ , and  $[K_4, K_5]$ .

(b). Determining the index value

Assuming that  $N$  experts are invited to rate index, the expert number is denoted as  $n (n = 1,2,\dots,N)$ . Through establishing the evaluation standards and rules above each index, each evaluation expert independently scores on the underline index according to the grading standards, obtaining the evaluation sample matrix of the role of ability.



(c). Building the center whitenization weight function

The central point albinism function  $f_{jm}(x)$  of observation value(x) of index j that belongs to the gray m ( $m = 1, \dots 5$ ) can be described as follows:

$$f_{jm}(x) = \begin{cases} \frac{x - k_{i-1}}{k_i - k_{i-1}}, & x \in (k_{i-1}, k_i) \\ \frac{k_{i+1} - x}{k_{i+1} - k_i}, & x \in (k_i, k_{i+1}) \end{cases}$$

(d). Calculating gray evaluation indexes

The gray evaluation coefficient  $l_{1jm}, l_{2jm}$  of the second level indexes  $C_{11j}, C_{12j}$  that belongs to gray m is computed as follows:

$$l_{1jm} = \sum_{n=1}^N f_{1jm}(d_{1jmn}), \quad l_{2jm} = \sum_{n=1}^N f_{2jm}(d_{2jmn})$$

Where  $d_{1jmn}$  denotes the score of the n expert about  $C_{11j}, C_{12j}$  indexes and at the same time the sums of  $l_{1jm}$  and  $l_{2jm}$ , the gray evaluation weight coefficient of index of evaluation  $C_{1ij}$  is shown as follows:

$$l_{1j} = \sum_{m=1}^5 l_{1jm}, \quad l_{2j} = \sum_{m=1}^5 l_{2jm}, \quad p_{1jm} = l_{1jm} / l_{1j}, \quad p_{2jm} = l_{2jm} / l_{2j}$$

Therefore, the gray evaluation weight vector of  $C_{1ij}$  belonging to five gray classes is shown as follows:

$$p_{1j} = [p_{1j1}, p_{1j2} \dots p_{1j5}]^T, \quad p_{2j} = [p_{2j1}, p_{2j2} \dots p_{2j5}]^T$$

Based on the above weight vector, the gray evaluation weight vector can be obtained through the weight matrix  $W1, W2$  and the evaluation weight vectors  $P_{1j}, P_{2j}$  are multiplied; the result will be denoted as  $V = [v1, v2]$ :

$$V = W \times P^T = (v_1, \dots v_5), \quad P = [p_{1j}, p_{2j}]$$

The evaluation of gray level according to the ‘‘gray level’’ is assigned; namely, each evaluation gray grade value is translated into vector  $S = [2, 4, 6, 8, 10]$ . The ability quantification value of detecting controller role can be expressed as follows:

$$C_{obs} = V \times S^T$$

The capability of attack situation analyzer, attack decision maker, and attack performer can be computed similarly based on the above method.

## 131.4 The Human Role of Cyber Defense

### 131.4.1 *The Definition and Function of Human Role*

After the analysis about the OODA loop model of defending, we propose four roles: defense supervisor, defense situation analyzer, defense decision maker, and defense performer. The function of each human role is defined in Table 131.3.

### 131.4.2 *The Capability Evaluation Index System of Human Role*

According to the function of the above human roles, we propose the tactical and technical performance, represented by a set of capacity values.

The capability of defense supervisor is evaluated by the sensitive information count collected within a certain time ( $N_{Dinfo}$ ) and the detection efficiency ( $V_{detec}$ ).

The capability of defense situation analyzer is evaluated by the time ( $S_{Di}$ ), the success ratio, and the efficiency of identifying attack threat behavior or event ( $V_{iden}$ ).

The capability of defense decision maker is evaluated by the time ( $T_{Dd}$ ) and the effectiveness of making defense plan ( $P_{De}$ ).

The capability of defense performer is evaluated by seven indexes; the detail index is shown in Table 131.4.

## 131.5 Conclusion

Cyber counterwork research is a new area and there are less research results about this area. Based on the analysis of cyber counterwork characteristic, we respectively establish eight human roles for attack side and defense side during the counterwork process. Then the index system of role's ability is established, evaluating the role's ability through Delphi evaluation method. Role model established in this paper can provide theoretical support for the future development of cyber counterwork experiment, such as providing the theory basis for the attacker and defender capacity requirements.

**Acknowledgements** This work is supported by advanced researched plan of the general armament department No.9140A04040113DZ38054.

## References

1. The President of the United States. (2003). *The national strategy to secure cyberspace*[R]. Washington, DC: The Whitehouse.
2. United States Army Training and Doctrine Command. (2010). *The United States Army's Cyberspace operations concept capability plan 2016-2028*[R]. Washington, DC: USDOD.
3. Le May Center for Doctrine Development and Education. (2010). *Cyberspace operations*[R]. Washington, DC: The United States AirForce.
4. Guang-xia, Z., & Xin, S. (2012). Study on cyberspace operations[J]. *Command Information System and Technology*, 3(2), 6–10 (In Chinese).
5. Chairman of the Joint Chiefs of Staff. (2006). *National military strategy for cyberspace operations*[R]. Washington, DC: USDOD.
6. Kuhl, M. E., & Sudit, M. (2007). Cyber attack modeling and simulation for network security analysis [C]. *Proceedings of the 2007 Winter Simulation Conference, Springer Berlin Heidelberg, Australia*. 2119(1), 1180–1189.
7. Rowe, N. C. (2004). A model of deception during cyber-attacks on information systems [J]. *IEEE First Symposium on Computing & Processing*, 1(2), 21–30.
8. Kottenko, I. (2007). Multi-agent modelling and simulation of cyber-attacks and cyber-defense for home-land security[J]. *IEEE International Workshop on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications.*, 4(1), 614–620.

# Chapter 132

## A Service Channel Assignment Scheme for IEEE 802.11p Vehicular Ad Hoc Network

Yao Zhang, Licai Yang, Haiqing Liu, and Lei Wu

**Abstract** IEEE 802.11p vehicular ad hoc network (VANET) applies multiple channels, including one control channel (CCH) and six service channels (SCHs); the enhanced distributor channel access (EDCA) mechanism is used to support wireless channel assignment and QoS requirements. But the method of SCH assignment is not proposed in IEEE 802.11p standard. We present a scheme to perform SCH assigning, previous transmission indicators of service channels are detected dynamically by service channel assignment controller set in medium access control (MAC) layer, service packets would be delivered into suitable SCH and EDCA access category (AC) queue according to SCH reservation probability and estimated transmission delay. Saturated throughput of our scheme in SCH is analyzed by theoretical model in different conditions; the results show that it can ensure higher SCH utilization and is an efficient way to improve performance of intelligent transportation system.

### 132.1 Introduction

Vehicular ad hoc network (VANET) has been considered an essential technology for future intelligent transportation system (ITS); the purpose of VANET is to provide vehicle to roadside unit (RSU) as well as vehicle to vehicle wireless communications. IEEE 802.11p is designed for wireless access in vehicular environments (WAVE) to support VANET, allocated 75 MHz bandwidth of licensed

---

Y. Zhang (✉)

School of Control Science and Engineering, Shandong University, Jinan 250061, China

School of Mechanical, Electrical and Information Engineering, Shandong University (Weihai),  
Weihai 264209, China

e-mail: [zhangyao@sdu.edu.cn](mailto:zhangyao@sdu.edu.cn)

L. Yang • H. Liu • L. Wu

School of Control Science and Engineering, Shandong University, Jinan 250061, China

spectrum at 5.9 GHz is divided into one control channel (CCH) and six service channels (SCHs), CCH is dedicated for broadcast of traffic safety messages, and SCHs are dedicated for transmission of various application messages [1–3]. In MAC layer, enhanced distributor channel access (EDCA) mechanism is used to perform wireless channel competition, EDCA classifies data traffic into different priorities to support varying QoS requirements [4]. To coordinate channel access on CCH and multiple SCHs *effectively*, a synchronized channel coordination scheme is proposed in IEEE 802.11p protocol; the channel access time is divided into synchronization intervals with a fixed length of 100 ms, consisting of 50 ms CCH interval and 50 ms SCH interval. During CCH interval, all OBUs must monitor the CCH for messages of traffic safety or SCH reservation. During SCH interval, OBUs can optionally switch to *preconcerted* SCH to perform service messages transmitting. Although this scheme can reduce the transmission delay of safety messages in CCH, SCH utilization cannot be ensured [5].

The rest of this paper is organized as follows: In Sect. 132.2, we present a dynamic SCH assignment scheme. In Sect. 132.3, saturated throughput of our scheme is analyzed in different conditions by theoretical model. Finally, the main conclusions and future research work are summed up in Sect. 132.4.

## 132.2 Service Channel Assignment Scheme

In our scheme, CCH is classified into four EDCA ACs: AC[3] is for safety-related urgent messages, AC[2] is for vehicles to advertise road traffic messages, AC[1] is for nonurgent traffic messages, and AC[0] is for SCH reservation. SCH is classified into three ACs: AC[2], AC[1], and AC[0] are for high-priority UDP packets, and AC[1] and AC[0] are for low-priority TCP packets. Setting service channel assignment controller in MAC layer, it is intended for detecting previous access delay of each AC queue in SCH interval, estimating transmission delay and packet error rate of SCHs, and then guiding AC queue assignment and service channel reservation.

The algorithm of service channel assignment scheme is as follows:

```
//  $n_j$ : the number of reserving SCH-j ( $j=1-6$ )
//  $adelay_{ji}$ : AC[i]'s access delay in SCH-j ( $i=0,1,2$ )
//  $time()$ : function of system time
//  $lenAC_{ji}$ : AC[i]'s queue length in SCH-j
//  $delay_{ji}$ : AC[i]'s transmission delay in SCH-j
//  $pe_j$ : packet error rate in SCH-j
//  $X_{ji}$ : the number of acknowledged (ACK) packets through AC[i] in SCH-j
//  $Y_{ji}$ : the total number of packets through AC[i] in SCH-j
//  $pr_j$ : SCH-j reservation probability
```

(continued)

(continued)

//*packet\_type*: denotes packet priority, the value of UDP packet is 1, TCP packet is 0

initialize: ( $i=0,1,2$ ;  $j=1-6$ )

$X_{ji}=0; Y_{ji}=0; pe_j=0; adelay_{ji}=0; delay_{ji}=0;$

//AC queue assignment during SCH-j interval

$lenAC_{ji}=0;$

when ACK from destination node is received by AC[i] in source nodes

$X_{ji}=X_{ji}+1; lenAC_{ji}=lenAC_{ji}-1;$

$$adelay_{j_{new}} = Time(receive\ ACK) - Time(frame\ begins\ to\ compete\ channel); \quad (132.1)$$

if  $adelay_{ji}=0$

$$adelay_{ji} = adelay_{j_{new}}; \quad (132.2)$$

else

$$adelay_{ji} = (1-\alpha) * adelay_{ji} + \alpha * adelay_{j_{new}}; \quad (132.3)$$

when a packet in AC[i] begins to compete channel

$Y_{ji}=Y_{ji}+1;$

when a packet arrives to SCH assignment controller of SCH-j:

$$delay_{ji} = lenAC_{ji} * adelay_{ji}; \quad (132.4)$$

switch (*packet\_type*)

{ case 1  $k=2$ ;

if  $delay_{j1} < delay_{j2} \& delay_{j1} < delay_{j0}$   $k=1$ ;

if  $delay_{j0} < delay_{j2} \& delay_{j0} < delay_{j1}$   $k=0$ ;

case 0  $k=1$ ;

if  $delay_{j0} < delay_{j1}$   $k=0$ ;

data packet  $\rightarrow$  AC[k] queue ;  $lenAC_{jk}=lenAC_{jk}+1;$  }

//SCH reservation through AC[0] during CCH interval

$n_j=0;$

when SCH reservation request is received by SCH assignment controller

(continued)

(continued)

$$pe_j = \left( \sum_{k=0}^2 X_{jk} \right) / \left( \sum_{k=0}^2 Y_{jk} \right), \quad pr_j$$

$$= \left[ 1 - \beta \frac{pe_j}{\sum_{k=1}^6 pe_k} - (1 - \beta) \frac{n_j}{\sum_{k=1}^6 n_k} \right] / 5 \quad (132.5)$$

*RN*=a random number from uniform distribution between [0,1];  
*low*=0; *high*=*pr<sub>1</sub>* ;  
 for *k*=1:6  
 {if *low*<*RN*<*high* { *j*=*k* ; break }  
*low*=*low*+*pr<sub>k</sub>*; *high*=*high*+*pr<sub>k</sub>*;}  
 sending SCH-*j* reservation request through AC[0]  
 when receiving ACK of SCH-*j* reservation request  
*n<sub>j</sub>*=*n<sub>j</sub>*+1;

Our algorithm has two steps:

Step 1: AC queue assignment and access delay detecting during SCH interval  
 AC[*i*]'s access delay includes time to compete service channels, packet sending delay, propagation delay, and time to send back ACK; it is detected according to formula (132.1), (132.2), and (132.3). In formula (132.3),  $\alpha$  is a predefined factor between 0 and 1. If  $\alpha$  is bigger, the new sample of access delay makes more influence on average access delay. When a packet arrives to SCH controller, transmission delay of AC[2], AC[1], and AC[0] is estimated by formula (132.4), and then UDP packet would be mapped into the AC queue that estimated transmission delay is smallest among AC[2], AC[1], and AC[0]. TCP packet would be mapped into the AC queue that estimated transmission delay is the smallest among AC[1] and AC[0].

Step 2: SCH reserving during CCH interval  
 SCH reservation messages are transmitted through AC[0] during CCH interval; SCH-*j* reserving probability is calculated by formula (132.5). With the increase of packet error rate or the number of nodes that has reserved SCH-*j*, reserving probability decreases. In formula (132.5),  $\beta$  is a factor between 0 and 1. If the value of  $\beta$  is bigger, packet error rate has more influence on SCH reserving probability than the number of nodes. For example, let  $\beta$  be 0.75; supposing packet error rate of SCH-1 to SCH-6 is  $1.25 \times 10^{-4}$ ,  $1.25 \times 10^{-4}$ ,  $1.25 \times 10^{-5}$ ,  $1.25 \times 10^{-5}$ ,  $1.25 \times 10^{-6}$ , and  $1.25 \times 10^{-6}$ , we calculate SCH reservation probability by formula (132.5), as shown in Table 132.1. Obviously, for service

**Table 132.1** SCH reservation probability (n1 = 10, n2 = 30, n3 = 15, n4 = 50, n5 = 25, n6 = 35)

SCH-1	SCH-2	SCH-3	SCH-4	SCH-5	SCH-6
0.1289	0.1233	0.1886	0.1781	0.1917	0.1887

channels of high packet error rate or the large number of users, their reservation probabilities are smaller. Our service channel reservation scheme accomplishes payload balance and adaptive SCH selection.

### 132.3 Performance Analysis of SCH Assignment Scheme

$N$  is defined as the number of nodes in SCH- $j$ ,  $\tau_i$  is the packet transmitting probability of AC[i] in a time slot,  $\tau$  is the packet transmission probability,  $p_{ic}$  is the packet transmission probability of AC[i],  $p_{si}$  is the successful probability of transmitting AC[i] packet in a time slot,  $p_{tr}$  is the channel utilization probability,  $CW_{min}$  is the minimum back-off windows of EDCA, and  $m_i$  is the maximum back-off stage of EDCA. We have [6, 7]

$$\tau_i = \frac{2(1 - 2p_{ic})}{(1 - 2p_{ic})(CW[i]_{min} + 1) + p_{ic}CW[i]_{min}(1 - (2p_{ic})^{m_i})}$$

$$\tau = 1 - \prod_{i=0}^2 (1 - \tau_i), \quad p_{si} = \frac{N\tau_i \prod_{j \neq i} (1 - \tau_j)(1 - \tau)^{N-1}}{1 - (1 - \tau)^{N-1}}$$

$$p_{ic} = 1 - (1 - \tau)^{N-1} \prod_{j \neq i} (1 - \tau_j), \quad p_{tr} = 1 - (1 - \tau)^N$$

Furthermore, if  $T_{si}$  is defined as the average time of transmitting AC[i] packets,  $T_{ci}$  is the average collision time,  $\delta$  is the propagation delay,  $\sigma$  is the time slot length, and  $E(L_f)$  denotes the average length of service packets, we can get saturated throughput of AC[i] [8, 9]:

$$S_i = \frac{p_{si}p_{tr}E(L_f)}{(1 - p_{tr})\sigma + \sum_{k=0}^2 (p_{sk}p_{tr}T_{sk}) + p_{tr} \sum_{l=0}^2 \left[ \left(1 - \sum_{k=0}^2 p_{sk}\right) T_{cl} \right]} \quad (132.6)$$

In formula (132.6), supposing RTS/CTS mechanism is used in EDCA, we have



$$T_{si} = RTS + 3SIFS[i] + 4\delta + CTS + T_{frame-head} + T_{frame} + ACK + AIFS[i]$$

$$T_{ci} = RTS + AIFS[i] + \delta + CTS$$

Let packet retry limit of AC[i] be  $RT_i$ , AC[i] packet arrival rate  $\lambda_i$ , UDP packet arrival rate  $\lambda_H$ , TCP packet arrival rate  $\lambda_L$ ,  $p(X = k)$  be the probability distribution of channel competing times, and  $p_i$  be the probability of AC[i] transmitting packet successfully. We have

$$p_i = \sum_{k=1}^{RT_i} p_{tr}(p_{si}/N) [(1 - p_{tr}(p_{si}/N))]^{k-1} \quad (i = 0, 1, 2)$$

$$p(X = k) = \frac{(1 - p_{tr}p_{si})^{k-1} p_{tr}p_{si}}{\sum_{j=1}^{RT_i} (1 - p_{tr}p_{si})^{j-1} p_{tr}p_{si}}$$

$$adelay_i = \frac{1}{\mu_i} = \left\{ \left[ \sum_{k=1}^{RT_i} kp(X = k) - 1 \right] T_{ci} + T_{si} \right\} / (1 - pe_j)$$

$$\lambda_2 = p_{H2}\lambda_H + (1 - p_2)\mu_2, \quad \lambda_1 = p_{H1}\lambda_H + p_{L1}\lambda_L + (1 - p_1)\mu_1$$

$$\lambda_0 = p_{H0}\lambda_H + p_{L0}\lambda_L + (1 - p_0)\mu_0, \quad d_i = lenAC[i] \times \{adelay_i\}$$

$$p_{Hi} = probability \left\{ \min(d_2, d_1, {}^1d_0) = d_i \right\} \quad (i = 0, 1, 2)$$

$$p_{Li} = probability \left\{ \min(d_1, {}^1d_0) = d_i \right\} \quad (i = 0, 1)$$

Hence, saturated throughput of UDP packets and TCP packets during SCH interval are

$$S_h = \sum_{j=1}^6 \left[ \left( \sum_{i=0}^2 p_{Hi} S_i \right) \times N \times T_{SCH}, \quad S_l = \sum_{j=1}^6 \left[ \left( \sum_{i=0}^1 p_{Li} S_i \right) \times N \times T_{SCH} \right]$$

(132.7)

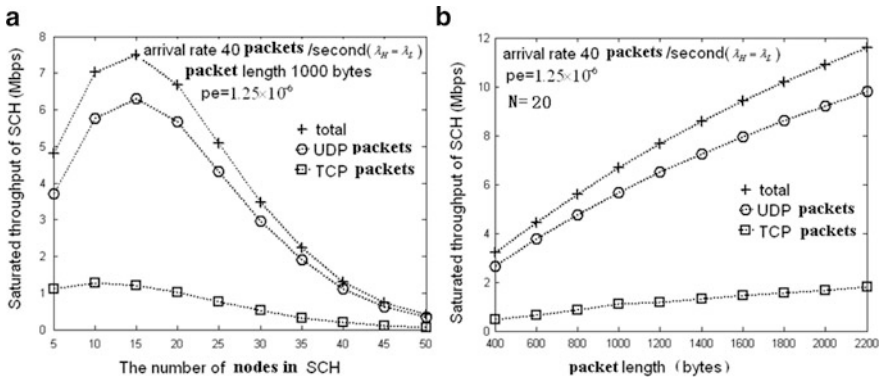
Total saturated throughput of SCH is

$$S_{SCH} = S_h + S_l \tag{132.8}$$

The values of IEEE 802.11 EDCA parameters are shown in Table 132.2; we analyze SCH saturated throughput by formulas (132.7) and (132.8), and the results are shown in Fig. 132.1. It is clear that:

**Table 132.2** IEEE 802.11 EDCA parameters

Parameter	Value	Parameter	Value	Parameter	Value
PHY header	192 $\mu$ s	ACK length	112 bit	AC[0]-AIFS	7
MAC header	272 bit	RTS length	160 bit	AC[1]-AIFS	3
SIFS	10 $\mu$ s	CTS length	112 bit	AC[2]-AIFS	2
Slot time	20 $\mu$ s	Propagation delay	1 $\mu$ s	m2	1
Retry limit	7	Transmission rate	15 M bit/s	AC[2]-CWmin	7
AC[0]-CWmin	15	m1	3	Maximum queue length	300
m0	5	AC[1]-CWmin	15	pe	1.25e-6



**Fig. 132.1** Analysis results of saturated throughput in SCH. (a) Saturated throughput in terms of the number of nodes. (b) Saturated throughput in terms of packet length

1. Saturated throughput of SCH increases with the number of nodes, but when  $N$  is larger than 15, saturated throughput begins to decrease. This is because VANET has little data traffic when  $N$  is small. However, when  $N$  is large, collision probability increases; it makes nodes having lesser chance to compete channel successfully.
2. Saturated throughput of SCH increases with the increase of packet length.
3. Saturated throughput of UDP packets is larger than TCP packets; it means QoS of high-priority services is ensured adequately.

### 132.4 Conclusion

In this paper, we present a service channel assignment scheme to improve SCH transmitting performance in VANET. SCH is classified into three ACs to support QoS requirements of two priorities (UDP packets and TCP packets). During SCH interval, instead of static queue management in EDCA, our scheme achieves delivering data packets into optimal AC queue dynamically according to packet type, transmission delay, and network scale. During CCH interval, the scheme

achieves adaptive SCH reservation according to SCH transmission quality and payload. Performance of our scheme is analyzed based on theoretical model; the result shows that our scheme is an efficient way to solve bottleneck of high-priority services and improve transmission performance of VANET.

However, with the increase of network scale or traffic load, performance of the scheme becomes worse, so the novel channel assignment scheme based on contention-free mechanism must be presented for high-density VANET. How to solve this problem is our further research work. In addition, we will extend the research to building simulation model and multi-hop wireless environment with routing protocols based on node location in VANET.

**Acknowledgements** This work is supported by National Natural Science Foundation of China (No. 61174175) and Natural Science Foundation of Shandong Province of China (No. ZR2010FM036).

## References

1. IEEE Std.1609.4. (2010). *IEEE standard for wireless access in vehicular environments (WAVE)-multiple channel operation*. Intelligent Transportation Systems Committee of the IEEE Vehicular Technology Society, the Institute of Electrical and Electronics Engineers, Inc. New York, USA.
2. Marica, A., Claudia, C., & Antonella, M. (2012). Enhancing IEEE 802.11p/WAVE to provide infotainment applications in VANETs. *Ad-Hoc Networks*, 10(2), 253–269.
3. Wang, Q., Leng, S., Fu, H., & Zhang, Y. (2012). An IEEE802.11p-based multichannel MAC scheme with channel coordination for vehicular Ad-Hoc networks. *IEEE Transactions on Intelligent Transportation Systems*, 7(2), 449–458.
4. IEEE Std 802.11e-2005. (2005). *Wireless LAN medium access control (MAC) and physical layer (PHY) specifications amendment 8: MAC quality of service enhancements*. LAN/MAN Committee of the IEEE Computer Society, the Institute of Electrical and Electronics Engineers, Inc. New York, USA.
5. Gallardo, J. R., Makrakis, D., & Mouftah, H. T. (2010). Mathematical analysis of EDCA's performance on the control channel of an IEEE 802.11p WAVE vehicular networks. *EURASIP Journal on Wireless Communications and Networking*, 2010(1), 1–15.
6. Mao, J.-B., Mao, Y.-M., Leng, S., & Bai, X. (2010). Research of the QoS-supporting IEEE 802.11 EDCA performance. *Journal of Software of China*, 21(4), 750–770.
7. Xiao, Y. (2005). Performance analysis of priority schemes for IEEE 802.11 and IEEE 802.11e wireless LANs. *IEEE Transactions on Wireless Communications*, 4(4), 1506–1515.
8. Huang, C.-L., & Liao, W. (2007). Throughput and delay performance of IEEE 802.11e enhanced distributed channel access (EDCA) under saturation condition. *IEEE Transactions on Wireless Communications*, 6(1), 136–145.
9. Bianchi, G. (2000). Performance analysis of the IEEE 802.11 distributed coordination function. *IEEE Journal on Selected Areas in Communications*, 18(3), 535–547.

# Chapter 133

## An Exception Handling Framework for Web Service

Hua Guan, Shi Ying, and Caoqing Jiang

**Abstract** According to the problems of exception handling for service-oriented software, this paper presents a framework for Web service exception handling (EHF-S) based on policy driven. The EHF-S processes the response message of invoking Web service and produces a response message which is added exception information and exception handling message. We introduce the realization principle, the component, and the key technology for EHF-S. This framework can support the development and integration of exception handling logic for Web service process, improve the exception handling capability, and simplify the exception handling process for Web service.

### 133.1 Introduction

Web services are rapidly becoming a fundamental program paradigm for the development of complex Web applications. Because of the dynamics and uncertainty during runtime as well as the autonomy and loose coupling in the service resources, there appears the diversity and complexity of exception. This paper focuses on the exception of receiving response message and real-time checks whether an exception has occurred during execution of Web services through a response message returned by the service node.

There exist the following problems in exception handling of Web service: They lack system approach and enough processing capacity to support quickly and efficiently the exception handling logic development. They can't resolve all kinds of

---

H. Guan (✉)

The State Key Lab of Software Engineering, Wuhan University, Wuhan 430072, China

Network Center, Wuhan Polytechnic University, Wuhan 430023, China

e-mail: [gh@whpu.edu.cn](mailto:gh@whpu.edu.cn)

S. Ying • C. Jiang

The State Key Lab of Software Engineering, Wuhan University, Wuhan 430072, China

exception, and exception handling cannot be reused. They lack the necessary exception handling set of resources. In response to these shortcomings, this paper proposes a framework of exception handling for Web services, referred to as the EHF-S, realized adding exception handling ingredients to the original Web service response message, and forming Web service response message with exception handling capabilities.

### 133.2 The Principle of EHF-S

The main function of EHF-S is to listen to response message of invoking Web service and identify exception of the Web service, using the exception handling logic provided by the framework to process the exception, resulting in response message of invoking Web service with exception handling capabilities. Response message with exception handling capabilities may still invoke the original service, but enrich exception description; it may invoke the new alternative Web services; it may also invoke the series of exception handling services, such as logging service and notification service. In order to use the logging service and exception notification service, we must produce exception information in detail, so we built the exception information and add exception handling information (message routing). Therefore, we specify the destination address and invoking information of the new Web services in the head tag of the SOAP message using WS-Addressing routing protocol. WS-Addressing provides a standard mechanism to identify Web services and Web services messages regardless of the transport protocol that is used. The SOAP protocol defines syntax specification of exception message, so we focus on SOAP message with the exception-associated fault tags, perfecting exception information (called SOAPFault), and extending the SOAP message v adding exception handling (WS-Addressing). Exception information of SOAP message is extracted at the process of exception handling for Web services, and based on the WS-Addressing we process the exception.

EHF-S detects exception information, collects the associated contents information in the exception listening and identifying module, and generates WS-Address and SOAPFault relevant content information in the exception handling module. We need to design a corresponding exception handling policy according to the type of the service layer exception. The implementation of exception handling policy will produce a corresponding action mode for exception handling (such as retry, replace, and skip); each action mode for exception handling will enrich the SOAPFault and WS-Address tag of SOAP message with exception handling capabilities. According to the corresponding WS Addressing format, the action mode adds information about the destination address. Finally, in the message processing module of EHF-S, the SOAPFault and WS-Addressing tag are added to the original SOAP message of invoking Web service, which generate a SOAP response message with exception handling capabilities, and the message is routed by EHF-S message processing module and then sent to the framework users or the appropriate service invoking.

### 133.3 The EHF-S Framework

EHF-S provides extension mechanism, supports developing exception handling for specific application exception, and registers in the framework as a new component for the exception handling resource sets. The input information of EHF-S is the response message of Web service. The output information of EHF-S is the fault-tolerant response message of Web service. At the same time, EHF-S provides logging and notification service; when an exception occurs, EHF-S will log the exception information and notify the framework's user to manually process; extension mechanism is also provided to support developers to develop exception handling services for particular application exception, and these services can be registered in the EHF-S as a new part of the resource set of exception handling.

#### 133.3.1 *The Implementation Principle of EHF-S*

In the following, we give an introduction of the work principle of EHF-S, as shown in Fig. 133.1, and we describe the detailed working process of the EHF-S:

1. Exception listening service intercepts request and response messages, real-time detects service invoking execution, captures message's exception of service layer, and checks whether exception information exists on output message of service. If it successfully detects exception, it produces a series exception handling characteristic information and exports to exception identifying service.
2. Exception identifying service receives exception information from exception listening service. Exception identifying service matches the exception characteristic information according to the exception identifying resource set and judges the exception type. The process may use exception transform service which can transform the current exception information to the specific exception type and transmit the exception category information to exception policy management service.
3. The service of exception policy management matches the exception handling resource according to the information of identifying exception type. Finding corresponding exception handling policy, it will invoke related exception handling action.
4. If execution service of exception handling action can't tackle the exception event that occurred, it will invoke the exception notification service and send the exception information to the service provider and service requestor. Meanwhile, it will invoke exception logging service and save the exception information to logging database.
5. If there exist untreated exception and newly throwing exception in the execution of exception action service, then it will return to step two and begin the exception logic handling again.

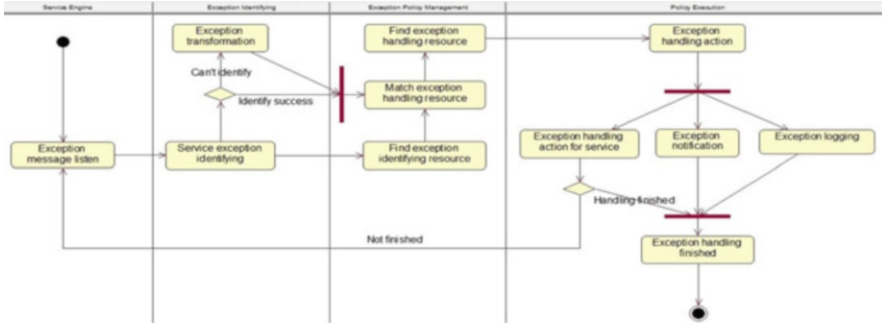


Fig. 133.1 The activity diagram of EHF-S

6. Service provider perfects exception log information exception handling policy repository, when the recurrence of such an exception, exception management processing.

### 133.3.2 The Architecture of EHF-S

As shown in Fig. 133.2, EHF-S is mainly composed of the following key components: service engine, message listener management, AOP message processing, message router management, exception identifying management, exception transformation management, exception notification service, exception logging service, exception policy management, execution service of exception action, exception identifying resource set, exception handling resource set, etc.

The foundation of EHF-S is service engine; all the exception handling is published and registered through Web service. The service engine is in charge of the execution of service component, service release and registration, service discovery, service invoking, etc. In the service engine, we can find and invoke the exception handling services, including the services of exception log and exception notification service.

The message listener management listens the response message of invoking service and determines whether there exists exception information in response message. If there exists Web service execution exception, check whether it contains custom exception in the specific parameters; if it finds, then extract the relevant exception information. The AOP message processor management extracts and analyzes the message that is accepted, identifies the format and content, and extracts the key values of exception information. It can modify the structure and add the exception handling content of the sending message, then composites to message format that fits to SOAP specification. In the message extracting function, it can support dividing the large message into small blocks and compositing to an integrated message at the destination.

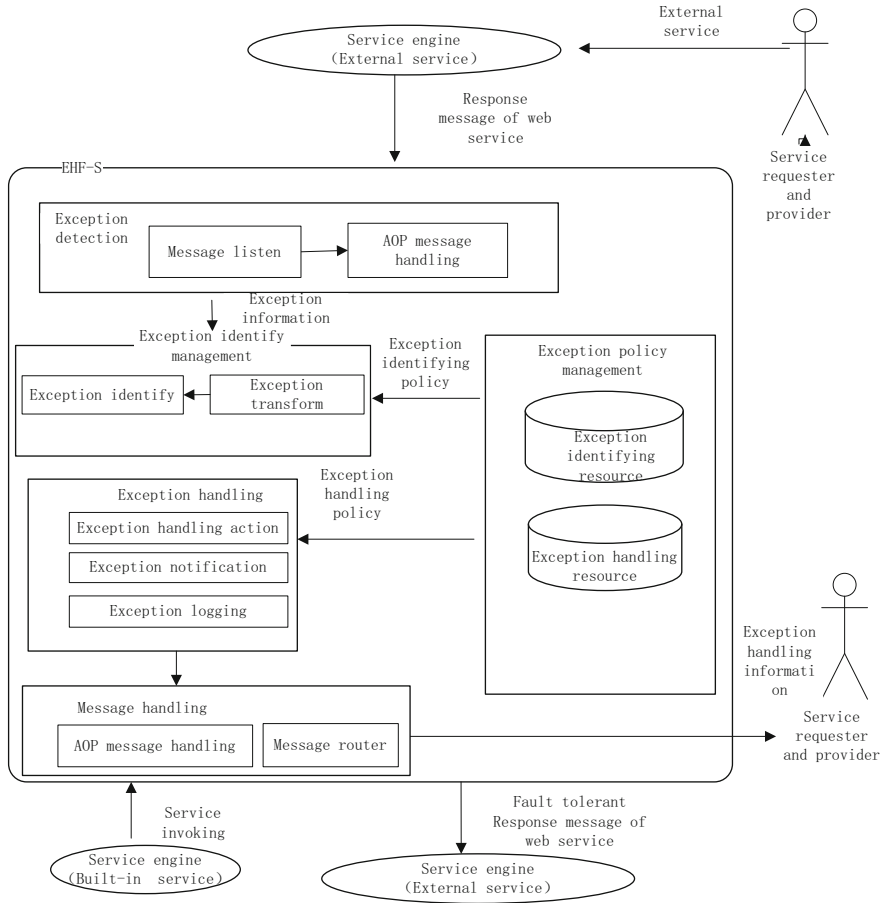


Fig. 133.2 The architecture of EHF-S

Exception transformation management realizes the conversion of exception types and transforms currently exception type into specific exception type. Exception identifying service captures the exception information of message, then matches the exception information and according to the exception identifying resource, judges the exception type of Web service. Exception identifying resource set saves all kinds of exception, and its characteristic of services provides related knowledge database for exception identifying service. Exception handling resource set provides all kinds of reusable exception handling resources. Such as ignore action, retry action for exception handling. At the same time, it has extensibility mechanism, allows adding the custom exception handling service made by developer into resource set.



Exception policy management realizes the integrated management of policy and publishes exception handling policy as Web service. It deploys and configures exception handling logic in runtime environment and deploys exception handling policy of exception handling logic into policy database; meanwhile, it adds the customizing exception handling made by developer into exception handling resource set. Exception action execution management executes the action that meets the exception handling rules; exception notification service sends the exception information to related EHF-S's users and notifies them to manual handling when the exception can't be processed. Exception logging service provides universal exception logging service.

### ***133.3.3 The Key Implementation Technology of EHF-S***

The implementation of service engine is based on Apache CXF which is an open source service-oriented framework; Apache CXF is based on a series of packaging and extended to make it able to support the deployment of service components, uninstall, start, stop, and configure. Message listening service is based on Apache CXF too; CXF provides message interception mechanism which builds on a general message layer. The implementation of message router is based on Apache Camel; message splitting divides message into multiple parts that is of fixed length. Through the aggregator component of Camel, the message composition aggregates message block based on related message ID. AOP message handling is based on AspectJ Development Tools (AJDT) of Eclipse Foundation; it can develop AOP programs. AOP message handling service utilizes relevant components of Camel through the splitter component of Camel.

The implementation of exception notification service is based on WS-Notification (WSN) criterion. The exception handling policy set uses WS-Policy criterion. We utilize the Apache Neethi implementation of the WS-Policy framework for editing and storing policies and convert the rules of policy set into rule engine's files. The implementation of rule engines is based on JBOSS Drools. We reference the rule file of WS-Policy files through WS-PolicyAttachment and use RuleFlow module of JBOSS Drools to orchestrate exception handling rules; it executes the exception handling action through Drools Fusion module. Policy management uses Guvnor module of Drools to edit and manage the policy files (or rules). Guvnor uses Web-based business rules management system, implements rules management and dynamic updates, provides a knowledge base of rules management, and enables developers and system administrators to manage the business rules online.

## 133.4 Related Research

Many scholars present exception handling method and framework for Web service. Sheng Quan et al. present SELF-SERV platform which uses configurable exception handling strategies, and based on the predefined exception handling strategies deal with the runtime exception of Web services with peer-to-peer network environment [1]. Giuliana Teixeira Santos et al. propose a Web service fault-tolerant infrastructure. The facility provides an agent that can be used on the interaction between client and server, and the agent adopts active fault-tolerant technology to achieve transparent fault-tolerant of client [2]. Hai L proposes a method to capture an exception of the outer layer in the process of Web services session and introduces enhanced Web service session context (CeWSC) mechanism to obtain the external context of the participants using the SOAP header and the confirmation message. An event-driven mechanism is presented to merge the context of an exception to the exception handling of composite Web services [3]. Gerald Kotonya et al. describe a differentiation-aware fault-tolerant framework for Web services; it supports for fault-tolerant framework of service-oriented computing, the framework uses asynchronous messaging agent LAMB to provide a news environment for different fault-tolerant protocol, it provides a plug-in way to express fault-tolerant protocol for a processing model, and fault-tolerant services container (sandbox) makes the Web service to be exposed and be found [4]. Liu Chen proposed a uniform rule-based exception handling of service-oriented software and the corresponding exception handling framework [5].

Most researches of exception handling mechanism focus on specific exception handling policy; there is little research on configurable and extendable exception handling framework. Our framework can significantly improve the ability of these aspects.

## 133.5 Conclusion

We propose a policy-based Web service's exception handling framework to handle exceptions in business processes, the framework realizes adding exception handling ingredients to the original Web service response message, thus forming a new Web service response message with exception handling capabilities. The framework provides exception characteristics resource set of exception type and exception handling resource set for each exception type and support developing exception handling service for particular application exception. Our framework simplifies the development and maintenance of business processes. Therefore, the developers could fast reuse the existing exception handling model.

**Acknowledgements** This work has been supported by the National 863 Program of China under Grant No. 2013AA102302 and the National Natural Science Foundation of China under Grant Nos. 61070012, 61272108, 61272113, and 61170022.

## References

1. Sheng Q. Z., Benatallah, B., Dumas, M., & Oi-Yan Mak, E. (2002). SELF-SERV: A platform for rapid composition of web services in a peer-to-peer environment. In *Proceedings of the 28th VLDB conference, VLDB endowment* (pp. 1051–1054), United States.
2. Santos, G. T., Cheuk Lung, L., & Montez, C. (2005). FTWeb: A fault tolerant infrastructure for web services. In *Proceeding of ninth IEEE international enterprise computing conference (EDOC 2005)* (pp. 95–105), Enschede, The Netherlands.
3. Liu, H., Li, Q., & Chiu, D. K. W. (2007). Enhancing web services conversation with exception contexts for handling exceptions of composite services. In *The fourth IEEE international conference on enterprise computing, E-commerce, and E-services* (pp. 39–46). Piscataway: IEEE Computer Society Press.
4. Kotonya, G., & Hall, S. (2010). A differentiation-aware fault-tolerant framework for web services. In *International conference on service oriented computing* (pp. 137–151). German: Springer.
5. Liu, C., Xu, Y., Deng, F., et al. (2010) A rule-based exception handling approach in SOA. In *2010 international conference on computer application and system modeling (ICCASM 2010)* (pp. 137–141). San Antonio, TX: IEEE CPS.

# Chapter 134

## Resource Congestion Based on SDH Network Static Resource Allocation

Fuyong Liu, Jianghe Yao, Gang Wu, and Huanhuan Wu

**Abstract** In order to reduce the operation blocking rate of static resource allocation in SDH Mesh network effectively, balance network traffic, optimize the allocation of network resources, enhance the success rate of multiline information routes, and improve the overall performance of the network. This chapter introduced resource congestion avoidance algorithm (RCAA) based on the adjustment, which can effectively solve the resource congestion in the static resource allocation. In order to prove the feasibility of this RCAA, three simulation examples of resource allocation theory were adopted. Through analysis validation of these three examples, this article proved that RCAA based on the adjustment proposed in this paper can effectively reduce the blocking rate and improve the overall performance of SDH network. RCAA based on adjustment is more superior to ANM. RCAA can avoid resource congestion problems caused by the allocation of resources effectively.

### 134.1 The Concept and Background of SDH Network

SDH (Synchronous Digital Hierarchy) [1] network consists of some basic network elements (NE), fuses the functions of multiple connection, line transmission, and exchange, and is a summarized information-transferring network operated by unified network management system. Thus, it can be a general technology system which is both suitable for optical fiber, microwave, and satellite transmission. It has the functions of effective management, real-time operation monitor, dynamic network maintenance, and interoperability of different manufacturers' equipments

---

F. Liu (✉) • J. Yao • G. Wu • H. Wu  
College of Information Engineering, Tarim University, A'er'la 843300, Xingjiang, China  
e-mail: [feng\\_yong2122@163.com](mailto:feng_yong2122@163.com); [417416506@qq.com](mailto:417416506@qq.com)

so as to greatly improve the utilization of network resource, reduce the cost in management and maintenance, and achieve flexible and reliable as well as efficient operation and maintenance. So it has been the hot issue for development and application in transmission technology and has attracted widespread attention.

PDH (Plesiochronous Digital Hierarchy) has been generally applied in transmission network before SDH. With the rapid development of information technology and the tremendous increase of user, users want to be provided with all kinds of circuits quickly, economically, and effectively by transmitting network. The inherent defects on PDH have dissatisfied the requirement of the modern information network transmission. In order to adapt to the rapid development of modern information society, the experts in American BELL Communications Research institute put forward synchronous optical network (SONET) and the corresponding standards which have been established as the new standards for the digital system in 1986. The Consultative Committee of International Telegraph and Telephone (CCITT) decided to make the modest modifications to SONET and renamed it as Synchronous Digital Hierarchy (SDH) system in 1988 [1, 2].

The allocation problem of communication network resources can be divided into four aspects: resource scarcity relatively, service diversification, decentralization of the resource distribution, and commercialization of the application. Increasing users made the network resources scarce relatively. Due to improper management of resource allocation, the entire network system cannot yet achieve good performance, and the users' investment also cannot get the corresponding returns. Such results may be caused by link fault, virus, or server trouble light reason. However, in fact, one of the main reasons is unreasonable allocation of resources which causes flow bandwidth contention.

Now we have a variety of SDH network routing algorithms. Key Link Routing algorithm is based on the static network resource allocation algorithm, and heuristic algorithm is based on the shared sets. Because these two algorithms only consider the current operation and network status rather than the future operation, they hold some limitations on improving the overall performance. Though dynamic resource allocation algorithms, such as minimum hop routing algorithm (MHA), the shortest path algorithm (MSP), and minimum interference routing algorithm (MIRA), can solve some certain problems, it does not take into account the influence on other panel points between in-and-out panel points and in the flow between in-and-out panel points or links so as to appear network bottleneck effects, cause congestion, and lead to low utilization. In this paper, the main research direction is the optimization of static network resources. And based on the predecessors' study of various static resource allocation algorithms, this paper proposed an improved static resource allocation algorithm [3, 4].

## 134.2 The Presence of the Resource Congestion Avoiding Algorithm

In any process of network communication, we have to face to network bandwidth resource allocation, while the static resource allocation in SDH network is just a branch of communication network resource allocation which refers to reserving required resources before transferring information. Static resource allocation of SDH Mesh network is a research hotspot currently. Planning network resource can effectively reduce blocking rate and improve the overall performance of the network.

When allocating the static resource, we count information by routes item by item. So we only consider the current information transmission and network status without considering the impact on the future information transmission during this process. So the earlier information is likely to occupy some resources, which results in the failed transmission of the later information. This is called resources allocation congestion [5]. In order to achieve the real equilibrium of network traffic, we should optimize the allocation of network resources to solve the problem of resource blocks.

Figure 134.1 is a simulative SDH network with seven links and six node points. Assume that the remaining available capacity of each link is 1; two users deliver information at the same time. They are respectively from node E to node F and node D to node B. The requesting bandwidth is 1. When calculating the first path, we get the shortest path B-E-C-F based on the current topology and link costs. Now, we figure out the second path. Because of the occupation of B-E and C-F, we cannot allocate the resources. Therefore, the second path cannot transfer or cause delay of information transmission.

It is a complex problem to make multichannel route information successfully. The resource congestion avoiding algorithm (RCAA) based on the adjustment put forward in this paper can effectively solve such problems. RCAA can adjust the information transmission path for the first user, make the business not take up B-C link, and get C-D-E-F link. Then, we need to calculate the second user's information to get a D-C-B link so as to achieve the equilibrium of the network traffic. The key point of the algorithm lies in adjusting the unreasonable take-up information resources to make the following delivery information calculate the path smoothly.

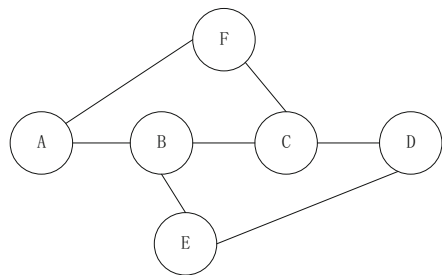


Fig. 134.1 Topology

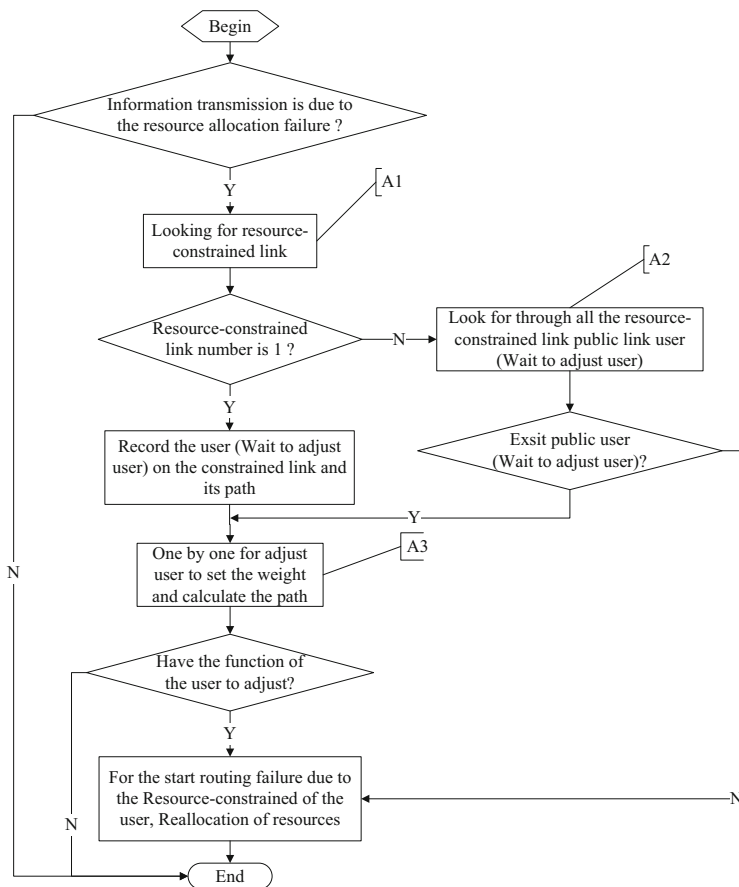
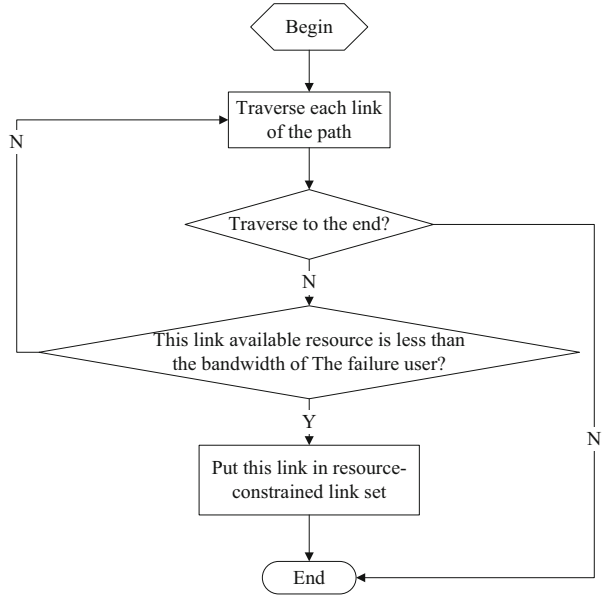


Fig. 134.2 The flow chart of resource congestion avoidance algorithm

### 134.2.1 Resource Congestion Avoidance Algorithm

The basic procedure of path calculation in simulative SDH network is as follows: weigh the design, calculate the path, and allocate the time slot. There may be some special cases in this process. If user A cannot route because the available resources of the link L is not enough, resource congestion may cause the failure of routing. If we move user B occupying in link L, the available resources on the link are greater than or equal to the bandwidth of the user As. Thus, user B will be able to succeed in allocating resources. This is the core idea of resource congestion avoidance algorithm—RCAA [6]. The procedure is as follows (Fig. 134.2).

**Fig. 134.3** The flow chart of looking for resource-constrained link

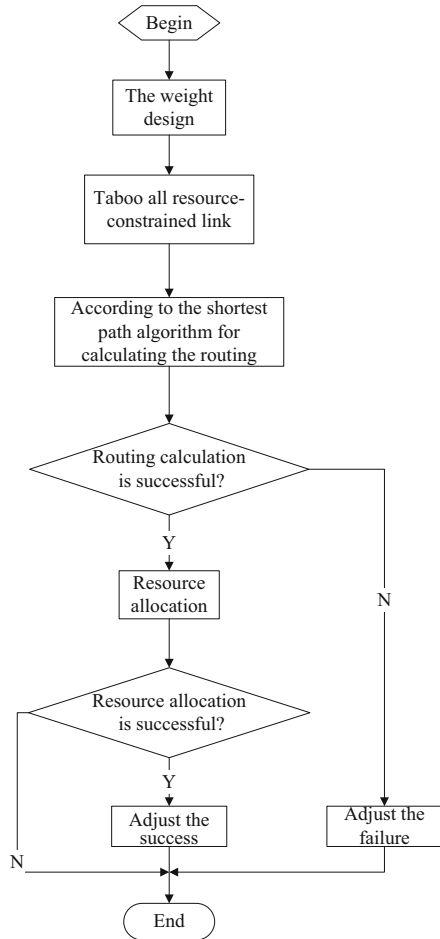


### 134.2.2 Algorithm Analysis of the Key Procedure of RCAA

1. A1: Search the resource-constrained link of the failed user in resource allocation. From the information of the static resource allocation algorithm, we can see that some users may meet with resource allocation failure after successful path calculation. For this kind of business, we traverse each link and judge whether the available capacity of each link is less than the bandwidth of the failed business. If it is, the link is resource-constrained; otherwise, the link is not resource-constrained. So we can get all the resource-constrained links. Specific process is shown in Fig. 134.3.
2. A2: When the resource-constrained link is greater than 1, search all users through the resource-constrained link. First, remove the users who go through the first two resource-constrained links respectively, and set out the intersection in this two. And then, remove the users who go through the third resource-constrained link, and set out the intersection with user collection. It doesn't stop by this method until all resource-constrained routes are went through.
3. A3: Weigh and calculate the path for the pending adjust users one by one. Weigh the path for the pending adjust users. Adopt the static design weighing method [4] in resources allocation, stop all resource-constrained links, and forbid the pending adjust users in going through those links.



**Fig. 134.4** The flow chart of link adjustment

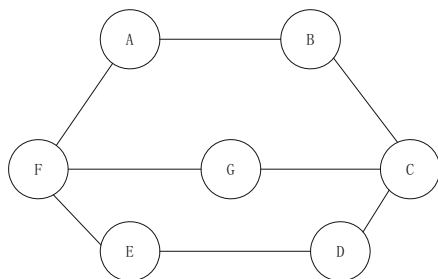


Apply the shortest path algorithm—Dijkstra Algorithm [7]—to calculate the path of the user. If the path calculation succeeds, the resources allocation and adjustment succeed. If not, the user will go back to the original path, so the adjustment fails. Specific process is shown in Fig. 134.4.

### 134.3 Theoretical Simulation and Result Analysis of RCAA

Several examples will be provided to verify the performance of the RCAA, and theoretical simulation comparison will be made to the method by business adjustment according to the weight calculation. To describe easily, we call this ANM (algorithm with no modulation) which does not consider business adjustment.

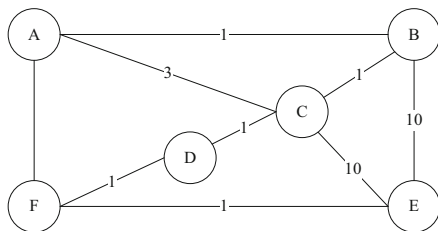
**Fig. 134.5** Topology of panel point



**Table 134.1** Routing calculation comparison

Algorithm	Calculation results
ANM (without adjustment)	User 1: Work way route: F→G→C Protect the road route: F→A→B→C
RCAA	User 1: Calculate failure User 1: Work way route: F→A→B→C Protect the road route: F→E→D→C User 2: Work way route: F→G→C

**Fig. 134.6** Topology of panel point



There are two service requests (shown in Fig. 134.5). The first service requests is user 1 from node F to node C with one VC4 bandwidth and without a hop count limit. The second service requests is user 2 from node F to node C with one VC4 bandwidth and without two hop count limits. Each link only has one VC4 bandwidth to be available. The optimization goal of the network is the smallest hop count limit.

Table 134.1 shows the simulation results for routing computation of the business. Seeing from this table, although ANM can guarantee two successful separate ways to user 1, at the same time, it takes the resources of path F→G→C, and only path F→E→D→C is free. But the path is not in conformity with the hop constraints of user 2. So the resource allocation calculation for user 2 fails. RCAA adjusts the links F→G and G→C and recalculates a work path for user 1, allowing user 2 to route successfully.

As shown in Fig. 134.6, the numbers on the link stand for the weights of the links. There is a user who has a request. He or she needs to compute a special

**Table 134.2** Routing calculation comparison

Algorithm	Calculation results
ANM (without adjustment)	Calculate failure
RCAA	Work way route: $A \rightarrow C \rightarrow D \rightarrow F \rightarrow E$ Protect the road route: $A \rightarrow B \rightarrow E$

**Table 134.3** Routing calculation comparison

Algorithm	Calculation results
ANM (without adjustment)	Calculate failure
RCAA	Work way route: $A \rightarrow B \rightarrow E \rightarrow D$ Protect the road route: $A \rightarrow F \rightarrow C \rightarrow D$

protection user from node A to E. He requests separated link with a VC4 bandwidth. The optimization goal of network is minimum weight and maximum business rate, and the network link capacity is 16 VC4.

Table 134.2 shows simulation results of the routing calculation. As we can see from the results, ANM cannot calculate the path to business successfully according to the weight calculation routing algorithm directly without adjustment, but it can successfully calculate by using RCAA with considerable adjustment. Mainly because after the ANM according to Dijkstra Algorithm calculates the first work path  $A \rightarrow B \rightarrow C \rightarrow D \rightarrow F \rightarrow E$ , link  $A \rightarrow B$  and link  $F \rightarrow E$  have no resources available, and the links cannot separate. This leads to the failure of protection route computation. And RCAA adjusts the link  $A \rightarrow B$  to ensure the calculation of protection route successfully.

As shown in Fig. 134.2, assume that we need to compute a special protection business from node A to node D, and request link separation and one VC4 bandwidth. The link capacity of network is 16 VC4. The optimization goal of network is the minimum hop count.

Table 134.3 shows the path calculation result. We can see that ANM is a failure for path calculation business, while RCAA can calculate both paths. Mainly because after the ANM according to Dijkstra Algorithm calculates the first work path  $A \rightarrow B \rightarrow C \rightarrow D$ , the second path for separation cannot calculate. So it fails. However, RCAA adjusts the links  $B \rightarrow C$  and  $C \rightarrow D$  so as to ensure the successful calculation of the protection route.

From the above three theoretical simulation experiments, we can conclude that RCAA based on adjustment is more superior to ANM. RCAA can avoid resource congestion problems caused by the allocation of resources, so as to improve the success rate of resource allocation in the SDH network planning.

## 134.4 Conclusion

This paper studied the resource congestion problem in static resource allocation of SDH network and then proposed an improved RCAA (resource congestion avoidance algorithm). RCAA can adjust resource-constrained links' path and avoid these links as much as possible, so the follow-up business would be able to collocate. These three examples with the theoretical analysis demonstrated that RCAA proposed in the paper can effectively reduce the blocking rate of network service routing as well as improve the overall performance of SDH network.

**Acknowledgement** *Fund project:* Date Sharing Platform Construction of Biological Science in Tarim University (TDZKPT201201).

## References

1. Sun, X., & Mao, M. (2009). *SDH Technology* (pp. 56–93). Beijing: Post & Telecom Press.
2. Xiao, P., & Wu, J. (2008). *Principle and application of SDH* (pp. 25–78). Beijing: Post & Telecom Press.
3. Guerin, R., Orda, A., & Williams, D. (1997). Qos routing mechanisms and OSPF extensions. *Proc. of IEEE GLOBECOM'97 IEEE*, 3, 1903–1908.
4. Kodialam, M., & Lakshman, T. V. (2000). Minimum interference routing with applications to MPLS traffic engineering. *Proc of IEEE INFOCOM, IEEE.*, 2, 884–893.
5. Gu, S. (2009). *Equipment principle and application of SDH* (pp. 61–179). Beijing: Beijing University of Posts and Telecommunications Press.
6. Wang, Y., & Li, L. (2001). Considering the link load balancing and capacity limit protection design of WDM Optical Transport Network. *Chinese Journal of Electronics.*, 29(10), 1319–1322.
7. Chen, B. (2005). *Optimization theory and algorithm* (pp. 101–203). Beijing: Tsinghua University press.

# Chapter 135

## Multilayered Reinforcement Learning Approach for Radio Resource Management

Kevin Collados, Juan-Luis Gorricho, Joan Serrat, and Hu Zheng

**Abstract** In this paper we face the challenge of designing self-tuning systems governing the working parameters of base stations on a mobile network system to optimize the quality of service and the economic benefit of the operator. In order to accomplish this double objective, we propose the combined use of fuzzy logic and reinforcement learning to implement a self-tuning system using a novel approach based on a two-agent system. Different combinations of reinforcement learning techniques, on both agents, have been tested to deduce the optimal approach. The best results have been obtained applying the Q-learning technique on both agents, clearly outperforming the alternative of using non-learning algorithms.

### 135.1 Introduction

The management of resources made on the radio interface for mobile access networks has traditionally followed a static approach [1, 2]. Any mobile operator, on pursuing a satisfactory quality of service, determines the amount of resources to be deployed on each base station, including the split in between those resources devoted to handovers and the remaining resources available to set up new connections [3]. Nevertheless, this working strategy seems to be too short-sighted for what will be necessary in the near future when upcoming optimization challenges will come into play. Key issues like minimizing the energy consumption, sharing the infrastructure among different operators on deploying the 4G mobile systems, or even borrowing radio resources among them are becoming desirable targets for the future mobile communication systems.

---

K. Collados • J.-L. Gorricho (✉) • J. Serrat  
Telematics department, Polytechnic University of Catalonia (UPC), Barcelona, Spain  
e-mail: [juanluis@entel.upc.edu](mailto:juanluis@entel.upc.edu)

H. Zheng  
Key Laboratory of Universal Wireless Communication (BUPT), Ministry of Education,  
Beijing University of Posts and Telecommunications (BUPT), Beijing, China

In this global scenario, one of the most promising approaches for an intelligent management of available resources comes from the use of machine learning techniques, learning from the system behavior to deduce suitable policies for managing those available resources [4], and policies pursuing different goals, ranging from an optimal quality of service for an individual operator to a global inter-domain efficient system. Ideally, the ultimate goal will be the implementation of self-tuning systems due to applying, for example, fuzzy neural methodologies for the radio resource management [5], similar to our present study, although our innovative contribution comes from using a two-agent approach.

## 135.2 Working Scenario

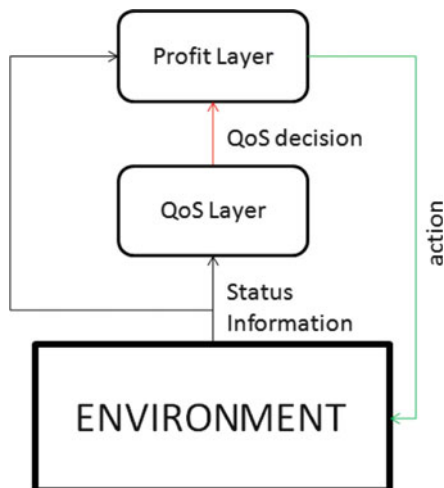
In our study case, we envisage a working scenario where several mobile operators are providing service in a completely overlapping fashion, as opposed to the approaches focusing on a single provider owning all the infrastructure [6]. More than that, as already pointed out for the future 4G mobile systems, we assume that all operators share the same base station infrastructure. This way we take a step forward from other approaches dealing with more than one access technology on any base station but still focusing on a single operator [7]. As usual, the geographical area is divided into many cells, a different base station provides service for each cell, and there is some overlapping on the coverage area provided by neighboring base stations.

Our aim is to allocate radio resources on the air interface for each base station to satisfy some given quality of service with a self-tuning system for the parameters governing the base stations' operation. To this end our strategy combines the use of fuzzy logic and reinforcement learning [4, 5], but, in our case, on each base station, two different agents will work together in order to manage the corresponding resources. It has been done in this way to separate two different goals, the quality of service and the operator's economic benefit. Both agents try to maximize their corresponding goals, although only one of the agents takes actions to modify the operating parameters on the base station, as shown in Fig. 135.1. As we can see, there is in practice a closed-loop control as the QoS layer influences the Profit layer jointly with the status information from the environment, and the Profit layer takes an action, which will influence the QoS layer on the next cycle. Hence, there is a mutual interaction between both agents, although not carried out directly from the Profit layer to the QoS layer due to the way we have implemented our approach.

In this scenario the operational parameters to work with are the following:

1. The coverage area per base station, configurable by tuning the power control mechanism.
2. The distribution of channels per base station, configurable by splitting up the channels in different categories on dealing with different types of services or establishing dedicated channels for handover.

**Fig. 135.1** Structural model



3. The total amount of channels per operator at each base station. In this case a trading mechanism is considered, so different operators exchange channels on their own benefit due to the irregular demand from their respective end users.

Using fuzzy logic we can easily cope with a continuous input space defining the possible system states and actions to be taken. The alternative would be the use of reinforcement learning alone; but, in this case, the disadvantage comes from working with a discrete number of states and actions [6, 7], producing an approach with a worse performance.

Two alternatives on implementing the reinforcement learning mechanism have been tested, the actor-critic and the Q-learning techniques, as can be shown in Figs. 135.2 and 135.3. For both algorithms, each action takes a different fuzzy logic weight  $\alpha_{R_n}$ , which is the output due to applying the fuzzy rule, but also an additional weight  $w_t$  to be learned [4].

For comparative reasons, a non-learning system has also been implemented using only the fuzzy logic technique. Regarding the time domain, all simulations work with the same sliding window to obtain the input for any approach. In the following sections, we describe in details how both agents of our architectural model work.

### 135.3 QoS Agent Design

The purpose of this agent is to achieve a given quality of service (QoS) regarding two basic parameters: the blocking rate (measuring the unavailability to set up a new service connection) and the dropping rate (measuring the unavailability to hand over a connection between two base stations due to the user's travelling

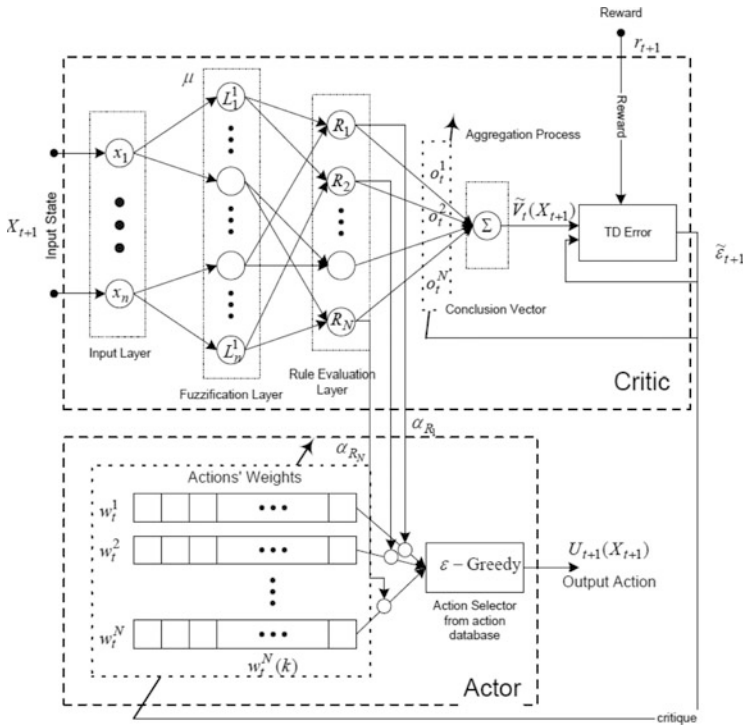


Fig. 135.2 Actor-critic technique

trajectory). Based on these two parameters, the reward expression needed to implement the reinforcement learning mechanism is formulated as

$$Reward = (T_B - B) + \beta(T_D - D)$$

where  $T_B - B$  is the difference between the actual blocking rate and a given target value and  $T_D - D$  is the difference between the actual dropping rate and a given target value. Usually the dropping rate is much more critical than the blocking rate on measuring the quality of service; consequently, a  $\beta$ -factor is added to emphasize this parameter in front of the other. The two inputs, the blocking and dropping rates, are labeled according to six fuzzification categories; for each category we obtain a membership degree of the input through a fuzzification stage as shown in Fig. 135.4.

The fuzzification-stage outputs are the inputs to a rule matrix (2 dimensions) as shown in Fig. 135.5; the rule matrix defines 5 rules (fuzzy rules) to produce the corresponding decision in a simple manner.

The rule-matrix output will be processed by the RL algorithm according to the reward definition, producing a decision due to the following procedure:



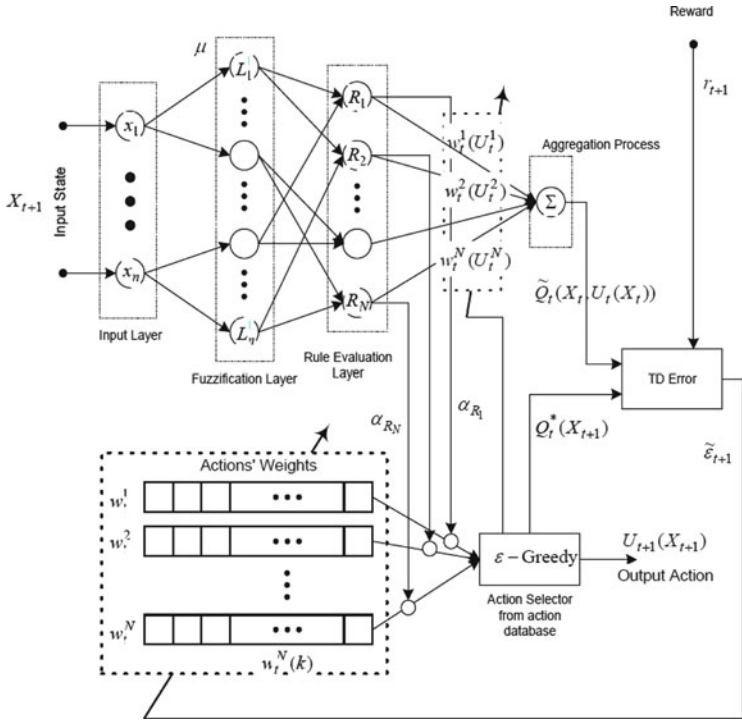


Fig. 135.3 Q-learning technique

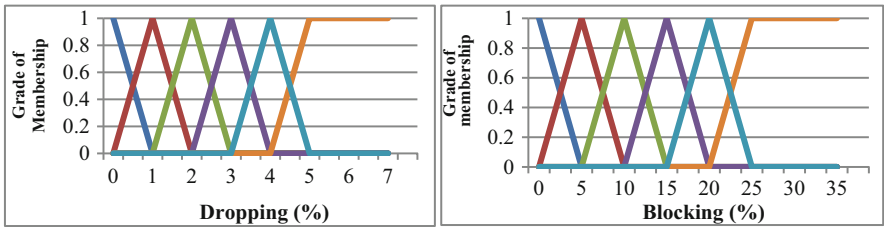
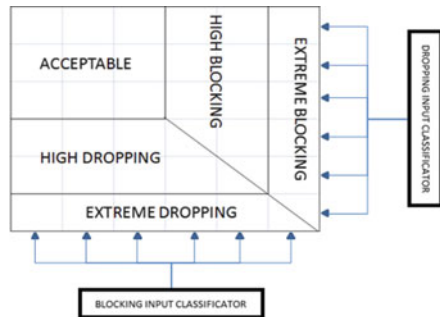


Fig. 135.4 Dropping and blocking rate labeling

Fig. 135.5 Rule matrix



1. The rule weight  $\alpha_{R_n}^i$  and action weight  $w_t^i$  are combined to produce a  $s_t^i$  selection weight; this way a compromise is acquired between the present needs and the learned behavior up until now.
2. Each weight  $s_t^i$  is evaluated following some predefined criteria; the criteria will be the requests of the agent for each rule; these are modification of the coverage area, modification of the distribution of channels in between handover and new service channels, and finally the request for extra channels; so for each  $s_t^i$  we will have three associated components.
3. The final decision is obtained by a combination of all  $s_t^i$ , deducing the definitive three-component request for the present input. This final decision is sent to the next agent, becoming one of its inputs.

### 135.4 Economical Benefit Agent Design

The objective of this agent is to obtain the maximum economical benefit for the operator; to accomplish this objective, the reward function used by the reinforcement learning mechanism is defined as

$$Reward = \frac{Load * Price * Num\_ChannelsOwned}{Call\_Duration} - Cost * Num\_ChannelsOwned$$

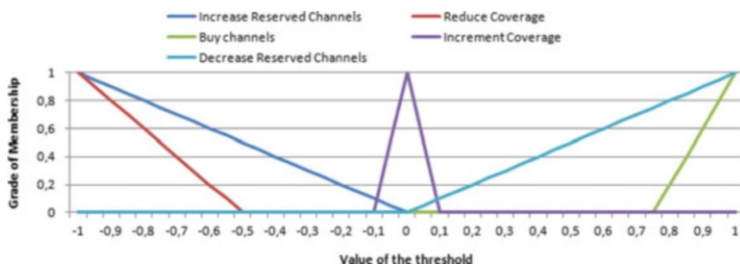
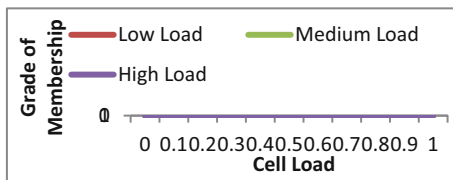
It is noteworthy that the QoS is not included in the reward function; this is because the QoS agent has already taken that into account, and it provides its input to the economical benefit agent. According to this, for the second agent, the inputs are:

- The system load
- The requests from the QoS agent
  - Modification of the coverage area
  - Modification of the distribution of channels
  - Need of extra channels

To avoid learning actions that are not feasible, some additional considerations must be applied invalidating the action. These considerations are the following: selling channels is forbidden if the minimum amount of channels that the operator must maintain is reached, reducing the coverage area is forbidden if the minimum cell radius that must be kept to assure some overlapping is reached, and increasing the coverage area is forbidden if the maximum transmitted power has been reached.

The system load is labeled according to three different categories, as shown in Fig. 135.6, low, medium, and high. When the system detects a low-load state, the weight for buying new channels is reduced and the weight for selling is increased and vice versa. This way the system tries to be led to a medium-load state.

**Fig. 135.6** Cell load labeling



**Fig. 135.7** Request of the QoS agent: modification of channel distribution

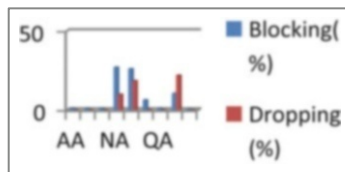
For labeling the requests of the QoS agent, we need to classify the modification of the channel distribution; in Fig. 135.7 we show the classification of this input; for this study the distribution of channels is split in between channels reserved for handover and channels available for new services. A negative value of this parameter implies an increase on the number of channels reserved for handover, but if the demand is high, the problem cannot be solved by only increasing the channels reservation; another solution is needed; in this case a reduction of the coverage area is applied, trying to delay the time for the handover execution and also to reduce the amount of handovers managed by the base station. A positive value of the parameter allows us to free channels for new connections. As it happened before, if the demand is high, the problem cannot be solved only by this approach; in this case the alternative will be to buy new channels if available. Finally, if the system performs properly in terms of QoS, it tries to increase the coverage area to benefit from more incoming calls.

The other requests of the QoS agent are evaluated directly without being classified, because there is no alternative to be applied. To avoid overacting on the system, a given time between actions is imposed, time defined by the skilled technician; this way the system can better learn the optimal action before acting again.

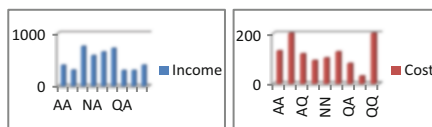
### 135.5 Simulation Results

To test the correctness of the algorithms by themselves without being affected by the behavior of another learning technique, each learning technique has been tested individually in the same scenario and under similar circumstances.

**Fig. 135.8** Simulation: dropping and blocking rate results



**Fig. 135.9** Simulation: income and cost results



To simplify the notation used in the following figures, any learning system will be labeled with two consecutive letters—the first one applies for the economical benefit agent and the second for the QoS agent. Also the names of the learning algorithms are labeled by their initials; hereafter A means the agent using actor-critic technique, Q means the agent using Q-learning technique, and N means the agent using a non-learning algorithm. Independently of the combination of learning techniques, we have run the simulations to assure that the overall system remains in a medium-load state, which is one of the inputs of the economic benefit agent.

As shown in Fig. 135.8, the average degradation in terms of QoS for the same client demand is higher when the non-learning technique is used on any or both agents. The QoS for the remaining learning combinations results in a dropping and blocking rate even below their target values, 1 % for dropping and 5 % for blocking.

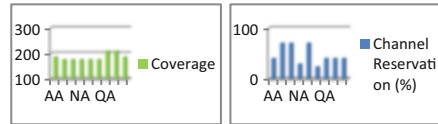
On the other hand, the operator’s profit is directly proportional to the amount of established calls, as in all simulations the price for call is kept constant; besides, the cost of maintenance is proportional to the amount of channels managed by the operator; this way more channels turn into more established calls, but the cost of maintenance is also increased. In Fig. 135.9 the cost and the income for each combination of learning techniques are shown.

Those systems having reached an inferior QoS usually have a better economic benefit; one of the reasons is that only the economic benefit agent is able to act on the environment.

One of the drawbacks on using a non-learning algorithm is its dependency with respect to the size of the sliding window used to obtain the sequence of inputs; if the window size is too small, it overacts on sudden peaks of clients’ demand, producing unnecessary changes on the system parameters to cope with the incoming demand; on the other hand, if the size is too big, it works with an unrealistic view of the environment behavior.

Figure 135.10 shows the alteration of the coverage area and the percentile of channels reserved for handover; usually those systems with a higher income also work with smaller radius of coverage per cell.

**Fig. 135.10** Simulation: coverage and channel reservation results



## 135.6 Conclusion

In this paper we have evaluated the combined use of fuzzy logic and reinforcement learning to implement a self-tuning system governing the working parameters of the base stations on a mobile network system. Targeting a double goal to assure some given quality of service and to maximize the economic benefit, the simulation results have shown that the two-agent approach that we have considered is suitable to handle both goals simultaneously; nevertheless, depending on the learning technique used on both agents, the behavior is slightly different. In any case, the obtained results applying reinforcement learning outperform the alternative of using non-learning algorithms.

**Acknowledgements** This work has been done in the framework of the EVANS project (PIRSES-GA-2010-269323), and with the support of projects TEC2012-38574-C02-02 and TEC2012-32531 from Ministerio de Ciencia y Educacion, and grant 2009-SGR-1242 from Generalitat de Catalunya.

## References

1. Tölli, A., Hakalin, P., & Holma, H. (2002). Performance Evaluation of Common Radio Resource (CRRM). In *IEEE International Conference on Communications* (pp. 3429–3433).
2. Pérez-Romero, J., Sallent, O., Agustí, R., Karlsson, P., Barbaresi, A., & Wang, L. (2005). Common radio resources management: Functional models and implementation requirements. In *IEEE Personal, Radio and Mobile Communications* (pp. 2067–2071).
3. Altman, Z., Dubreil, H., Nasri, R., Nawrocki, M. J., Dohler, M., & Hamid Aghvami, A. (2006). *Understanding UMTS radio network modelling, planning and automated optimisation*. Chichester: Wiley.
4. Naeeni, A. F. (2004). *Advanced multi-agent fuzzy reinforcement learning*. Master Dissertation, Computer Science Department, Dalarna University College, Sweden.
5. Giupponi, L., Agustí, R., Pérez-Romero, J., & Sallent, O. (2008). A novel approach for joint radio resource management based on fuzzy neural methodology. *IEEE Transaction on Vehicular Technology*, 57(3), 1789–1805.
6. Stefan, A. L., Ramkumar, M., Nielsen, R. H., & Prasad, N. R. (2011). A QoS aware reinforcement learning algorithm for macro-femto interference in dynamic environments. In *International Congress on Ultra Modern Telecommunications and Control Systems* (pp. 1–7).
7. Vucevic, N., Pérez-Romero, R., Sallent, O., & Agustí, R. (2011). Reinforcement learning for joint radio resource management in LTE-UMTS scenarios. *Computer Networks*, 55(7), 1487–1497.

# Chapter 136

## A Network Access Security Scheme for Virtual Machine

Mingkun Xu, Wenyuan Dong, and Cheng Shuo

**Abstract** Virtual machines have been widely adopted as servers nowadays. They have essential difference with physical machine. We can utilize the feature of virtual machine to let them be safer and resist an attack from Trojan and hackers. This paper introduces a kind of network access security scheme, which deploys the execution of security strategy outside virtual machine and monitors virtual machine's access to security-sensitive device. The measurements above can transfer the control for key hardware from upper Guest OS to host a platform. Even if Guest OS is affected by virus or Trojan, host can still effectively monitor the network communication of upper virtual machine. In this project, software running in Host OS is programmed to realize the scheme introduced above, it monitors the network communication of virtual machine according to the rules written in XML format. The software can prevent Guest OS or an application running on the virtual machine from communicating with designated domain or IP address successfully, which verifies the effectiveness of the proposed security scheme.

### 136.1 Introduction

Virtual machines have been widely adopted as ordinary servers or cloud servers nowadays [1, 2]. This paper researches and realizes a kind of security scheme for network access based on virtual machine, to provide a high-standard security solution of network access.

Let us introduce virtualization first. Virtualization allows several virtual machines to run on the same physical computer. Each virtual machine has a set

---

M. Xu (✉) • W. Dong  
Beijing University of Posts and Telecommunications, Beijing 100876, China  
e-mail: [henry7120@hotmail.com](mailto:henry7120@hotmail.com); [thisisapollo@163.com](mailto:thisisapollo@163.com)

C. Shuo  
The University of Macau, Macau 999078, China  
e-mail: [xcsbruce@163.com](mailto:xcsbruce@163.com)

of hardware of its own, such as RAM, CPU, and NIC card, which supports the OS and applications to run. No matter what practical physical hardware is mounted, they are viewed as a set of standard hardware. Virtual machine is sealed in files, so that it can be quickly saved, copied, and deployed. The whole virtual system including application with complete configuration, OS, BIOS, and virtual hardware can be move from one server to another, to maintain service without interruption and keep continuous balance of computation load [3].

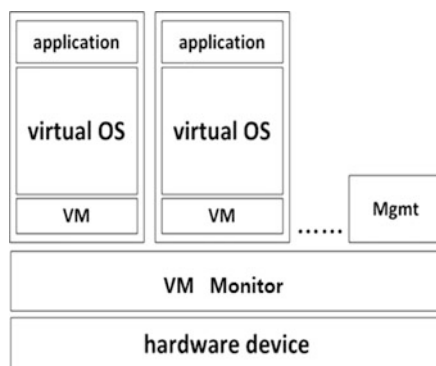
This paper arranges the execution of security scheme out of Guest OS, the scheme runs independent of Guest OS, and monitors the access of virtual machine to security-sensitive hardware device; this prevents virtual machine from external attack.

No matter whether virtual machine is attacked passively by outside Trojan and virus or virtual machine actively communicates with network, its communication packet is monitored by the network access security filtration software VFirewall introduced in this paper. Meanwhile the security measurements in original OS environment are still applicable [4]. Thus, any kind of communication between hacker and virus-affected Guest OS will be detected out.

## 136.2 General Design of a Network Access Security Scheme for Virtual Machine

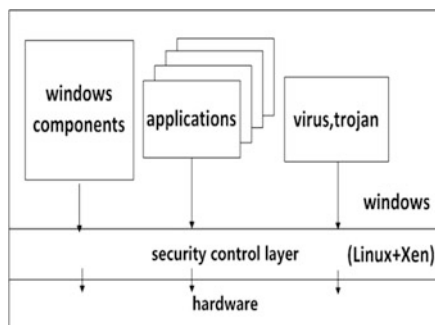
### 136.2.1 The Construction of Virtual System

VMware or XEN is a well-known virtual machine software which is illustrated as VM monitor in Fig. 136.1. They can be installed on Linux to support Guest OS noted as virtual OS in Fig. 136.1. Configuration of VMware is relatively simple, here take XEN as an example, to illustrate the construction of virtual system [4].



**Fig. 136.1** The principle of virtual system

**Fig. 136.2** The network access control scheme for virtual machine



Here take CentOS, the open source version of Red Hat Linux as Host OS. Pay attention to that only if CPU chip supports VT (Intel) or AMD-V (AMD) functionality; Windows can be installed on XEN. Under Linux Gnome graphic interface, XEN Virtual Machine Manager should be installed first, and then Windows OS can be installed on XEN [5].

### ***136.2.2 The Principle of Network Access Control for Virtual Machine***

This paper takes Linux as Host OS, which directly runs on hardware, and takes VMware or XEN as virtual machine software. Network security control layer is realized as application of Linux [3]. The architecture of the network access control scheme for virtual machine is illustrated in Fig. 136.2.

## **136.3 The Realization of Network Access Control Scheme for Virtual Machine**

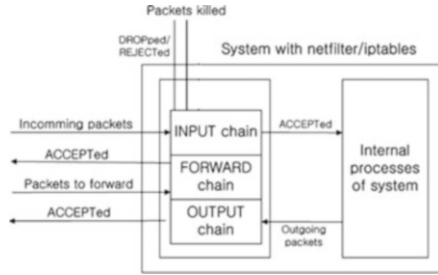
### ***136.3.1 The Architecture of Netfilter Software***

VFirewall software introduced in this paper calls the functionality supplied by Netfilter software. Netfilter is a firewall framework designed by Rusty Russell in Linux kernel, which can realize many functionality of security process, such as data packet filtration, data packet process, address disguising, transparent proxy, dynamic Network Address Translation (NAT), Media Access Control, address filtering based on user identity, filtering based on status, and speed control of packet [6].

Netfilter has defined five hook functions for IP protocol. These hook functions may be called on five key locations of the protocol stack that data packets pass



**Fig. 136.3** The packet process model by Netfilter module in Linux kernel



through. In every key location of the five, data packet passed by will be captured and compared with corresponding rule chain in IP tables. Packet's destination depends on the comparison result: some packet is put back into IPv4/v6 protocol stack and transmitted continually; some packet is modified and put back into IPv4/v6 stream; some packet is dropped directly.

Figures 136.3 and 136.5 show the packet selection system in Netfilter: IP tables model. In the figure above, IP packet process procedure includes tracing connection, packet content modification, packet filtration, and so on. IP tables' model demonstrates the packet process functionality of Linux kernel, each of the functionality should be registered on corresponding hook in Netfilter, and the corresponding functionality is ultimately realized by the program.

Combination of above functionality flexibly can achieve comprehensive network security strategies. This paper mainly makes use of the packet-filtering ability of Netfilter.

### 136.3.2 The Architecture of VFirewall Software

In this paper, the network access control software VFirewall is developed in C++ [7]. It parses the network access control strategies described in XML file, checks and records the communication between the Guest OS and special IP segment or domain. VFirewall interacts with Netfilter in Linux kernel space through IP tables in Linux user space, to set network access rule and to capture network access event [8].

The data flow chart in VFirewall is showed below.

As shown in Fig. 136.4, the network security strategies are described in a XML file. The XML file, which is written in pure text mode, can be edited with any text editor. Network access rules described in XML file can be updated through human-machine interface of VFirewall.

The following access security strategy configuration in XML format can be recognized and accepted by VFirewall.

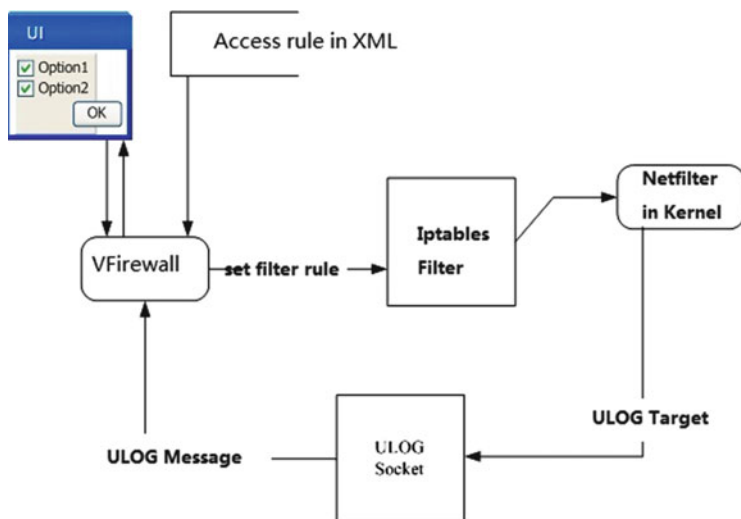


Fig. 136.4 The data flow chart in VFirewall software

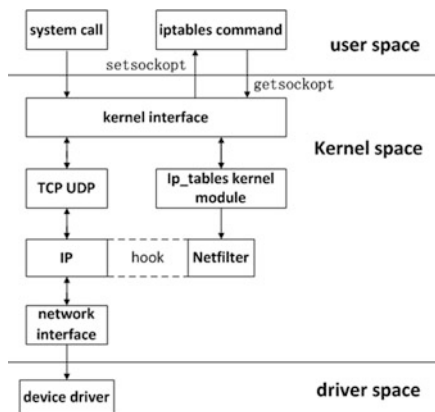
- Single IP:  
<block IP="210.25.132.38"/>
- Single domain:  
<block NS="www.bupt.edu.cn"/>
- IP address segment:  
<block IPRANGE="210.23.132.0-210.25.132.255"/>

Having obtained network access control rules from human-machine interface or XML file, VFirewall writes corresponding rules into rule chain in Netfilter. Meanwhile, VFirewall registers a log target (ULOG). If Netfilter captures the IP data packet in line with the security access rule, it will send corresponding report message to VFirewall through ULOG. Then, VFirewall writes the message into log file.

### 136.3.3 Related Work

- The IP address of virtual machine can be hidden to external network; thus, a lot of attack based on IP address will not work [9].
- As shown in Fig. 136.5, the device driver of Linux may be modified to add data stream inspection mechanism in it, so that it can also monitor and filter data received and sent by upper Guest OS. In fact this work has already been done by the same author [10]. Since the functionality of program instruction set in driver

**Fig. 136.5** Guest OS data stream can be inspected in device driver space



is less than that in user space, packet inspection in driver is incomplete compared with the VFirewall software proposed in this paper.

However, VFirewall is only suitable for virtual machine, while packet inspection in driver is practicable for both virtual machine and physical machine.

### 136.4 Conclusion

The running result of the software demonstrates that VFirewall successfully prevents the connection between Guest OS and designated domain or IP address without mounting any firewall in Guest OS, no matter whether unique Guest OS or several Guest OSs are installed in host platform; thus, the effectiveness of the network access security scheme is verified.

The communication efficiency, however, declines a little, because of the security check to access a network. We observed that the packet transmission speed may delay for about a few hundred milliseconds or seconds correspondent with the number of virtual machines launched on Linux.

**Acknowledgements** This paper is partially supported by the National High Technology Research and Development Program, No. 2011AA010704: the Key Technology and Verification of Network Security Based on IPv6 in Designated Scope.

### References

1. Green, M. (2013). The threat in the cloud. *IEEE Security & Privacy*, 11(1), 86–89.
2. Popa, L., & Kumar, G. (2012). FairCloud: Sharing the network in cloud computing. *Computer Communication Review*, 42(4), 187–198.

3. Smith, J. E., & Nair, R. (2006). *Virtual machines-versatile platforms for systems and process* (pp. 10–35). Singapore: Elsevier Pte Ltd.
4. Qin, Z. Y. (2012). Survey on virtual system security. *Application Research of Computers*, 29(5), 1620–1622 (In Chinese).
5. Jang, M. (2011). *Security strategies in Linux platforms and applications* (pp. 62–65). Sudbury, MA: Jones & Bartlett Learning.
6. Russel, R. (2002). *Linux 2.4 Packet Flitering HOWTO*. Retrieved from <http://www.netfilter.org/documentation/HOWTO//packet-filtering-HOWTO.html>
7. Lippman, S. B. (2005). *C++ primer* (pp. 20–150). New York, NY: Pearson Education.
8. Cheng, S. L. (2009). *Network access control system model research based on virtualization*. Beijing, China: Beijing University of Posts and Telecommunications (In Chinese).
9. Nestler, V. J. (2006). *Computer security lab manual* (pp. 255–261). New York, NY: McGraw-Hill.
10. Xu, M. K. (2005). Encrypt data through streams module in kernel. *Computer Engineering and Design*, 26(7), 1710–1711 (In Chinese).

# Chapter 137

## Light Protocols in Chain Network

Ying Wang, Yifang Chen, and Lenan Wu

**Abstract** Aiming at some special applications, such as monitoring of high-speed rail and monitoring of large farm field, a wireless sensor network based on chain structure is proposed. Considering of simplicity and energy saving, two light protocols, which are based on time slot and competition, respectively, are applied in the above network. Finally, the two light protocols are compared with IEEE802.15.4 protocol by OPNET simulation, and the results show that the proposed light protocols have good reliability and low energy consumption.

### 137.1 Introduction

Wireless sensor network (WSN) has caught great attention and achieved great development in recent years. Theoretically, as long as the nodes are distributed densely enough, the WSN can implement communication by self-organization, relay, and multiple hops [1, 2]. But due to the limitations of volume, cost, and battery energy, the communication distance of WSN is greatly limited for some typical applications. When the sensor nodes in the WSN are sparsely distributed, for example, in the road subsidence monitoring in the highway or high-speed rail, in the stress monitoring of the super-large bridge, as well as in the hydrological monitoring of the rivers, the network topology is often a simple chain type, and the nodes are not randomly distributed. Therefore, the traditional protocols, which can support self-organization and are applied more often in the networks with dense nodes,

---

Y. Wang (✉)

School of Information Science and Engineering, Southeast University, Nanjing 210096, China

Hunan Post and Telecommunication College, Changsha 410015, China

e-mail: [wangying\\_only@163.com](mailto:wangying_only@163.com)

Y. Chen • L. Wu

School of Information Science and Engineering, Southeast University, Nanjing 210096, China

may not be optimum in sparse networks [3]. Therefore, this paper proposes a chain-type structure in WSN. By OPNET simulation, two light protocols in the MAC layer, based on time slot and competition, respectively, are analyzed. In this condition, “light” means simpler and more energy saving. Finally, we compared the proposed protocols with IEEE802.15.4, and the results show that the proposed protocols are superior to IEEE802.15.4 protocol in chain network with good reliability and low energy consumption.

### 137.2 The Light Protocol Based on Time Slot

As shown in Fig. 137.1, this is a chain network. All the nodes can produce and forward data frames, but each node could communicate with adjacent nodes, and the remote nodes could only communicate with the help of mid-nodes.

First, we propose the light protocol based on time slot for the above network, the “coordinator0” in the middle of the network is coordinator and the other nodes are ordinary nodes. The coordinator contains the information of time slot distribution in a beacon frame and regularly broadcasts it to the ordinary nodes in the network. By receiving the beacon frame and extracting the information, the ordinary nodes can know which timeslot belongs to them, so they can send the packets within the time slot. Therefore, in the light protocol based on time slot, conflicts can be avoided by unified beacon frame sent by the coordinator.

We design three kinds of frame formats in the light protocol based on time slot: beacon frame, data frame, and confirmation frame. Beacon frame is sent by the coordinator through broadcasting, which does not need to be confirmed by ACK. Confirmation frame is sent by nodes after receiving data frames. Confirmation frame itself does not need to be confirmed. Data frames are needed to be confirmed, and each data frame has a unique ID, which can be used to ensure the reliability.

The data flow and time slot allocation are shown in Fig. 137.2, the numbers in the figure represent that in which time slot will the nodes transmit data, and the arrows represent the direction of data transmission. Through this time slot allocation, conflict can be avoided, and the demand of real time can be guaranteed.

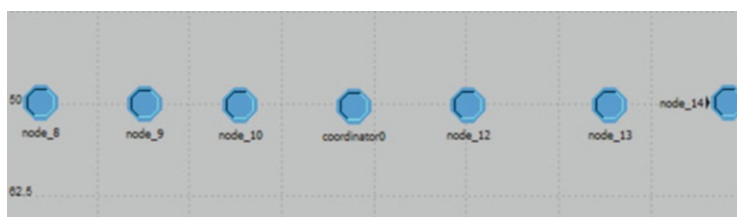


Fig. 137.1 Chain network

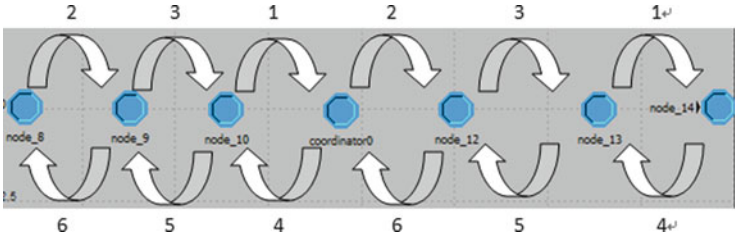


Fig. 137.2 Time slot allocation

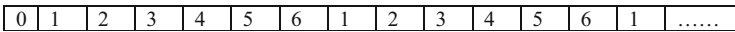


Fig. 137.3 Optimized allocation of time slot

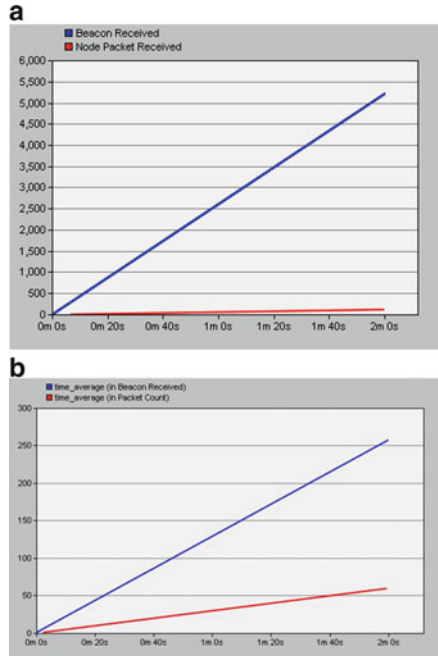
In this condition, a specific time slot no longer belongs to the packets produced by the node itself (also including the forwarding data produced by other nodes). Therefore, flexible allocation of time slot according to the amount of packets is not very suitable. So, in this protocol, the time slot belonging to each node is fixed. The coordinator sends a beacon frame, carrying only synchronization information.

In fact, in the chain network with a relatively stable environment, there is no such severe motion in the nodes that the WSN will lose synchronization in a very short period of time. Therefore, the protocol can be optimized: ordinary nodes will no longer receive a beacon frame each time after sending data frame; instead, the time slot can be allocated *n* times by one beacon frame (shown in Fig. 137.3). This can greatly reduce the number of beacon frames, which can lower the system energy consumption. According to simulation (shown in Fig. 137.4, the upper line in each figure represents the number of beacon frame a node received, while the lower one represents the number of data packets), when time slot allocates for 50 times in one superframe, the number of beacon frame can be reduced to only about 5 % of the original, so the energy cost on the beacon frame can save 95 %.

### 137.3 Light Protocol Based on Competition

In the light protocol based on time slot, whether the node has data to send or not, the coordinator always preserves certain time slots for it. When some nodes have large data amount, the others have small, it is difficult for the time slot protocol to coordinate. In order to guarantee the reliability of the protocol, the nodes need to intercept the channel continuously, which causes large energy consumption, especially in low-rate WSN. Therefore, this paper further proposes the light protocol based on competition to solve the above problems.

**Fig. 137.4** Results of optimized allocation of time slot. **(a)** Before optimization. **(b)** After optimization



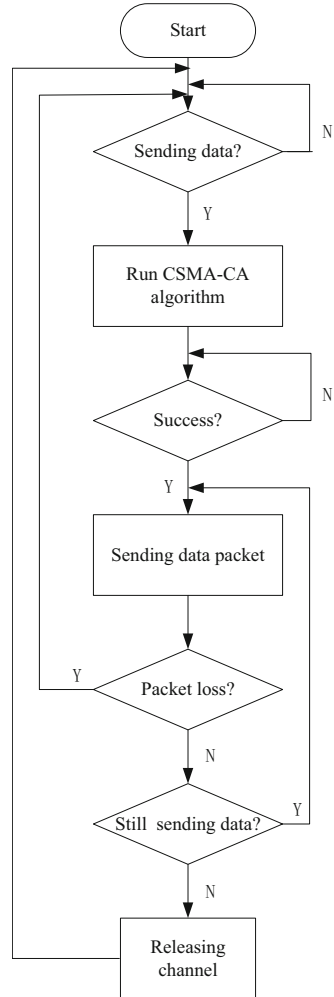
The light protocol based on competition relies on CSMA/CA algorithm [4]. In this chain-type network, synchronization is no longer need to be considered, so all the nodes are set as common nodes.

Figure 137.5 shows the flow chart of the light protocol based on competition. After initialization, if there are packets need to be sent, CSMA/CA algorithm will be performed. After successful access to the channel, the node occupies the channel alone to send the packets. Once the packet is lost, no matter whether the node has other data packets to send or not, the node is forced to give up the channel. Otherwise, if the packet has been successfully received and there are other data packets in the queue, the procedure repeats until the queue is empty. Then, the node releases the channel, and the other nodes can compete for the channel again.

In this flow chart, optimization has also been made to the traditional competitive protocol: a statement is added to judge whether there is a packet lost. Without this statement, the protocol can also work, but it is not good enough. Because when packet loses, the traditional protocol always removes the packet from the queue without trying more times. On the other hand, resending the packets may bring more dramatic conflict and cause more packet loss. Thereby, when a node hasn't received the confirmation frame after expectant time, it may suggest that this channel is not suitable for the node to transmit the packets right now. In this consideration, the node is forced to abdicate the channel. But, the lost packet is not deleted from the queue, which can be sent again after accessing the channel next time. Simulation shows that the optimization reduced the amount of lost packets significantly.



**Fig. 137.5** Flow chart of light protocol based on competition



### 137.4 Comparison Between Light Protocols and the Standard Protocol

In this section, the light protocols are compared with IEEE802.15.4 [5]. First of all, IEEE802.15.4 is realized in OPNET [6]. The chain-type topology is shown in Fig. 137.6; the coordinator sends a beacon frame first, which is received by the neighboring node “node\_9” and transmits to the next adjacent node; the procedure repeats until “node\_4” receives the beacon. The nodes that have already received the beacon frame will immediately step into the CAP (contention access period). In this period, if they have the GTS (guaranteed time slot) data, the nodes will send

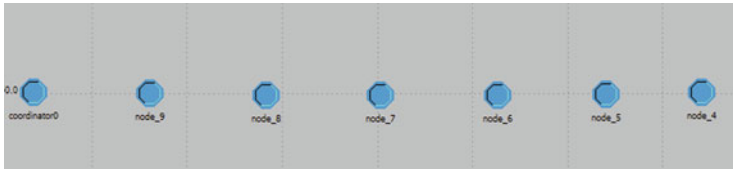


Fig. 137.6 Chain-type topology for IEEE802.15.4

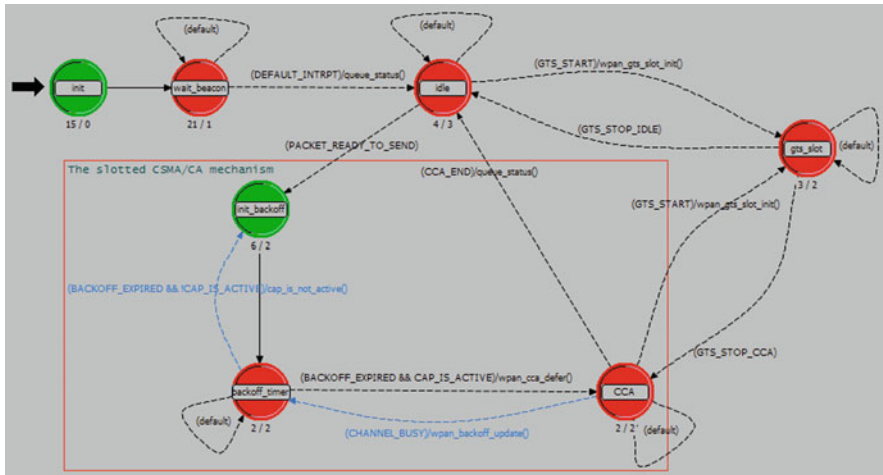


Fig. 137.7 State diagram of IEEE802.15.4 in OPNET

command frame to request the allocation of time slot. After receiving the command frame, the coordinator will contain the information of time slot allocation in a beacon frame and send it out. Then, the nodes can get their slot and send the packets during the time slot. If the GTS data needs to be forwarded, it will be forwarded as GTS data in the time slot of the transmitting node.

According to the analyses above, IEEE802.15.4 protocol is built in OPNET, and the state diagram is shown in Fig. 137.7.

Below is the comparison between IEEE802.15.4 protocol and the two kinds of light protocol in the chain network.

Simulation Settings (see Fig. 137.6): The packets are generated in “node\_9” and the generated time interval of CAP packets is 1 s, the data load is 160 bits, the destination address is “node\_4,” the simulation time is 20 min, the code rate is 250 kbps, and the modulation mode is BPSK. The collected statistics are the number of packets successfully received and the energy consumption in the network.

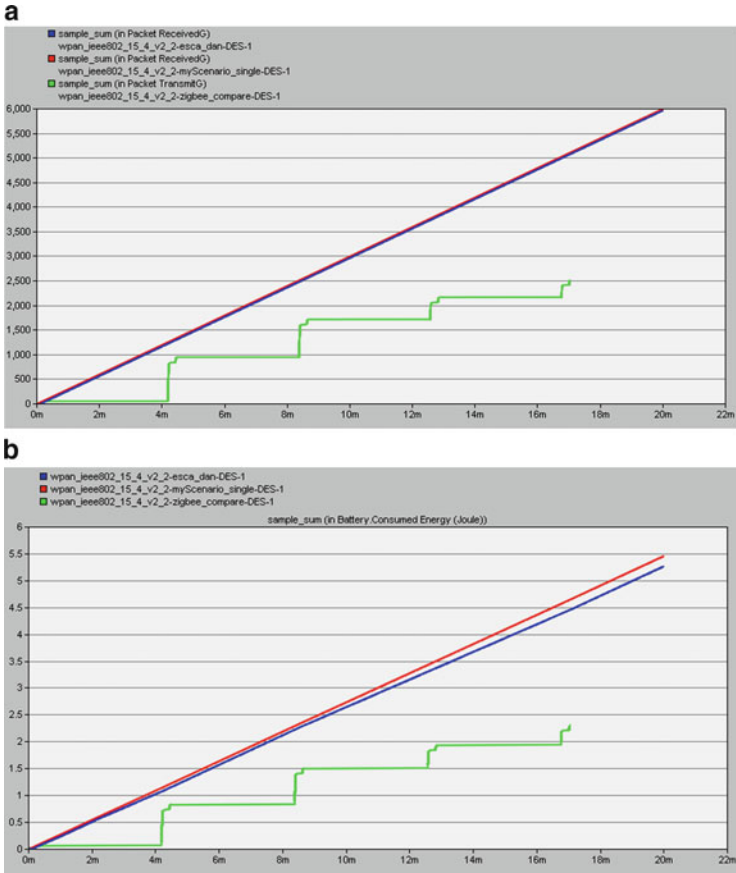


Fig. 137.8 Comparison results. (a) Packets received. (b) Energy consumption

The results are shown in Fig. 137.8; the curves of IEEE802.15.4, the light protocol based on competition, and the light protocol based on time slot are respectively represented from the bottom to the top in each figure.

By calculating, the energy consumption per packet by using the three protocols respectively is

IEEE802.15.4:  $2.35/2500 = 9.4e-4$  (J)

Light protocol based on time slot:  $5.5/6000 = 9.2e-4$  (J)

Light protocol based on competition:  $5.25/6000 = 8.75e-4$  (J)

From the calculation, the energy consumption of the light protocol based on competition is the lowest. Two light protocols have sent the same number of packets with better real-time performance. IEEE802.15.4 protocol costs the most energy, and the number of packages received successfully is less than half of that generated.

This is partly due to packet loss; more importantly, more packets are failed to be sent out in time. According to the results, the two light protocols are more reliable and more energy saving than IEEE802.15.4 protocol in chain network.

## 137.5 Conclusion

For chain-type WSN in the actual engineering, there is no mature protocol. This paper designs two kinds of light protocols: protocol based on time slot and protocol based on competition, which are superior to IEEE802.15.4 protocol both in reliability and energy efficiency. The research and simulation in this paper have built the theoretical foundation for the engineering application of the two light protocols.

## References

1. Cui, L., Hailing, J., Yong, M., Tianpu, L., Wei, L., & Ze, Z. H. A. O. (2005). Overview of wireless sensor network. *Journal of Computer and Development*, 42(1), 163–174.
2. Limin, S., Jianzhong, L., Chen, Y., & Hongsong, Z. (2012). *Wireless sensor network* (pp. 4–11). Beijing, China: Tsinghua University Press.
3. Weili, X., Mengna, T., & Baoguo, X. (2010). Research of MAC protocol based on OPNET in wireless sensor networks. *Chinese Journal of Sensors and Actuators*, 23(1), 139–143.
4. Yueping, W., Xiaojun, G., Chun-tao, L. I. U., & Jian-de, L. U. (2009). Low energy consumption research on MAC protocol of IEEE 802.15.4. *Computer Technology and Development*, 19(12), 139–142.
5. Yang, H., Yang, G., et al. (2005). Performance research of sensor network based on IEEE802.15.4. *Journal of Signal Processing*, 21(1), 444–447.
6. Chen, Y. (2013). *Research on light protocols using digital interphone as sparse nodes of wireless sensor network*. Nanjing, China: Master dissertation of Southeast University.

# Chapter 138

## Research and Implementation of a Peripheral Environment Simulation Tool with Domain-Specific Languages

Maodi Zhang, Zili Wang, Ping Xu, and Yi Li

**Abstract** The importance to build relevant peripheral environment in the testing process for complex embedded software is becoming higher. This paper discussed the current design method of simulation test environment for the embedded software and then presented a modelling method which is used to build peripheral simulation environment for the SUT (system under test) through ICD (interface control data) documents and the software requirement specification. Using this method, the peripheral environment simulation tool which consisted of relevant database and simulation model was set up with Ruby program language. This tool could provide necessary control commands and data support just like in a real running environment for the SUT. Furthermore, a DSL (domain-specific languages) design method for this domain was researched on the basis of the model. The experiment result has demonstrated that it's feasible to set up a peripheral environment for embedded system with our simulation tool.

### 138.1 Introduction

Most of complex embedded system consists of many subsystems, and there is a large amount of data interaction between each subsystem. This feature makes the software testing different from those in a simple embedded system. When we want to test the core software in the system, we must prepare a complete runtime environment. The cost to do this may be very huge, and in some cases, it's difficult or even impossible to establish a physical system for the testing. Therefore, using simulation method to establish the simulation environment is worth considering.

---

M. Zhang (✉) • P. Xu • Y. Li

Key Laboratory of Science and Technology on Reliability and Environment Engineering,  
School of Reliability and System Engineering, Beihang University, Beijing 100191, China  
e-mail: [canbi007@163.com](mailto:canbi007@163.com)

Z. Wang

Institute of Reliability Engineering, Beihang University, Beijing 100191, China

In our previous work, there is an embedded system emulator designed by our laboratory which can run operating system and has the function of fault injection. It also needs peripheral environment if we want to test the software running on the emulator. Some researchers have provided a real-time test script for automated test equipment [1]. They use the test script to drive the simulation test model and put forward a method to simulate the peripheral environment through ICD (interface control data) documents. But its focus is just on the design of the test script, not the establishment of the simulation models, and like most of the other simulation environment for embedded system, their application object isn't suitable to a simulation system [2]. So we also need to explore a method of simulation modelling.

The establishment of simulation model needs the analysis for the specific requirement. So the extensibility of the model is a question that worth considering. Domain-specific languages are designed for domain requirement, which abstract from domain entities and operations [3]. Establishing the simulation environment through domain-specific languages can make the model have better extensibility in this domain and promote its application values. Otherwise, the DSL could make it easy for the professionals in this domain to set up simulation models.

Ruby is an object-oriented dynamic programming language, with simple and beautiful syntax. Ruby has a powerful meta-programming capability. It can support dynamic method, introspection, code block, etc., which can make Ruby suitable to design a DSL as mother language [4, 5].

The rest of this paper is organized as follows. Section 138.2 gives a design method of database and simulation model. The implementation of the simulator consists of database; simulation models are described in Sect. 138.3. Section 138.4 shows a preliminary experiment, and Sect. 138.5 concludes this paper and our future work.

## 138.2 Design Requirement of the Peripheral Environment Simulation

The peripheral environment simulation is mainly based on the ICD and the software requirement specifications, so that this simulation environment does not support the data processing ability just like in the real environment. It just provides necessary incentive information and the response data according to the requirements of the real operation. So we can conclude that the most important two elements to build peripheral simulation environment are the following:

1. Database: provides the necessary control and data information
2. Simulation model: helps to realize the basic functions and running process

The rest of this paper will give the detail method for database and simulation model.

### ***138.2.1 Design Method for Data File***

The establishment of the database is mainly based on the ICD document. In the ICD document, it provides the whole data which is used in the real running process, with the specification of the data format, size, type, and so on.

Now, through the analysis of the ICD document, there are two methods to build the database.

First, if all of the data has the feature of clear rules, simple structure, and single type, we can adopt the method of mathematical simulation and get the required data through some programmed algorithm. This method can help us to simplify the process of data processing.

Second, when the data doesn't have obvious characteristics as in the first condition, we can store the data which is needed to the test in the form of data files. In this process, we need to design a storage format according to the actual needs. For example, the *message type*, the *subtype*, the *size*, the *corresponding model*, and the data subject are useful to identify the uniqueness of the data in the database. We can also combine with the first method to simplify the data structure.

The implementation of the data files may be diverse. We can store them in a database like Oracle, MySQL, SQLite, or store them as the CSV format, or even in an array or Hash.

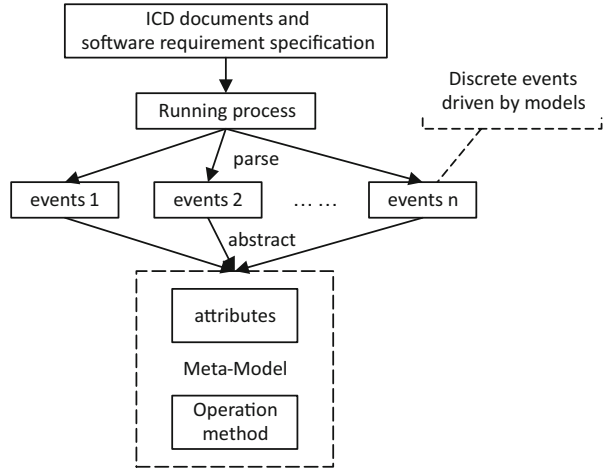
### ***138.2.2 Design Method for Domain Meta-Model***

Meta-model is a kind of model which has a higher abstraction degree and has the characteristics and commonalities of the general models. The relationship between meta-model and general model is similar to the relationship between class and object in the programming language.

The general theory of DSL shows that domain modelling is the first step to design a DSL [6]. In our condition, we can abstract the running process to be some small models which could present a specific task in the process. And these small models are the instances of a meta-model which has a higher abstract level. So the meta-model could be the basis of our DSL designed in the future.

To design the domain meta-model for peripheral environment, we parse the running process through the software requirements specification and the ICD documents and make the running process be some discrete events. All these discrete events were driven by their corresponding models, which could have their attributes and operation method. We can conclude the meta-model from the abstraction of these models. The detail process of domain modelling is shown in Fig. 138.1.

**Fig. 138.1** Domain modelling process for peripheral environment simulation



### 138.3 Implementation of the Simulation Model with Ruby

Figure 138.2 shows our framework of the simulator; now we have 3 main parts in the peripheral environment simulator, Ruby interpreter, simulation models, and data transceiver. The simulator obtains the data from the database and, driven by the test script, finally sends the data to the SUT through pipe file.

#### 138.3.1 Model Structure

In our simulation meta-model, we have designed four attributes: *name*, *state*, *type*, and *subtype*. The details for each attribute are showed in Table 138.1. These four attributes will be initialized in the process of the model instantiation.

Then we designed some basic method for the meta-model according to the requirement of the simulation:

1. Data transmission method: open or create a pipe file, send or receive data through the pipe file.
2. Data match method: obtain the data which is matched to the model attribute according to the initialized model.
3. Data modification method: according to different data type and structure, using different algorithms to modify the original data.



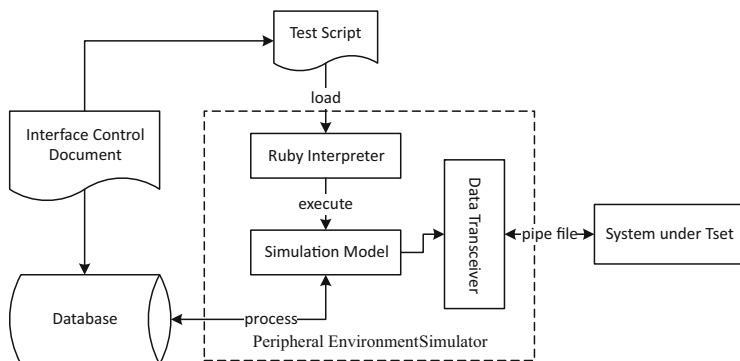


Fig. 138.2 Main framework of the peripheral environment simulator

Table 138.1 The main attributes in the meta-model

Attribute name	Description	Note
Name	Define the name of this model	Specific by keywords
State	Define the state of this model	Only to be “normal” or “fault”
Type	Define the type of data in this model	
Subtype	Branch of the data type	This is an optional attribute

### 138.3.2 Data Structure and Processing

In this section, we conclude the standard data format from the ICD document. Most of the data we need to process are the BIT information and some other control command (Fig. 138.3).

There are no commonalities between these data in addition to follow the above data structure. The data length may range from 3 to 700 bytes. Especially in the BIT answer data, it contains a lot of subtypes (Table 138.2).

From the above analysis, it’s impossible to design an algorithm to generate data automatically or store all kinds of data in detail. So, we have only stored the necessary control information and the normal data of each type of the BIT information. And these data are stored together with their data ID and data type, which may be important parameters in the process of data matching and processing.

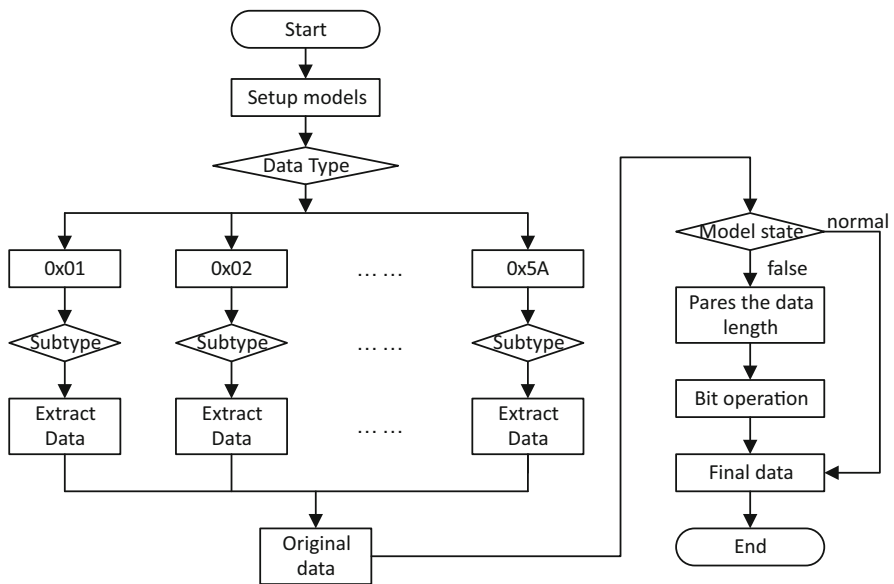
The process of the data matching and processing is mainly based on the data ID, subtypes, and their length. Figure 138.4 shows a flow diagram of the data processing.

**Fig. 138.3** Format of the data information

Data ID	Data Length	Data Details
---------	-------------	--------------

**Table 138.2** Data information

Data ID	Data label	Description	Size (bytes)
0x01	BitAnsProc	BIT response	8–14
0x02	InfoReq	Request for information	3
0x1D	CanRtr	CAN node status	5
0x41	SysSw	Master–slave switch command	3
0x42	BitReq	BIT request	3
0x43	MSListAll	Module state synchronization	724
0x49	FunListAll	Function state synchronization	154
0x5A	MSListAck	Ack for module state update	3



**Fig. 138.4** Flow diagram of data processing

From the data structure, we can know that the low 16 bits of every data present the data ID and data length. So we can make bit operation to the high data bits of the original data, for example:

1. Control command of 3 bytes:  $final\_data = original\_data \mid 0x010000$
2. BIT information of 8 bytes:  $final\_data = original\_data \mid 0xFFFFFFFF00000000$

Then, the final data would be sent to the SUT through the data transceiver method.

## 138.4 Experiments and Analysis

In our current experimental situation, we make a simple data receiver for the test. This data receiver will create a new pipe file at the run time and will always be monitoring this pipe file.

In this experiment, we set up two models through the meta-model: (“SysSw,” “normal,” “0x41”) and (“BITAnsProc,” “fault,” “0x01,” “ZHXD1\_M\_JK1”). The first model presents a control model, which would send a normal excitation signal for the software under test. The second model presents BIT response model; it could send a fault BIT information of the subtype named “ZHXD1\_M\_JK1” to the SUT. The result of this experiment shows in Fig. 138.5.

From the running result, we can see that the “SysSw” model sends a normal control command “0x000341” to the data receiver. And the “BITAnsProc” model sends a fault BIT information “0xffffffffffffffff0000320e01” of a subtype named “ZHXD1\_M\_JK1.” The data receiver has received the control command and the BIT information.

We can conclude that our simulation tool has the following advantages compared with other simulation tools:

1. This is a full virtualized peripheral environment simulation tool; it could be an effective test method for this equipment which is difficult and high cost to be tested.
2. This simulation tool is designed for another virtual platform which can run an OS and their software; we haven’t found a similar tool that can realize peripheral environment simulation functions for virtual platforms.
3. Our design is supported by DSL design ideal, which makes the communication between the programmer and domain experts, and this test script could be very easy to understand and write for the domain experts.
4. We have proposed the modelling and simulation method for peripheral environment, and our target system is mainly avionics systems. The meta-model is designed through the standard ICD documents in the avionics systems. We can

```

amao@ubuntu:~/workplace/extern-env-emu$ ruby machinemodel_v1.rb
The normal BIT data for SysSw model: 000341
The final data for this normal model: 000341
Sending data via the pipe file...
The normal BIT data for BITAnsProc model: 000000000000000000000000320E01
The final data for this fault model: ffffffffffffffff0000320e01
Sending data via the pipe file...
amao@ubuntu:~/workplace/extern-env-emu$ 

```

```

amao@ubuntu:~/workplace/extern-env-emu$ ./master.rb
000341
ffffffffffffffff0000320e01

```

Fig. 138.5 The experiment result

change the meta-models or the model attributes easily through this method and make it suitable for other models and systems. So this tool has a good generality for avionics systems or other systems.

## 138.5 Conclusion and Further Work

In this paper, we summarized the current methods of running environment simulation and then we proposed a domain modelling method for peripheral environment simulation through ICD document and software requirement specification. The modelling and simulation methods have a good generality for the avionics systems and other systems. After all, we setup a small simulator with a data source and some models through this method. The feasibility was confirmed in a small experiment.

In the future, we still have to optimize the model structure, and add more functions according to the actual requirement, make our simulator have the ability to send and receive data at the same time, and in the last communicate with the QEMU simulator.

**Acknowledgements** This research was supported by the Technological Foundation Project of China National Defence Science and Engineering Bureau under grant No. Z132012A004.

## References

1. Jiang, C., Liu, B., Yin, Y., & Liu, C. (2009) Study on real-time test script in automated test equipment. *In 8th International Conference on Reliability, Maintainability and Safety* (pp. 738–742). Chengdu, China: IEEE.
2. Zainzinger, H.J., & Austria, S.A. (2002). Testing embedded systems by using a C++ script interpreter. *Proceedings of the 11th Asian Test Symposium* (pp. 380–385). Guam, USA: IEEE.
3. Gunther, S., Haupt, M., & Splieth, M. (2010). Agile engineering of internal domain-specific languages with dynamic programming languages. *In 5th International Conference on Software Engineering Advances* (pp. 162–168). Nice, France: IEEE.
4. Gunther, S. (2010). Multi-DSL application with Ruby. *IEEE Software*, 27(5), 25–30.
5. Cuadrado, J., & Molina, J. (2007). Build domain-specific languages for model-driven development. *IEEE Software*, 24(5), 48–55.
6. Dinkelaker, T., & Mezini, M. (2008). Dynamically linked domain-specific extensions for advice languages. *ACM Proceedings of the 2008 AOSD Workshop on Domain-Specific Aspect Languages*, New York (pp. 1–7).

# Chapter 139

## Probability Model for Information Dissemination on Complex Networks

Juan Li and Xueguang Zhou

**Abstract** In order to analyze the regulation of information dissemination on the complex network, SIR probability model has been built to represent the peoples' interaction during information dissemination on complex networks. By introducing and computing the state transiting probabilities of the net nodes, we can effectively analyze and update the nodes' states at each step in information dissemination. Accordingly, the evolution algorithm of information dissemination is designed and realized by simulation. Simulation experiments of information dissemination on ER network and BA network with different parameters reveal that the density of final awareness will not be affected by the total of nodes, but increase progressively following the increase of average degree until a certain value. Different degree distributions can also be effect on the density of final awareness. SIR probability model can accurately reflect the process of information dissemination on complex networks. It can be used for the description and analysis of information dissemination on complex networks.

### 139.1 Introduction

How many persons will know the message during information dissemination in the realistic society or Internet space? Which factors will influence or determine the results of information dissemination? Whether the different network structure will affect information dissemination? How will these factors affect information dissemination? These are all that interest us for research on the information dissemination on complex networks.

---

J. Li (✉) • X. Zhou  
Naval University of Engineering, WuHan 430033, China  
e-mail: [lijuan770107@126.com](mailto:lijuan770107@126.com)

The information dissemination is a typical evolution process of complex system [1]. To understand the information dissemination in the complex network well, first we need a proper model to describe the people's interaction during information dissemination on complex networks. Basing on SIR (susceptible, infective, recovered) model, this paper presents SIR probability model by introducing the state transiting probability of the net node. First, we can compute the state transiting probabilities of all nodes in the complex network. Then we will analyze and update the states in the next moment accordingly. Thus, we can describe the process of information dissemination accurately.

Furthermore, to discover the regulation of information dissemination on the complex networks, experiments of information dissemination on different networks with different parameters are needed. In this paper, many emulation experiments have been done for the information dissemination on ER network and BA network with different parameters. Many parameters of the network can be effect on the result of information dissemination. We have analyzed the total of nodes, average degree, and topology of complex networks affecting the density of final awareness.

## 139.2 SIR Probability Model

### 139.2.1 The Basic SIR Model

SIR model [2] can be used for information dissemination. Here, "S" is the state for who have not known the information yet. "I" is the state for who have known the information and would transmit it. "R" is the state for who have known the information but will not transmit it. Many persons including Sudbury, Zanette, and Zonghua Liu had researched on this model. The dissemination rules they made can be expressed with the following formula:

$$\begin{cases} I(i) + S(j) \rightarrow I(i) + I(j) \\ I(i) + I(j) \rightarrow R(i) + I(j) \\ I(i) + R(j) \rightarrow R(i) + R(j) \end{cases} \quad (139.1)$$

Disseminator staying in "I" state would select a neighbor randomly to spread the information. If his selection is in "S" state, this neighbor will obtain the information and become a new disseminator. His state changes from "S" to "I". Otherwise, when the selection is in "I" or "R" state, the disseminator will lose the interest to spread the news, he will change his state to "R," and the neighbor he selected has no change.

### 139.2.2 State Changing Probability in SIR Model

The value  $k_v$  is defined for the degree of the node  $v$  in the spreading network. It is the scalar of his neighbors.  $k_{vS}$ ,  $k_{vI}$ , and  $k_{vR}$  are defined respectively for the scalar of his “S,” “I,” and “R” state neighbors. By all appearances,

$$k_{vS} + k_{vI} + k_{vR} = k_v \quad (139.2)$$

If  $v$  is a node in “I” state, we define  $p_{vI}$  as the probability for node  $v$  to keep “I” state and  $p_{vR}$  as the probability for node  $v$  changing to “R” state next time. So, according to dissemination rules of SIR model, we can calculate  $p_{vI}$  and  $p_{vR}$ :

$$p_{vI} = \frac{k_{vS}}{k_v} \quad (139.3)$$

$$p_{vR} = 1 - p_{vI} = \frac{k_v - k_{vS}}{k_v} \quad (139.4)$$

If  $u$  is a node in “S” state, we define  $p_{uS}$  as the probability for node  $u$  to keep “S” state and  $p_{uI}$  as the probability for node  $u$  changing to “I” state next time. We can calculate  $p_{uS}$  and  $p_{uI}$  with the following expressions:

$$p_{uS} = \prod_{v=I} p_{vR} = \prod_{v=I} \frac{k_v - k_{vS}}{k_v} \quad (139.5)$$

$$p_{uI} = 1 - p_{uS} = 1 - \prod_{v=I} \frac{k_v - k_{vS}}{k_v} \quad (139.6)$$

In (139.5) and (139.6),  $v$  is the neighbor with “I” state of the node  $u$ .

In the information dissemination network,  $N$  is defined for the total of nodes as well as the numbers of crowds.  $C_t$  is defined for the numbers of crowds who known the information at  $t$  moment. Then, we can estimate it with the following expression:

$$C_t = N - \sum_{u=S} p_{uS} \quad (139.7)$$

Here,  $u$  is the node with “S” state in  $t-1$  moment.

Unitarily,  $c_t$  is defined for the density of the known:

$$c_t = \frac{C_t}{N} \quad (139.8)$$

### 139.2.3 Evolving of SIR Probability Model

In the SIR model, three states are existing: “S,” “I,” and “R.” In the beginning, the number of nodes with “I” state is  $C_0$ . Others are in “S” state. None is in “R” state. Node  $u$  is in “S” state. If there are some neighbors in “I” state, node  $u$  will change its state to “I” according to  $puI$ . The evolving algorithm is the following:

1. Initializing the states and information of all nodes in the complex network. There into,  $C_0$  is the initialization of dissemination nodes. The initialization of time  $t$  is 0. The terminate time is set with  $T$ .
2. Calculating probability value  $pvI$  and  $pvR$  for each node  $v$  with “I” state and calculating probability value  $puS$  and  $puI$  for each node  $u$  with “S” states.
3. Changing the next time state of the node  $v$  with “I” state to “R” according to  $pvR$ , as well as changing the next time state of the node  $u$  with “S” state to “I” according to  $puI$ .
4.  $t = t + 1$ . If  $t = T$  or there has no node in “I” state, it is end. Otherwise, go to the step (2).

## 139.3 Emulation of SIR Probability Model in Complex Network

SIR probability model and its evolvement algorithm are independent of the structure of network. However, information spreading in the network is nearly correlative to the network topology [3]. No matter whether information spreads in social network or internet, all these networks are typical complex networks. They have complicated dynamics action. To study the effect of complex network structure on information dissemination, in this paper, the two mature complex network models, ER (Erdős–Rényi) stochastic network and BA (Barabási–Albert) scale-free network, are selected for emulating and analyzing.

### 139.3.1 Creating of Complex Network

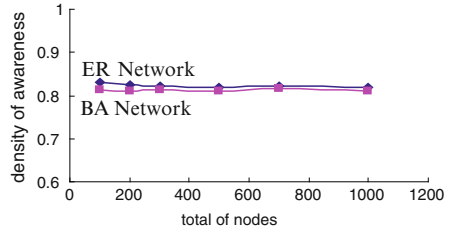
#### 1. ER Model

The following is the algorithm for creating ER network:

- (a). Creating a network only including  $N$  nodes.  $N$  is the total of nodes.
- (b). For each pair of nodes, the edge is created following the probability value  $p$ .  
 $p = \langle k \rangle / (N - 1)$ ,  $\langle k \rangle$  is the average degree of the nodes in the network.



**Fig. 139.1** The relation between total of nodes and density of awareness



## 2. BA Model

Network like Internet, WWW, and so on, their degree distributions obey power law. They have the characteristic of scale-free. BA network is a typical scale-free network [4–6]. We can create BA network with the method of dynamic growth as the following:

1. Creating the initial network including  $m_0$  nodes.
2. Adding new node  $v$ .  $m$  edges from the new node link to the existing nodes according to preference rule. The node with great degree has more probability to get new edge, i.e., 3. repeating 2. until the total of nodes is enough.

$N$  is setting for total of nodes in network.  $\langle k \rangle$  is setting for the average degree. Then, we can set the initial network as a complete network with  $m_0$  ( $m_0 = \langle k \rangle / 2$ ) nodes.  $\langle k \rangle / 2$  ( $m = \langle k \rangle / 2$ ) edges will be created for every new node. Accordingly, we can create BA network with  $N$  nodes.

### 139.3.2 Experiment and Analysis

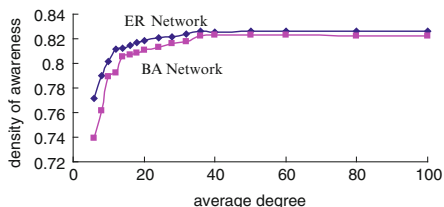
Setting the initial value  $C_0$  for density of known crowd, we can process the evolution of SIR probability model in ER network and BA network. We can review the effects of various facts on information dissemination.

#### 1. Total of Nodes Affecting Information Dissemination

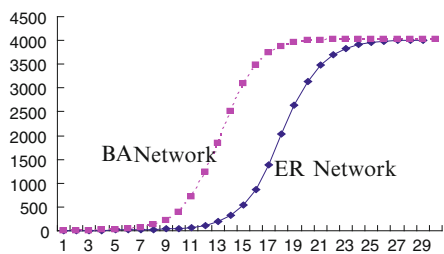
Assuming only one node with “I” state in the first, emulation experiments have been processed for  $\langle k \rangle = 20, N = 100, 200, 300, 500, 700, 1,000$ . Through these experiments, we can examine the effect degree of total of nodes on information dissemination. To assure the correctness of the result, each experiment has been done 100 times independently. The average of the results is shown in Fig. 139.1.

From Fig. 139.1, we can know that the total of nodes affects little on the density of awareness at end.

**Fig. 139.2** The relation between average degree and density of awareness



**Fig. 139.3** Increasing peoples with “R” state



## 2. Average Degree Affecting Information Dissemination

Assuming only one node with “I” state in the first, emulation experiments have been processed for  $N = 500$ ,  $\langle k \rangle = 12, 16, 20, 24, 28, 32, 36, 40, 50, 60$ . Through these experiments, we can examine the effect degree of average degree on information dissemination. To assure the correctness of the result, each experiment has been done 100 times independently. The average of the results is shown in Fig. 139.2.

From Fig. 139.2, we can know that density of awareness at end will increase to a certain degree following augment of average degree. While  $\langle k \rangle$  reaches about 36, the density of awareness will keep tranquilization.

## 3. Network Structure Affecting Information Dissemination

The structures of ER network and BA network are different. From Figs. 139.1 and 139.2, we can realize that the density of awareness end in BA network is always less than that in ER network generally. To analyze the network structure affecting information dissemination more, emulation experiments have been processed for  $N = 5,000$ ,  $\langle k \rangle = 10$ ,  $C0 = 0.002$ . During the processes of information dissemination in these two types of network, the totals of people in “R” state (who know the information but will not transmit it) are changing as shown in Fig. 139.3.

From Fig. 139.3, we can find that increasing speed of peoples with “R” state is rapid during the processes of information dissemination. Because there are more nodes with great degree in scale-free network than in stochastic network, the information is easy to be send to the nodes with great degree. When these nodes

are in “I” or “R” state, the others will connect them with big probability and then change themselves in “R” state. So, information dissemination ends more quickly in BA network than in ER network. It leads that density of awareness at end in BA network is less than in ER network. Figs. 139.1, 139.2, and 139.3 all show the characteristic.

## 139.4 Conclusion

The paper uses probability method to develop the SIR model. We apply this probability SIR model in the ER network and BA network to study the information dissemination in complex networks. The results of emulation are same as the theoretic existed. It indicates the validity of the model we created.

## References

1. ShiZhe, G., Zheming, L. (2012). *Complex network basic theory*[M]. Beijing: Science Press. 6, 143–146.
2. Guanrong Chen, Xiaofan Wang, & Xiang Li. (2012). *Introduction to complex networks: Models, structures and dynamics*[M]. Beijing: Higher Education Press. 5, 187–198.
3. Jie, Z., Zonghua, L., & Baowen, L. (2007). Influence of network structure on rumor propagation [J]. *Physical Letters A*, 368, 458–463.
4. Yi, C. X. (2012). Weighted BA scale-free random graph model [J]. *Advanced Materials Research*, 1566, 2780–2783.
5. Kaiying, D., Jingwei, D., & Yingxing, L. (2012). Research and application of improved BA networks evolving models[J]. *Computer Systems & Applications*., 8, 116–119.
6. Barabási, A.-L. (2003). Linked: The new science of networks[J]. *American Journal of Physics*, 71, 409–410.

# Chapter 140

## Verification of UML Sequence Diagrams in Coq

Liang Dou, Lunjin Lu, Ying Zuo, and Zongyuan Yang

**Abstract** The UML is a semiformal modeling language which only has syntax and static semantics precisely defined. The dynamic semantics for the UML is specified neither formally nor algorithmically. When using UML at the design phase, there does not exist a systematic way that allows the model designer to specify its formal semantics and automatically verify correctness properties of the described model. The UML sequence diagrams are widely used to describe the behaviors of software. Reasoning about properties of sequence diagrams at the analysis and design phase may reveal software faults before software implementation. We propose to use the theorem proof assistant—Coq to verify syntax and semantics constrains of sequence diagrams. The verification and proof process are useful for improving the correctness of sequence diagrams and hence increases the software quality.

### 140.1 Introduction

In practical software engineering, the UML (Unified Modeling Language) has become the de facto modeling language. The UML sequence diagrams [1] provide a graphical notation to describe dynamic aspects of software during the design phase and have been proven very useful for modeling system behavior in model-based development.

However, the UML is a semiformal modeling language which only has syntax and static semantics precisely defined. There should not be a systematic way that allows the model designer to specify formal semantics for UML so that their

---

L. Dou (✉) • Y. Zuo • Z. Yang  
Department of Computer Science and Technology, East China Normal University,  
Shanghai 200241, China  
e-mail: [ldou@cs.ecnu.edu.cn](mailto:ldou@cs.ecnu.edu.cn)

L. Lu  
Department of Computer Science and Engineering, Oakland University, Rochester,  
MI 48309, USA

behavior can be precisely and unambiguously understood, and safety-critical properties of system specifications can be verified automatically. As a result, it is highly desirable to use formal methods to analyze UML models in the design phase.

Coq [2] is an interactive theorem proof assistant based on the Calculus of Inductive Constructions and widely used in program verification and theorem proving. In this chapter, we explore the possibility of using Coq to formalize the semantics of sequence diagram and verify the desired properties. The verification and proof process will provide increased reliability guarantee in the design phase, thereby improving software quality and reducing development cost. In our solution, the syntax of a sequence diagram is represented as inductive types and its denotational semantic is represented as a recursive function in Coq. Based on trace semantics, we could verify the syntax and semantics constraint on models. The desired properties of the semantics can be stated as lemmas and the proof ensures that the semantics satisfy their corresponding properties.

The chapter is organized as follows. In Sect. 140.2 the trace semantics of sequence diagrams is introduced formally, together with a brief introduction to Coq. In Sect. 140.3 we describe the representation for sequence diagrams in Coq. We then describe how to verify the syntax and semantic restrictions and prove desired properties in Coq in Sect. 140.4. Section 140.5 uses a case study to illustrate our approach. Section 140.6 discusses the related work and we conclude in Sect. 140.7. All the proofs and samples can be found at <https://github.com/lisa-dou/VerifySD>.

## 140.2 Background

### 140.2.1 Abstract Syntax

Based on a previous work [4], we give the abstract syntax of sequence diagrams. *Name* is a denumerable set of names. An event  $e$  in *Evt* has the following structure. An event sending a message with signal  $S \in Name$ , transmitter  $T \in Name$ , and receiver  $R \in Name$  is written as  $(!,S,T,R)$ , and its corresponding receiving event is written as  $(?,S,T,R)$ . We abstract from details of guard conditions  $c$  in *Cnd* and require that the collection of guard conditions is closed under classical logical negation ( $\neg$ ), conjunction ( $\wedge$ ), and disjunction ( $\vee$ ) operations. Let  $\tau$  represent unobservable events. The abstract syntax for sequence diagrams in  $D$  is given below:

$$D := \tau | e | \text{strict}(D_1, D_2) | \text{alt}(c, D_1, D_2) | \text{opt}(c, D) | \text{par}(D_1, D_2) | \text{seq}(D_1, D_2) | \text{loop}(c, D)$$

### 140.2.2 Semantics

Let  $\Sigma$  be an alphabet.  $\Sigma^*$  denotes the set of all strings over  $\Sigma$ . The interleave of two strings is the set of strings obtained by interleaving the two strings in all possible ways. Let  $x, y \in \Sigma$  and  $u, v \in \Sigma^*$ . We define of the interleave operator  $\parallel$  as in [5]:

$$\begin{aligned} \varepsilon \parallel \mu &= \mu \parallel \varepsilon = \mu \\ x\mu \parallel yv &= \{x\}(\mu \parallel yv) \cup \{y\}(x\mu \parallel v) \end{aligned}$$

where  $\cdot$  is the language concatenation operator and  $\varepsilon$  represents the empty string. Let  $\oplus$  be a binary operation on domain  $S$ . Then  $\oplus^\#$  defined below is a binary operation on the power set  $\wp(S)$  of  $S$ :

$$X \oplus^\# Y = \{x \oplus y \mid x \in X \wedge y \in Y\}$$

*Semantic Domain.* A sequence diagram is a partial specification of required and prohibited behaviors of an application. Our work is concerned only with required behaviors. The semantics of sequence diagram is trace based. A trace is a sequence of tokens, each of which is either an event or a guard condition. The domains of tokens and traces are respectively:

$$\begin{aligned} Tk &= Evt \cup Cnd \\ Tr &= Tk^* \end{aligned}$$

An obligation may contain more than one trace. Once an obligation is chosen, all traces in the obligation are required in that for each trace  $t$  in the obligation, there is an interaction that produces  $t$ . The domain of the obligation  $Ob$  and the semantic model  $Mo$  are respectively:

$$\begin{aligned} Ob &= \{o \in \wp(Tr)\} \\ Mo &= \{m \in \wp(Ob)\} \end{aligned}$$

*Semantic Function.* The semantics of a sequence diagram  $D$  is denoted as  $\llbracket D \rrbracket$  follows:

1. Unobservable and observable events:

$$\begin{aligned} \llbracket \tau \rrbracket &= \{\{\varepsilon\}\} \\ \llbracket e \rrbracket &= \{\{e\}\} \end{aligned}$$

2. *Strict* fragments:

$$\llbracket \text{strict}(D_1, D_2) \rrbracket = \llbracket D_1 \rrbracket \cdot^\# \llbracket D_2 \rrbracket$$

3. *Alt* fragments: *Alt*, *opt*, and *loop* fragments introduce guards to traces. Define  $c \triangleright \varepsilon = \varepsilon$ ,  $c \triangleright \sigma = \sigma$  where  $c \in \text{Cnd}$  and  $\sigma \in \text{Tr}$  such that  $\sigma \neq \varepsilon$ . Let  $c \triangleright \#M = \{\{c \triangleright \sigma \mid \sigma \in O\} \mid O \in M\}$ . We have:

$$\llbracket \text{alt}(c, D_1, D_2) \rrbracket = (c \triangleright^\# \llbracket D_1 \rrbracket) \cup (\neg c \triangleright^\# \llbracket D_2 \rrbracket)$$

4. *Opt* fragments: The semantics of *opt* fragments is obtained similarly:

$$\llbracket \text{opt}(c, D) \rrbracket = (c \triangleright^\# \llbracket D \rrbracket) \cup \{\{\varepsilon\}\}$$

5. *Par* fragments: Parallel interleaving produces a set of alternative obligations from  $O_1$  and  $O_2$ . Define

$$O_1 \uparrow O_2 = \{O \mid \forall \sigma_1 \in O_1, \forall \sigma_2 \in O_2 : \exists \sigma \in O. (\sigma \in \sigma_1 \parallel \sigma_2)\}.$$

The semantics of  $\text{par}(D_1, D_2)$  is defined:

$$\llbracket \text{par}(D_1, D_2) \rrbracket = \cup_{O_1 \in M_1, O_2 \in M_2} O_1 \uparrow O_2$$

6. *Seq* fragments: The interaction operator *seq* combines traces from component sequence diagrams via weak sequencing. By  $t_1 \hat{\ } t_2$  we denote the trace consisting of  $t_1$  immediately followed by  $t_2$ . By  $A\theta t$  we denote the trace obtained from the trace  $t$  by removing all elements that are not in the set of elements  $A$ . Let  $e \cdot l$  denote the set of events that may take place on the lifeline  $l$ ; weak sequencing of trace sets is defined as follows [6]:

$$s_1 \succ s_2 = \{h \in H \mid \exists h_1 \in s_1, \exists h_2 \in s_2 : \forall l \in L, e \cdot l \theta h = e \cdot l \theta h_1 \hat{\ } e \cdot l \theta h_2\}$$

The semantics of *seq* fragments is defined:

$$\llbracket \text{seq}(D_1, D_2) \rrbracket = \{t_1 \succ t_2 \mid t_1 \in \llbracket D_1 \rrbracket \wedge t_2 \in \llbracket D_2 \rrbracket\}$$

7. *Loop* fragments: We use a natural number  $n$  to indicate the counter in the *loop* fragment. The semantics of *loop* fragments is defined:

$$\begin{aligned} \llbracket \text{loop}(0, c, D) \rrbracket &= \{\{\varepsilon\}\} \\ \llbracket \text{loop}(n+1, c, D) \rrbracket &= c \triangleright^\# \llbracket \text{seq}(D, \llbracket \text{loop}(n, c, D) \rrbracket) \rrbracket \end{aligned}$$

*Example* Let  $m, n$  be signals,  $l_1, l_2$  be lifelines,  $c$  be a condition,  $f_1 = (!, m, l_1, l_2)$ ,  $f_2 = (? , m, l_1, l_2)$ ,  $f_3 = (!, n, l_1, l_2)$ , and  $f_4 = (? , n, l_1, l_2)$ . Put  $D_1 = \text{strict}(f_1, f_2)$  and  $D_2 = \text{strict}(f_3, f_4)$ . Then we have:

$$\begin{aligned} \llbracket D_1 \rrbracket &= \{\{f_1 f_2\}\} \\ \llbracket D_2 \rrbracket &= \{\{f_3 f_4\}\} \\ \llbracket \text{strict}(D_1, D_2) \rrbracket &= \{\{f_1 f_2 f_3 f_4\}\} \\ \llbracket \text{alt}(c, D_1, D_2) \rrbracket &= \{\{c f_1 f_2\}, \{\neg c f_3 f_4\}\} \\ \llbracket \text{opt}(c, D_1) \rrbracket &= \{\{c f_1 f_2\}, \{\varepsilon\}\} \\ \llbracket \text{par}(D_1, D_2) \rrbracket &= \{\{f_1 f_2 f_3 f_4\}, \{f_1 f_3 f_2 f_4\}, \{f_1 f_3 f_4 f_2\}, \{f_3 f_4 f_1 f_2\}, \{f_3 f_1 f_4 f_2\}, \{f_3 f_1 f_2 f_4\}\} \\ \llbracket \text{seq}(D_1, D_2) \rrbracket &= \{\{f_1 f_2 f_3 f_4\}, \{f_1 f_3 f_2 f_4\}\} \\ \llbracket \text{loop}(2, c, D_1) \rrbracket &= \{\{c c f_1 f_2 f_1 f_2\}, \{c c f_1 f_1 f_2 f_2\}, \{c f_1 c f_2 f_1 f_2\}, \{c f_1 c f_1 f_2 f_2\}, \{c f_1 f_2 c f_1 f_2\}\} \end{aligned}$$

### 140.3 Theorem Proof Assistant: Coq

Coq is an interactive theorem proof assistant [3]. All data is represented with inductive data types in Coq. The Coq library contains elementary data types, including natural numbers, strings, and lists, among others. They are “predefined” types and equipped with many useful functions. As an example, Coq defines natural numbers and list of (parametric) type  $A$  using the following inductive definitions:

```
Inductive nat : = | O : nat | S : nat -> nat .
Inductive list A : =
| nil : list A
| cons : A -> list A -> list A .
```

Here,  $O$  and  $S$  are the *constructors* of type  $\text{nat}$ .  $\text{nil}$  and  $\text{cons}$  are the *constructors* of type  $\text{list } A$  for any type  $A$ . Type  $\text{nat}$  and  $\text{list}$  are inductive types, which means their elements are obtained as finite combinations of the constructors. For example, there is a number  $O$  (zero), and for every number  $n$ , there is another number  $S\ n$  (the successor of  $n$ ). So we could write number 3 as  $S(S(S\ O))$ . In Coq, the empty list is denoted by  $\text{nil}$  while  $a::b$  is a notation for  $\text{cons } a\ b$ , so the list 1, 2, 3 can be written as  $1::2::3::\text{nil}$ . Inductive types are at the core of powerful programming and reasoning techniques.

The keyword *Definition* is used to define new types, for example:

```
Definition pair string : Set : = string * string
```



The *pair string* is the type of pairs of strings. The *\** is the pair type constructor.

We can use the keyword *Fixpoint* to define well-founded recursion for inductive types. For example:

```
Fixpoint fact (n: nat) {struct n} := match n with
| 0 => 1
| S p => n * fact p
end.
```

The *fact* is a simple function that calculates a number's factorial using recursion. The *struct n* annotation states that it is structurally recursive on its *n* parameter, and therefore guaranteed to terminate. The *match with* construct represents pattern-matching on the shape of the inductive type and handle all possible cases.

The keyword *Lemma* can be used to signify proofs. Coq enters interactive proving mode to assist us in building a proof. When proving properties of the inductive type, we could do an induction on the structure of the type.

## 140.4 Formalize the Trace Semantics in Coq

In this section, we show how to formalize the trace semantics of sequence diagrams in Coq. Due to space limitations, only important definitions and functions are shown.

The first step is to map sequence diagrams to the formal semantic models. We begin by defining the semantic model in Coq as follows:

```
Inductive kd := | Send : kd | Receive : kd. (* Define the kind
of event *)
Definition sg := string. (*Define the signal *)
Definition lf := string. (*Define the lifeline *)
Definition evt := kd * sg * lf * lf. (*Define the event *)
(*Define the syntax of condition *)
Inductive id := Id : nat -> id.
Inductive cnd := | Bvar : id -> cnd | Btrue : cnd | Bfalse : cnd
| Bnot : cnd -> cnd
| Band : cnd -> cnd -> cnd | Bor : cnd -> cnd -> cnd | Bimp : cnd
-> cnd -> cnd.
Inductive tk := | ev : evt -> tk | cd : cnd -> tk. (* Inductive
type for token *)
Definition tr := list tk. (* A trace is a list of tokens *)
Definition ob := set tr. (* An obligation is a set of traces *)
Definition mo := set ob. (* A model is a set of obligations *)
Definition state := id -> bool. (* States associate values to
variables *)
```

Then we define the sequence diagrams as an inductive type:

```
Inductive sd := | Dtau : sd | De : evt -> sd | Dstrict : sd ->
sd -> sd
| Dopt : cnd -> sd -> sd | Dalt : cnd -> sd -> sd -> sd | Dpar :
sd -> sd -> sd
| DSeq : sd -> sd -> sd | Dloop : cnd -> nat -> sd -> sd.
```

We can now write the denotational functions for each inductive constructor of the sequence diagram *sd*. Because the semantic model is the power set of obligations, and the obligation is the power set of traces, the denotational function needs to compute on the three-tier model.

When defining the denotational functional for the par operator, the interleave operator  $\parallel$  fulfills the foundational computation. We implement the interleave operator as a recursive function *intlev*:

```
Fixpoint intlev (t1 : tr) {struct t1}: tr -> set tr :=match
t1 with
| nil => fun t2 => (t2 :: nil)
| x :: u => fix aux (t2 : tr) {struct t2} : (set tr) :=
    match t2 with
    | nil => (t1 :: nil)
    | y :: v => set_union tr_dec (addEvtOb x (intlev u t2))
    (addEvtOb y (aux v))
end
end.
```

Here the *addEvtOb* is a function that adds a particular event before each trace of an obligation. The *aux* is a local recursive function.

When implementing the denotational function for the *seq* operator, the ordering of events on each lifelines and the ordering between transmission and receipt are all preserved, while all other ordering of events are arbitrary. The following two functions *filter* and *isWeak* are the most important ones:

```
Fixpoint filter (evs : set evt) (t : tr) {struct t} : tr :=
match t with
| nil => nil
| ev e :: tail => if (evtMem e evs) then (ev e) :: (filter evs
tail) else (filter evs tail)
| cd c :: tail => filter evs tail
end.
Fixpoint isWeak (st : state) (es : set evt) (h1 h2 t : tr) (ls :
set lf) : bool :=match ls with
| nil => true
| l :: tail => (andb (beq_tr st (filter (projLf l es) t) ((fil-
ter (projLf l es) h1) ++
(filter (projLf l es) h2)))) (isWek st es h1 h2 t tail))
end.
```

The *evtMem* returns a Boolean value, checking whether an event belongs a set of events. The filtering function *filter* is used to filter away elements from a trace, corresponding to operator  $\emptyset$ . The function *projLf* is corresponding to operator  $e \cdot l$  while the *isWeak* checks whether trace  $t$  is a member of weak sequencing of trace  $h1$  and  $h2$ .

Finally, the denotational semantics is presented as a recursive function *interp* that associates a sequence diagram to the semantic model  $mo$  in the state  $st$ . We use the corresponding function to match the shape of the sequence diagram  $d$  for each case. The *interp* works as an interpreter, which can evaluate a sequence diagram in a known environment:

```
Fixpoint interp (st : state) (d : sd) {struct d}: mo := match
d with
| Dtau => ((nil) :: nil) :: nil | De e => ((eve :: nil) :: nil)
:: nil
| Dstrict d1 d2 => strMo (interp st d1) (interp st d2)
| Dopt c d => optMo st c (interp st d)
| Dalt c d1 d2 => altMo st c (interp st d1) (interp st d2)
| Dpar d1 d2 => parMo (interp st d1) (interp st d2)
| Dseq d1 d2 => seqMo st (getEvts (DSeq d1 d2)) (getLfs (DSeq
d1 d2)) (interp st d1) (interp st d2)
| Dloop c n d => loopMo st c n (interp st d)
end.
```

## 140.5 Verification of Sequence Diagrams

Based on the trace semantics presented in previous section, we could impose some restrictions on the syntactically correct sequence diagrams and verify them directly in Coq. As an example, if both the transmitter and the receiver lifelines of a signal are present in a diagram, then the corresponding receive event of any transmit event must be in the diagram, and vice versa. The function *checkEvt* in Coq is used to check whether an event  $e$  satisfies this restriction:

```
Definition chkEvt (d : sd) (e : evt) :=
if (andb (set_mem lf_dec (getTrLf e) (getLfs d)) (set_mem
lf_dec (getReLf e) (getLfs d))) then set_mem evt_dec
(revEvt e) (getEvts d) else true.
```

Here, the function *getTrLf* and *getReLf* are used to get the lifeline of transmitter and receiver from event, respectively. The function *revEvt* is used to get the corresponding receive event of a transmit event or the transmit event of a receive event.

Furthermore, we could verify the semantic constraint on traces. As an example, regarding all traces, if at a point in a trace we have a receive event of a signal, then

up to that point we must have had at least as many transmits of that message as receives. We write the following function *chkTr* to check whether trace *t1* satisfies this restriction, in which the *countKd* is used to count the number of kind *k* in trace *t*:

```
Fixpoint chkTr (t1 : tr) : bool :=match t1 with
| nil => true
| ev (?, s, t, r) :: tail => if (ble_nat (countKd Receive
(filter ((!,s,t,r)::(? ,s,t,r)::nil) t1))
(countKd Send (filter ((!,s,t,r)::(? ,s,t,r)::nil) t1)))
then chkTr tail else false
| _ :: tail => chkTr tail
end.
```

Finally, the denotation function *interp* can also be used to prove properties of the semantics. For example:

```
Lemma eventNil : forall (st : state) (d : sd), getEvts d =
nil-> interp st d=(nil::nil)::nil.
```

Our definition of sequence diagrams as an inductive type enables the proof of this lemma to be straightforward. The proof of this lemma is a simple induction on the structure of *d* and use of the inversion tactic of Coq for each case.

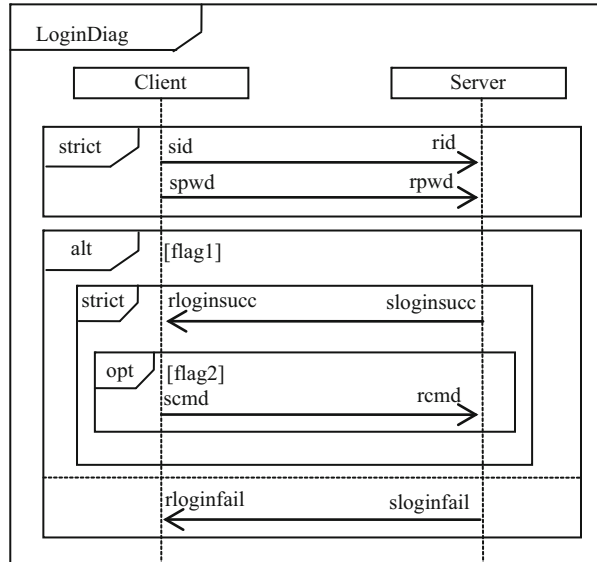
## 140.6 Case Study

Figure 140.1 shows a UML sequence diagram *LoginDiag* that describes a common login interaction. A client sends its id and password to the server and gets a login reply (*loginsucc* or *loginfail*). If the client successfully login, it can send a command to the server. The diagram *LoginDiag* can be transformed to Coq representations as follows:

```
Definition LoginDiag := Dstrict (Dstrict (Dstrict (De sid)(De rid)) (Dstrict
(De spwd)(De rpwd))) (Dalt (Bvar flag1) (Dstrict (Dstrict (De sloginsucc)
(De rloginsucc)) (Dopt (Bvar flag2) (Dstrict (De scmd) (De rcmd)))) (Dstrict
(De sloginfail) (De rloginfail))).
```

By executing the denotational function *interp* mentioned before, Coq prints the computation result which is the right interactive traces between object:  $((sid::rid::spwd::rpwd::flag1::sloginsucc::rloginsucc::flag2::scmd::rcmd::nil)::(sid::rid::spwd::rpwd::(notflag1)::sloginfail::rloginfail::nil)::nil)::((sid::rid::spwd::rpwd::flag1::sloginsucc::rloginsucc)::(sid::rid::spwd::rpwd::(notflag1)::sloginfail::rloginfail::nil)::nil)::nil$ .

**Fig. 140.1** An example of UML sequence diagram



Moreover, we could verify the syntax and semantic constraint by executing the verification functions in Coq. For example, when executing *Compute checkSd (LoginDiag)*, Coq returns *true*, which represents all the interactive traces from the diagram meet the restriction of *chkTr* defined before.

## 140.7 Related Work

The researchers survey the representative formal semantics proposed for UML sequence diagrams [7, 8]. However, the existing approaches mainly focus on getting a formal representation of sequence diagrams. The area is lack of a systematic methodology to analyze and verify the important properties.

The sequence diagrams have been translated to existing formalisms. The researchers use the formal specification language Z to present the semantics of sequence diagrams, well-formed rules, and consistent constraints and demonstrate that the work supports the validation [9]. The ASM semantic model for sequence diagram is proposed to describe the model's characteristics and improve the testing process of large systems [10]. A sequence diagram has been translated to a Petri net [11, 12]. These approaches can take advantage of the existing well-established tools to reason after translation. However, unlike the approach presented here, these techniques are not intuitive, and the interactive traces between objects are not clear.

There has been some work on formalize UML using a higher-order proof assistant. The HOL-OCL system is an interactive proof environment for UML

(mainly focus on the class diagrams) and OCL specifications, supporting object-oriented modeling and reasoning [13]. A corrected version for the UML 2.3 metamodel relating to state machine is provided with the help of the Coq [14]. In contrast to these approaches, here we apply Coq to support model-based development by offering the capability of analyzing the relations on models or specifications.

## 140.8 Conclusion and Future Work

We propose to use the theorem proof assistant—Coq to formalize the syntax and semantics of UML sequence diagrams and verify the syntax and semantics constraint on models. The desired properties can be stated as lemmas and the proof of the lemma ensures that the semantics satisfy their corresponding properties. The semantics is well constructed and takes into account the most popular combined fragments. This work will provide increased reliability guarantee for the future design and implementation in model-based development. A tool which can transform sequence diagram models into the Coq representations has been developed.

In the future, we will extend the syntax and semantics to cover a larger subset of UML sequence diagrams. Another topic is to conduct and evaluate larger case studies.

**Acknowledgement** This work is supported by National Natural Science Foundation of China (No.61070226 and No.61003181).

## References

1. OMG. (2005). Unified modeling language: superstructure. Version 2.0.
2. Coq, [EB/OJ]. Retrieved from <http://coq.inria.fr/>.
3. Yves, B., & Castéran, P. (2004). *Interactive theorem proving and program development—Coq'Art: The calculus of inductive constructions* (pp. 1–496). Berlin: Springer.
4. Lu, L., & Kim, D. (2011). Required behavior of sequence diagrams: Semantics and refinement. In: *Proceedings of the 2011 16th IEEE international conference on engineering of complex computer systems (ICECCS)* (pp. 127–136). Washington, DC: IEEE Computer Society.
5. Störrle, H. (2003). Semantics of interactions in UML 2.0. In: *Proceedings of the 2003 I.E. symposium on human centric computing languages and environments* (pp. 129–136). Washington, DC: IEEE Computer Society.
6. Lund, M. S. (2008). *Operational analysis of sequence diagram specifications*. Ph.D. Thesis, Faculty of Mathematics and Natural Sciences, University of Oslo.
7. Micskei, Z., & Waeselynck, H. (2011). The many meanings of UML 2 sequence diagrams: A survey. *Software and Systems Modeling*, 10(4), 489–514.
8. Lund, M. S., Refsdal, A., & Stølen, K. (2010). Semantics of UML models for dynamic behavior: A survey of different approaches. In *Proceedings of the 2007 international Dagstuhl conference on model-based engineering of embedded real-time systems* (pp. 77–103). Berlin: Springer.

9. Jingfeng, L., Yan, L., & Ping, C. (2003). The Z specification-based method for the semantic analysis of UML sequence diagrams. *Journal of Xidian University (Natural Science)*, 30(4), 519–524.
10. Xiang, Z., & Zhi-qing, S. (2009). ASM semantic modeling and checking for sequence diagram. In *Proceedings of the fifth international conference on natural computation* (pp. 527–530). Washington, DC: IEEE Computer Society.
11. Eichner, C., Fleischhack, H., Schrimpf, U., & Stehno, C. (2005). Compositional semantics for UML 2.0 sequence diagrams using Petri Nets. In *Proceedings of the 12th international conference on model driven* (pp. 133–148). Berlin: Springer.
12. Fernandes, J. M., Tjell, S., Jorgensen, J. B., & Ribeiro, O. (2007). Designing tool support for translating use cases and UML 2.0 sequence diagrams into a Coloured Petri Net. In *Proceedings of the sixth international workshop on scenarios and state machines* (p. 2). Washington, DC: IEEE Computer Society.
13. Brucker, A. D., & Wolff, B. (2008). HOL-OCL—a formal proof environment for UML/OCL. In *Proceedings of the 11th international conference on fundamental approaches to software engineering* (pp. 97–100). Berlin: Springer.
14. Barbier, F., & Ballagny, C. (2010). Proved metamodels as backbone for software adaptation. In *Proceedings of the 12th international symposium on high-assurance systems engineering* (pp. 114–121). Washington, DC: IEEE Computer Society.

# Chapter 141

## Quantitative Verification of the Bounded Retransmission Protocol

Xu Guo, Ming Xu, and Zongyuan Yang

**Abstract** In order to verify the reliability of the bounded retransmission protocol, probabilistic model checking technology is used in this paper. The integer semantics approach is introduced, which allows working directly at the level of the original probabilistic timed automaton (PTA). In such a method, clocks are viewed as counters storing nonnegative integer values, which increase as time passes. The PTA modeling the system can then be seen as a discrete-time Markov chain. Based on this fact, the protocol is modeled directly with DTMC. Properties are described in probabilistic computation tree logic. By making an analysis of the quantitative properties of the protocol, a threshold is obtained. Experimental result shows that no matter how many chunks to be transmitted, if the maximum retransmitted time is greater than or equal to 3, the protocol can be considered reliable. Method in this paper can not only verify the correctness of a system but also make analysis of nonfunctional indices of a system such as reliability or performance.

### 141.1 Introduction

The bounded retransmission protocol (BRP) is a data link layer protocol. It has been designed to transmit messages which are divided into small frames over unreliable channels. The protocol does not rely on fairness of data transmission channels, i.e., repeated transmission of a frame does not guarantee its eventual arrival. The number of retransmission attempts is also limited. So the pressure for reliability of the protocol involved poses an important challenge to verification techniques.

---

X. Guo (✉)  
East China Normal University, Shanghai 200241, China

Shanghai Dianji University, Shanghai 200092, China  
e-mail: [neuguox@126.com](mailto:neuguox@126.com)

M. Xu • Z. Yang  
East China Normal University, Shanghai 200241, China



This paper applies probabilistic model checking techniques to the quantitative verification of the protocol. Following D'Argenio's work [1, 2], we model the protocol in the framework of discrete-time Markov chain (DTMC) and then use PRISM [3, 4], a probabilistic model checking tool, to analyze the resulting model of the BRP. We establish quantitative properties of the protocol in order to verify its reliability, such as the minimal correctness of the protocol and the minimum number of retransmissions that satisfies our probabilistic requirements.

Related work. D'Argenio investigates what extent real-time aspects are important to guarantee the protocol correctness by using model checker SPIN and UPPAAL [2]. Ravn uses labeled transition systems to specify behavior of BRP and compositional reachability of the protocol based on its software architecture [5]. Forejt presents a framework for analyzing multiple quantitative objectives of system that exhibit both nondeterministic and stochastic behavior [6]. Kwiatkowska captures some quantitative properties of zero-conf protocol using probabilistic computation tree logic (PCTL) [7].

Organization of the paper. After a brief description of the BRP and related model technology (Sect. 141.2), the paper proceeds by giving the modeling of BRP in the framework of DTMC (Sect. 141.3). In the rest of the paper (Sects. 141.4 and 141.5), we present the PRISM tool dedicated to the model checking of probabilistic systems with a brief description of its theoretical framework and the results of experiments.

## 141.2 Preliminaries

### 141.2.1 *Sketch of the Protocol*

The protocol control procedures will be described by means of a sender S, a receiver R, and two lossy communication channels K and L. The sender sends elements of a file one by one over K to the receiver. After sending the frame, the sender waits for an acknowledgement or for a timeout. In case of acknowledgement, if it corresponds to the last chunk, the sending client is informed of correct transmission (signal OK); otherwise the next element of the file is sent. If a timeout occurs, the frame is resent, or the transmission of the file is broken off.

The receiver waits for a frame to arrive. This frame is delivered at the receiving client informing whether it is the first (FST), an intermediate (INC), or the last one (OK). Afterwards, an acknowledgement is sent over L to the sender. Then the receiver simply waits for more frames to arrive.

### 141.2.2 Basic Models

We introduce two basic models that will be used later in the paper.

**Definition 1.** A discrete-time Markov chain (DTMC) is a tuple  $D = (S, s, P, L)$  where  $S$  is a finite set of states,  $s$  is the initial state, and  $P: S \times S \rightarrow [0,1]$  is the transition probability matrix; this gives the probability  $P(s; s')$  that a transition will take place from state  $s$  to state  $s'$ . And  $L : S \rightarrow 2^{AP}$  is function labeling states with atomic propositions.

Classically, analysis of DTMCs often focuses on transient or steady-state behavior, i.e., the probability of being in each state of the chain at a particular instant in time or in the long run, respectively. Probabilistic model checking adds to this the ability to reason about path-based properties, which can be used to specify constraints on the probability that certain desired behaviors are observed. Properties are then expressed using temporal logic. For DTMCs, specifications can be written in PCTL, a probabilistic extension of the temporal logic CTL [7].

**Definition 2.** The syntax of PCTL is given by

$$\begin{aligned} \Phi &::= \text{true} \mid a \mid \Phi \wedge \Phi \mid \neg \Phi \mid P_{\sim p}[\Psi] \\ \Psi &::= X\Phi \mid \Phi \cup^{\leq k} \Phi \mid \Phi \cup \Phi \end{aligned}$$

where  $a$  is an atomic proposition,  $p \in \{0,1\}$  is a probability bound,  $\sim \in \{<, >, \leq, \geq\}$ ,  $k \in \mathbb{N}$ .

The key operator in PCTL is  $P_{\sim p}[\Phi]$  which means that the probability of a path formula  $\Phi$  being true in a state satisfies the bound  $\sim p$ .

### 141.2.3 Probabilistic Model Checking and PRISM

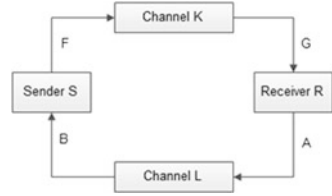
Probabilistic model checking is based on the construction of a probabilistic model from a precise, high-level description of a system's behavior. The model is then analyzed against one or more formally specified quantitative properties, usually expressed in temporal logic. These properties not only contain the correctness of the system but also a wide range of indices such as reliability or performance [8].

PRISM is a probabilistic model checker which provides support for analysis of DTMC and performs verification of PCTL formulae for DTMC.

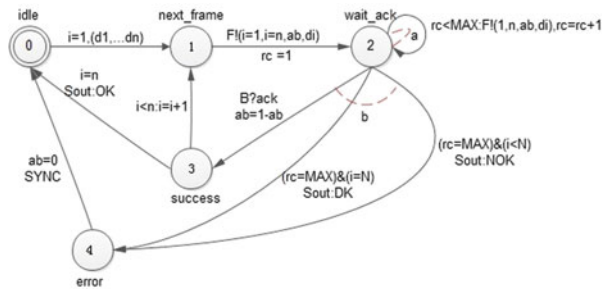
## 141.3 Modeling

We focus on a restricted set of reachable properties. For verifying probabilistic reachable properties, we must derive an equivalent finite-state (probabilistic) system. In the non-probabilistic framework, possible methods of reduction are region

**Fig. 141.1** Schematic view of the BRP



**Fig. 141.2** Model of sender



equivalence [9], forward exploration [10], and integer semantics [11]. These methods have been extended to the probabilistic framework. The first two methods require in practice the preliminary construction of an abstraction of the original probabilistic timed automaton [3]. For the sake of simplicity, we choose the third approach (integer semantics), which allows us to work directly at the level of the original probabilistic timed automaton. In such a method, clocks are viewed as counters storing nonnegative integer values, which increase as time passes. The PTA modeling the system can then be seen as a finite-state DTMC.

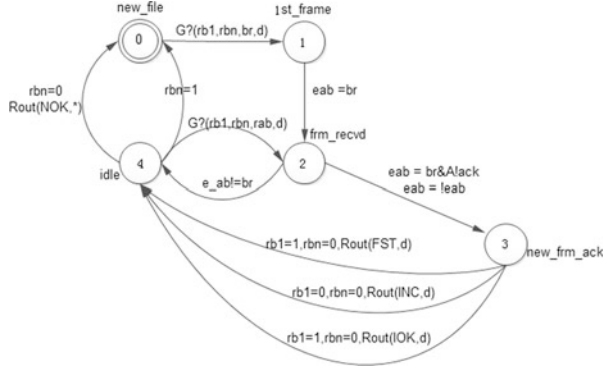
The model of the BRP protocol consists of four components operating in parallel, namely, sender, receiver, channel K, and channel L (see Fig. 141.1).

### 141.3.1 Sender

The sender S (see Fig. 141.2) has three system variables:  $ab \in \{0,1\}$ , indicating the alternating bit that accompanies the next chunk to be sent;  $i, 0 \leq i \leq n$ , indicating the subscript of the chunk currently being processed by S; and  $rc, 0 \leq rc \leq MAX$ , indicating the number of attempts undertaken by S to transmit a certain chunk.

On receipt of a new file, S sets  $i$  to 1. Going from state  $n\_frame$  to  $wait\_ack$ , chunk  $d_i$  is transmitted with corresponding information and  $rc$  is reset. In state  $wait\_ack$ , there are several probabilities: in case the maximum number of retransmissions has been reached, S moves to an error state while resetting  $x$  and emitting an DK or NOK indication depending on whether  $d_i$  is the last chunk or not.

**Fig. 141.3** Model of receiver



If  $rc < \text{MAX}$ , either an ACK is received and S moves to the success state while alternating ab or a retransmission is initiated. If the last chunk has been acknowledged, S moves from state success to state idle indicating the successful transmission of the file.

**141.3.2 Receiver**

The receiver R (see Fig. 141.3) has one system variable:  $e\_ab \in \{0,1\}$  indicating the alternating bit. In the state new\_file, R is waiting for the first chunk of the new file to arrive. On receiving the chunk, e\_ab is set to the just received alternating bit and R enters the state frame\_received. If e\_ab agrees with the just received alternating bit (which, due to the former assignment to ab, is always the case for the first chunk), then an ACK is sent via A and e\_ab is toggled. R is now in the state idle and waits for the next frame to arrive. It moves to the states frame\_received if such a frame arrives in time. And the above described procedure repeats. If timeouts occur, in case R has not just received the last chunk of a file, then an indication NOK is sent, and in case R has just received the last chunk, no failure is reported.

**141.3.3 Channels**

The channels K and L are lossy channels. In this model we assume that a frame is lost with probability 0.02, and acknowledgement is lost with probability 0.01.

### 141.4 Experiments

The experiment we perform is to try to find the minimum number of retransmissions that satisfies our probabilistic requirements for these properties when the transmitted file length  $N$  is an input parameter.

Properties 1–3 are concerned with transmissions that the sender does not consider successful, while property 4 considers an attempt for transmission with no reaction at the receiver side. The four main properties we have verified are as follows:

Property 1 represents the sender does not report a successful transmission. This is written (in PCTL):  $P=? [true \ U \ s=4]$ .

Figure 141.4 shows the probability that satisfies property 1 for different values of the transmitted file length ( $N$ ) and the maximal retransmission times ( $MAX$ ). The value of parameter  $N$  equals to 16, 32, 48, 64 and that of  $MAX$  equals to 2, 3, 4, 5, respectively. When  $MAX$  is greater than or equal to 3, no matter how many chunks to be transmitted, a failed transmission is rather slim, for the probability is less than 0.00025. When  $MAX$  is 2, as  $N$  increases, the probability of a failed transmission increases fast. So if  $MAX$  is greater or equal to 3, we can consider that the transmission is reliable.

Property 2 represents the sender that reports an uncertainty on the success of the transmission. The property written in PCTL is  $P=? [true \ U \ s=4 \ \& \ srep=2]$ .

Figure 141.5 shows the probability that satisfies property 2 for different values of  $N$  and  $MAX$ . When  $MAX$  is greater than or equal to 3, we can ignore this phenomenon that the sender receives a DK instead of OK acknowledgement after a success transmission for the probability nearly equals to 0. So if  $MAX$  is greater than or equal to 3, we can consider that the transmission is reliable.

Property 3 represents the sender that reports a certain unsuccessful transmission after transmitting half a file. Written in PCTL is  $P=? [true \ U \ s=4 \ \& \ srep=1 \ \& \ i>N/2]$ .

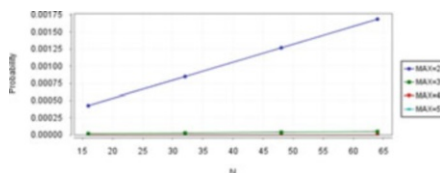


Fig. 141.4 PRISM measurement for property 1

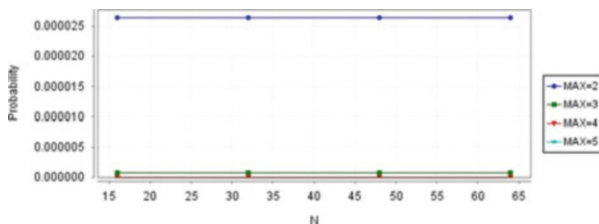


Fig. 141.5 PRISM measurement for property 2

Property 4 represents the receiver does not receive any chunk of a file. Written in PCTL is  $P=? [\text{true} \cup !(\text{srep}=0) \ \& \ !\text{recv}]$ .

From the result, we can see that if the value of MAX is greater than or equal to 3, the probability of the sender reports a certain unsuccessful transmission after transmitting half a file is less than 0.00025. And the probability that the receiver does not receive any chunk of a file is less than 0.00001. Due to space constraints, experiment results of properties 3 and 4 are emitted.

So we can conclude that the protocol is reliable no matter how many chunks to be transmitted; if we set a proper value of maximal retransmission time, the threshold is 3.

## 141.5 Conclusion

In this paper, we applied probabilistic model checking technology to the verification of quantitative properties of BRP.

A crucial problem in our work is the assumption of 0 delay in the transmission channels K and L. This reduces the size of region space significantly. As a result of these zero delays, the receiver may detect a transmission failure, while the sender has not aborted the transmission.

Our work provides a fully automatic verification of the parameterized version of the BRP. With transmitted file length N and the maximal retransmission counter MAX as parameters, PRISM can measure the properties. We can extend the reachable properties of the protocol with probability. Such properties are expressed in terms of a probability bound, for example, the property the sender reports an uncertainty on the success of the transmission can be expressed as a probability bound.

Future work can address the application of PCTL\* to describe the properties of one model. As PCTL\* subsumes PCTL and LTL, it is a more expressive logic. Application of PRISM to model and analysis mobile ad hoc networks is underway.

**Acknowledgements** This work is supported by National Natural Science Foundation of China (No. 61070226).

## References

1. D'Argenio, P. R., Jeannot, B., Jensen, H. E., & Larsen, K. G. (2001). *Reachability analysis of probabilistic systems by successive refinements. Process algebra and probabilistic methods. Performance modelling and verification* (pp. 39–56). Berlin: Springer.
2. D'Argenio, P. R., Katoen, J. P., Ruys, T. C., & Tretmans, G. J. (1997). *The bounded retransmission protocol must be on time!* (pp. 416–431). Berlin: Springer.
3. Kwiatkowska, M., Norman, G., & Parker, D. (2011). *PRISM 4.0: Verification of probabilistic real-time systems. Computer aided verification* (pp. 585–591). Berlin: Springer.

4. Hinton, A., Kwiatkowska, M., Norman, G., & Parker, D. (2006). *PRISM: A tool for automatic verification of probabilistic systems. Tools and algorithms for the construction and analysis of systems* (pp. 441–444). Berlin: Springer.
5. Ravn, A. P., Srba, J., & Vighio, S. (2011). *Modelling and verification of web services business activity protocol. Tools and algorithms for the construction and analysis of systems* (pp. 357–371). Berlin: Springer.
6. Forejt, V., Kwiatkowska, M., Norman, G., Parker, D., & Qu, H. (2011). *Quantitative multi-objective verification for probabilistic systems. Tools and algorithms for the construction and analysis of systems* (pp. 112–127). Berlin: Springer.
7. Kwiatkowska, M., Norman, G., & Parker, D. (2010). Advances and challenges of probabilistic model checking. In *48th Annual Allerton Conference on Communication, Control, and Computing (Allerton), IEEE*, Los Alamitos (pp. 1691–1698).
8. Filieri, A., Ghezzi, C., & Tamburrelli, G. (2011). Run-time efficient probabilistic model checking. In *Proceedings of the 33rd International Conference on Software Engineering, ACM*, New York (pp. 341–350).
9. Aceto, L., Ingólfssdóttir, A., & Larsen, K. G. (2007). *Reactive systems: Modelling, specification and verification* (pp. 121–136). Cambridge: Cambridge University Press.
10. David, A., Illum, J., Larsen, K. G., & Skou, A. (2010). Model-based framework for schedulability analysis using uppaal 4.1. *Model-Based Design for Embedded Systems*, 7(4), 93–119.
11. Hartmanns, A., & Hermanns, H. (2009). A Modest approach to checking probabilistic timed automata. In *Sixth International Conference on the Quantitative Evaluation of Systems, 2009. QEST'09, IEEE*, Los Alamitos (pp. 187–196).

# Chapter 142

## A Cluster-Based and Range-Free Multidimensional Scaling-MAP Localization Scheme in WSN

Ke Xu, Yuhua Liu, Cui Xu, and Kaihua Xu

**Abstract** As using traditional MDS-MAP algorithm to locate nodes' position in irregular WSN leads to low positional accuracy, based on this fact, this chapter presents an improved algorithm named MDS-MAP(C, RF). The algorithm can effectively divide a WSN into several clusters, and each cluster locates all nodes' position in it and forms a local position map. After all clusters get local position maps, the algorithm merges all the local position maps together using the information of inter-cluster nodes. Simulations demonstrate the proposed algorithm yields smaller accuracy error in irregular WSN.

### 142.1 Introduction

A sensor network is composed of a large number of sensor nodes, which are densely deployed either inside the phenomenon or very close to it. The goal of a sensor network is to perceive, collect, and process the information of specific objects and send the information to observers [1]. The localization refers to computing all nodes' position based on a few nodes' position information in a sensor network. Localization plays a key role within the application of WSN. Localization is a hot research topic [2–4]. Localization techniques can be classified in a different way [4]. MDS-MAP algorithm is first proposed in the year of 2003 [5]; the noted advantage of MDS-MAP is that with a few anchor node information, it can compute

---

K. Xu • Y. Liu (✉) • C. Xu  
School of Computer, Central China Normal University, Wuhan 430079, China  
e-mail: [yhliu@mail.ccnu.edu.cn](mailto:yhliu@mail.ccnu.edu.cn)

K. Xu  
College of Physical Science and Technology, Central China Normal University,  
Wuhan 430079, China



the position of each node, and its localization accuracy is higher than most known localization technology. The main disadvantage of MDS-MAP is that it does not work well in irregular sensor networks, and the position error is high.

This chapter proposes the MDS-MAP(C, RF) algorithm. The core idea is dividing a given network into several clusters; each cluster computes its local position map through MDS-MAP algorithm, respectively, then merges all the local position maps together to get a global map.

## 142.2 MDS-MAP

### 142.2.1 MDS

MDS (Multidimensional Scaling) is a set of data analysis technology which can transform the given data into geometry model, thus problems can be visually solved. Torgerson had firstly given the terminology MDS based on the work of Richardson and proposed the first MDS method [6].

### 142.2.2 Procedure of MDS-MAP

MDS-MAP consists of three steps [5]:

*Step 1:* Calculate the shortest distances between nodes in WSN. The time complexity of this step is  $O(n^3)$ , where  $n$  is the number of nodes.

*Step 2:* Calculate the first  $r$  maximum eigenvalues of  $r$ -dimensional space to construct the relative location map of nodes. The time complexity of this step is  $O(n^3)$ , where  $n$  is the number of nodes.

*Step 3:* Based on the location information of anchor nodes (it needs at least  $r + 1$  nodes in  $r$ -dimensional space), the algorithm transforms the relative location map to absolute location map. The time complexity of this step is  $O(m^3)$ , where  $m$  is the number of anchor nodes.

## 142.3 Improved MDS-MAP

Traditional MDS-MAP could not work well in irregular WSN because the location error is quite large.

The procedure of MDS-MAP(C, RF) is shown below:

*Step 1:* Divide a wireless sensor network into several clusters; the method of dividing the cluster is k-hop clustering.

Step 2: Use traditional MDS-MAP algorithm to build the location map of each cluster which is produced in step 1.

Step 3: Merge the location map of each cluster together to form a global location map.

### 142.3.1 Clustering

#### 142.3.1.1 Selection of Cluster Heads

At the initial stage, all sensor nodes are randomly placed in a large sensor field and broadcast Hello messages. Then each node discovers its node degree (the number of nodes it can reach by one hop), the IDs of all neighboring nodes, and distances to these neighboring nodes. A node with the lowest ID becomes an initiator of cluster generation step and starts a timer  $T_1(i)$ ; formula (142.1) is the computation of  $T_1(i)$ :

$$T_1(i) = \alpha_{deg} \max \left( 0, 1 - \frac{\deg(i)}{\theta_{deg}(i)} \right) + t_{BCD} \quad (142.1)$$

where  $t_{BCD}$  is a small broadcasting random delay,  $\deg(i)$  is the connectivity degree of node  $i$ ,  $\theta_{deg}(i)$  denotes the largest degree of neighbor nodes, and  $\alpha_{deg}$  is a weight factor for the degree.

The initiator, node  $i$ , becomes a cluster head node when  $T_1(i)$  is time out; name the cluster as Cluster- $i$ . However, if it receives a cluster head declaration message from other nodes before  $T_1(i)$  expires, the node will stop the timer and become a member of the cluster.

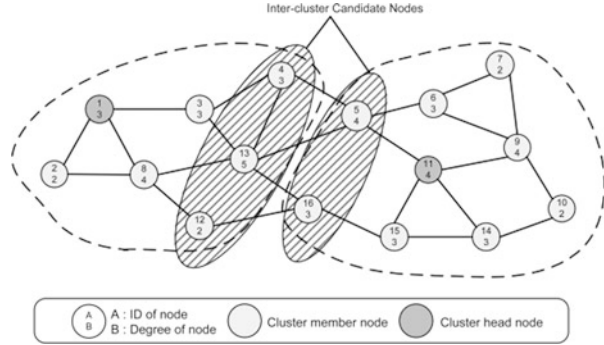
The cluster head floods declaration messages to nodes within  $2k$  hops. Any node within  $k$  hops to Cluster- $i$  head nodes becomes a member of Cluster- $i$ ; any node located between  $(k + 1)$  hops and  $2k$  hops to node  $i$  becomes a candidate of a new cluster head node. Each candidate starts a timer  $T_2(j)$  ( $j$  is the candidate node's ID):

$$T_2(j)^2 = \alpha_{deg} \max \left( 0, 1 - \frac{\deg(j)}{\theta_{deg}(j)} \right) + \alpha_{dist} \left( 1 - \frac{chdist(j)}{2k} \right) + t_{BCD} \quad (142.2)$$

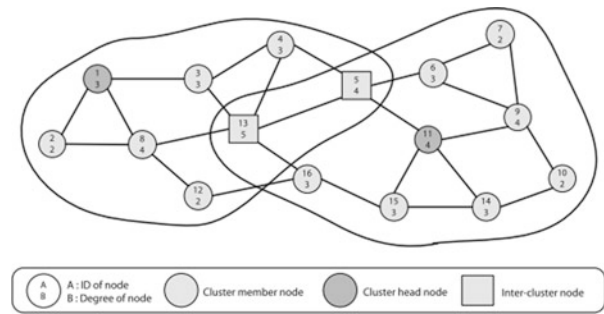
where  $\alpha_{dist}$  is a weight factor for the distance and  $chdist(j)$  denotes the number of hops between current node  $i$  and the cluster head node ( $k + 1 \leq chdist(j) \leq 2k$ ). A node with an expired timer  $T_2(j)$  becomes a new cluster head node and floods cluster head declaration messages declare- $j$ , continuing the cluster generation process. Using random delay values  $t_{BCD}$  prevents a broadcast storm happen.

Figures 142.1 and 142.2 give an example of  $k$ -hop clustering ( $k = 2$ ). In Fig. 142.2, nodes 2, 3, 4, 5, 8, 12, and 13 are members of Cluster = 1 (node 1 is its cluster head). Nodes 5, 6, 7, 9, 10, 11, 13, 14, 15, and 16 are members of Cluster-11, where node 11 is its cluster head.

**Fig. 142.1** Selection of cluster heads



**Fig. 142.2** Result of k-hop clustering



### 142.3.1.2 Inter-cluster Nodes

There exists a special kind of nodes between two adjacent clusters: they receive more than one declares messages. As shown in Fig. 142.2, nodes 4, 12, 13, 5, and 16 receive declare from Cluster-1 and Cluster-11. These nodes are called inter-cluster candidate nodes. Nodes 4, 12, and 13 are inter-cluster candidate nodes of Cluster-1, and nodes 5 and 16 are inter-cluster candidate nodes of Cluster-11, respectively. Each cluster head node randomly selects one of the inter-cluster candidate nodes as an inter-cluster node to the neighboring clusters. These nodes will be used in the phase of merging local map together to construct a global map. As shown in Fig. 142.2, nodes 5, 13, and 16 are the inter-cluster nodes of Cluster-1 and Cluster-11.

### 142.3.2 Building Cluster Location Map

The step of this process is shown below:

*Step 1:* After k-hop clustering is completed, each cluster head node calculates distance based on RSSI and the IDs of neighboring nodes [7]. Using the distance information which is expressed in a distance matrix and the shortest path algorithm, Dijkstra or Floyd, the cluster head node constructs a shortest distance matrix **D**.

*Step 2:* Using this shortest distance matrix **D**, the MDS-MAP algorithm produces a relative location map within the cluster.

*Step 3:* Refine cluster location map.

Step 1 and step 2 are clearly described [8]. This chapter only gives the detail of step 3.

Define  $NB(i)$  as a set of nodes in Cluster- $i$ ; the cluster head node is node  $i$ ,  $v_r, v_s \in NB(i)$ ,  $v_r(ID) \leq v_s(ID)$ , and  $\epsilon (> 0)$  is a threshold which has a small value for the purpose of high accuracy. The process of refinement is illustrated below:

1. Define  $\vec{X}_{ir} = (x_{r1}, x_{r2}, \dots, x_{rm})$  to indicate row vector of node  $r$  within Cluster- $i$  in  $m$ -dimensional space; set any values to nodes' initial coordinates in Cluster- $i$ .
2. Calculate Euclidean distance between nodes by formula (142.3):

$$d_{rs} = \sqrt{(\vec{X}_r - \vec{X}_s)(\vec{X}_r - \vec{X}_s)^T}, V_r(ID) \leq V_s(ID) \tag{142.3}$$

3. Use PAV technique to compute the difference  $\hat{d}_{rs}$  of  $d_{rs}$ .
4. Compute Stress 1:

$$Stress1 = \sqrt{\frac{\sum (d_{rs} - \hat{d}_{rs})^2}{\sum d_{rs}^2}} \tag{142.4}$$

If  $Stress1 < \epsilon$ , end the process, and  $\mathbf{X}_{cluster-i} = [\vec{X}_{i1}, \vec{X}_{i2}, \dots, \vec{X}_{in}]^T$  is the coordinate matrix of Cluster- $i$ , else go to step 5.

5. Compute the new coordinate of each node using formula (5):

$$\vec{X}_r^* = \vec{X}_r + \frac{\alpha}{M-1} \sum_{V_s \in NB(i)} (1 - \hat{d}_{rs}/d_{rs})(\vec{X}_s - \vec{X}_r) \tag{142.5}$$

where  $\alpha$  is step factor and  $M$  depicts the number of  $NB(i)$ .

6. Update **X**; go to step 2.

### 142.3.3 Merging Cluster Location Map

If  $NB(i) \cap NB(j)$  have more than three nodes, then using formula (6) can transform the coordinates of nodes in Cluster-j to coordinate in Cluster-i:

$$\mathbf{X}_{cluster-i} = sR(\mathbf{X}_{cluster-j}) + \mathbf{X}_0 \tag{142.6}$$

where  $s$  is zoom factor,  $R(\cdot)$  depicts rotation transformation,  $\mathbf{X}_0$  indicates translation transformation; the technique of compute  $s, R(\cdot), \mathbf{X}_0$  is shown below.

Suppose  $\mathbf{S}_{cluster-i}, \mathbf{S}_{cluster-j}$  are coordinate matrices of nodes of  $NB(i) \cap NB(j)$  in Cluster-i and Cluster-j. The row vector of  $\mathbf{S}_{cluster-i}, \mathbf{S}_{cluster-j}$  is depicted by  $\vec{S}_{cluster-i,i}, \vec{S}_{cluster-j,i}$ . Matrix  $\mathbf{D}$  is the nodes' coordinate matrix of  $NB(j) - NB(i) \cap NB(j)$  in Cluster-i,  $|NB(i) \cap NB(j)| = n, (n \geq 3)$ . Expand the nodes' coordinate in  $\mathbf{S}_{cluster-i}$  and  $\mathbf{S}_{cluster-j}$  to 3-dimensional space, and set 1 to the third coordinate value of each node. The central points of  $\vec{S}_{cluster-i}, \vec{S}_{cluster-j}$  are computed through formula (142.7):

$$\vec{S}_{cluster-i} = \frac{1}{n} \sum_{k=1}^n \vec{X}_{cluster-i,k}, \quad \vec{S}_{cluster-j} = \frac{1}{n} \sum_{k=1}^n \vec{X}_{cluster-j,k} \tag{142.7}$$

Assume  $\vec{S}'_{cluster-i,k} = \vec{S}_{cluster-i,k} - \vec{S}_{cluster-i}, \vec{S}'_{cluster-j,k} = \vec{S}_{cluster-j,k} - \vec{S}_{cluster-j}$ ; the zoom factor  $s$  in formula (142.6) is calculated through formula (142.8):

$$s = \sqrt{\sum_{k=1}^n \|\vec{S}'_{cluster-i,k}\|^2 / \sum_{k=1}^n \|\vec{S}'_{cluster-j,k}\|^2} \tag{142.8}$$

The covariance matrix  $\mathbf{C}$  of  $\vec{S}'_{cluster-j}, \vec{S}'_{cluster-i}$  can be computed by using formula (142.9).

$$\mathbf{C} = \sum_{k=1}^n \vec{S}'_{cluster-j,k} \vec{S}'_{cluster-i,k}{}^T = \begin{bmatrix} C_{xx} & C_{xy} & C_{xz} \\ C_{yx} & C_{yy} & C_{yz} \\ C_{zx} & C_{zy} & C_{zz} \end{bmatrix} \tag{142.9}$$

Define matrix  $\mathbf{U}$  as formula (142.10):

$$\mathbf{U} = \begin{bmatrix} C_{xx} + C_{yy} + C_{zz} & C_{yz} - C_{zy} & C_{zx} - C_{xz} & C_{xy} - C_{yx} \\ C_{yz} - C_{zy} & C_{xx} - C_{yy} - C_{zz} & C_{xy} + C_{yx} & C_{zx} + C_{xz} \\ C_{zx} - C_{xz} & C_{xy} + C_{yx} & -C_{xx} + C_{yy} - C_{zz} & C_{yz} + C_{zy} \\ C_{xy} - C_{yx} & C_{zx} + C_{xz} & C_{yz} + C_{zy} & -C_{xx} - C_{yy} + C_{zz} \end{bmatrix} \tag{142.10}$$

Compute  $\vec{\lambda}_m$  which is composed by eigenvalue of matrix  $\mathbf{U}$ :  $\vec{\lambda}_m = (\lambda_0, \lambda_1, \lambda_2, \lambda_3)$   
 Compute  $\mathbf{R}(\cdot)$  of formula (142.6) using formula (142.11):

$$\mathbf{R} = \begin{bmatrix} \lambda_0^2 + \lambda_1^2 - \lambda_2^2 - \lambda_3^2 & 2(\lambda_1\lambda_2 - \lambda_0\lambda_3) & 2(\lambda_1\lambda_3 + \lambda_0\lambda_2) \\ 2(\lambda_2\lambda_1 + \lambda_0\lambda_3) & \lambda_0^2 - \lambda_1^2 + \lambda_2^2 - \lambda_3^2 & 2(\lambda_2\lambda_3 - \lambda_0\lambda_1) \\ 2(\lambda_3\lambda_1 - \lambda_0\lambda_2) & 2(\lambda_3\lambda_2 + \lambda_0\lambda_1) & \lambda_0^2 - \lambda_1^2 - \lambda_2^2 + \lambda_3^2 \end{bmatrix} \quad (142.11)$$

Compute  $X_0$  through formula (142.12):

$$\vec{X}_0 = \vec{S}_{cluster-i} - \vec{S}_{cluster-j} \mathbf{R}^T \quad (142.12)$$

Using formula (142.6), (142.8), (142.11), and (142.12), the matrix  $\mathbf{D}$  (the nodes' coordinate matrix of  $NB(j) - NB(i) \cap NB(j)$ ) can be transformed to the coordinate system of Cluster- $i$ ; the technique is shown as formula (142.13).

$$\mathbf{F} = \mathbf{D}\mathbf{R}^T + \mathbf{I}X_0 \quad (142.13)$$

where  $\mathbf{I}$  is an all-1 matrix which has an equal row number as matrix  $\mathbf{D}$ .

Using this technique, all clusters' location map can be merged to one location map.

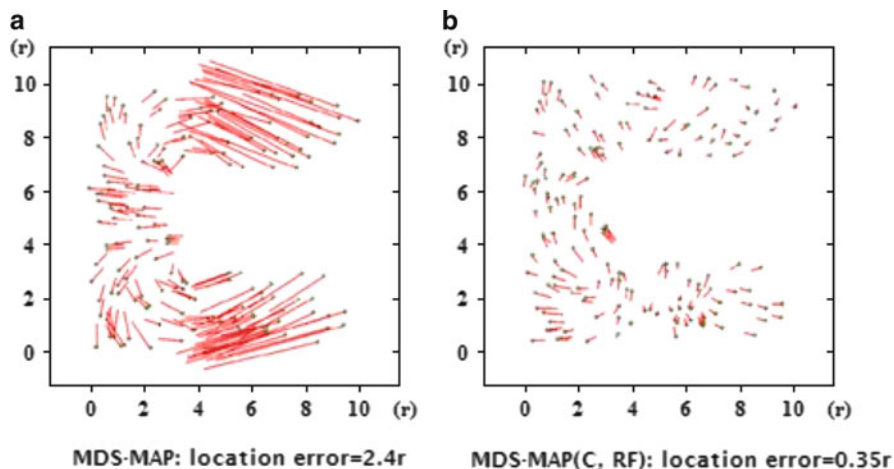
## 142.4 Experiments and Simulations

### 142.4.1 Simulations of Two Types of WSN

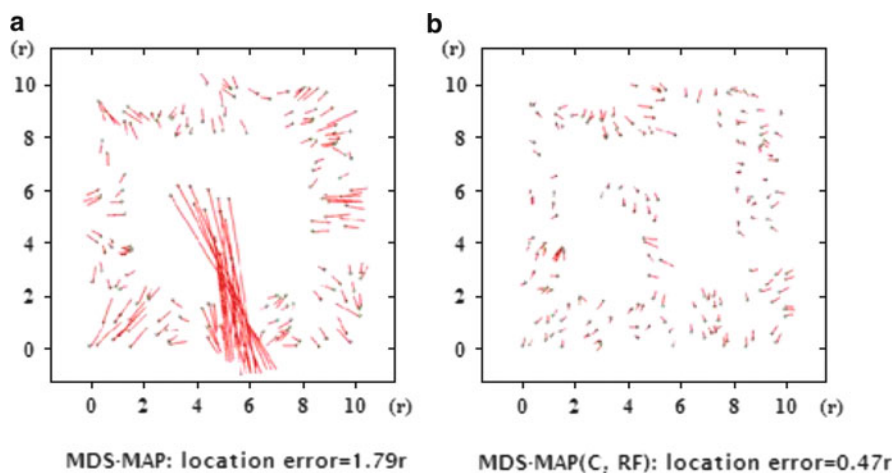
Suppose the transmission radius of sensor node is  $1.5r$ , all nodes are deployed in a  $10r \times 10r$  area. The simulations have two types. Type 1, 170 nodes are randomly distributed within a sensing field that is C shaped in the middle. Type 2, 170 nodes are randomly distributed in an environment with a horseshoe-shaped hole in the middle of the sensing field.

Figures 142.3 and 142.4 show the results of two types of simulations. Figure 142.3 depicts localization results where a sensor topology is configured in a C shape. Figure 142.3a indicates large errors are produced by conventional MDS-MAP in such environments. The average error of conventional MDS-MAP is  $2.4r$ . Figure 142.3b is the localization result of MDS-MAP(C,RF); its average error is quite small compared to (a), it is  $0.35r$ . MDS-MAP(C,RF) improves localization accuracy up to 585 % in a C-shaped sensor topology compared to the conventional MDS-MAP.

Figure 142.4 shows localization results in a sensor topology with a horseshoe-shaped hole. The localization errors are  $1.79r$  for conventional MDS-MAP and



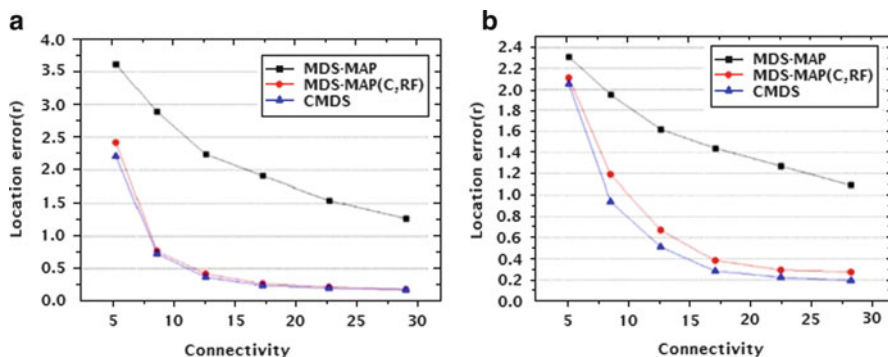
**Fig. 142.3** Localization results where a sensor topology is configured in a C shape



**Fig. 142.4** Localization results in a sensor topology with a horseshoe-shaped hole

0.47r for MDS-MAP(C,RF). MDS-MAP(C,RF) improves localization accuracy up to 280 % in a horseshoe-shaped sensor topology compared to the conventional MDS-MAP.

From the simulation results, we conclude that the MDS-MAP(C,RF) produces better result than conventional MDS-MAP in irregular WSNs.



**Fig. 142.5** Localization error with connectivity in C-shaped WSNs (a), horseshoe-shaped WSNs (b)

### 142.4.2 Localization Error with Connectivity

Figure 142.5 shows the localization errors with connectivity in two types of environments. Figure 142.5a shows that the localization errors of the HMDS and MDS-MAP(C,RF) are quite smaller than the conventional MDS-MAP under the different connectivity in a sensor network within a C-shaped hole. In sensor networks where a horseshoe-shaped hole is in the middle, as is shown in Fig. 142.5b, the localization errors of HMDS and MDS-MAP(C,RF) are not hugely different, but conventional MDS-MAP's localization error is quite larger than the former ones'. Figure 142.5b illustrates that MDS-MAP(C, RF) is better than conventional MDS-MAP. MDS-MAP(C, RF) improves a bit compared to HMDS.

It can be concluded from the simulation results that MDS-MAP(C, RF) works better in irregular WSNs and can gain a high precision localization result.

## 142.5 Conclusion

This chapter proposes MDS-MAP(C, RF) algorithm based on conventional MDS-MAP. Simulation results indicate that MDS-MAP(C, RF) is better than MDS-MAP for it can improve localization accuracy. Since MDS-MAP(C, RF) is a cluster-based technique which can reduce the burden of center nodes, it can extend the life span of WSN. The next stage work is focused on the effect to WSN's life span of MDS-MAP(C, RF).



## References

1. Mao, G., Fidan, B., & Anderson, B. (2007). Wireless sensor network localization techniques. *Computer Networks.*, 51(10), 2529–2553.
2. Pal, A. (2010). Localization algorithm in wireless sensor networks: Current approaches and future challenges. *Network Protocols and Algorithms.*, 2(1), 45–73.
3. Stefano, G. D., & Petricola, A. (2008). A distributed AOA based localization algorithm for wireless sensor networks. *Journal of Computers.*, 3(4), 1–8.
4. Niewiadomska-szynkiewicz, E. (2012). Localization in wireless sensor networks: Classification and evaluation of techniques. *International Journal of Applied Mathematics and Computer Science.*, 22(2), 281–297.
5. Shang, Y., Ruml, W., Zhang, Y., & Fromherz, M. (2003). Localization from mere connectivity. In *Proceeding of the fourth ACM international symposium on Mobile ad hoc networking & computing* (pp. 201–212). New York: ACM.
6. Torgeson, W. S. (1965). Multidimensional scaling of similarity. *Psychometrika.*, 30(4), 379–393.
7. Xu, K., Wang, Y., & Liu, Y. (2007). A clustering algorithm based on power for WSNs (Vol. 4489, pp. 153–156). *Lecture Notes in Computer Science*, Springer.
8. Shang, Y. & Ruml, W. (2004). Improved MDS-based localization. In *INFOCOM 2004, twenty-third annual joint conference of the IEEE computer and communications societies* (pp. 2640–2651), Hong Kong, China.

# Chapter 143

## A Resource Information Organization Method Based on Node Encoding for Resource Discovering

Zhuang Miao, Qianqian Zhang, Songqing Wang, Yang Li, Weiguang Xu, and Jiang Xiao

**Abstract** In order to discover a variety of network resources of structured P2P, resource information organization methods are required, which should have scalability and robustness. However, structured P2P has bad performance because of churn, so it cannot be widely used currently. To solve the problem, a resource information organization method based on node encoding is provided in this chapter. A node group-based resource information organization and resource distribution-based node encoding algorithm are presented. Redundancy tables are established based on the overlay of the node. The proposed algorithm can decrease the burst of transmission and reduce the traffic load of transited information. The experiment results show that the presented method is tolerant to churn.

### 143.1 Introduction

The structured P2P-based resource discovery protocol is the best choice for building a resource discovery system [1]. It has the best performance in scalability and robustness. The route algorithm that the protocol adopts can keep high performance and low cost. However, it cannot be widely used in the present world with bad performance for churn.

Traditional resource information organization modes of structured P2P need to transfer a great deal of information in a short period under churn. The high burst of transmission and big amount of information will dispend all the bandwidth for a long time. There will be no bandwidth to deal the users' requests. The route algorithm which the structured P2P-based resource discovery protocol adopts has low success rate of route with churn.

---

Z. Miao (✉) • Q. Zhang • S. Wang • Y. Li • W. Xu • J. Xiao  
Institute of Command Information System, PLA University of Science and Technology,  
Nanjing 210000, China  
e-mail: [emiao\\_beyond@163.com](mailto:emiao_beyond@163.com)

In this chapter, a resource information organization method based on node encoding is presented. The method can solve the churn problem of structured P2P. In the method, resource information is organized to be distributed according to geographic position. Nodes of the structured P2P are encoded and grouped according to the overlay of the node. The new information organization method can decrease the burst of transmission and reduce the amount of transited information.

## 143.2 Related Works

To reduce the traffic load of transportation, data redundancy strategy [2] and node selection strategy [3] are used. Multiple backups of the resource information are stored in some different nodes in data redundancy strategy. The node selection strategy uses geographic position and node load and so as the standard to select node as the neighbor. The success rate of route can be improved by mending the route table maintain strategy [4] and route configuration strategy [5]. However, these strategies lack pertinence for structured P2P-based resource discovery protocol. Ou presented a performance evaluation of a structured communication-oriented P2P system in the presence of churn [6].

To eliminate churn, a general solution is proposed which makes a P2P network need not pay much attention to churn problem by introducing a logic layer named Dechurn [7]. Adding the churn resilience, which is achieved by employing the properties of threshold secret sharing schemes, an existing anonymous peer-to-peer network design is improved [8]. The impact of churn on object management policies in CDN-P2P systems is studied and the effectiveness of buffer enlargement and replicating control in reducing the effects of churn on these policies is analyzed [9]. A peer-churn resistant video multicast system over P2P networks, which takes into account both the link delay and the peer stability in order to achieve a seamless video streaming multicast service with a low delay, is presented [10]. There is no overall design for churn-tolerant resource protocol in existing strategies. So to solve churn is an open-ended problem [11].

## 143.3 Node Group-Based Resource Information Organization

Resource information organization consists of overlay structure, resource information storing, and transporting rules. The structured P2P resource information organization has good performance in scalability and robustness. It should be used widely, but it cannot adapt to churn. There will be a great deal of information needed to transfer in a short period. This makes cost out of existing bandwidth. There will be no more bandwidth to deal users' requests.

The main reason of the problem is that the node encoding is unique and random. The uniqueness makes that there is only one node on one position in code space. So if the node leaves, the information stored in one must be moved to other nodes immediately. The randomness makes that the resource information cannot be stored in local node. These information have different actions with the node that stores them, so they need to be moved frequently.

To solve the problems, we present a resource distribution-based node encoding algorithm. It breaks the limitation of the uniqueness and randomness of node encoding. Based on this algorithm and FreePastry [12], a redundancy table is provided and the redundancy table-based overlay is designed to decrease the burst of transmission of route information. To reduce the amount of resource information, a new resource information storing rule is given.

### ***143.3.1 Resource Distribution-Based Node Encoding Algorithm***

The resource distribution-based node encoding algorithm encodes node depending on the node's resource distribution and the local bandwidth and breaks the limitation of the uniqueness and randomness of traditional ones. The node codes will have relationship with the local resource distribution according to the new algorithm.

The main idea of the node encoding algorithm is as follows: *Limen* is called as threshold of encoding and equal to the local bandwidth; if the amount of one local resource information exceeds *Limen*, the node will have the code of this resource; if there is no resource information out of threshold of encoding in one node, this node will get a code randomly. The pseudo code of the node encoding algorithm is shown in Table 143.1.

After encoding there will be some difference from traditional algorithm: (1) one node may have more than one code, and (2) one code may be mapped to more than one node. The codes of one node will be adjusted with the change of amount of resource information.

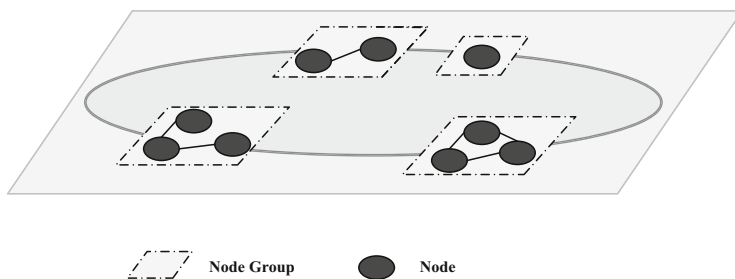
### ***143.3.2 Redundancy Table-Based Overlay***

The resource information organization needs to maintain three tables, leaf table, route table, and neighbor set, because it is based on FreePastry. However, there may be more than one node on one position in code space, so a redundancy table is designed. Some or all table entries of one table can be mapped to multiple appropriate nodes with strategy of redundancy table. Then the overlay is provided based on the redundancy table, and Fig. 143.1 shows the structure of it.

**Table 143.1** Pseudo code of resource distribution-based node encoding algorithm*Algorithm:* Resource distribution-based node encoding algorithm**Input:** encoding the *node**R*: resource set of *node* $c_k$ : amount of information of  $k$ -th resource of *R**W*: bandwidth of *node***Output:** codes of *node**CODE*: code set of *node***Algorithm:**

```

1 Allocate_Id(node){
2   CODE ← ∅,  $k = 0$ , Limen = W;
3   while( $k < |R|$ ){
4     if( $c_k > \textit{Limen}$ ){
5       CODE = CODE ←  $\textit{code}_k$ ; //  $\textit{code}_k$ : code of the  $k$ -th resource
6     }
7      $k++$ ;
8   }
9   if( $|CODE| = 0$ ){
10    CODE = CODE ← Random(); // Random(): producing node code randomly
11  }
12  return CODE;
13 }
```

**Fig. 143.1** Node group-based overlay

There may be more than one node on one position, so the construction rule of overlay is that all the nodes which have the same code are organized into a group, called node group. Nodes are constructed according to unstructured P2P in one group, called group intra-network. And one node group is connected to other groups according to the three redundancy tables, called group inter-network.

Because of multiple codes for one node, it can belong to more than one group. Node has independent action in each group which it belongs to. When constructing the group inter-network, we should select the node that is not itself.

There are multiple links between one node and one corresponding group. If one of the links is destroyed, protocol will not reconstruct it immediately. This work will be dispersed in a long time. The burstiness of transmission is abased.

### 143.3.3 Resource Information Storing Rule

Resource information storing rule defines the strategies on how to store, transfer, and renew the information. In the present protocol, the resource information cannot be stored in local node, unless the local node regards code of this resource as its code. When the amount of resource information exceeds the threshold, the resource information should be stored in local node. So the resource information of large amount is not transferred to far end and the information stored in local node will not be moved. The amount of information needed to be transferred under churn is reduced.

Information which doesn't exceed *Limen* will be transferred to other node group and is backed up by redundancy. The process of transferring is as follows:

**Step1:** Original node selects one node from the corresponding group randomly, and the information is transferred to it; go to **Step2**.

**Step2:** The node which stores information checks whether there are other nodes in the corresponding group; if so go to **Step3**, or else go to **Step5**.

**Step3:** The node selects a part of nodes which are in the same group to back up the information randomly; go to **Step4**.

**Step4:** The node in the same group with the last node checks whether the information is existing in local; if so go to **Step5** immediately; otherwise, store it and go to **Step5**.

**Step5:** End.

With the process, the information that is stored in far end will have multiple backups. When one node storing the information leaves, it is not necessary to transfer new information to other nodes. Further, the traffic load is reduced.

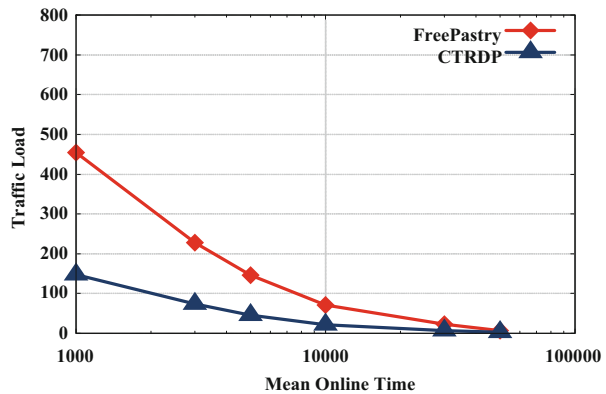
Resource distribution-based node encoding algorithm, redundancy table-based overlay, and resource information storing rule compose the node group-based resource information organization. In theory, the presented resource information organization can decrease the burst of transmission and reduce the load of information. The experiments will be given in Sect. 143.4.

## 143.4 Experiments

We verify the theoretical results by measuring the performance. The method is simulated by PeerSim [13] which is a simulator that can be extended easily. The simulator adopts the event-based simulation of PeerSim. The traffic load of the resource information organization is compared with FreePastry to prove that the burst and traffic load of transmission are both reduced under churn. Then, the simulator is presented to show that the route algorithm has good performance in route success rate with high efficiency and low cost as FreePastry.

**Table 143.2** Main parameters of simulator

Parameter	Value
Number of node	1,000
Number of kind of resource for each node	$U(1, 64)$
Number of one resource for each node	$U(10, 100)$
Amount of each information	10
Bandwidth of each node	$U(10, 500)$
Length of code	128

**Fig. 143.2** Traffic load under churn

The churn model of our experiment is exponential distribution [14]. The churn level is adjusted by changing average online time of nodes in simulator. The longer the time is, the lower the churn level is.

Some main parameters of simulator are shown in Table 143.2, where  $U$  is the uniform distribution on the unit interval of value in the bracket. There are 1,000 nodes in the simulator. When one node leaves, a new node must join immediately. The size of each information is 10, including resource information and route table information. The code of resource and node is 128-bit BigInteger of Java.

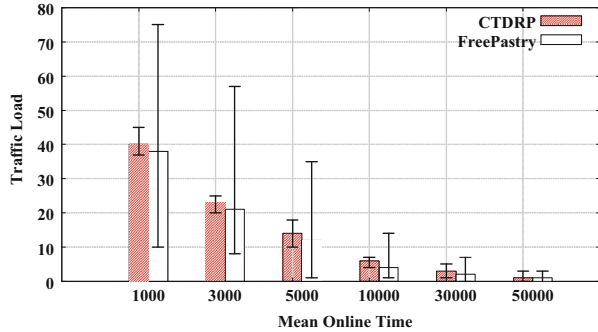
Because our method is extended on FreePastry, the parameter of FreePastry is as follows:  $b = 4$ ,  $M = 32$ ,  $L = 16$ .

The experiments of transmission give comparison between our method and FreePastry on burst and traffic load under churn.

*Traffic Load.* The values of mean online time are as follows: 1,000, 3,000, 5,000, 10,000, 30,000, and 50,000. The comparison between our method and FreePastry on traffic load under churn is shown in Fig. 143.2.

The result shows that the traffic load of our method is less than FreePastry under churn. The higher the churn level, the more obvious is the advantage. The reason of the phenomenon is that a great deal of resource information is stored in local nodes. These resource information need not to be moved with node's leaving or joining. Although there are multiple backups for some resource information that is stored

**Fig. 143.3** Traffic load of route information under churn



in far end, the amount of information doesn't increase dramatically. At the same time, the redundancy table doesn't make the traffic load increasing obviously. So the traffic load of our method is lower than FreePastry.

*Route Information.* This experiment compares our method and FreePastry on route information under churn. The same mean online times as the last one are adopted. The result can be found in Fig. 143.3.

In the figure, the bins represent the average traffic load of route information in 10,000 cycles for each node. The upper extreme points are the max value and lower extreme points are the min value in the period. The result shows that the average traffic load of our method is little more than FreePastry's. However, the stability of our method's traffic load is better than the old one's obviously. So we can get the conclusion that our method decreases the burst of transmission with a little more traffic load.

Summing up the above, node group-based resource information organization that our method adopts can reduce the traffic load and decrease the burst of transmission.

## 143.5 Conclusion

In this chapter, a resource information organization mode based on node encoding is provided to solve the churn problem. The encoding algorithm breaks the limitation of uniqueness and randomness of traditional ones. After encoding, a code may be mapped to more than one node. The redundancy table is designed on this character of encoding. A table entry can be mapped to multiple nodes in redundancy table. The overlay is constructed according to the redundancy table. The group intra-network is based on unstructured P2P and the group inter-network on structured P2P, FreePastry. The redundancy table-based overlay can abase the burstiness of transmission. The resource information storing rule is given, which makes a great deal of resource information stay in local node. So the traffic load of our method is lower than FreePastry.



**Acknowledgements** The authors are supported by the China Postdoctoral Science Foundation 2012T50844, by Provincial Nature Science Foundation of Jiangsu China BK2012512, and by the Advanced Research Foundation of PLA University of Science and Technology KYZYZLXY1205.

## References

1. Ranjan, R., Harwood, A., & Buyya, R. (2008). Peer-to-peer based resource discovery in global grids: A tutorial. *IEEE Communications Surveys and Tutorials.*, 10(2), 6–33.
2. Tian, J., & Dai, Y. F. (2007). Study on durable peer-to-peer storage techniques. *Journal of Software*, 18(6), 1379–1399.
3. Li, J. Y., Stribling, J., Morris, R., & Kaashoek, M. F. (2005). Bandwidth efficient management of DHT routing tables. In *Proceedings of second conference on symposium on networked systems design & implementation* (pp. 99–114), Berkeley, CA, USA.
4. Lam, S., & Liu, H. (2006). Failure recovery for structured P2P networks: Protocol design and performance under churn. *Computer Networks.*, 50(16), 3083–3104.
5. Stutzbach, D., & Rejaie, R. (2006). Improving lookup performance over a widely-deployed DHT. In *Proceedings of 25th IEEE international conference on computer communications* (pp. 1–12). Piscataway, NJ: IEEE Press.
6. Ou, Z. H., Harjula, E., Kassinen, O., & Ylianttila, M. (2010). Performance evaluation of a Kademia-based communication-oriented P2P system under churn. *Computer Networks.*, 54(5), 689–705.
7. Meng, X. F., Chen, X. L., & Ding, Y. L. (2013). Using the complementary nature of node joining and leaving to handle churn problem in P2P networks. *Computers & Electrical Engineering.*, 39(2), 326–337.
8. Alexandrova, T., Huzsak, G., & Morita, H. (2012). Churn resilience in network coding-based anonymous P2P system. In *Proceedings of 2012 international symposium on information theory and its Applications (ISITA)* (pp. 270–274). Piscataway, NJ: IEEE Press.
9. Melo, C. A. V., Vieira, D., & Liborio, J. M. (2012). Impact of churn on object management policies deployed in CDN-P2P systems. *IEEE (Revista IEEE America Latina) Latin America Transactions*, 10(3), 1811–1816.
10. Kwon, O. C., Song, H. J. (2012). A peer-churn resistant video multicast system over P2P networks. In *Proceedings of 2012 international symposium on communications and information technologies (ISCIT)* (pp. 1003–1008). Piscataway, NJ: IEEE Press.
11. Zhang, Y. X., Yang, D., & Zhang, H. K. (2009). Research on churn problem in P2P networks. *Journal of Software.*, 20(5), 1362–1376 (In Chinese).
12. FreePastry protocol v2.1. (2012). Retrieved from <http://www.freepastry.org/FreePastry>.
13. PeerSim. (2012). Retrieved from <http://wenku.baidu.com/view/>.
14. Ou, Z. H., Harjula, E., & Ylianttila, M. (2009). Effects of different churn models on the performance of structured peer-to-peer networks. In *Proceedings of IEEE 20th international symposium on personal, indoor and mobile radio communications* (pp. 2856–2860). Piscataway, NJ: IEEE Press.

# Chapter 144

## The Implementation of Electronic Product Code System Based on Internet of Things Applications for Trade Enterprises

Huiqun Zhao and Biao Shi

**Abstract** In order to solve the EPC codec problems based on Internet of Things (IOT) applications for trade enterprises, in this chapter an EPC codec system is designed for enterprise applications. According to the “Tag Data Standards,” we design the encoding and decoding algorithm/schema of SGTIN-96. On the basis of the algorithm/schema, the system has improved its coding and decoding algorithm, making the coding algorithm more simple and practical and improving the efficiency. Besides, it also has realized the transformation between SGTIN-96 and GTIN-14, which makes the final printed electronic tag contain both a bar code and the EPC code, and realizes the compatibility of bar code and EPC code. The coding and decoding system can code and decode well for the products of the trade enterprise. Through the system, SGTIN-96 labels can be generated and printed. And the content of SGTIN-96 labels can be decoded. Finally, we test the codec system to prove that it can achieve our established requirements.

### 144.1 Introduction

IOT is based on the Internet and uses the radio frequency identification (RFID), wireless data communications, and computer technology to construct an Internet that covers everything in the world [1, 2]. IOT EPC aims to build an open global network, where anything and its position in the logics chain can be identified [3]. IOT can apperceive the EPC code within the electronic tag on goods. When the perception of information is transmitted to the high-level, it will be identified, transferred and integrated by using the information processing software in the high-level. At last, the processing power of the Internet is taken advantage of to

---

H. Zhao • B. Shi (✉)

Department of Computer Science, North China University of Technology,  
Beijing 100041, China

e-mail: [zhaohq6625@sina.com](mailto:zhaohq6625@sina.com); [shibiao462@163.com](mailto:shibiao462@163.com)

integrate the perception of information, forming the effective management and control capability to meet user demand for a variety of applications.

To achieve IOT, EPC encoding within the electronic tag on goods is vitally crucial. Moreover, IOT will be widely used firstly in the trade logistics management. In this chapter, the EPC codec system based on IOT applications for trade enterprise has been designed and implemented.

The chapter is organized as follows: Sect. 144.2 introduces EAN.UCC standards of GID (General Identifier) and its derivatives GTIN and SGTIN [4, 5] and proposes an encoding and decoding scheme for implementing the codec system, and the converting algorithm for mapping barcode into EPC is discussed. In Sect. 144.3, in order to support our technique, a system is designed and implemented. In Sect. 144.4, we test the encoding and decoding system. Finally as a key technique we mainly introduce related work and conclude our research work in Sect. 144.5.

## 144.2 The Encoding and Decoding Schema

In this part, we design the proposals. It includes some algorithms, such as the transformation algorithm from SGTIN-96 to GTIN-14, SGTIN-96 encoding algorithm, and so on. The purpose is to provide a general method for constructing the EPC codec system based on IOT applications for trade enterprises.

### 144.2.1 Introduction of the GTIN and the SGTIN

The Serialized Global Trade Item Number (SGTIN) is a new identity type based on the EAN.UCC Global Trade Item Number (GTIN) code defined in the General EAN.UCC Specifications [6]. A GTIN by itself does not fit the definition of an EPC pure identity, because it does not uniquely identify a single physical object.

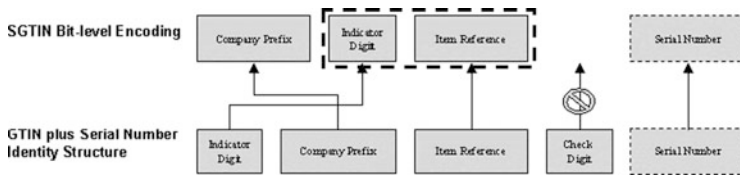
The SGTIN is an EPC encoding scheme that permits the direct embedding of EAN.UCC system standard GTIN and serial number codes on EPC tags [7]. There are two launched encoding schemes: SGTIN-64 (64 bits) and SGTIN-96 (96 bits) now.

For instance, in the SGTIN-96 encoding, the limited number of bits prohibits a literal embedding of the GTIN. As a partial solution, a Company Prefix *Index* is used. This Index is assigned to companies that need to use the 96-bit tags, in addition to their existing EAN.UCC Company Prefixes [8]. The Index is encoded on the tag instead of the Company Prefix and is subsequently translated to the Company Prefix at low levels of the EPC system components (i.e., the Reader). While this means that only a limited number of Company Prefixes can be represented in the 96-bit tag, this is a transitional step to full accommodation in 96-bit and additional encoding schemes (Table 144.1).

The relationship between GTIN and SGTIN is shown in Fig. 144.1.

**Table 144.1** The structure of SGTIN-96

Header	Filter value	Partition	Company Prefix <i>Index</i>	Item reference	Serial number
8	3	3	20–40	24–4	38



**Fig. 144.1** The encoding of GTIN and SGTIN

The SGTIN-96 consists of the following information elements:

- The *Company Prefix*, assigned by GS1 to a managing entity. The Company Prefix is the same as the Company Prefix digits within an EAN.UCC GTIN decimal code.
- The *Item Reference*, assigned by the managing entity to a particular object class. The Item Reference for the purposes of EPC Tag Encoding is derived from the GTIN by concatenating the Indicator Digit of the GTIN and the Item Reference digits and treating the result as a single integer.
- The *Serial Number*, assigned by the managing entity to an individual object. The serial number is not part of the GTIN code, but is formally a part of the SGTIN.

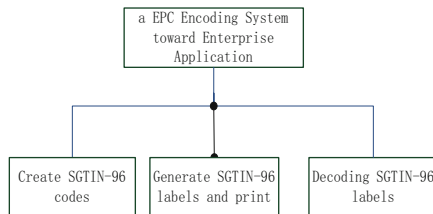
### 144.2.2 Transformation of GTIN-14 and SGTIN-96

In Fig. 144.2, the process of the transformation from GTIN-14 to SGTIN-96 is demonstrated, and the converse process is the transformation from SGTIN-96 to GTIN-14. The following procedure creates an SGTIN-96 encoding from GTIN-14:

Given:

1. A GS1 GTIN-14 consisting of digits  $d_1d_2 \dots d_{14}$
2. The Length  $L$  of the Company Prefix portion of the GTIN
3. A Serial Number  $S$  where  $0 \leq S < 2^{38}$ , or a GS1-128 Application Identifier 21 consisting of character  $S_1S_2 \dots S_K$
4. A Filter Value  $F$  where  $0 \leq F < 8$

**Fig. 144.2** The structure of the system



Yields:

5. An SGTIN-96 as a 96-bit string

Procedure:

Look up the Length  $L$  of the Company Prefix in the “Company Prefix Digits” column of the Partition Table;

If (the Length  $L$  exist){

determine the Partition Value,  $P$ ;

the number of bits  $M$  in the Company Prefix field;

the number of bits  $N$  in the Item Reference and Indicator Digit field.;

Construct the Company Prefix by concatenating digits  $d_2d_3 \dots d_{(L+1)}$  and considering the result to be a decimal integer,  $C$ ;

Construct the Indicator Digit and Item Reference by concatenating digits  $d_1d_{(L+2)}d_{(L+3)} \dots d_{13}$  and considering the result to be a decimal integer,  $I$ ;

If (Serial Number is provided & an integer  $S$  where  $0 \leq S < 2^{38}$ ){

Construct the final encoding =

Header 00110000 (8 bits)+Filter Value  $F$  (3 bits)+Partition Value  $P$  (3 bits)+Company Prefix  $C$  ( $M$  bits)+Item Reference ( $N$  bits)+Serial Number  $S$  (38 bits);

}else if (the Serial Number is provided as a GS1-128,  $s_1s_2 \dots s_K$ ){

If ( $s_1s_2 \dots s_K$  not a digit)

Return;

If ( $K > 1$  and  $s_1 = 0$ )

Return;

consider the result to be a decimal integer,  $S$ ;

If ( $S \geq 2^{38}$ ){

this Serial Number cannot be encoded in the SGTIN-96 encoding;

Return;

}else{

Construct the final encoding =

Header 00110000 (8 bits)+Filter Value  $F$  (3 bits)+Partition Value  $P$  (3 bits)+Company Prefix  $C$  ( $M$  bits)+Item Reference ( $N$  bits)+Serial Number  $S$  (38 bits);

}

```

    }
}else{
    this GTIN cannot be encoded in an SGTIN-96;
    return;
}

```

### ***144.2.3 Algorithm of SGTIN-96 Encoding and Decoding***

The SGTIN-96 encoding algorithm is to form an EPC code, according to the given and known information, such as the filter value, manufacturer identification code, packing instructions, and commodity types of product information. It includes obtaining the bar code by SGTIN-96 and GTIN-14. The following procedure creates an SGTIN-96 encoding:

Given:

6. Filter value
7. Company Prefix
8. Indicator Digit
9. All kinds of information of products
10. Information and the number to generate code for product

Yields:

All kinds of products and the specified number of SGTIN-96 code and GTIN-14 code

Procedure:

1. According to the company prefix, we look up the partition table to find out the partition value and obtain figures of item reference. If there are no corresponding values in the table, the digit company prefix is not valid and stops with a prompt error message.
2. According to the given filter value, we query the SGTIN-96 filter table to see if the selected filter value is in practical sense, i.e., not retention value. If it is reserved, the filter value is not valid and stops with a prompt error message.
3. Also we can query that the indicator digit is legitimate. If not legal, we prompt the error message and stop.
4. According to the number of commodity product information, the item reference field is assigned to the different parts, which is the kind of distribution of several fields. It means that the field assigned preserves the different information of commodities.
5. According to the number of kinds of goods code given, we encode for the commodities. If the number is not lengthy enough, we add leading zeros to achieve a specified number of bits. Then the code is stored in the database.

6. According to the generated code types, the number of product information, and the SGTIN-96 tag field format, we splice each field to generate EPC tags. According to the EPC label and the formation of GTIN-14 bar code, ITF-14 is structured (the ITF-14 is a kind of bar code of GTIN-14).
7. By means of the replacement encryption on the formation of the SGTIN-96 tag, we form 24 bits 16 hexadecimal code.
8. The labels are printed by a special printer, forming effective labels that contain IOT EPC and bar code.

### **144.3 Design and Implement**

We design an EPC codec system according to the analysis of the proposal. It is mainly to form the proposal of SGTIN-96 encoding, based on the specific application coding information provided by company. At the same time, it can transform SGTIN-96 encoding to GTIN-14 encoding. After the label is formed, we can use the label printer to print. We use BarTender to complete the mission. Then, we send the label data generated by the system, including SGTIN-96 and ITF, to the driver of the BarTender, and the driver drives the printer to print the label. Besides, the system also realizes the decoding of specified label, that is, obtains the information contained in the code by the SGTIN-96. The structure diagram of the system is shown in Fig. 144.2.

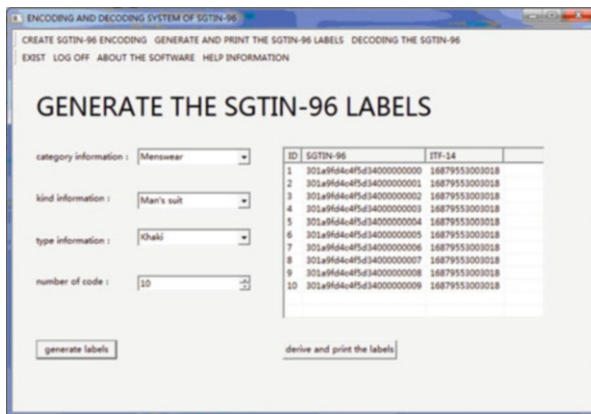
#### ***144.3.1 Encoding Module of IOT Labels***

Encoding module of the EPC label is to provide the unique assigned company prefix, the trade type of the product, the indicator value of the project reference field, and the information of the product to the system, according to the distribution of the SGTIN-96 field. The system will form the SGTIN-96 label based on this basic information and store the obtained data in the database.

#### ***144.3.2 Printing Module of IOT Labels***

The printing module of IOT label, based on tag encoding, will need to print out the tags. When in printing, all kinds of information need to choose to print labels, as well as the number of print labels, in order to produce continuous tags and labels. Figure 144.3 shows the printing module interface of the system.

**Fig. 144.3** The generating label interface of system



### 144.3.3 Decoding Module of IOT Labels

For the decoding module of IOT label, according to some parameters of each field in an SGTIN-96 distribution and encoding algorithm, we can respectively get the fields of the IOT label from SGTIN-96 such as the company prefix, item reference, indicator, and kinds of products field. Finally, we analyze the practical significance which the fields represent.

## 144.4 Testing the System

The encoding module can generate the same encoding results with the respected by inputting all kinds of the related testing data in the encoding module of the system. It also can export and print EPC tag efficiently. The decoding module can analyze the input label well. The system can complete the function well, put out the correct result, and have good operability and better fault tolerance. Besides, the system has good expansibility which is good for the updating and maintenance of the system.

Taking clothing industry as an example, we input the information needed about the clothing and formulate the information about the labels in the database. The printing module will form tags that contain both SGTIN-96 and ITF. Enter a valid label information in the decoding module; the information about the product will be resolved within the label. Finally, Fig. 144.4 depicts the printed label, which contains both SGTIN-96 in the inner chip and GTIN-14 on the surface of the label.

Given that the encoding and decoding system is the basis of the IOT and will be planted into other system, this system has good portability. However, because of being developed in the Windows 7 and uses SWT/JFace package, the system has a higher operational efficiency in the Windows 7 system than in the Linux system. According to the principle of the EPC encoding, EPC tag should have the



**Fig. 144.4** The printing labels

\* \* \* incorporated company 1



confidentiality and security, but it is not considered much in the system. The system just does a simple substitution encryption to the generated EPC encoding, so the security of encoding needs to be strengthened.

The improved proposal: Encoding and decoding of the EPC label should be published as a service, so the system will have better cross-language, cross-platform, expansibility, support, and so on. Besides, the system will have higher confidentiality and security by combining the security and encryption technology.

## 144.5 Related Works and Conclusion

EPC and its application have become a popular research topic in the last decade, so some implementations of this EPC system have been applied in the business world. GS1, a management and development body for EPC standards, has been developing universal standards to regulate its encode and decode. However, some changes should be made to satisfy the specific requirements for trade enterprises. Thus we put forward our own algorithms/schemes for the codec system based on the customary EPC system. Because of the improvement of coding and decoding algorithm, the system can be used easily and the efficiency has been improved. Combined with practical work, an EPC codec system is designed. Our system can provide a unique physical marking for retail or trade items industry. The efficient implementation of the design and printing label provides the basis for the popularization of IOT.

## References

1. Staake, T., Thiesse, F., & Fleisch, E. (2005). Extending the epc network—the potential of RFID in anti-counterfeiting. *Symposium on Applied Computing—SAC* (pp. 1607–1612).
2. Price, J., Jones, E., Kapustein, H., Pappu, R., Pinson, D., Swan, R., & Traub, K. (2003). *Auto-ID reader protocol 1.0* (pp. 29–33).
3. Hoag, J. E., & Thompson, C. W. (2006). Architecting RFID middleware. *IEEE Internet Computing*, 10(5), 88–92.
4. Harrison, M. (2004). EPC information service (EPCIS). *Auto-ID Labs Research Workshop* (pp. 29–30).

5. Ding, Z. H., Li, J. T., Zheng, W. M. & Feng, B. (2007). A filter design of RFID middleware in the progress of updating barcode to RFID. *1st Annual RFID Eurasia*, 18(2), 56–66 (In Chinese).
6. Rivest, R. L., Shamir, A., & Adleman, L. (1978). A method for obtaining digital signatures and public-key cryptosystems. *Communications of the ACM*, 21(2), 120–126.
7. Juels, A., Rivest, R. L., & Szydlo, M. (2003). The blocker tag: selective blocking of RFID tags for consumer privacy. *Proceedings of 10th ACM Conference on Computer and Communications Security, CCS 2003* (pp. 103–111).
8. Ham, Y. H., & Kim, N. S. (2005). A study on establishment of secure RFID network using DNS security extension. *2005 Asia-Pacific Conference on Communications, Perth, Western Australia* (pp. 3–5).

# Chapter 145

## The Characteristic and Verification of Length of Vertex-Degree Sequence in Scale-Free Network

Yanxia Liu, Wenjun Xiao, and Jianqing Xi

**Abstract** Many natural large-sized complex networks exhibit a scale-free, power-law distribution of vertex degree. To better understand the formation mechanism of power law in the real network, we analyze the general nature in scale-free network based on the vertex-degree sequence. We show that when the power exponent of scale-free network is greater than 1, the number of degree- $k_1$  vertices, when nonzero, is divisible by the least common multiple of 1,  $k_2^\gamma/k_1^\gamma, \dots, k_l^\gamma/k_1^\gamma$ , and the length of vertex-degree sequence  $l$  is of order  $\log N$ , where  $1 \leq k_1 < k_2 < \dots < k_l$  is the vertex-degree sequence of the network and  $N$  is the size of the network. We verify the conclusion by the coauthorship network DBLP and many other real networks in diverse domains.

### 145.1 Introduction

Complex networks are widespread in nature and human society, such as the Internet, the World Wide Web, protein networks, scientific collaboration networks, and transportation networks. As a key technology of depicting and studying the topology and behaviors of complex systems, complex network has become a hot spot of the multidisciplinary research of common concern in recent years.

---

Y. Liu (✉)

School of Computer Science and Engineering, South China University of Technology, Guangzhou 510006, China

School of Software Engineering, South China University of Technology,

Guangzhou 510006, China

e-mail: [cslyx@scut.edu.cn](mailto:cslyx@scut.edu.cn)

W. Xiao • J. Xi

School of Software Engineering, South China University of Technology, Guangzhou 510006, China

**Table 145.1** Scale-free networks with  $\gamma < 2$  ( $N, M, \gamma$ )

Network	$N$	$M$	$\gamma$
E-mails [7]	56,969	84,314	1.81
Gnutella [8]	1,026	3,752	1.4
Word Web [9]	478,773	$1.8 \times 10^7$	1.5
Software package [4]	1,439	1,723	1.6/1.4
Coauthorship in HEP [10]	56,627	9,796,471	1.2

The basic theory of complex network research is small-world network model [1] and scale-free network model [2]. In 1998, Watts and Strogatz proposed WS small-world network model by randomly reconnecting a small number of the edge of regular network [1]. As the transition from completely regular network to random network, the small-world model features localized clusters connected by sparse long-range edges, leading to a short average distance between vertices that grow logarithmically with the network size. Scale-free networks, on the other hand, show heterogeneous vertex connectivity, in which a fraction of the vertices is highly connected. In 1999, Barabási and Albert proposed a scale-free network model BA [2]. It demonstrated that the scale-free power-law distribution of vertex degrees in many large-sized networks is a direct consequence of two generic mechanisms that govern the network formation: (1) network expansion over time through addition of new vertices and (2) preferential attachment of new vertex to those existing ones that are already highly connected.

In this chapter, we focus on scale-free networks. Many real networks belong to scale-free networks which obey an approximate power-law distribution, i.e.,  $P(k) \propto k^{-\gamma}$ . Usually the scale-free networks need to satisfy power exponent  $\gamma \geq 1$ ; the extra  $\gamma \geq 1$  requirement ensures that  $P(k)$  can be normalized. Furthermore, a large number of empirical research have shown that the exponent  $\gamma$  in real network is between 2 and 3. Consequently many of the scale-free network research are on the premise of  $\gamma \geq 2$  by default [3–5], thus  $\gamma \geq 2$  have been a precondition to study the general nature and dynamic behaviors of scale-free networks. In other words,  $1 < \gamma < 2$  network has been ignored and few research pay attention to such networks. In fact, such networks are also widely used in the real world, such as Gnutella P2P network and HEP collaborators network. Table 145.1 lists the parameters for several real scale-free networks with  $\gamma < 2$ .

It's very meaningful to study the similarity and difference between such network and network with  $\gamma \geq 2$ , which can help us better understand the laws of the generation of different power law in the real network.

In the previous works [5, 6], we have presented necessary conditions for scale freedom in complex networks, which is based on the assumption  $\gamma \geq 2$ . Here, we extend the research results to the case  $\gamma < 2$  and pay our attention to the general characteristic in scale-free networks with the exponent  $\gamma > 1$ . We proved that the length of vertex-degree sequence  $l$  of scale-free networks with the exponent  $\gamma > 1$  is of order  $\log N$ , where  $N$  is the size of the network. We further verify the conclusion and some applications are given.

## 145.2 Scale-Free Networks

Complex networks can be abstracted as undirected or directed graph  $G(V, E)$ , where  $V$  is the set of vertices or nodes and  $E$  is the set of edges or links. The parameters that are of interest in this chapter appear in the following list:

$M$  Number of edges;  $M = |E|$

$N$  Number of vertices;  $N = |V|$

$d(v)$  Degree of the vertex;  $v \in V$

$\bar{d}$  Average vertex degree of the network;  $\bar{d} = 2M/N = \sum_{v \in V} d(v)/N$

$n_k$  Number of degree- $k$  vertices;  $n_k = \{v | d(v) = k\}$

$P(k)$  Degree distribution or fraction of vertices that are of degree  $k$ ,  $P(k) = n_k/N$

$l$  Length of vertex-degree sequence;  $\{k_1, k_2, \dots, k_l\}$  is the vertex-degree sequence of the network where  $1 \leq k_1 < k_2 < \dots < k_l$

Most real complex networks are scale-free network. The degree distribution of scale-free networks is not like random network which is presented in the form of the Poisson distribution. Scale-free networks comparatively have the nature of the power-law distribution, which is presented in the following form:

$$P(k) = ck^{-\gamma}, \quad \gamma > 1$$

where  $\gamma$  is power exponent, also known as degree distribution exponent or scaling exponent. The extra  $\gamma > 1$  requirement ensures that  $P(k)$  can be normalized. Power laws with  $\gamma < 1$  rarely occur in nature. The term  $c$  is normalizing constant, defined as  $c = (\sum_{k \in K} k^{-\gamma})^{-1}$ , and  $K$  is the set of all node degrees occurring in the network.

Scale-free network aroused great interest of researchers. Despite extensive research on scale-free network within multiple scientific disciplines over the past decade, there are still gaping holes in our understanding of such networks. New results that shed light on the static structure and dynamic properties of different classes of large-scale networks are needed to facilitate further progress. In the previous works [5, 6], we have exposed some characteristics of scale-free network with the exponent  $\gamma \geq 2$ . We proved that when the vertex degree of a large-sized network follows a scale-free power-law distribution with exponent  $\gamma \geq 2$ , the number of degree-1 vertices, when nonzero, tends to be of the same order as the network size  $N$  and that the average degree is of order lower than  $\log N$ . Our method provides an analytical tool that helps to answer the question of whether a network is scale-free because it relies on conditions that are static and easily verified for any network. Furthermore, we showed that the number of degree-1 vertices is divisible by the least common multiple  $k_1', k_2' \dots k_l'$ , where  $k_1 < k_2 < \dots < k_l$  is the vertex-degree sequence of the network, which leads a remodeling method to equip a scale-free network with small-world features.

Next, based on our previous research [5, 6], we will further investigate the general characteristic of scale-free networks by extending the exponent condition. We mainly focus on the characteristic of vertex-degree sequence in scale-free networks with the exponent  $\gamma > 1$ .

### 145.3 New Characteristic of Scale-Free Network and Its Derivation

In this section, we present the new characteristic of scale-free network and its mathematical derivation. Supposing that the vertex-degree sequence of network can be represented as  $1 \leq k_1 < k_2 \dots < k_l$ , according to the definition of strict scale-free networks, we have

$$P(k_i) = \frac{n_{k_i}}{N} = ck_i^{-\gamma} \tag{145.1}$$

Here,  $c$  is a constant. When  $i = 1$ , we have

$$P(k_1) = \frac{n_{k_1}}{N} = ck_1^{-\gamma} \tag{145.2}$$

Thus, we can get the value of  $c$

$$c = \frac{n_{k_1}k_1^\gamma}{N} \tag{145.3}$$

By substituting Eq. 145.3 into Eq. 145.1, we have

$$n_{k_i} = n_{k_1} \left(\frac{k_1}{k_i}\right)^\gamma \tag{145.4}$$

Recall that  $n_k$  denotes the number of vertices of degree  $k$ , apparently, we have

$$N = \sum_i n_{k_i} \tag{145.5}$$

The following identities follow from Eq. 145.4 and Eq. 145.5

$$N = \sum_i n_{k_i} = n_{k_1}k_1^\gamma \sum_i \frac{1}{k_i^\gamma} \tag{145.6}$$

Therefore, assuming  $\gamma > 1$ , according to Riemann- $\zeta$  function, we know that  $\sum_i \frac{1}{k_i^\gamma}$  in Eq. 145.6 is convergent, which satisfied

$$\sum_i \frac{1}{k_i^\gamma} \leq \sum_{k=1}^\infty \frac{1}{k^\gamma} < +\infty \tag{145.7}$$

Supposing that  $\sum_i \frac{1}{k_i^\gamma}$  converge to some constant  $a$ , i.e.,  $a = \sum_i \frac{1}{k_i^\gamma}$ , clearly we have the following conclusion:

$$N = n_{k_1} k_1^\gamma a \tag{145.8}$$

Thus, we have shown that for scale-free networks with power-law exponent  $\gamma > 1$ , the size of scale-free network is related to some constant  $a$ . Furthermore, if  $k_1=1$ , then  $N = n_{k_1} a$ . It means that when  $\gamma$  gets smaller, constant  $a$  gets greater, thereby the ratio of the number of degree-1 vertices and the size of network gets smaller. Note that  $k_1=1$  is not the necessary condition for the scale-free networks.

Following from the above analysis, we can obtain interesting results on the degree sequence of scale-free networks. In the preceding derivation, we have shown that Eq. 145.4 holds which leads to  $n_{k_1}/n_{k_i} = k_i^\gamma/k_1^\gamma$ . We may assume that  $\gamma$  is a rational number since any real number can be approximated infinitely by rational number sequence. Taking this equality to be exact and noting that the left-hand side is a rational number, we can easily prove that either  $k_i^\gamma/k_1^\gamma$  is an integer with  $(k_i^\gamma/k_1^\gamma) | n_{k_1}$  or  $k_i^\gamma$  is an integer with  $k_i^\gamma | n_{k_1} k_1^\gamma$ . Thus, recall that  $1 \leq k_1 < k_2 \dots < k_l$  is the degree sequence of the network, we can obtain that either  $n_{k_1}$  must be divisible by the least common multiple of  $1, k_2^\gamma/k_1^\gamma, \dots, k_l^\gamma/k_1^\gamma$ , denoted as  $n_{k_1} = c [1, k_2^\gamma/k_1^\gamma, \dots, k_l^\gamma/k_1^\gamma]$ , or  $n_{k_1} k_1^\gamma = c [k_1^\gamma, k_2^\gamma, \dots, k_l^\gamma] \neq 0$  holds for some constant  $c$ . Using the method as above, in the case of  $\gamma > 1$ , we have the following estimate:

$$n_{k_1} k_1^\gamma \geq 2 \times 2 \times \dots \times 2 \geq 2^{(l-2)} \tag{145.9}$$

Then taking all logarithms to be in base 2, Eq. 145.9 yields

$$O(\log N) > \log n_{k_1} k_1^\gamma \geq l - 2 \tag{145.10}$$

Therefore, we obtain our result that for scale-free network with  $\gamma > 1$ , the length of degree sequence is of order  $\log N$ , i.e.,  $l \leq O(\log N)$ .

This conclusion is very meaningful, which means the characteristic of the length of degree sequence is the general nature in scale-free network. All scale-free networks have a very small degree sequence compared with the size of the network. Using the conclusion, we can reconstruct the scale-free network presenting apparent small-world feature and further improve the current maximal degree search algorithm.

In addition, we must also stress that the above conclusion holds based on the premise that the network obeys strict power-law distribution. Actually many real

scale-free networks are not exact which only have approximate scale-free characteristic, so there are subtle differences between the real network and the theoretical estimate. We further pointed out that for many real networks the length of degree sequence is also of order  $\log^b N$  at the very most, namely,  $l \leq O(\log^b N)$ , where  $b$  is a very small constant, which means  $l$  is very small compared with the size of network. Next, we will verify our conclusion in some applications.

## 145.4 Analysis and Verification in DBLP and Many Other Real Networks

We start our analysis by investigating coauthorship network of scientists using bibliographic data drawn from DBLP. DBLP is a well-known computer science bibliography website hosted at [Universität Trier](#) in [Germany](#) which provides bibliographic information on major computer science journals and proceedings. Until now DBLP has indexed more than 2.1 million articles and DBLP data is available from the website which is stored as xml records.

For our research, we downloaded the latest `dblp.xml` file and parsed it using SAX parser in Java. Then we constructed the network of scientists in which a link between two scientists is established by their coauthorship of one or more scientific papers. Due to the large amount of data, the coauthorship network has to be constructed by year. We produced several separate networks by respectively extracting the articles generated in 2009, 2010, and 2012. Among them, the data in year 2012 only contains the articles generated before the March. Subsequently, we computed the parameters of each network which are listed in the following table.

From Table 145.2, we can see that the three networks have consistent statistical values which reflect the general characteristic in the whole DBLP data to a certain extent. First, the ratio of degree-1 vertex and the size of network approximately equal to 0.1, which means that about 10 % of scientists have only one coauthor when they wrote papers during 1-year period. Second, the average degree of the networks is roughly 5, which means that, on average, scientists wrote articles with five other people each year. Third, most importantly, we can see that  $l \leq \log^2 N$  holds which means that compared with the size of network, the range of the possible numbers of collaborators per author is extremely small. Here, the reason why the length of degree sequence is not of order  $\log N$  is that DBLP network doesn't strictly obey power-law distribution.

**Table 145.2** The parameters in DBLP networks

Network	$N$	$M$	$n_1$	$\bar{d}$	$l$	$\log N$	$b$
DBLP_2009	57,623	160,456	5,266	5.569	83	15.814	2
DBLP_2010	74,899	197,849	7,048	5.283	88	16.193	2
DBLP_2012	17,846	49,443	1,820	5.541	56	14.123	2

Notes: For DBLP data, see <http://dblp.uni-trier.de/>



**Table 145.3** The length of degree sequence in some real networks

Network	Type	$N$	$M$	$\bar{d}$	$l$	$\log N$	$b$
TG city	TP	18,263	23,797	2.606	7	14.157	1
OL city	TP	6,105	7,029	2.303	5	12.576	1
US Air	TP	332	2,126	12.807	58	8.375	2
Linux	SW	5,285	11,352	4.296	51	12.368	2
Mysql2	SW	1,480	4,190	5.662	43	10.531	2
Helico	Bio	710	1,396	3.932	31	9.472	2
Elegans	Bio	314	363	2.312	17	8.295	2
Ncstrlwg2	SC	6,396	15,872	4.963	42	12.643	2

Notes: For data on transportation networks, see <http://www.cs.fsu.edu/~lifeifei/SpatialDataset.htm>; for software networks, see [www.tc.cornell.edu/~myers/Data/SoftwareGraphs/index.html](http://www.tc.cornell.edu/~myers/Data/SoftwareGraphs/index.html); for biological networks, see [www.cosin.org/extra/data](http://www.cosin.org/extra/data); the data of Ncstrlwg2 is provided by M.E.J Newman [11]

Besides DBLP network which belongs to scientific collaboration networks, we have found that various types of real network without exception exhibit the same characteristic, i.e.,  $l \leq O(\log N)$  or  $l \leq O(\log {}^bN)$ , where  $b$  is a very small constant. So far, we have done a lot of data verification and observed that the length of the degree sequence in the real network is less than the  $\log^2N$ , i.e.,  $b$  value is 2. Table 145.3 shows some real examples from the scale-free networks in diverse domains, including scientific collaboration networks, transportation networks, software packages networks, and biological networks.

### 145.5 Conclusion

In this chapter we show that when the vertex degree of a large-sized network follows a scale-free power-law distribution with exponent  $\gamma > 1$ , the number of degree- $k_1$  vertices is divisible by the least common multiple of  $1, k_2^\gamma/k_1^\gamma, \dots, k_i^\gamma/k_1^\gamma$  and the length of vertex-degree sequence  $l$  is of order  $\log N$ , where  $k_1 < k_2 < \dots < k_l$  is the vertex-degree sequence of the network. Furthermore, we generalized the conclusion taking into account that many real networks are approximately scale-free. We pointed out that for many real networks the length of degree sequence is also of order  $\log {}^bN$  at the very most, where  $b$  is a very small constant. We verify the conclusion by the coauthorship network DBLP and other real networks in different domains. Next, we will further research on searching problem based on the vertex-degree sequence in scale-free networks.

## References

1. Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684), 440–442.
2. Barabási, A. L., & Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439), 509–512.
3. Albert, R., & Barabási, A. L. (2006). Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1), 47–91.
4. Newman, M. E. J. (2003). The structure and function of complex networks. *SIAM Review*, 45(2), 167–256.
5. Xiao, W. J., Chen, W. D., & Parhami, B. (2011). On necessary conditions for scale-freeness in complex networks with applications to computer communication systems. *International Journal of Systems Science*, 42(6), 951–995.
6. Xiao, W. J., Jiang, S. Z., & Chen, G. R. (2011). A small-world model of scale-free networks: Features and verifications. *Applied Mechanics and Materials*, 50–51(2011), 166–170.
7. Ebel, H., Mielsch, L. I., & Bornholdt, S. (2002). Scale-free topology of e-mail networks. *Physical Review E*, 66(3), 1–4.
8. Jovanovic, M.A., Annexstein, F.S., & Berman, K.A. (2001). *Scalability issues in large peer-to-peer networks—A case study of gnutella* (Technical Report). Cincinnati: University of Cincinnati.
9. Cancho, R. F. I., & Solé, R. V. (2001). The small world of human language. *Proceedings of the Royal Society of London B*, 268(1482), 2261–2265.
10. Newman, M. E. J. (2001). Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Physical Review E*, 64(1), 016132.
11. Newman, M. E. J. (2001). Scientific collaboration networks. I. Network construction and fundamental results. *Physical Review E*, 64(1), 016131.

# Chapter 146

## A Preemptive Model for Asynchronous Persistent Carrier Sense Multiple Access

Lin Gao and Zhijun Wu

**Abstract** In order to analyze the problem of packet collision in the asynchronous mode of persistent carrier sense multiple access (p-CSMA), in which there is no time slot different from synchronous mode and propagation delay have a heavy effect on the probability of packet collision, a preemptive asynchronous p-CSMA probability model is established for the first time in the chapter. From sub-cycle conditional probability, the model gives the expectations of an idle and busy period. On the basis, performance targets, e.g., throughput/delay/success rate and channel efficiency, are gotten. For illustration, VDL2 (a typical asynchronous p-CSMA network) simulation model is set up on OPNET platform and experiments are also carried out to verify the correctness of this model in diverse scenarios. Through simulation, the results of fixed position distribution have the good consistency with the preemptive probability model. Finally, the conclusion is achieved that packet collisions will aggravate with the stations distribution becoming more uneven.

---

L. Gao (✉)

School of Electrical and Information, Tianjin University, Tianjin 300072, China

School of Science and Technology, Tianjin Economic and Financial University,  
Tianjin 300000, China

e-mail: [gavingao71@sohu.com](mailto:gavingao71@sohu.com)

Z. Wu

School of Electronics & Information Engineering, Civil Aviation University of China,  
Tianjin 300000, China

e-mail: [zjwu@cauc.edu.cn](mailto:zjwu@cauc.edu.cn)

## 146.1 Introduction

### 146.1.1 *p*-CSMA Protocol

As a widely used channel access method, *p*-CSMA is evolved by ALOHA/CSMA. In the early random media access modes, such as ALOHA, each user is independent to send packet. The conflict probability increases exponentially with the payload increasing, which leads to low channel efficiency and throughput. In contrast to ALOHA, CSMA can sense the channel and judge if it is busy or not before transmitting packet. By this means, the user only has the opportunity to transmit when a channel is idle. As a result, the system performance is improved obviously.

With the function of probability decision, the basic principles of *p*-CSMA are as follows:

1. In the case of sensing that a channel is busy, the transmitting station keeps monitoring until it senses the channel to be free. Then, it will send packet with probability  $p$  or will not send with the probability  $1 - p$ .
2. In the state of channel being free, if there is at least one packet waiting to send and the packet is not sent out due to out of coverage of  $p$ -value, the station will not sense for an instant and will re-sense the channel after  $2\tau$  (in the chapter,  $\tau$  represents the delay for radio propagation between stations). Then, it will continue to monitor, repeating the principle 1.

Furthermore, *p*-CSMA can be divided into synchronous mode and asynchronous mode according to whether or not time slot is applied. In synchronous mode, a channel is divided into time slots (all the nodes are synchronized by master clock) and a message can only be sent at the beginning moment of a slot. In contrast, asynchronous mode has no function both of time slot and of synchronization.

### 146.1.2 *Related Research*

The theoretical research into CSMA series has a long history. Early in the middle of the 1970s, Kleinrock, one of the Internet founders, published his landmark papers on packet switching by radio. As the theoretical basis, throughput–delay characteristics are given for ALOHA, CSMA, and *p*-CSMA and the large advantage which CSMA provides is shown as compared to the random ALOHA [1]. Furthermore, Kleinrock successfully discussed the problem of system stability [2]. The classic literatures supply us the basic principle and analytic methods for CSMA, e.g., the definition of system stability and the data stream modeling by means of Poisson distribution, which are applied in the chapter as the precondition of analysis.

In the premise of saturation, Bianchi model is established at the beginning of this century. The model is based on the two-dimensional Markov chain in analyzing the

DCF (CDMA/CA) performance [3]. In some sense, the model can be regarded as a specializing and deepening one of the Kleinrock model.

In recent years, Richard MacKenzie expounds how the throughput and delay for the traffic flows are affected by the relative  $p$ -persistence [4]. Taking the capture effect into account, Salim Abukharis discussed the analysis method of throughput and packet delay in the Rayleigh fading channel [5]. In Yayu Gao's paper, a view is pointed out: with the small transmitting probability, throughput has nothing to do with retransmission strategy, but it is only relevant to propagation delay [6]. Deqing Wang et al. discussed the long-delay hidden terminal problem which is decided by propagation delays between finite nodes [7]. Recently, the advantage of asynchronous  $p$ -CSMA is paid more attention: it saves the time of waiting in time slot and facilitates the realization. As a result, the access mode is applied in distributed network, especially in the next generation of aeronautical data link-VDL 2 [8].

## 146.2 Theory and Methodology

Different from others, this chapter establishes an asynchronous preemptive  $p$ -CSMA model, with full consideration of the propagation delay.

### 146.2.1 Basic Conception

For the preemptive multiplexing, the probability of collision plays a vital role: in the ideal channel, no collision means that a packet transmits successfully in a cycle. Otherwise, all of the transmissions fail in the cycle. Although  $p$ -CSMA have the ability of channel detection and probability decision, packet collision is still inevitable. The reasons why packet collision is destined to happen are as follows:

Reason 1: When more than one station both have packets to send and have detected the channel to be free within the coverage of  $p$  decision, they will send messages at the same time, which will cause packets to collide.

Reason 2: Due to the existence of propagation between the stations, a sent packet should only be sensed by a station after the propagation delay. Therefore, in the period of propagation delay, the station will think the channel to be idle and may send message, which will lead to packet collision.

It is obvious that collision probability of reason 2 is far greater than that of reason 1. Therefore, in the following analysis, collision from reason 1 is neglected and only collision from reason 2 is taken into consideration.

### 146.2.2 Assumptions

To facilitate the discussions and statements, we give two assumptions as follows:

1. With the same as Kleinrock model [1], we assume that the new-generated packets comply with Poisson distribution and a new-generated packet will be rescheduled to send in a future time if it can't be sent immediately. If the interval to retransmit is long enough, according to Kleinrock's assumption, the two kinds of packets can combine into a Poisson flow with a new intensity. Here, the combined Poisson intensity of station  $i$  is marked as  $\lambda_i$ .
2. In the chapter, we assume that packet length ( $l$ ) complies with uniform distribution and we have  $l \sim U(l_a, l_b)$ , where  $l_a$  and  $l_b$ , respectively, represent the lower and upper bound of the distribution.

### 146.2.3 Modeling

In the chapter, idle period ( $T_0$ ) is defined as the duration from the moment of channel being idle to  $t_0$ . In contrast, busy period ( $T_1$ ) is defined as the duration from  $t_0$  to the moment that all the  $M$  senders have just finished sending message. An idle period is followed by a busy period, vice versa. For  $S_1$ , we have:

$$P(S_i = S_1) = p_i \sum_{L=0}^{N-1} P(L) \prod_{j \neq i}^{L-1} P \left[ \left( \frac{d_{0i}(0)}{C} + \tau_i \right) < \left( \frac{d_{0j}(0)}{C} + \tau_j \right) \right] \quad (146.1)$$

Supposing  $a_{ij} = \frac{d_{0j}(0) - d_{0i}(0)}{C}$ , the result is as shown in Eq. (146.2):

$$\begin{aligned} P \left[ \left( \frac{d_{0i}(0)}{C} + \tau_i \right) < \left( \frac{d_{0j}(0)}{C} + \tau_j \right) \right] &= \iint_{\tau_i - \tau_j < a_{ij}} \lambda_i e^{-\lambda_i \tau_i} \lambda_j e^{-\lambda_j \tau_j} d\tau_i d\tau_j \\ &= \frac{\lambda_i}{\lambda_i + \lambda_j} e^{a_{ij} \lambda_j} \end{aligned} \quad (146.2)$$

$$P(S_i = S_1) = p_i \sum_{L=0}^{N-1} \left( P(L) \prod_{j \neq i}^{L-1} \frac{\lambda_i}{\lambda_i + \lambda_j} e^{a_{ij} \lambda_j} \right) \quad (146.3)$$

Taking  $T_0 = t_0$  into account, we have:

$$E[t_0 | S_i = S_1] = E \left[ \frac{d_{0i}(0)}{C} + \tau_i \right] = \frac{d_{0i}(0)}{C} + E[\tau_i] = \frac{d_{0i}(0)}{C} + \frac{1}{\lambda_i} \quad (146.4)$$

$$\begin{aligned}
 E[T_0] &= \sum_N E[t_0|S_i = S_1]P(S_i = S_1) \\
 &= \sum_{i=0}^{N-1} p_i \left[ \frac{d_{0i}(0)}{C} + \frac{1}{\lambda_i} \right] \left[ \sum_{N-1} P(L) \prod_{j \neq i}^{L-1} \frac{\lambda_i}{\lambda_i + \lambda_j} e^{a_{ij}\lambda_j} \right] \tag{146.5}
 \end{aligned}$$

Similarly, if we suppose  $b_i = \frac{d_{1i}(t_0) - d_{0i}(0)}{C}$ , we have For  $S_2$ :

$$\begin{aligned}
 P(S_i = S_2) &= \sum_N P(S_i = S_2|L)P(L) \\
 &= p_i \left[ e^{-\lambda_i \left( t_0 - \frac{d_{0i}(0)}{C} \right)} - e^{-\lambda_i (b_i + t_0)} \right] \sum_{N-1} P(L) \prod_{j \neq i, j \neq 1}^{L-2} \frac{\lambda_i}{\lambda_i + \lambda_j} e^{a_{ij}\lambda_j} \tag{146.6}
 \end{aligned}$$

$$\begin{aligned}
 E[t_1] &= \sum_N E[t_1|S_i = S_2]P(S_i = S_2) = \sum_{N-1} p_i \left[ \frac{d_{0i}(0)}{C} + \frac{1}{\lambda_i} - E[T_0] \right] \\
 &\left[ e^{-\lambda_i \left( E[T_0] - \frac{d_{0i}(0)}{C} \right)} - e^{-\lambda_i (b_i + E[T_0])} \right] \left[ \sum_{N-2} P(L) \prod_{j \neq i, j \neq 1}^{L-2} \frac{\lambda_i}{\lambda_i + \lambda_j} e^{a_{ij}\lambda_j} \right] \tag{146.7}
 \end{aligned}$$

Due to  $E[t_{N-1}] \ll \dots \ll E[t_2] \ll E[t_1]$ , we have:

$$E[T_1] = E \left[ l + \sum_{i=0}^{N-1} t_i \right] \approx E[t_1] + E[l] \tag{146.8}$$

P-CSMA performance can be valued by the targets as follows:

$$P_s = \prod_{i \neq 1} [1 - P(S_i = S_2)] \tag{146.9}$$

$$\begin{aligned}
 \bar{S} = E[S] &= E \left[ \frac{P_s}{v} \lim_{n \rightarrow \infty} \frac{\sum_n l_i}{\sum_n T_i} \right] \approx \frac{P_s}{v} \lim_{n \rightarrow \infty} \frac{E \left[ \sum_n l_i \right]}{E \left[ \sum_n T_i \right]} \\
 &= \frac{P_s}{v} \frac{E[l]}{E[T_0] + E[T_1]} \tag{146.10}
 \end{aligned}$$

$$\bar{\tau}_s = E[T_0] + (1 - P_s)E[T_1] \tag{146.11}$$

$$\eta \approx \frac{E[T_1]}{E[T_0] + E[T_1]} \tag{146.12}$$

In Eq. (146.10),  $v$  represents channel capacity, which is the maximum bit rate of a communication system. With limited bandwidth, it will satisfy the Nyquist theory to guarantee no inter-symbol interference.

From Eqs. (146.5) to (146.8), we can calculate packet success rate Eq. (146.9), throughput expectation Eq. (146.10), packet delay expectation Eq. (146.11), and channel efficiency Eq. (146.12). Based on the targets, performance can be evaluated in a quantified way.

### 146.3 Simulation Based on VDL2

Asynchronous p-CSMA is typically applied in VDL2 air-ground communication which can provide reliable network services for future aviation network—ATN—and has been tested in Europe and identified as the main stream in the future system.

In VDL2 communication, there are two sides: ground station (GS) and aircraft (AC), by which a peer entity is composed. The communication can only be realized in half-duplex manner by peer entities. VDL2 applies hierarchical structure and conforms to ISO standard. From the bottom to the upper, there are three layers in VDL2 protocol stack [9]: physical layer, data link layer, and network layer. MAC, which is a sub-layer of data link layer, can realize Media access control and channel multiplexing with asynchronous p-CSMA. In order to simulate the performance of asynchronous p-CSMA, a model has been constructed and programmed for VDL2 protocol on OPNET 14.5 platform.

In the simulation, 16 ACs are equipped with scattered position around GS, which can closely meet the condition of the theory in the chapter, as shown in Fig. 146.1.

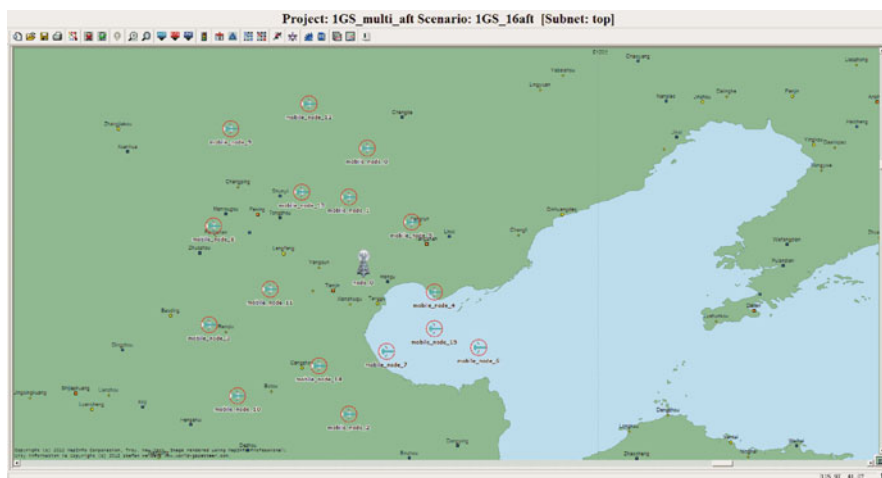


Fig. 146.1 Multi-ACs scenario



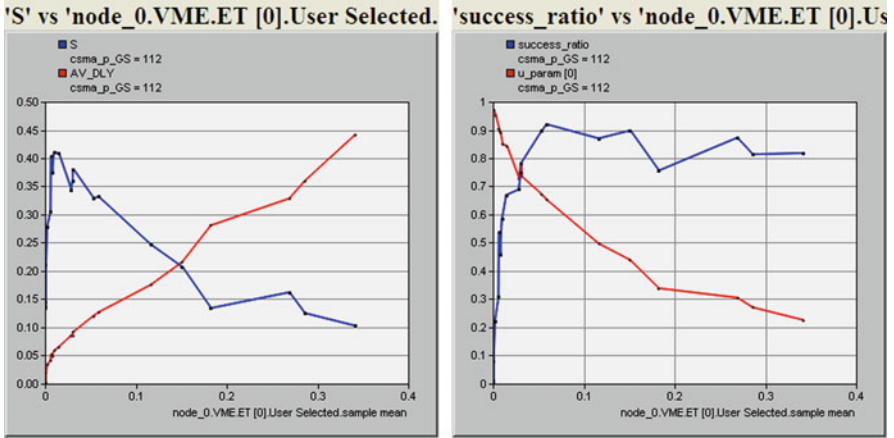


Fig. 146.2  $E[T_0]$  trends with parameters

In simulation, parameter  $\text{iat}$  is defined as the mean of exponential distribution of the packet interval and we have  $\text{iat} = 1/\lambda$  ground on Poisson theory. For packet length, we set the lower and upper limits to be  $l_a = 128$  bits and  $l_b = 8, 320$  bits according to VDL2 protocol and assume that the random variable abides by uniform distribution.

From Eqs. (16), (18), and (19), we can see that the expectations of idle and busy period are keys to calculate the targets of both throughput and packet delay as well as channel efficiency. From Eqs. (146.7) and (13), we know that  $E[T_0]$  and  $E[T_1]$  change in opposite direction. Therefore, we can focus on  $E[T_0]$ , which is decided by multiple parameters. The simulation results are shown in Fig. 146.2. From the figure, we can see  $E[T_0]$  has positive correlation with packet delay and has negative correlation with throughput and channel efficiency ( $u\_param$ ). There is no obvious correlation between success rate and  $E[T_0]$ . The results are in consistent with the theory we deduced.

With the method of numerical analysis, the simulation results of fixed position distribution have the good consistency with preemptive probability model we set up, which can verify the correctness of the mathematic model.

Furthermore, different scenarios are built by changing the distribution of nodes' positions. The results of simulation become complicated and variable. But with the multiple experiments, we can see that distribution pattern has a significant impact on the system performance. Specially, packet collisions will aggravate as the locations of stations become more uneven. The conclusion conforms to the sequitur from the probability model that success rate is negatively related to the sum of the absolute value of  $d_{ij}(t) - d_{ik}(t)$ , where  $i \neq j \neq k$  and  $i, j, k \in (0, N - 1)$ .

## 146.4 Conclusion

Taking propagation delay into account, a preemptive probability model is built for asynchronous p-CSMA. Network performance, e.g., throughput and packet delay, can be achieved with the idle and busy period expectation calculated in the model. Through simulation, the probability model is verified in scenarios with both fixed position distribution and varied position distributions.

The model can also be simplified by specialization. For instance, it can be specialized on Bianchi's critical saturation condition. Under the condition, the actor of packet rate is eliminated and P-CSMA performance Eqs. (146.9)–(146.12) can be expressed in a simplified form. Through specialization, further certain relationships between performances and parameters (especially position distribution of nodes) will be expected, which will benefit the network optimizations.

**Acknowledgements** This work is financially supported by the National Natural Science Foundation of China (No. 61170328) and Natural Science Foundation of Tianjin (No. 12JCZDJC20900).

## References

1. Kleinrock, L., & Tobagi, F. A. (1975). Packet switching in radio channels: Part I-carrier sense multiple-access modes and their throughput-delay characteristics. *IEEE Transactions on Communications*, 23(12), 1400–1416.
2. Tobagi, F. A., & Kleinrock, L. (1977). Packet switching in radio channels: Part IV-stability considerations and dynamic control in carrier sense multiple accesses. *IEEE Transactions on Communications*, 25(10), 1103–1119.
3. Bianchi, G. (2000). Performance analysis of the IEEE 802.11 distributed coordination function. *IEEE Journal on Selected Areas in Communications*, 18(3), 535–547.
4. MacKenzie, R., & O'Farrell, T. (2010). Throughput and delay analysis for p-persistent CSMA with heterogeneous traffic. *IEEE Transactions on Communications*, 58(10), 2881–2891.
5. Abukharis, S., MacKenzie, R., & O'Farrell, T. (2011). Throughput and delay analysis for a differentiated p-persistent CSMA protocol with the capture effect. In *Proceedings of the 2011 I. E. 73rd vehicular technology conference* (pp. 1–5), Piscataway, NJ, USA.
6. Gao, Y., & Dai, L. (2011). On the throughput of CSMA, information sciences and systems. In *45th annual conference of IEEE Computer Society* (pp. 1–4), Piscataway, NJ, USA.
7. Wang, D., Hu, X., Xu, F., Chen, H., & Wu, Y. (2012). Performance analysis of P-CSMA for underwater acoustic sensor networks. In *IEEE OCEANS 2012 MTS* (pp. 1–6), Piscataway, NJ, USA.
8. Wargo, C. A., D'Arcy, J. F. (2011). Performance of data link communications in surface management operations. In *IEEE aerospace conference* (pp. 1–10), Piscataway, NJ, USA.
9. ICAO Annex 10 Aeronautical Telecommunications. (2007). Volume III part I, digital data communication systems.

# Chapter 147

## Extended Petri Net-Based Advanced Persistent Threat Analysis Model

Wentao Zhao, Pengfei Wang, and Fan Zhang

**Abstract** In order to display the attack scene in the description of the multistep process-oriented attack—advanced persistent threat, a specific model on advanced persistent threat behavior analysis—EPNAM is proposed, which is based on the Petri net and combined with the characteristics of APT. Firstly we carry out hierarchical analysis on the attack scene with AHP method to build the APT architecture and extract scene factors, then associate the attack scene with Petri net to construct extended Petri net, and finally, traverse the extended Petri net to generate the formal expression. The proposed model can achieve the combination of the attack scene, attack process, and state space, and its feasibility is proved by the application on actual case analysis of the RSA SecurID theft attack.

### 147.1 Introduction

APT (advanced persistent threat) is a kind of network attack launched by organizations (especially governments) or small groups, which aims at specific targets and works persistently by means of advanced techniques. Generally believed, APT is a particular type of attack on the network infrastructure, aiming to obtain the key information of a particular organization or government, including energy, power grid, financial organization, and national defense. APT utilizes a variety of approaches to get the authorization of an inner organization step by step, including advanced techniques and social engineering methods. Attackers launch a series of attacks to a specific target; what those attackers really want is usually not getting benefit in a short period of time but keeping on searching through the compromised host, until they thoroughly grasp the key information of the target person or thing [1–3].

---

W. Zhao • P. Wang (✉) • F. Zhang  
School of Computer, National University of Defense Technology, Changsha 410073, China  
e-mail: [wpengfei\\_nudt@163.com](mailto:wpengfei_nudt@163.com)

## 147.2 Related Works

APT brings the threat that continues to strengthen, but for the relative lack of research, systemic knowledge has not yet formed, nor has the practical research method been found.

Currently the most sophisticated models of network attack include the attack tree model, attack graph model, and attack net model based on Petri net. The attack tree model proposed by B. Schneier [4] is intuitive and hierarchical, but hard to distinguish attack state from attack behavior, and has poor scalability [5].

Swiler et al. proposed the attack graph model [6, 7] from the graph theory, but analysis of large-scale network by the attack graph model will lead the state space of the attack graph algorithms into exponential growth, which is difficult to search or has the flexibility to add and modify [8, 9].

McDermott proposed the attack net model based on Petri net [10], which can separate attack state from attack process, and has the hierarchical nature of the attack tree model. But as a general attack model, it is insufficient in the description of the characteristics of the APT, such as persistence and penetrability, especially it cannot reflect the scene features and milestones of APT.

Based on the above status, we propose an extended Petri net-based modeling approach, combined with persistence and penetrability of APT, as well as the depiction of the attack scene, which can completely include the key factors of APT, including the object, intention, preconditions, vulnerability, penetration, and state space.

## 147.3 APT Modeling Analysis Based on Extended Petri Net

The attack net model based on Petri net was first proposed by McDermott [10]. Places of Petri net represent the path attack stage or state, indicated as “O,” transitions represent the attack behavior, indicated as “□”; paths represent the attack process, indicated as “→”; and Token represents the current state, indicated as “•.”

The extended Petri net-based APT model proposed in this chapter combines with APT characteristics. Firstly do hierarchical analysis on the attack scene to extend the classical Petri net model according to the extracted scene factors. Then take advantage of the extended Petri net for modeling analysis to generate formal expression. And finally achieve the target of the attack scenes factors reflected in a multistep process-oriented attack description so as to complete the combination of attack scenes, attack process, and time states.

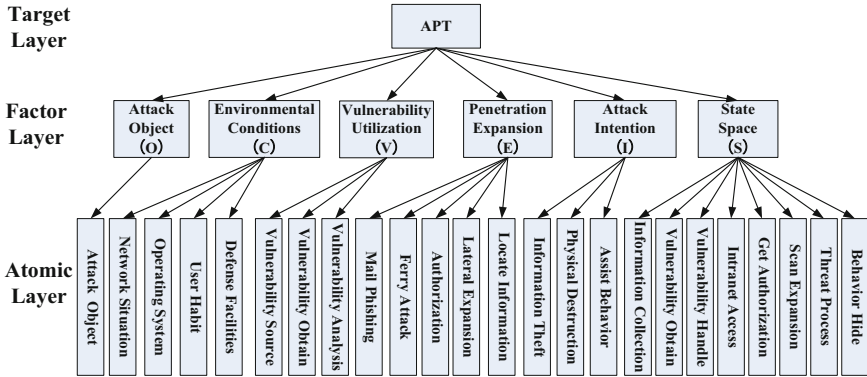


Fig. 147.1 Hierarchical analysis structure

### 147.3.1 Attack Scene Hierarchical Analysis

First analyze the attack scene of APT hierarchically by analytical hierarchy process. As shown in Fig. 147.1, target layer is defined as the top layer, then the factor layer forms 6-tuple  $\{O, C, V, E, I, S\}$  to characterize a specific APT scene, and the underlying atomic layer represents the specific operation.

1. *Attack object*: Potential target of an APT, such as the corporation intranet, mail server, or industrial control systems, denoted by **O**
2. *Environmental conditions*: Surrounding environment and preconditions to launch an APT, denoted by **C**
3. *Vulnerability utilization*: Vulnerability (0 day) obtaining and utilization is prerequisite for the implementation of APT, denoted by **V**
4. *Penetration expansion*: Locate important information by implementing lateral extension, which is the key step in APT implementation, denoted by **E**
5. *Attack intention*: The ultimate goal of the attacker, such as information theft, physical damage, or simply obtaining authorization to conceal, denoted by **I**
6. *State space*: Set of critical state in the procedure of APT, denoted by **S**

### 147.3.2 Extended Petri Net Modeling Principle

The proposed extended Petri net model is extended to a 6-tuple  $\{O, C, V, E, I, S\}$  on the basis of the original 3-tuple. State space set **S** comes from the place node set (**P** in the 3-tuple) of the classical Petri net, which describes the critical state in the procedure of APT. The first state from **S** is extracted to represent the attack object, denoted as **O**. Then further divide the original transition node set (**T** in the 3-tuple) into a set of environmental condition set, vulnerability utilization set, penetration

**Table 147.1** Extended Petri net system

Scene factor	Atomic index	Meaning	Candidate value
<b>O</b>	<i>object</i>	Attack object	object_intranet, object_server, object_ctrlSystem
<b>C</b>	<i>net</i>	Network situation	net_connect, net_isolated
	<i>os</i>	Operating system	os_windows, os_linux, os_mac, os_embed
	<i>habit</i>	Habits characteristic	habit_office, habit_photo, habit_music, habit_movie
	<i>defense</i>	Defense facilities	defense_ids, defense_audit
<b>V</b>	<i>source</i>	Vulnerability source	source_office, source_os, source_app
	<i>obtain</i>	Vulnerability obtain	obtain_buy, obtain_dig, obtain_exchange
	<i>analysis</i>	Vulnerability analysis	analysis_0day, analysis_1day
<b>E</b>	<i>mail</i>	Mail phishing	mail_server, mail_pc, mail_social, mail_sqlInject
	<i>ferry</i>	Ferry attack	ferry_worm, ferry_Trojan
	<i>authorize</i>	Gain authorization	authorize_malWare, authorize_backDoor, authorize_bruteForce, authorize_phishing
	<i>expand</i>	Laterally expansion	expand_scan User, expand_sendMsg
	<i>locate</i>	Information locating	locate_info, locate_asset, locate_person
<b>I</b>	<i>info</i>	Information theft	info_pack, info_compress, info_encrypt
	<i>destroy</i>	Physical destruction	destroy_modify, destroy_delete, destroy_cheat
	<i>assist</i>	Assist behavior	assist_conceal, assist_maintain
<b>S</b>	<i>s1</i>	Informationcollection	s_infoCollect
	<i>s2</i>	Vulnerability obtain	s_vulnerabilityObtain
	<i>s3</i>	Vulnerability handle	s_vulnerabilityHandle
	<i>s4</i>	Intranet entrance	s_intranetAccess
	<i>s5</i>	Get authorization	s_getAuthorization
	<i>s6</i>	Scan expansion	s_scanExpand
	<i>s7</i>	Threat process	s_threatProcess
	<i>s8</i>	Behavior conceal	s_behaviorConceal

expansion set, and attack intention set. In the need to characterize the dynamic attack scene, Token, which from the original Petri net is retained, the state node where Token at indicates the current state. The connection set ( $F$  in the 3-tuple) is omitted in the 6-tuple. Extended Petri net system is shown in Table 147.1.

**Definition 1:** Extended Petri net-based APT model.  $EPNAM = \{ob, co, vu, ex, it, st \mid ob \in O, co \in C, vu \in V, ex \in E, it \in I, st \in S\}$ . Where

$O = \{object\}$ ;  $C = \{net, os, habit, defense\}$

$V = \{source, obtain, analysis\}$ ;  $E = \{mail, ferry, authorize, expand, locate\}$

$I = \{info, destroy, assist\}$ ;  $S = \{s1, s2, s3, s4, s5, s6, s7, s8\}$

Extended Petri net is generated, as shown in Fig. 147.2.

**Definition 2:** State change. State changes of EPNAM can be expressed as  $O \times C \times V \times E \times I \times S \rightarrow S$ . APT process can be classified into stages as environmental conditions, vulnerability utilization, penetration expansion, and attack intention; state change in each stage can be expressed as

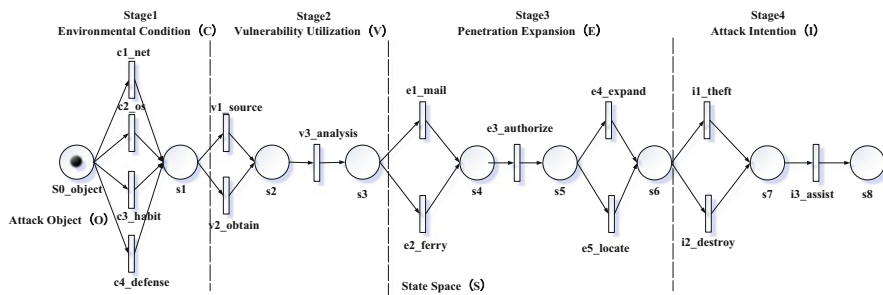


Fig. 147.2 Extended Petri net

Environmental conditions:  $O \times C \rightarrow S$ ; Vulnerability utilization:  $S \times V \rightarrow S$   
 Penetration expansion:  $S \times E \rightarrow S$ ; Attacks intention:  $S \times I \rightarrow S$

### 147.3.3 The EPNAM Generation Algorithm

**Step 1:** Scene hierarchical analysis

Analyze APT scene with AHP, extract each layer, and determine the 6-tuple  $\{O, C, V, E, I, S\}$  of the factor layer, and the atomic indicators of each factor.

**Step 2:** Scenes association with Petri net

Extract key characteristics of the attack scene, then associate them with atomic layer indicators. Determine the value of each atom layer indicator from the existing values within the selecting domain to complete the 6-tuple.

**Step 3:** Construction of extended Petri net

The Petri net is constructed in accordance with the 6-tuple stage order with a serial structure. In each stage, the parallel structure is used between the atomic indicators and uses serial structure between the chosen values for each indicator.

**Step 4:** Formal expression generation

Traverse the constructed Petri net by the mobile of Token to generate formal expression. *attack\_sqc* represents the attack sequence generated at the last stage, *stage\_sqc* represents the attack subsequence produced at a certain stage in APT, and *atom\_sqc* is the collection of value of a certain atomic indicator.

---

```

Algorithm 1 Generation of attack sequence
Input:  $S = \{s1, s2, s3, s4, s5, s6, s7, s8\}$ 
Output: attack_sqc
PROCEDURE GENERATE_SEQUENCE(S)
BEGIN:
attack_sqc  $\leftarrow \emptyset$ ; section_sqc  $\leftarrow \emptyset$ ; atom_sqc  $\leftarrow \emptyset$ 
token  $\leftarrow s1$ ; section  $\leftarrow S1$ 
FOR token  $\leftarrow s1$  TO  $s9$  DO
    
```

---

(continued)

(continued)

---

```

section_sqc ← ∅
WHILE nextIndex(token) ≠ ∅
  atom_sqc ← ∅
  WHILE isSelected(atom) IN AtomDomain
    atom_sqc ← atom_sqc ∪ atom
  section_sqc ← section_sqc ∪ atom_sqc
  token ← token + 1
IF changSection(token) THEN
  nextSection(section)
attack_sqc ← (attack_sqc → section_sqc)
END

```

---

## 147.4 Case Study

In this chapter, we choose the famous RSA SecurID theft attack as the analysis case to prove the feasibility of the proposed model, which is the most representative one that contains the universal characteristics of APT.

RSA, which is the subsidiary company of EMC, suffered APT in March 2011, and part of the SecurID technology and customer information was stolen. Its consequences led to the theft of important information of the company which uses SecurID as the authentication credentials to establish a VPN network. The following analysis is implemented using the proposed EPNAM approach on the background of RSA SecurID theft attack.

1. Scene hierarchical analysis (Table 147.2)
2. Scene association with Petri net (Table 147.3)
3. Construction of extended Petri net (see Fig. 147.3)
4. Formal expression generation

**Table 147.2** RSA SecurID theft attack scene analysis

Scene factor	Description
<i>O</i>	Intranet of RSA, mail server
<i>C</i>	Collect real-time information on RSA corporation. Imitate communication characteristics and habits of branch company. Intranet web server, windows operating system, and MS office, equipped with IDS and auditing system
<i>V</i>	Sending malicious e-mail to the chief manager in the name of the branch company, attachment (2011Recruitment Plan.xls) of which is embedded with a new 0-day vulnerability, Adobe Flash CVE-2011-0609
<i>E</i>	Trojan download from the C&C server; expand rapidly, get access to PC with high authority, raise authorization, locate important files
<i>I</i>	Gather the SecurID technique and customer information of RSA corporation, then pack compress, encrypt, and send out through FTP to a remote PC. Clear trace and maintain the current permission of the network
<i>S</i>	Information collection, vulnerability obtain, vulnerability utilization, mail penetration, lateral expansion, information theft, conceal maintain



**Table 147.3** Association

Scene factor	Atomic index	Candidate values
<i>O</i>	<i>object</i>	object_intranet, object_server
<i>C</i>	<i>net</i>	net_connect
	<i>os</i>	os_windows
	<i>habit</i>	habit_office, habit_photo
	<i>defense</i>	defense_ids, defense_audit
<i>V</i>	<i>source</i>	source_office, source_app
	<i>obtain</i>	obtain_dig
	<i>analysis</i>	analysis_0day
<i>E</i>	<i>mail</i>	mail_server, mail_pc, mail_social
	<i>authorize</i>	authorize_malWare, authorize_phishing
	<i>expand</i>	expand_scan User, expand_sendMsg
	<i>locate</i>	locate_info, locate_asset, locate_person
<i>I</i>	<i>info</i>	info_pack, info_compress, info_encrypt
	<i>assist</i>	assist_conceal, assist_maintain
<i>S</i>	<i>s1</i>	s_infoCollect
	<i>s2</i>	s_vulnerabilityObtain
	<i>s3</i>	s_vulnerabilityHandle
	<i>s4</i>	s_intranetAccess
	<i>s5</i>	s_getAuthorization
	<i>s6</i>	s_scanExpand
	<i>s7</i>	s_threatProcess
	<i>s8</i>	s_behaviorConceal

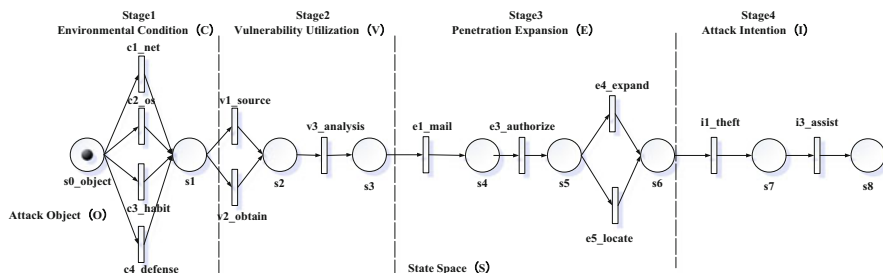
According to algorithm 1 to traverse the preconstructed extended Petri net, the generated formal expression is described as follows:

$$\begin{aligned}
 &O(\text{object\_inNet} \wedge \text{object\_server}) \rightarrow C(\text{net\_connect} \vee \text{os\_windows} \vee (\text{habit\_office} \wedge \\
 &\text{habit\_photo}) \vee (\text{defense\_ids} \wedge \text{defense\_audit})) \rightarrow \\
 &V(((\text{source\_office} \wedge \text{source\_app}) \vee \text{obtain\_dig} \vee \text{analyse\_0day}) \rightarrow \\
 &E((\text{mail\_server} \wedge \text{mail\_pc} \wedge \text{mail\_social}) \vee \\
 &\quad (\text{authorize\_malWare} \wedge \text{authorize\_phishing}) \vee \\
 &(\text{expand\_scanUser} \wedge \text{expand\_sendMsg}) \\
 &\quad \vee (\text{locate\_info} \wedge \text{locate\_asset} \wedge \text{locate\_person})) \rightarrow \\
 &I((\text{info\_pack} \wedge \text{info\_compress} \wedge \text{info\_encrypt}) \vee (\text{assist\_conceal} \wedge \text{assist\_maintain}))
 \end{aligned}$$

Generation of attack states transformation is described as below:

$$\begin{aligned}
 &s\_infoCollect \rightarrow s\_vulnerability\ Obtain \rightarrow s\_vulnerability\ Handle \rightarrow s\_intranet \\
 &\quad \text{Access} \rightarrow \\
 &s\_getAuthorization \rightarrow s\_threatProcess \rightarrow s\_behaviorConceal
 \end{aligned}$$

The above analysis proves the feasibility of the proposed model. Furthermore, the use of the model can do a better depiction on the attack scene, attack process as well as state change of APT, and the formal expression can be used as basis for further behavioral analysis and quantitative assessment.



**Fig. 147.3** Extended Petri net of RSA SecurID theft attack

## 147.5 Conclusion

In this chapter, an APT behavior analysis model based on extended Petri net—EPNAM model—is proposed, which can do hierarchical analysis on attack scene, refine the APT hierarchy and scene factors, associate the attack scene with Petri net to construct extended Petri net, and generate the formal expression by traversing the Petri net. EPNAM model can completely portray the characterization of an APT by embodying the attack scene in a multistep process-oriented attack description and finally complete the combination of the attack scene, attack process, and state space. The feasibility of the proposed model is proved by the analysis of actual case.

The following work focuses on refining the scene analysis and scene association, improving the adaptability of the proposed model on different APT cases, and implementing behavior analysis and quantitative assessment on the basis of the generation of formal expression.

**Acknowledgements** This work is supported by the National Natural Science Foundation of China (No. 61271252).

## References

- Huang, D., & Xue, Z. (2012). Research and analysis on advanced persistent threat behavior. *Information Security and Communication Secrecy*, 1(5), 94–96 (In Chinese).
- Bo, N. (2012). APT: Conceal, aims at theft of enterprise secret. *Journal of China Computer*, 1(1), 1–3 (In Chinese).
- Chen, J., Wang, Q., & Wu, M. (2012). Network APT attack and defense policy. *Information Security and Communication Secrecy*, 1(7), 30–33 (In Chinese).
- Schneier, B. (1999). Attack trees-modeling security threats. *Journal of Software Tools*, 24(12), 21–29.
- Li, W. (2010). Research on network attack modeling method. *Computer CD Software and Applications*, 1(13), 28–29.
- Phillips, C., & Swiler, L. (1998). A graph-based system for network vulnerability analysis. In *Proceedings of the 1998 Workshop on New Security Paradigms* (pp. 71–79). New York, NY: ACM.

7. Swiler, L., Philips, C., Ellis, D., & Chakerian, S. (2001). Computer-attack graph generation tool. In *Proceedings of DARPA Information Survivability Conference and Exposition* (pp. 307–321). Los Alamitos, CA: IEEE.
8. Wang, G., Wang, H., Chen, Z., & Xian, M. (2009). An attack graph-based computer network attack modeling method. *Journal of National University of Defense Technology*, 31(4), 78–84.
9. Cheng, K. (2010). *Cyber attacks based on attack graphs and Petri nets model study*. Xi'an, China: Xi'an University of Architecture University of Science and Technology.
10. McDermott, J. (2000). Attack net penetration testing. In *Proceedings of the 2000 Workshop on New Security Paradigms* (pp. 15–21). New York, NY: ACM.

# Chapter 148

## Energy-Efficient Routing Protocol Based on Probability of Wireless Sensor Network

Kaiguo Qian

**Abstract** This chapter mainly discusses the problem of wireless sensor network routing protocols. Based on analysis of the disadvantages of information implosion and overlapping caused from implementation mechanisms of the flooding protocol, energy-efficient routing protocol based on probability of wireless sensor network (ERPBP) is proposed and evaluated. It uses the node distance and residual energy as the weights to calculate the forwarding probability of neighbor nodes and chooses some of maximum forwarding probability nodes as router. It saves the energy by avoiding redundancy packet copies produced and improves the disadvantage of flooding routing protocol. Performance analysis and simulation experiment show that the new protocol effectively reduces the data redundancy, reduces the energy consumption, and prolongs the network lifetime.

### 148.1 Introduction

Wireless sensor network (WSN) [1] is a self-organization network system that is widely used in environmental monitoring, medical care, urban traffic management, warehouse management, and military reconnaissance and is now becoming the hottest research field. Wireless sensor network (WSN) [2] has the following characteristics: node energy is limited and large-scale network of nodes is without uniform distribution, node mobility, and network topology dynamic change. Routing protocol [3] is designed to transmit the data from the source node to the base station by intermediate forwarding nodes. Primary goal of routing design is to meet the energy constraint of sensor node, to reduce energy consumption, and to prolong the network life cycle. Second, it has higher scalability because of the topological structure dynamic change. Third, routing protocol implementation is

---

K. Qian (✉)  
Department of Physics Science and Technology, Kunming University,  
Kunming 650031, China  
e-mail: [qiankaiguo@qq.com](mailto:qiankaiguo@qq.com)

simple on demands of weakness of computing power and storage capacity. Researchers have proposed hundreds of solutions, in which the flooding routing protocol originated in the ad hoc network is the simplest and reliable algorithm. Flooding does not need the information of the whole sensor network and has the advantages of high scalability, simple calculation, and high reliability. Whereas, it causes the problem of redundant packet copies, information implosion and overlap. With the analysis of Gossip [4] and comprehensively considering the influence of the node residual energy and communication distance, the chapter presents a new routing protocol base nodes forwarding probability calculated according to the node residual energy and communication distance. Performance analysis and simulation experiment show that the new protocol effectively reduces the data redundancy and the energy consumption and prolongs the network lifetime.

## 148.2 Related Work

In flood-routing protocol, source node sends collected data packet to all of the own neighbors. Every neighbor node receives data from other nodes and broadcasts to its neighbors, so that it goes on until the data transmitted to the destination sink. It produces a large amount of redundant information, causing an information implosion and overlap. Gossip [4] randomly chooses a neighbor node as the data receiver if a node has the data to transmit to sink, so do the neighbors. This approach decreases redundant data copies and relieves the information implosion and overlapping. It causes longer delay time for the packet transmission. SPIN [5] and DD [6] are two data-centric protocols. Clustering routing protocol is a kind of hierarchy that divides the nodes into the cluster head nodes and member nodes. Cluster heads manage member nodes in the native. Leach [7] and TEEN [8] are not reliable as flooding, which are representatives for the clustering routing protocol.

## 148.3 Problem Description

The energy-efficient routing protocol based on probability (ERPBP) routing protocol runs in the network application scenario as the following:  $N$  sensor node is randomly deployed in  $M \times M$  square area. Each sensor node is stationary after completion of WSN deployment. All nodes in the network are homogeneous and energy constrained. We take the radio energy model [4] in order to verify energy efficiency of the ERPBP routing protocol:

$$E_{Tx}(k, d) = \begin{cases} kE_{elec} + k\epsilon_{fs}d^2 & : d < d_0 \\ kE_{elec} + k\epsilon_{mp}d^4 & : d > d_0 \end{cases} \quad (148.1)$$

Equation (148.1) is energy dissipation equation for transmitting k-bit message and for receiving k-bit message is shown Eq. (148.2):

$$E_{Rx}(k) = kE_{elec} \quad (148.2)$$

In above model,  $E_{elec}$  is the radio dissipate to run the transmitter or receiver circuitry.

## 148.4 Energy-Efficient Routing Protocol Based on Probability

There are two ways to improve energy efficiency for running routing protocol from the radio energy model. The first is decreasing packets on the data transmission process, that is, decrease k in the radio energy model. The second is shorting the communication distance. It rotates the routing node to realize balancing energy consumption. The ERPBP protocol calculates the neighbor node forwarding probability based on node residual energy and communication distance if nodes may transmit packets and selects a few nodes of the maximum probability as forwarding nodes.

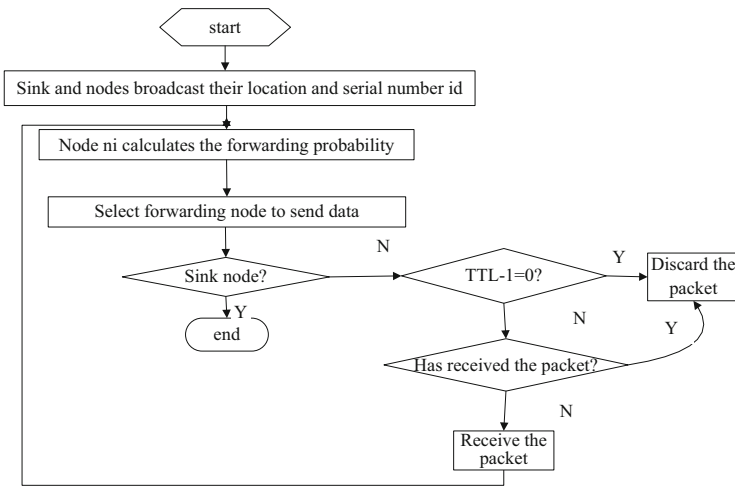
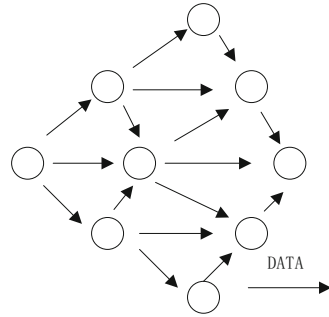
### 148.4.1 The Principle of ERPBP Protocol

In flooding protocol, if the node S wants to transmit data to the sink, it sends the data copies to every neighbor node, and so does every neighbor except sending back to source node until that sink node receives the data. The data is discarded if the TTL (time to live) becomes 0. As can be seen from Fig. 148.1, it produces redundant data and information implosion.

In ERPBP protocol, Source node  $S_i$  calculates the forwarding probability  $p_i$  of neighbors according to Eq. (148.3) and chooses a few of nodes of the maximum probability to send packets. The forwarding probability is proportional to the node residual energy. The greater the residual energy of the node, the greater the forwarding probability. The forwarding probability is not a simple inverse ratio relationship with the communication distance between nodes. We assume that the forwarding probability will decrease quickly when node distance is greater than the threshold  $d_0$ . We put forward the probability calculation mathematical model as shown in Eq. (148.3):

$$p(e, d) = \begin{cases} ae + \frac{1}{(d+1)^a} (0 < d \leq d_0) \\ ae - bd + c (d > d_0) \end{cases} \quad (148.3)$$

**Fig. 148.1** Principle of flooding



**Fig. 148.2** Process of ERPBP protocol

**148.4.2 The ERPBP Protocol Description**

Data message contains TTL field and DATA field. TTL field in the message prevents unrestricted forwarding packet in the network. If a node  $n_i$  needs to send the data message, it does as shown in Fig. 148.2:

- (a) Sink and nodes broadcast their location and serial number id.
- (b) Node  $n_i$  calculates the forwarding probability of neighbors according to the node's residual energy and distance.
- (c) Node  $n_i$  chooses a few nodes with maximum forwarding probability as next hops.

- (d) If the received message node is sink node, the message has already arrived the destination; otherwise, jump to (e).
- (e) If the  $TTL - 1$  is equal to zero or the destination node has already received the packet, jump to (f); otherwise, turn to (b).
- (f) Discard the packet.

## 148.5 The Performance Analysis and Performance

### 148.5.1 Performance Analysis

The ERPBP protocol carries the advantage that algorithm implementation mechanism is simple. It is suitable for topological dynamic change. In the process of packet routing, the neighbors of larger residual energy and near distance are preferentially selected for the forwarding nodes. Nodes, in which the distance is far and residual energy is small, do not participate in data forwarding. For this reason, it reduces energy consumption and realizes power load balance to prolong the system life. At the same time, the ERPBP protocol limits the number of nodes to avoid the generation of a large amount of redundant information and alleviates the information implosion in flooding algorithm.

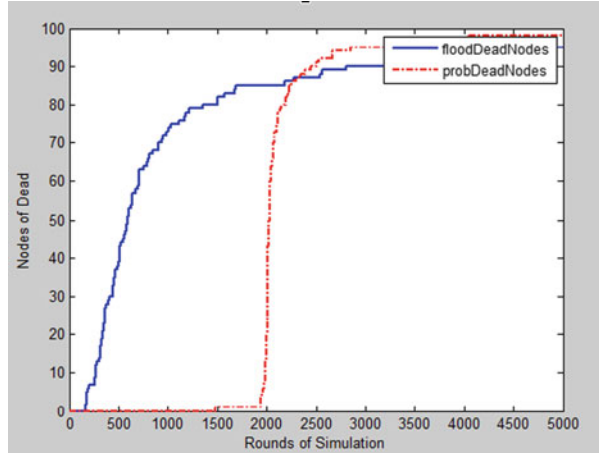
### 148.5.2 Performance

An experiment is designed in MATLAB to simulate and estimate the performance. 100 sensor nodes are randomly deployed in the area of  $100\text{ m} \times 100\text{ m}$ . Sink node is located at point  $100 \times 100$ .

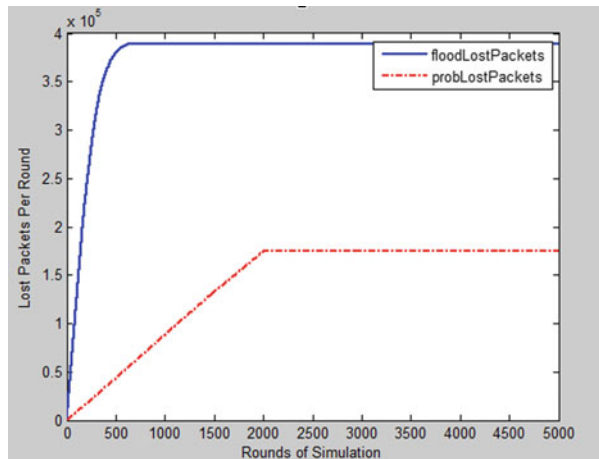
1. Lifetime of system. Experiment running 5,000 rounds, the results of dead nodes for the traditional flooding protocol and the ERPBP algorithm is shown as Fig. 148.3. In the flooding protocol, the sensor nodes begin to die at about 250 rounds and the dead nodes reach to more than 70% at about 1,000 rounds. Ninety percent of nodes are dead at the end of running. When ERPBP protocol runs 1,500 rounds, the first dead node emerges. Eighty percent nodes die between 1,950th round and 2,000th round. The result proves that the energy consumption of the RPBEN protocol is balanced. At the same time, it shows that the RPBEN protocol has higher energy utilization rate than flooding and has longer life cycle.
2. Loss packets. Loss packet experimental results are as shown in Fig. 148.4. In traditional flooding protocol, the total number of lost packets is probably  $4.4 \times 10^5$  when sensor network runs 1,000 rounds. The reason is that redundancy data copies are produced in the process of the operation of the network with flooding protocol. After 1,000 rounds, there is no data to produce because



**Fig. 148.3** Dead nodes with rounds



**Fig. 148.4** Lost packets with rounds



sensor network is dead. In the ERBPB protocol, the total number of lost packets grows a lot less than the flooding protocol. Running at 1,950th round, loss packets achieve  $1.8 \times 10^5$ . There are no loss packets after 1,950th round because almost all nodes are dead.

Combining the experimental results of lifetime of system and loss packets show that the network life cycle of flooding lasted about 500 rounds and the ERBPB protocol life cycle lasted for 2,000 rounds. It is proved that the improved protocol is a more efficient flooding algorithm.

## 148.6 Conclusion

This chapter presents a new energy-efficient routing protocol based on neighbors' probability calculated for the weights of the node's residual energy and distance. Source node chooses a few neighbor as forwarding hop routing nodes according to probability. This mechanism balances energy consumption in sensor networks and reduces a large number of redundant packets of flood process. It improves the disadvantages of information implosion, overlapping and the blinding usage of resources in flooding algorithm. Performance analysis and simulation results show that the new protocol is more energy-efficient than the flooding protocol.

**Acknowledgements** This research was supported by Foundation of Yunnan Educational Committee (2011Y238), China. The authors thank the anonymous reviewers whose comments have significantly improved the quality of this chapter.

## References

1. Akyildiz, L. F., Su, W. L., Sankarasubramaniam, Y., & Cayirci, E. (2002). A survey on sensor networks. *IEEE Communications Magazine*, 40(8), 102–114.
2. Ren, F. Y., Huang, H. N., & Lin, C. (2003). Wireless sensor networks. *Journal of Software*, 14(7), 1282–1291 (in Chinese).
3. Tang, Y., Zhou, M. T., & Zhang, X. (2006). Overview of routing protocols in wireless sensor networks. *Journal of Software*, 17(3), 410–421 (In Chinese).
4. Haas, Z., Joseph, Y., & Li, H. L. (2006). Gossip-based ad hoc routing. *IEEE/ACM Transactions on Networking (TON)*, 14(3), 479–491.
5. Kulik, J., Rabiner, W., Balakrishnan, H. (1999). Adaptive protocols for information dissemination in wireless sensor networks. In Proceedings of the fifth annual ACM/IEEE international conference on mobile computing and networking (pp. 174–185). New York, NY: ACM.
6. Ntanagonwivat, I. C., Govindan, R., & Estrin, D. (2000). Directed diffusion: A scalable and robust communication paradigm for sensor networks. In Proceedings of the sixth annual international conference on mobile computing and networking (pp. 56–57). New York: ACM.
7. Heinzelman, W. R., Anantha, C., & Hari, B. Energy-efficient communication protocol for wireless microsensor networks. In IEEE proceedings of the Hawaii international conference on system sciences (pp. 3005–3014), Maui, Hawaii.
8. Manjeshwar, A., & Agrawal, D. P. (2001). TEEN: A routing protocol for enhanced efficiency in wireless sensor networks. In Proceedings of the 15th parallel and distributed processing symposium (pp. 2009–2015), San Francisco, USA.

# Chapter 149

## A Dynamic Routing Protocols Switching Scheme in Wireless Sensor Networks

Zusheng Zhang, Tiezhu Zhao, and Huaqiang Yuan

**Abstract** Many sensor query processing systems have been developed to acquire, process, and aggregate data from wireless sensor networks. The energy consumption of query processing is significantly impacted by routing protocol. In this chapter, we propose a dynamic routing protocols switching scheme for query processing. The scheme supports multiple kinds of routing protocols coexisting in a single sensor node, and these protocols can be switched according to query tasks. Simulation results show that the dynamic scheme is more energy efficient than single routing protocol.

### 149.1 Introduction

The data query processing systems, such as TinyDB [1], are promising for wireless sensor networks. With these systems, user can inject SQL-style queries into a network through a PC. The networked sensor nodes then work together to process the queries and send results back to the PC. The performance of these sensor systems is greatly affected by the routing protocol, because routing protocol is responsible for result gathering.

Traditionally, in wireless sensor network each routing protocol is designed, developed, and evaluated separately, and the network layer uses a single routing protocol to deal with all kinds of queries in a query system [2]. However, a routing protocol is optimum for special task and network condition. For acquisition, such as “select temperature from sensors,” a node doesn’t aggregate data packets from neighbors, and all data packets are forwarded to BS through multi-hop. MintRoute [3], CTP [4], and other collection protocols [5] are energy efficient to establish flows up a tree and pull data out of a network. Another example, for aggregation,

---

Z. Zhang (✉) • T. Zhao • H. Yuan  
Dongguan University of Technology, Dongguan 523808, China  
e-mail: [zushengzhang@126.com](mailto:zushengzhang@126.com)

such as “select average (temperature) from sensors,” a node aggregates its sensed data with data packets received from neighbors into one packet. Clustering technique is energy efficient for aggregation. Many clustering protocols [6–8] that have been proposed for sensor networks, such as LEACH [6] and HEED [7], are the most popular. Jun and Julien [9] demonstrated that an optimal routing protocol can be selected for a particular application in ad hoc networks. He and Raghavendra [10] described a framework to build programmable routing services for sensor networks.

This chapter proposes the protocols switching scheme for reducing energy consumption of query processing. We decompose the network layer into shared modules and private modules. The design allows several routing protocols to coexist without burdensome memory requirement. BS decides protocols switching according to query tasks, attaches the information in the query, and disseminates the query to whole network. Then the network executes protocol switching. Simulation results show that dynamic switching scheme is more energy efficient than single routing protocol.

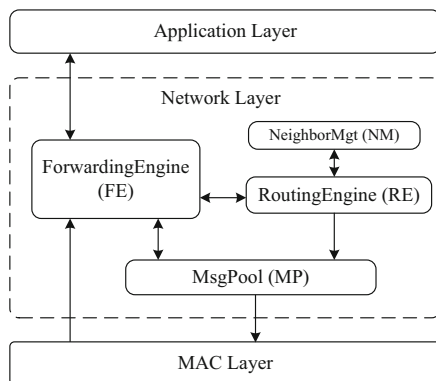
## 149.2 A Switching Scheme

### 149.2.1 *Modular Design*

Unlike traditional monolithic way, an overall modular design can increase code reuse and runtime sharing. Code reuse will foster more rapid protocol and application development. Runtime sharing refers to the sharing of code and resources such as memory and radio. Based on the TinyOS [11] code of MintRoute and the chapter of [12], we decompose the network layer into separate four parts: ForwardingEngine (FE), RoutingEngine (RE), MsgPool (MP), and NeighborMgt (NM), as shown in Fig. 149.1. The FE and MP modules support information exchanging for application and MAC layers by interfaces. The MP and NM are shared modules for all routing protocols. This is in contrast to the FE and RE modules; different routing protocols have different algorithms implemented in these modules.

The main function of an FE module is to obtain the next hop to which the packet is to be forwarded. Additional functions include detection of cycles, network level retransmission, and duplicate packet elimination. RE is responsible for creating and maintaining the routing topology. Examples of topologies are trees, geographic coordinates, or clusters. NM manages information about node’s neighbors. MP is the buffer place for outgoing packets, and it supports packet scheduling for all network protocols.

**Fig. 149.1** Modular design of the network layer



**Table 149.1** Polices of protocol selection

Keywords	Task	Routing protocol
SUM, AVE, MAX, MIN	Aggregation	HEED
WHEN	Event	CTP
Others	Collection	MintRoute

### 149.2.2 Protocol Selection

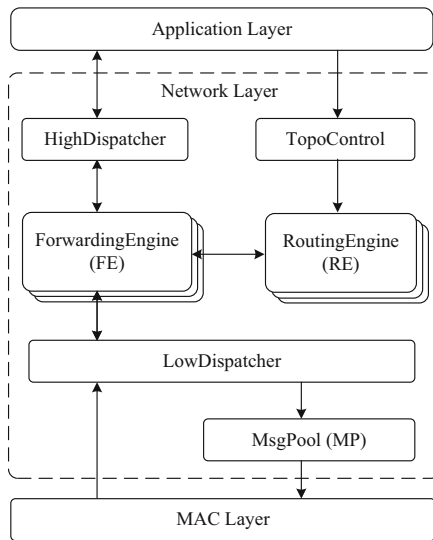
We adopt a centralized control mode for routing protocol selection. BS maintains a protocol selection table, as shown in Table 149.1. User defined the table entries, i.e., routing switching criteria, according to the overall knowledge of the network performance by theoretic analysis, simulation, and experiment results. Procedure at BS is as follows:

(a) Upon receiving a query from user, BS matches the query command to task. As shown in Table 149.1, for SQL language, BS checks the query task according to SQL keywords. (b) BS selects routing protocol according to query task. Then it attaches the routing protocol information on the query command. Such as “select temperature from sensors using MintRoute,” it informs the network switching to MintRoute for the query. Finally, BS disseminates the query to the whole network.

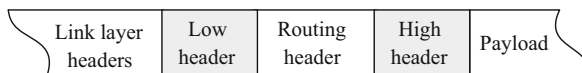
### 149.2.3 Protocols Switching

To achieve the compatibility and expansibility, the dynamic scheme provides uniform interface for application and MAC layers. Small changes of the code inside of modules of routing protocols are required, while other layers’ code need not be modified. The modular design of the dynamic scheme is shown in Fig. 149.2. Because FE and RE are private modules of each routing protocol, multiple

**Fig. 149.2** Architecture of the dynamic scheme



**Fig. 149.3** Network packet header format



instances of FEs and REs simultaneously exist on the single node. So correctly switching between FEs and REs modules is the key issue of the dynamic scheme. We add three modules: HighDispatcher, LowDispatcher, and TopoControl for routing protocols switching.

The TopoControl module controls the routing topology switching. When a node receives a query, explains the query, and starts the corresponding application program, TopoControl matches applications to REs. According to the routing protocol identifier (RpID for short) carried by the query command, when the query task starting, TopoControl starts the RpID's RE module. When the query task is finished, TopoControl stops the RpID's RE module. The HighDispatcher and LowDispatcher modules hide the differences of routing protocols for application and MAC layers. FE module is responsible for data packet sending and forwarding. As shown in Fig. 149.2, the dynamic scheme is competent for the switching between FEs by the HighDispatcher and LowDispatcher modules. The network header of a data packet is shown in Fig. 149.3. The dynamic scheme adds two fields: low header and high header. The low header is added and explained by HighDispatcher module. The HighDispatcher module provides a bridge for data packets correctly flowing between application layer programs and FEs. LowDispatcher manages the low header field in the packet header, and it provides a bridge for data packets correctly flowing between MAC layer and FEs.

### 149.3 Simulation

To evaluate the feasibility of our approach and to make the proposal concrete, we implemented the dynamic scheme in TinyOS [11]. We used two routing protocols: MintRoute and HEED for our implementation. Based on our previous work [13], BS selects routing protocol according to query tasks: for aggregation, BS chooses HEED as routing protocol; for acquisition, it chooses MintRoute. BS attaches the routing protocol information on the query command and sends it to the network using flooding algorithm.

We adopt the simple energy model [7]. The energy consumption for sending  $l$  bit data with distance  $d$  is  $E_{Tx}$ , and energy consumption for receiving  $l$  bit data is  $E_{Rx}$ . The energy model can be expressed as

$$E_{Rx}(l) = lE_{elec} \quad (149.1)$$

$$E_{Tx}(l, d) = \begin{cases} lE_{elec} + l\xi_{fs}d^2, & d < d_0 \\ lE_{elec} + l\xi_{mp}d^4, & d > d_0 \end{cases} \quad (149.2)$$

where  $E_{elec}$  is the circuitry power consumption, which depends on factors such as the digital coding, modulation, and filtering.  $\xi_{fs}d^2$  and  $\xi_{mp}d^4$  are the transmission power which depends on the distance to the receiver and acceptable bit-error rate. In our simulation, 100 nodes are randomly distributed in a  $100 \times 100$  m area, node communication range is 50 m, and the sink node is located at (0, 0 m). Energy consumption of data fusion is 5 nJ/bit. Energy consumption for sleep and sense mode is ignored. Simulation parameters are shown in Table 149.2.

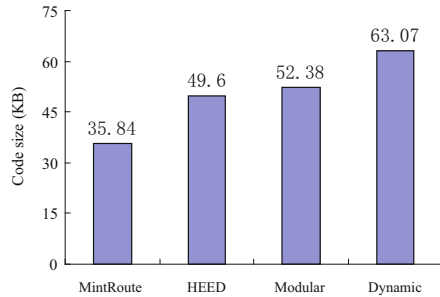
#### 149.3.1 Code Size

One of the main objectives of creating a dynamic scheme for sensor networks is to increase code reuse, thereby allowing for multiple protocols to coexist cleanly and efficiently on a single node. Compared to the monolithic implementations, Fig. 149.4 compares code size between the different implementations. Modular is

**Table 149.2** Simulation parameters

Parameter	Value
Threshold distance ( $d_0$ )	75 m
Data packet size	500 bytes
Control packet size	25 bytes
Initial energy	2 J
$E_{elec}$	50 nJ/bit
$E_{fusion}$	5 nJ/bit
$E_{fs}$	10 pJ/bit
$\xi_{mp}$	0.0013 pJ/bit

**Fig. 149.4** Code size comparison



the modular design of coexist MintRoute and HEED, and Dynamic refers to dynamic scheme based on coexist of MintRoute and HEED. When we combine different protocols, we observe clear gains: the modular design use up 40 % less code size.

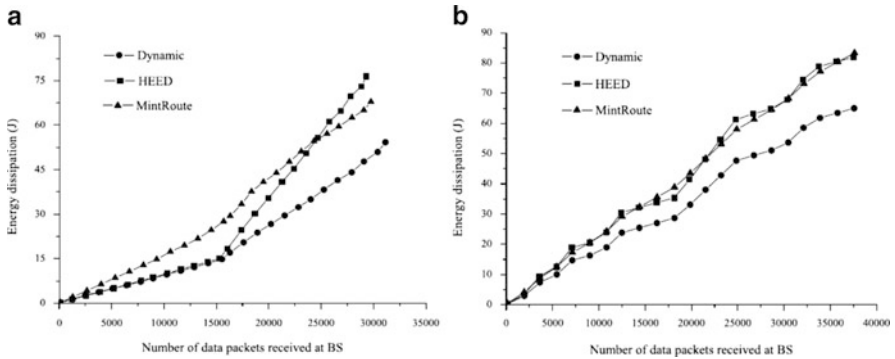
### 149.3.2 Continuous Query

BS sends an aggregation query to the network at first; when the aggregation query is finished, BS sends an acquisition query to the network continuously. The sample period of each query is 10 s and the duration is 30 min. We simulated the dynamic scheme, HEED, and MintRoute, respectively. The query system adopting dynamic scheme can switch routing protocols according to query tasks. As shown in Fig. 149.5a, we record the number of data packets received at BS and energy dissipation. When the number of data packets is about between 0 and 16,000, the network is running an aggregation query. In the interval, HEED is more energy-efficient than MintRoute, so the dynamic scheme selects HEED as routing protocol. When the number of data packets is about 16,000, the aggregation query is finished. And when the number of data packets is about between 16,000 and 32,000, the acquisition is executing. In the interval, the performance of MintRoute is better than HEED. The dynamic scheme can get the best of both worlds by switching from HEED to MintRoute. So it is more energy efficient than both MintRoute and HEED.

### 149.3.3 Random Query

The number of queries is randomly generated, query type is randomly selected from aggregation and acquisition, and query duration is a random integer between 1 and 30 min. Using these constraints, we generate the random queries scene, which contains 16 aggregation queries and 14 acquisition queries. Queries are injected into a network one by one, and there is only single query running in the network at a time.





**Fig. 149.5** Energy dissipation vs. the number of data packets received at BS (a) In the case of continuous query. (b) In the case of random query

We simulate MintRoute, HEED, and dynamic scheme using the same random queries scene. As shown in Fig. 149.5b, the dynamic scheme is more energy efficient than HEED and MintRoute.

## 149.4 Conclusion

This chapter describes the routing protocols switching scheme in sensor network that can switch between routing protocols according to query tasks. Simulation results show that it is more energy efficient than single routing protocol for query processing. The scheme decides protocols switching only based on query tasks. The performance of routing protocol also is significantly affected by network conditions. So as future work, we intend to let BS decide protocols switching based on knowledge by automatic learning.

## References

1. Madden, S., Franklin, M. J., Hellerstein, J. M., & Hong, W. (2005). Tinydb: An acquisitional query processing system for sensor networks. *ACM Transactions on Database Systems*, 30(1), 122–173.
2. Luo, Q., & Wu, H. (2007). System design issues in sensor databases. In *Proceeding of the 2007 ACM SIGMOD international conference on management of data (SIGMOD)* (pp. 1182–1185). New York: ACM Press.
3. Woo, A., Tong, T., & Culler, D. (2003). Taming the underlying challenges of reliable multihop routing in sensor networks. In *Proceedings of the 1st international conference on embedded networked sensor systems (SenSys)* (pp. 14–17). New York: ACM Press.

4. Gnawali, O., Fonseca, R., Jamieson, K., Moss, D., & Levis, P. (2009). Collection tree protocol. In *Proceeding of the 7th ACM conference on embedded networked sensor systems (SenSys)* (pp. 1–14). New York: ACM Press.
5. Borsani, L., Guglielmi, S., Redondi, A., & Cesana, M. (2011). Tree-based routing protocol for wireless sensor networks. In *Proceedings of the IEEE international conference on Wireless On-Demand Network Systems and Services (WONS)* (pp. 158–163). Washington, DC: IEEE.
6. Heinzelman, W., Chandrakasan, A., & Balakrishnan, H. (2002). An application specific protocol architecture for wireless microsensor networks. *IEEE Transactions on Wireless Communications*, 1(2), 660–670.
7. Younis, O., & Fahmy, S. (2004). HEED: A hybrid, energy-efficient, distributed clustering approach for ad hoc sensor networks. *IEEE Transactions on Mobile Computing*, 3(4), 366–379.
8. Khan, A., Madani, S., Hayat, K., & Khan, S. (2012). Clustering-based power-controlled routing for mobile wireless sensor networks. *International Journal of Communication Systems*, 25(4), 529–542.
9. Jun, T., & Julien, C. (2007). Automated routing protocol selection in mobile ad hoc networks. In *Proceedings of ACM symposium on applied computing* (pp. 906–913). New York: ACM Press.
10. He, Y., & Raghavendra, C. S. (2005). Building programmable routing service for sensor networks. *Computer Communications*, 28(6), 664–675.
11. Levis, P., Madden, S., Polastre, J., Szewczyk, R., Whitehouse, K., Woo, A., et al. (2005). TinyOS: An operating system for sensor networks. In W. Weber, J. Rabaey, & E. Aarts (Eds.), *Ambient intelligence* (pp. 115–148). Berlin: Springer.
12. Cheng, T., Fonseca, R., Kim, S., Moon, D., Tavakoli, A., Culler, D., et al. (2006). A modular network layer for sensornets. In *Proceedings of the 7th USENIX symposium on operating systems design and implementation* (pp. 249–262). Berkeley, CA: USENIX Association.
13. Zhang, Z., & Yu, F. (2010). Performance analysis of cluster-based and tree-based routing protocols for wireless sensor networks. In *Proceedings of the international conference on Communications and Mobile Computing (CMC)* (pp. 418–422). Washington, DC: IEEE.

# Chapter 150

## Incipient Fault Diagnosis in the Distribution Network Based on S-Transform and Polarity of Magnitude Difference

Jinqian Zhai and Xin Chen

**Abstract** It is difficult for conventional relaying algorithms to detect incipient faults, such as insulator current leakage, electrical faults due to tree limbs, and transient or intermittent earth faults, which are frequent in distribution networks. With the time, they may lead to a catastrophic failure. In order to avoid this situation, S-transform technique is proposed to extract the suitable features of incipient fault in this chapter. A least square support vector machine (LS-SVM) classifier is developed utilizing the features so that incipient fault is distinguished from the normal disturbances. Then the polarity of magnitude difference of residual current is used to determine the fault section of distribution network. The proposed technique has been investigated by ATP/EMTP simulation software. Simulation results show that this technique is effective and robust.

### 150.1 Introduction

Incipient faults in power lines are normally characterized as the faulty phenomena with the relatively low fault currents, such as high-impedance faults, insulator leakage current faults, and intermittent/transient faults [1]. These faults represent little threat of damage to power system equipment. But with time, they may lead to a catastrophic failure (i.e., a permanent damage beyond repair). So, early detection of power line fault would undoubtedly be a great benefit to the utilities enabling them to avoid catastrophic failures, unscheduled outages, and thus loss of revenues.

Various methods have been proposed by researchers and protection engineers. Among them, harmonic analysis [2], randomness detection [3], artificial neural networks [4], Hilbert-transform-based [5], wavelet transform [6–9], etc. are used to extract the feature of incipient fault signals in the distribution line. But due to high

---

J. Zhai (✉) • X. Chen  
Zhengzhou Power Supply Company, Zhengzhou 450000, China  
e-mail: [jinqianzhai@163.com](mailto:jinqianzhai@163.com)

time resolution and low frequency resolution for high frequencies and high frequency resolution and low time resolution for low frequencies, wavelet transform can achieve a better solution. To overcome the shortcomings associated with the existing techniques, novel methods need to be developed. The purpose of this chapter is to develop an online system that uses voltage and current during a period to diagnosis incipient faults of power line; the proposed algorithms are as follows: The incipient fault features are extracted by S-transform. All these features are then used to train LS-SVM to enable it to make prediction of possible occurrences of incipient faults in the distribution network. Then the polarity of magnitude difference of residual current is used to determine the faulty section in the distribution network.

## **150.2 Incipient Fault Characterization and Simulation System**

### ***150.2.1 The Characterization of Incipient Fault***

The term “incipient faults” refers to certain pre-fault “symptoms” or electrical activities taking place prior to a power system failure or blackout. Power line incipient faults are the primary causes of catastrophic failures in the distribution network. For underground cable, these faults develop in the extruded cables from gradual deterioration of the solid insulation due to the persisting stress factors. For overhead lines, incipient faults are associated with degraded equipment (insulators, arresters, transformer insulation and bushings, etc.) and the gradual intrusion of tree limbs as they grow into the overhead power line. Therefore, the present development aims at detecting these defects through prediction. The investigation searches out defect patterns or signatures that are still hidden, so they can be solved before they effectively interrupt the transmission system.

### ***150.2.2 Simulation System***

The 10-kV distribution system is supplied with power by a 110-kV grid via a 40-MVA transformer as shown in Fig. 150.1; the isolated neutral 10-kV system consists of several overhead lines and cables as radial feeders, which are simulated using ATP/EMTP, in which the processing is created by ATPDraw [10]. L2 and L3 are cables, 7 and 5 m, respectively; other feeders are overhead lines. L1 is 15 m, AB is 10 m, BC is 5 m, BD is 4 m, CE is 5 m, CF is 7 m, CG is 4 m, DH is 5 m, and DI is 6 m. The feeder overhead line and cable are represented using the frequency-dependent JMarti model. The high-impedance arc fault is simulated, and the several wireless sensors are installed in specific measuring sites [11, 12].

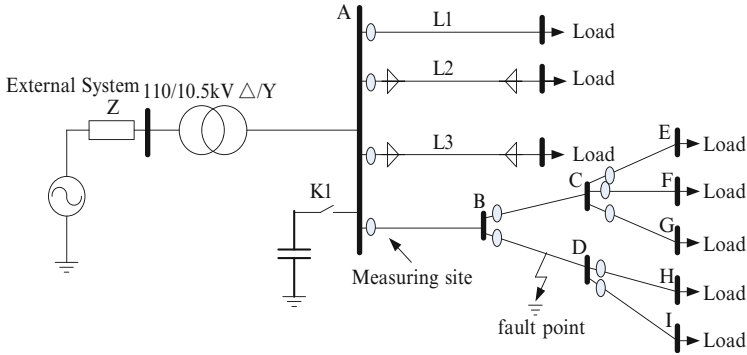


Fig. 150.1 Configuration of simulated system

### 150.3 S-Transform-Based Feature Extraction Method for Incipient Fault in the Distribution Network

#### 150.3.1 The Principle of S-Transform

The S-transform, put forward by Stockwell in 1996, is considered as one of the most recent signal processing techniques, which is a kind of time-frequency analysis method based on continuous wavelet transformation and short-time Fourier transformation, and overcomes some of disadvantages of the wavelet transforms [13]. The basis function for the S-transform is the Gaussian modulation cosinusoids. The consinusoid frequencies are used for the interpretation of a signal that will result in the time-frequency spectrum. The output of the S-transform is an  $N \times M$  matrix called the S-matrix whose row pertains to the frequency and columns to time [14]. Each element of the S-matrix is complex valued and is used for extracting features to classify the incipient faults. The S-transform performs multi-resolution analysis on a time-varying power signal as its window width varies inversely with the frequency. The S-transform for a function  $x(t)$  is defined as [15]

$$S(\tau, f) = \int_{-\infty}^{\infty} x(t) \frac{|f|}{\sqrt{2\pi}} e^{-\frac{(\tau-t)^2 f^2}{2}} e^{-i2\pi ft} dt \tag{150.1}$$

The S-transform will generate time-frequency contours; these contours can provide excellent features, which can be used by a pattern recognition system for classifying the incipient faults.

**Table 150.1** Features from the S-transform of incipient fault signal

Features	Description
F1	The maximum of S-matrix for residual current
F2	The minimum of S-matrix for residual current
F3	Standard deviation of residual current
F4	Mean value of residual current
F5	The maximum of S-matrix for residual voltage
F6	The minimum of S-matrix for residual voltage
F7	Standard deviation of residual voltage
F8	Mean value of residual voltage

### 150.3.2 Feature Extraction from the S-Transform

Since the residual voltage and current are real-time summation of the three phase voltages and currents, respectively, any single-phase change is reflected in the residual voltage and current. Therefore, in this study, both residual voltage and current were analyzed. They are computed as

$$u_r = u_a + u_b + u_c \quad (150.2)$$

$$i_r = i_a + i_b + i_c \quad (150.3)$$

where  $u_r$  and  $i_r$  are the residual voltage and current, respectively.  $u_a$ ,  $u_b$ , and  $u_c$  are the phase voltages.  $i_a$ ,  $i_b$ , and  $i_c$  are the phase currents.

Characterization of the measurements of voltage and current is proposed to obtain features to be used as inputs of classifier. In this study, S-matrix is obtained by S-transform of residual voltage and current. According to S-matrix of incipient fault signal, several important information in terms of magnitude, phase, and frequency can be extracted. Many features such as amplitude, variance, mean, standard deviation, and energy of the transformed signal are widely used for proper classification. In this study, features are as follows.

The first set of features: The maximum value of the absolute S-matrix of residual current and voltage, which are F1 and F5. The second set of features: The minimum value of the absolute S-matrix of residual current and voltage, which are F2 and F6. The third set of features: Standard deviation of the data set comprising of the elements corresponding to maximum magnitude of each column of the S-matrix, which are F3 and F7. The fourth set of features: Mean value of the data set comprising of the elements corresponding to maximum magnitude of each column of the S-matrix, which are F4 and F8.

The overall features selected for characterizing the residual voltage and current waveforms are shown in Table 150.1.

## 150.4 Classification for Incipient Fault and Normal Switching in the Distribution Network

In this section, the LS-SVM is used to classify the incipient fault and normal disturbance, such as capacitor switching and load switching. The LS-SVM model is finally given as follows [16]:

$$y(x) = \text{sign} \left[ \sum_{i=1}^N \alpha_i y_i K(x, x_i) + b \right] \quad (150.4)$$

$K(x, x_i)$  is the kernel function. In our experiment, we choose the radial basis function (RBF) kernel where  $\delta^2$  is the bandwidth of the RBF kernel as our kernel, because it tends to achieve better performance. In determining the kernel bandwidth,  $\delta^2$  and the margin  $\gamma$  are set at 0.4 and 10, respectively. In order to fit the requirements of LS-SVM classifier, the training sample data must be collected and preprocessed before inputted into the classifier. By simulation using the ATP/EMTP, 600 numbers of residual voltage and current waveforms for incipient faults and normal disturbances were obtained for training LS-SVM. The features for all the waveforms were extracted from the S-transform. These features were used in developing the training database for the LS-SVM. Then, 300 numbers of residual voltage and current waveforms for incipient faults and normal disturbances were used for testing the training model. The results for test using LS-SVM classifier for input feature data from S-transform and wavelet transform are shown in Tables 150.2 and 150.3.

From Table 150.2, it is shown that LS-SVM classifier has good performance for these features extracted from S-transform; the correct classification rate is 90 %, which is higher than wavelet transform in Table 150.3. So S-transform-based incipient fault detection demonstrates the validity compared with wavelet transform widely used in power system.

**Table 150.2** The classification result using features from S-transform

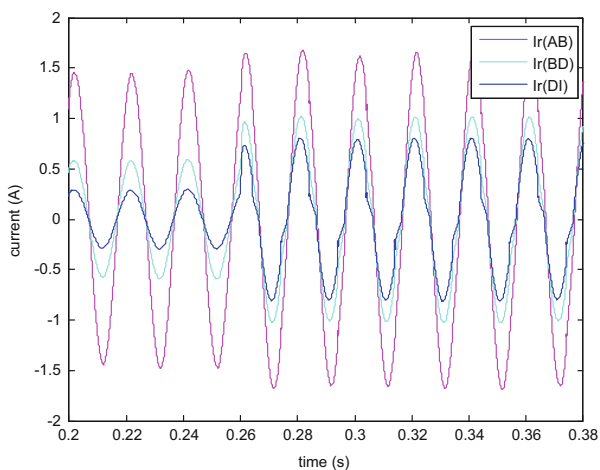
Disturbance type	No. of data	Correct classification	Accuracy (%)
Incipient fault	100	90	90
Capacitor switching	100	96	96
Load switching	100	93	93

**Table 150.3** The classification result using features from wavelet transform

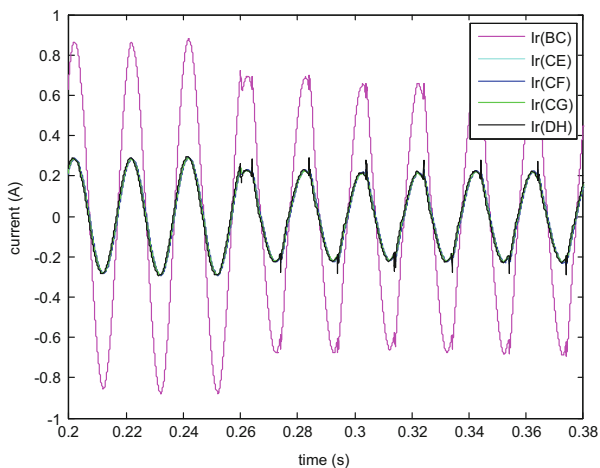
Disturbance type	No. of data	Correct classification	Accuracy (%)
Incipient fault	100	87	87
Capacitor switching	100	95	95
Load switching	100	94	94

## 150.5 Fault Section Discrimination for Incipient Fault in the Distribution Network

Once incipient fault is detected, it is necessary to find the position in power distribution network. In this section, a location method for incipient fault is proposed. Suppose the incipient fault occurs in the section DI. When the incipient fault occurs at 0.26 s, the corresponding residual current waveforms for these branches are shown in Figs. 150.2 and 150.3. We know that pre-fault residual current magnitude of some sections is lower than that during the fault, which is drawn in Fig. 150.2, and pre-fault residual current magnitude of other sections is larger than that during the fault, which is drawn in Fig. 150.3.



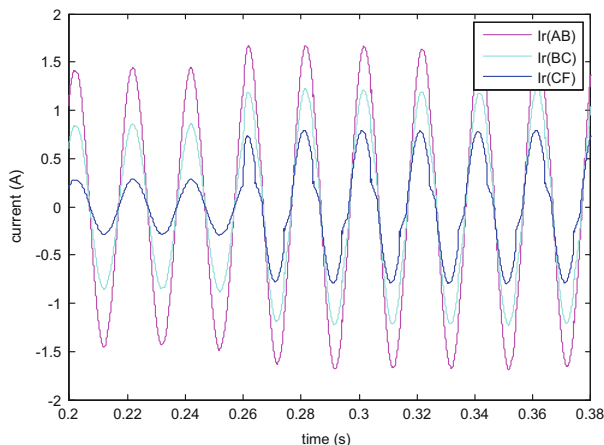
**Fig. 150.2** The section of residual current magnitude during fault large than pre-fault when fault in DI



**Fig. 150.3** The section of residual current magnitude during fault lower than pre-fault



**Fig. 150.4** The section of residual current magnitude during fault large than pre-fault when fault in CF



When the incipient fault occurred at 0.26 s, from Figs. 150.2 to 150.3, it is obvious that the residual current magnitude of every section has changed so much. Some magnitudes are bigger than pre-fault, and others are smaller than pre-fault. So this change is the difference between the residual current magnitude during and pre-fault measurements, which is described as

$$\Delta i_r = i_{r\_during} - i_{r\_pre} \quad (150.5)$$

Through the simulation in Fig. 150.1, it is well known that if  $\Delta i_r > 0$ , like the case of Fig. 150.2, it is the case with faulty section, and otherwise, if  $\Delta i_r < 0$ , like the case of Fig. 150.3, it is the case of healthy section. So, the positive or negative of  $\Delta i_r$  is considered as the discrimination rule to the status of power line.

The case was studied for incipient fault occurring at section CF. The corresponding residual current magnitude difference between during and pre-fault higher zero are shown in Fig. 150.4. By analyzing the Fig. 150.4, when the incipient fault occurred in section CF, the residual current magnitude of section AB is higher, and it is slightly reduced for each downstream section BC and CF. But for the healthy section, the residual current magnitude during the fault is lower than that of pre-fault.

## 150.6 Conclusion

In this chapter, a complete incipient fault diagnosis scheme is proposed. A novel technique based on S-transform of residual voltage and current is proposed to extract the features of incipient faults. According to these features, LS-SVM classifier is used to distinguish between fault status and normal disturbance. By analyzing the magnitude difference between during the fault and pre-fault, an estimation rule is

proposed to determine the faulty section in the distribution network. The proposed scheme performance showed its high reliability and robustness to a wide variety of simulated tests.

## References

1. Sidhu, T. S., & Xu, Z. H. (2010). Detection of incipient faults in distribution underground cables. *IEEE Transactions on Power Delivery*, 25(3), 1363–1371.
2. Kim, C. J., Jeong Hoon Shin, Myeong-Ho Yoo, & Gi Won Lee. (1999). A study on the characterization of the incipient failure behavior of insulators in power distribution line. *IEEE Transactions on Power Delivery*, 14(2), 519–524.
3. Benner, C. L., & Russell, B. D. (1997). Practical high-impedance fault detection on distribution feeders. *IEEE Transactions on Industry Applications*, 33(3), 635–640.
4. Al-Dabbagh, M., & Al-Dabbagh, L. (1999). Neural networks based algorithm for detecting high impedance faults on power distribution lines. *Proceedings of the International Joint Conference on Neural Networks*, 5(1), 3386–3390.
5. Cui, T., Dong, X. Z., Bo, Z. Q., & Andrzej Juszczyk (2011). Hilbert transform based transient/intermittent earth fault detection in noneffectively grounded distribution systems. *IEEE Transactions on Power Delivery*, 26(1), 143–151.
6. Butler, K. L. (1996). An expert system based framework for an incipient failure detection and predictive maintenance system. In *Proceedings of the 1996 intelligent systems application to power systems conference* (pp. 321–326). Orlando, FL.
7. Miri, S. M., & Privette, A. (1994). A survey of incipient fault detection and location techniques for extruded shielded power cables. In *26th annual Southeastern symposium on system theory* (pp. 402–405). Athens, OH.
8. Kim, C.J., Seung-Jae Lee, & Sang-Hee Kang (2004). Evaluation of feeder monitoring parameters for incipient fault detection using Laplace trend statistic. *IEEE Transactions on Industry Applications*, 40(6), 1718–1724.
9. Huang, S. J., & Hsieh, C. T. (1999). High-impedance fault detection utilizing a Morlet wavelet transform approach. *IEEE Transactions on Power Delivery*, 14(4), 1401–1410.
10. Prikler, L., & Hoildalen, H. (1998). *ATPDraw users' manual*. SINTEF TR A4790.
11. Darwish, H., & Elkalashy, N. (2005). Universal arc representation using EMTP. *IEEE Transactions on Power Delivery*, 20(2), 774–779.
12. Kizilcay, M., & Pniok, T. (1991). Digital simulation of fault arcs in power systems. *European Transactions on Electrical Power*, 1(1), 55–59.
13. Stockwell, R. G., Mansinha, L., & Lowe, R. P. (1996). Localization of the complex spectrum: The S-transform. *IEEE Transactions on Signal Processing*, 44(4), 998–1001.
14. Faisal, M.F., Azah Mohamed, & Aini Hussain (2009). S-transform based support vector regression for detection of incipient faults and voltage disturbances in power distribution networks. In *Proceedings of the 11th WSEAS international conference on mathematical methods, computational techniques and intelligent systems* (pp. 139–145). La Laguna.
15. Li, B., Zhang, P.L., Tian, H., Mi, S.S., Liu, D.S., & Ren, G.Q. (2011). A new feature extraction and selection scheme for hybrid fault diagnosis of gearbox. *Expert Systems with Applications*, 38(8), 10000–10009.
16. Suykens, J. A. K. (2002). *Least square support vector machines* (pp. 64–75). Singapore: World Scientific.

# Chapter 151

## Network Communication Forming Coalition

### S4n-Knowledge Model Case

Takashi Matsuhisa

**Abstract** This paper is to introduce the new concept of coalition Nash equilibrium of a strategic game. A coalition Nash equilibrium for a strategic game consists of (1) a subset  $S$  of players, (2) independent mixed strategies for each member of  $S$ , and (3) the conjecture of the actions for the other players not in  $S$  with the condition that each member of  $S$  maximizes his/her expected payoff according to the product of all mixed strategies for  $S$  and the other players' conjecture. Let us consider that each player communicates privately not only his/her belief about the others' actions but also his/her rationality as messages according to a protocol and then the recipient updates their private information and revises her/his prediction. Then we show that the conjectures of the players in a coalition  $S$  regarding the future beliefs converge in the long run communication, which lead to a coalition Nash equilibrium for the strategic game.

### 151.1 Introduction

For few decades, researchers in economics, AI, and computer science become entertained lively concerns about relationships between knowledge and actions. At what point does an economic agent sufficiently know to stop gathering information and make decisions? There are also concerns about cooperation and knowledge. What is the role of sharing knowledge to making cooperation among agents?

Considering a coalition among agents, we tacitly understand that each agents in the coalition shares their individual information and so they commonly know each other. In mathematical point of view yet a little is known what structure they have to know commonly. The aim of this paper is to fill the gap. Our point is that in a

---

T. Matsuhisa (✉)

Department of Natural Sciences, Ibaraki National College of Technology, Nakane 866,  
Hitachinaka-shi, Ibaraki, 312-8508, Japan  
e-mail: [mathisa@ge.ibaraki-ct.ac.jp](mailto:mathisa@ge.ibaraki-ct.ac.jp)

coalition, the members do not necessarily have common-knowledge about each other but they communicate his/her own beliefs on the others to each other through messages.

The purposes of this paper are to introduce the concept of coalition Nash equilibrium of a strategic game and to show that a communication among the players in a coalition leads to the equilibrium through messages. A coalition Nash equilibrium for a strategic game consists of (1) a subset  $S$  of players, (2) independent mixed strategies for each member of  $S$ , and (3) the conjecture about the actions for the players outside of  $S$  maximizes his/her expected payoff according to all mixed strategies for all members in  $S$  together with the conjectures of all the players outside of  $S$ .

This paper analyzes the solution concept from the Bayesian point of view: The players start with the same prior distribution on a state-space. In addition they have private information which is given by a reflexive and transitive binary relation on the state-space. Each player in a coalition  $S$  predicts the other players' actions as the posterior of the others' actions given his/her information. He/she communicates privately their beliefs about the other players' actions through messages among all members in  $S$  according to the communication network in  $S$ , which message is information about his/her individual conjecture about the others' actions. The recipients update their belief by the messages. Precisely, at every stage each player communicates privately not only his/her belief about the others' actions but also his/her rationality as messages according to a protocol, and then the recipient updates their private information and revises her/his prediction. In this circumstance, we shall show that

**Main theorem** *In a communication process of the game according to a protocol with revisions of their beliefs about the other players' actions, the profile of their future predictions converges to a coalition Nash equilibrium of the game in the long run.*

This paper is organized as follows. In Sect. 151.2, after recalling the knowledge structure associated with an RT-information structure, we present a game on knowledge structure. The communication process for the game is introduced where the players send messages about their conjectures about the other players' action. In Sect. 151.3 we give the formal statement of the main theorem (Theorem 1) with an illustrated example, and we conclude with some remarks.

## 151.2 The Model

Let  $\Omega$  be a *state-space* which is a nonempty *finite* set,  $N$  a set of finitely many *players*  $\{1, 2, \dots, n\}$  at least two ( $n \geq 2$ ), and let  $2^\Omega$  be the family of all subsets of  $\Omega$ . Each member of  $2^\Omega$  is called an *event* and each element of  $\Omega$  called a *state*. Let  $\mu$  be

a common probability measure on  $\Omega$  for all players. For simplicity it is assumed that  $(\Omega, \mu)$  is a *finite* probability space with  $\mu$  *full support*.<sup>1</sup>

### 151.2.1 Information and Knowledge<sup>2</sup>

By an *RT-information structure*<sup>3</sup> we mean  $\langle \Omega, (\Pi_i)_{i \in \bar{N}} \rangle$  in which  $\Pi_i : \Omega \rightarrow 2^\Omega$  satisfies the three postulates: For each  $i \in \bar{N}$  and for any  $\omega \in \Omega$ ,

**Ref**  $\omega \in \Pi_i(\omega)$ ;    **Trn**  $\xi \in \Pi_i(\omega)$  implies  $\Pi_i(\xi) \subseteq \Pi_i(\omega)$ ;

This structure is equivalent to a Kripke semantics for the multi-modal logic **S4n**. The set  $\Pi_i(\omega)$  will be interpreted as the set of all the states of nature that  $i$  knows to be possible at  $\omega$ , or as the set of the states that  $i$  cannot distinguish from  $\omega$ . We will therefore call  $\Pi_i$   $i$ 's *possibility operator* on  $\Omega$  and also will call  $\Pi_i(\omega)$   $i$ 's *information set* at  $\omega$ .

If the RT-information structure satisfies the below postulate

**Sym** If  $\xi \in \Pi_i(\omega)$ , then  $\omega \in \Pi_i(\xi)$ ,

it is called an *information partition*, which is equivalent to a Kripke semantics for the multi-modal logic **S5n**.

Our interpretation is given as a player  $i$  for whom  $\Pi_i(\omega) \subseteq E$  knows, in the state  $\omega$ , that some state in the event  $E$  has occurred, and then we say that in the state  $\omega$  the player  $i$  knows  $E$ . It follows

**Definition 1** The *knowledge structure*  $\langle \Omega, (\Pi_i)_{i \in N}, (K_i)_{i \in N} \rangle$  consists of an RT-information structure  $\langle \Omega, (\Pi_i)_{i \in N} \rangle$  and a class of  $i$ 's *knowledge operator*  $K_i$  on  $2^\Omega$  such that  $K_i E = \{\omega \in \Omega \mid \Pi_i(\omega) \subseteq E\}$ .

The set  $K_i E$  will be interpreted as the set of states of nature for which  $i$  knows  $E$  to be possible.

We record the properties of  $i$ 's knowledge operator<sup>4</sup>: For every  $E, F$  of  $2^\Omega$ ,

**N**  $K_i \Omega = \Omega$  and  $K_i \emptyset = \emptyset$ ;    **K**  $K_i(E \cap F) = K_i E \cap K_i F$ ;

**T**  $K_i F \subseteq F$ ;    **4**  $K_i F \subseteq K_i K_i F$ ;

$i$ 's possibility operator  $\Pi_i$  is uniquely determined by  $i$ 's knowledge operator  $K_i$  satisfying the above four properties as  $\Pi_i(\omega) = \bigcap_{\omega \in K_i E} E$ . If  $\Pi_i$  satisfies further **Sym**, then the below property is also true:

**5**  $\Omega \setminus K_i(E) \subseteq K_i(\Omega \setminus K_i(E))$ .

<sup>1</sup> That is,  $\mu(\omega) \neq 0$  for every  $\omega \in \Omega$ .

<sup>2</sup> C.f.; Binmore [2] for information and knowledge.

<sup>3</sup> RT-information stands for a reflexive and transitive information.

<sup>4</sup> According to these we can say the structure  $\langle \Omega, (K_i)_{i \in N} \rangle$  is a model for the multi-modal logic **S4n**.

### 151.2.2 Game and Knowledge<sup>5</sup>

In this paper, by a *game*  $G$  we always mean a *finite* strategic form game  $\langle N, (A_i)_{i \in N}, (g_i)_{i \in N} \rangle$  with the following structure and interpretations:  $N$  is a finite set of players  $\{1, 2, \dots, i, \dots, n\}$  with  $n \geq 2$ ,  $A_i$  is a finite set of  $i$ 's *actions* (or  $i$ 's pure strategies), and  $g_i$  is an  $i$ 's *payoff function* of  $A$  into  $\mathbb{R}$ , where  $A$  denotes the product  $A_1 \times A_2 \times \dots \times A_n$ ,  $A_{-i}$  the product  $A_1 \times A_2 \times \dots \times A_{i-1} \times A_{i+1} \times \dots \times A_n$ . We denote by  $g$  the  $n$ -tuple  $(g_1, g_2, \dots, g_n)$  and by  $a_{-i}$  the  $(n - 1)$ -tuple  $(a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_n)$  for  $a$  of  $A$ . Furthermore we denote  $a_{-I} = (a_i)_{i \in N \setminus I}$  for each  $I \subset N$ . A probability distribution  $\sigma_i$  on  $A_i$  is called an  $i$ 's *mixed strategy* for a game  $G$ . We denote by  $\Delta(A_i)$  the set of all  $i$ 's mixed strategies, so we will denote  $\Delta(A) = \prod_{i=1}^n \Delta(A_i)$  and  $\Delta(A_I) = \prod_{i \in I} \Delta(A_i)$ .

A profile  $(\sigma_i)_{i \in N}$  of mixed strategies is called a *Nash equilibrium* if for each  $i \in N$  and for every  $b_i \in A_i$ , we have

$$\sum_{a_{-i} \in A_{-i}} g_i(a_i, a_{-i}) \prod_{j \in N \setminus \{i\}} \sigma_j(a_j) \geq \sum_{a_{-i} \in A_{-i}} g_i(b_i, a_{-i}) \prod_{j \in N \setminus \{i\}} \sigma_j(a_j)$$

By  $i$ 's *overall conjecture* (or simply  $i$ 's *conjecture*) we mean a probability distribution  $\varphi_i \in \Delta(A_{-i}) = \Delta(A_{N \setminus \{i\}})$ . For each player  $j$  other than  $i$ , the marginal distribution on  $j$ 's actions is called  $i$ 's *individual conjecture* about  $j$  (or simply  $i$ 's *conjecture about  $j$* .) Functions on  $\Omega$  are viewed like random variables in the probability space  $(\Omega, \mu)$ . If  $\mathbf{x}$  is a such function and  $x$  is a value of it, we denote by  $[\mathbf{x} = x]$  (or simply by  $[x]$ ) the set  $\{\omega \in \Omega \mid \mathbf{x}(\omega) = x\}$ .

The information structure  $(\Pi_i)$  with a common prior  $\mu$  yields the distribution on  $A \times \Omega$  defined by  $\mathbf{q}_i(a, \omega) = \mu([\mathbf{a} = a] \mid \Pi_i(\omega))$ ; and the  $i$ 's overall conjecture defined by the marginal distribution  $\mathbf{q}_i(a_{-i}, \omega) = \mu([\mathbf{a}_{-i} = a_{-i}] \mid \Pi_i(\omega))$  which is viewed as a random variable of  $\varphi_i$ . We denote by  $[\mathbf{q}_i = \varphi_i]$  the intersection  $\bigcap_{a_{-i} \in A_{-i}} [\mathbf{q}_i(a_{-i}) = \varphi_i(a_{-i})]$  and denote by  $[\varphi]$  the intersection  $\bigcap_{i \in N} [\mathbf{q}_i = \varphi_i]$ . Let  $\mathbf{g}_i$  be a random variable of  $i$ 's payoff function  $g_i$  and  $\mathbf{a}_i$  a random variable of an  $i$ 's action  $a_i$ .

According to the Bayesian decision theoretical point of view, we assume that each player  $i$  absolutely knows his/her own actions; i.e., letting  $[a_i] := [\mathbf{a}_i = a_i]$ ,  $[a_i] = K_i([a_i])$  (or equivalently,  $\Pi_i(\omega) \subseteq [a_i]$  for all  $\omega \in [a_i]$  and for every  $a_i$  of  $A_i$ .)  $i$ 's action  $a_i$  is said to be *actual* at a state  $\omega$  if  $\omega \in [a_i]$ ; and the profile  $a_I$  is said to be actually played at  $\omega$  if  $\omega \in [\mathbf{a}_I = a_I] := \bigcap_{i \in I} [\mathbf{a}_i = a_i]$  for  $I \subset N$ . The payoff functions  $g = (g_1, g_2, \dots, g_n)$  is said to be *actually played* at a state  $\omega$  if  $\omega \in [\mathbf{g} = g] := \bigcap_{i \in N} [\mathbf{g}_i = g_i]$ . Let **Exp** denote the expectation defined by  $\mathbf{Exp}(g_i(b_i, \mathbf{a}_{-i}); \omega) := \sum_{a_{-i} \in A_{-i}} g_i(b_i, a_{-i}) \mathbf{q}_i(a_{-i}, \omega)$ .

<sup>5</sup> C.f., Aumann and Brandenburger [1].

By a *coalition*  $S$  we mean  $S$  is a non empty subset of  $N$ . Let  $(\sigma_i)_{i \in S}$  be the profiles of mixed strategies of  $G$  for a coalition  $S$ . By *S-expectation* of  $i$ 's payoff function  $g_i$  at  $\omega$  we mean

$$\mathbf{Exp}_S(g_i(a_S, \mathbf{a}_{-S}); \omega) := \sum_{a_{-S} \in A_{-S}} g_i(a_S, a_{-S}) \left( \prod_{i \in S} \sigma_i \right) \mathbf{q}_i(a_{-S}, \omega).$$

**Definition 2** A profile  $(\sigma_i)_{i \in S}$  is called a *coalition S-Nash equilibrium* of  $G$  if each member  $i$  in  $S$  maximizes his/her  $\mathbf{Exp}_S(g_i(a_S, \mathbf{a}_{-S}); \omega)$  for every  $\omega \in \Omega$ ; i.e.:  $\mathbf{Exp}_S(g_i(a_S, \mathbf{a}_{-S}); \omega) \geq \mathbf{Exp}_S(g_i(b_S, \mathbf{a}_{-S}); \omega)$  for every  $b_S$  in  $A_S$ .

A coalition  $S$  is called *rational* at  $\omega$  if for every  $i \in S$ , each  $i$ 's actual action  $a_i$  maximizes the expectation of his actually played payoff function  $g_i$  at  $\omega$  when the other players' actions are distributed according to his conjecture  $\mathbf{q}_i(\cdot; \omega)$ . Formally, letting  $g_i = \mathbf{g}_i(\omega)$  and  $a_i = \mathbf{a}_i(\omega)$ ,  $\mathbf{Exp}(g_i(a_i, \mathbf{a}_{-i}); \omega) \geq \mathbf{Exp}(g_i(b_i, \mathbf{a}_{-i}); \omega)$  for every  $b_i$  in  $A_i$ . Let  $R_i$  denote the set of all of the states at which  $i$  is rational.

### 151.2.3 Protocol<sup>6</sup>

We assume that the players communicate by sending *messages*. Let  $T$  be the time horizontal line  $\{0, 1, 2, \dots, t, \dots\}$ . A *protocol* on a coalition  $S$  of a game  $G$  is a mapping  $\text{Pr}_S : T \rightarrow S \times S, t \mapsto (s(t), r(t))$  such that  $s(t) \neq r(t)$ . Here  $t$  stands for *time* and  $s(t)$  and  $r(t)$  are, respectively, the *sender* and the *recipient* of the communication which takes place at time  $t$ . Simply we call it an *S-protocol*. We consider the protocol as the directed graph whose vertices are the set of all members in  $S$  and such that there is an edge (or an arc) from  $i$  to  $j$  if and only if there are infinitely many  $t$  such that  $s(t) = i$  and  $r(t) = j$ .

In this paper a protocol  $\text{Pr}_S$  is assumed to be *fair*; that is, the graph is strongly connected: In words, every player in this protocol communicates directly or indirectly with every other player infinitely often. It is said to contain a *cycle* if there are players  $i_1, i_2, \dots, i_k$  with  $k \geq 3$  such that for all  $m < k$ ,  $i_m$  communicates directly with  $i_{m+1}$ , and such that  $i_k$  communicates directly with  $i_1$ . The communications are assumed to proceed in *rounds*.

### 151.2.4 Communication on Coalition

Let  $S$  be a coalition of  $G$ . A *coalition S-communication process*  $\pi_S(G)$  with revisions of players' conjectures  $(\varphi_i^t)_{(i, t) \in S \times T}$  according to a protocol for a game  $G$  is a tuple

<sup>6</sup> C.f.: Parikh and Krasucki [4].

$$\pi_S(G) = \langle G, (\Omega, \mu), \text{Pr}_S, (\Pi_i^t)_{i \in S}, (K_i^t)_{i \in S}, (\varphi_i^t)_{(i,t) \in S \times T} \rangle$$

with the following structures: the players have a common prior  $\mu$  on  $\Omega$ , the protocol  $\text{Pr}_S$  among  $N$ ,  $\text{Pr}_S(t) = (s(t), r(t))$ , is fair and it satisfies the conditions that  $r(t) = s(t + 1)$  for every  $t$  and that the communications proceed in rounds. The revised information structure  $\Pi_i^t$  at time  $t$  is the mapping of  $\Omega$  into  $2^\Omega$  for player  $i \in S$ . If  $i = s(t)$  is a sender at  $t$ , the *message* sent by  $i$  to  $j = r(t)$  is  $M_i^t$ . An  $n$ -tuple  $(\varphi_i^t)_{i \in S}$  is a revision process of individual conjectures. These structures are inductively defined as follows:

- Set  $\Pi_i^0(\omega) = \Pi_i(\omega)$ .
- Assume that  $\Pi_i^t$  is defined. It yields the distribution  $\mathbf{q}_i^t(a, \omega) = \mu([\mathbf{a} = a] | \Pi_i^t(\omega))$ . Whence
  - $R_i^t$  denotes the set of all the state  $\omega$  at which  $i$  is *rational* according to his conjecture  $\mathbf{q}_i^t(\cdot; \omega)$ <sup>7</sup>;
  - The message  $M_i^t : \Omega \rightarrow 2^\Omega$  sent by the sender  $i$  at time  $t$  is defined by

$$M_i^t(\omega) = \bigcap_{a_{-i} \in A_{-i}} \{ \xi \in \Omega \mid \mathbf{q}_i^t(a_{-i}, \xi) = \mathbf{q}_i^t(a_{-i}, \omega) \}.$$

Then:

- The revised knowledge operator  $K_i^t : 2^\Omega \rightarrow 2^\Omega$  is defined by  $K_i^t(E) = \{ \omega \in \Omega \mid \Pi_i^t(\omega) \subseteq E \}$ .
- The revised partition  $\Pi_i^{t+1}$  at time  $t + 1$  is defined as follows:  $\Pi_i^{t+1}(\omega) = \Pi_i^t(\omega) \cap M_{s(t)}^t(\omega)$  if  $i = r(t)$ , and  $\Pi_i^{t+1}(\omega) = \Pi_i^t(\omega)$  otherwise,
- The revision process  $(\varphi_i^t)_{(i,t) \in S \times T}$  of conjectures is inductively defined as follows: Let  $\omega_0 \in \Omega$ , and set  $\varphi_{s(0)}^0(a_{-s(0)}) := \mathbf{q}_{s(0)}^0(a_{-s(0)}, \omega_0)$ . Take  $\omega_1 \in M_{s(0)}^0(\omega_0) \cap K_{r(0)}([\varphi_{s(0)}^0] \cap R_{s(0)}^0)$ ,<sup>8</sup> and set  $\varphi_{s(1)}^1(a_{-s(1)}) := \mathbf{q}_{s(1)}^1(a_{-s(1)}, \omega_1)$ . Take  $\omega_{t+1} \in M_{s(t)}^t(\omega_t) \cap K_{r(t)}([\varphi_{s(t)}^t] \cap R_{s(t)}^t)$ ,<sup>9</sup> and set  $\varphi_{s(t+1)}^{t+1}(a_{-s(t+1)}) := \mathbf{q}_i^{t+1}(a_{-s(t+1)}, \omega_{t+1})$ .

The specification is that a sender  $s(t)$  at time  $t$  informs the recipient  $r(t)$  his/her prediction about the other players' actions as information of his/her individual conjecture. The recipient revises her/his information structure under the information. She/he predicts the other players' action at the state where the player knows

<sup>7</sup> That is, each  $i$ 's actual action  $a_i$  maximizes the expectation of his payoff function  $g_i$  being actually played at  $\omega$  when the other players actions are distributed according to his conjecture  $\mathbf{q}_i^t(\cdot; \omega)$  at time  $t$ . Formally, letting  $g_i = \mathbf{g}_i(\omega)$ ,  $a_i = \mathbf{a}_i(\omega)$ , the expectation at time  $t$ ,  $\mathbf{Exp}^t$ , is defined by  $\mathbf{Exp}^t$

$(g_i(a_i, \mathbf{a}_{-i}); \omega) := \sum_{a_{-i} \in A_{-i}} g_i(a_i, a_{-i}) \mathbf{q}_i^t(a_{-i}, \omega)$ . A player  $i \in S$  is said to be  $S$ -rational according to his conjecture  $\mathbf{q}_i^t(\cdot, \omega)$  at  $\omega$  if for all  $b_i$  in  $A_i$ ,  $\mathbf{Exp}^t(g_i(a_i, \mathbf{a}_{-i}); \omega) \geq \mathbf{Exp}^t(g_i(b_i, \mathbf{a}_{-i}); \omega)$ .

<sup>8</sup> We denote  $[g_i] := [\mathbf{g}_i = g_i]$ .

<sup>9</sup> It is noted that  $\omega_{t+1} \in M_{s(t)}^t(\omega_t) \cap K_{r(t)}([\varphi_{s(t)}^t] \cap R_{s(t)}^t) \neq \emptyset$  for every  $t = 0, 1, 2, \dots$ .



that the sender  $s(t)$  is rational, and she/he informs her/his the predictions to the other player  $r(t + 1)$ .

We denote by  $\infty$  a sufficiently large  $\tau \in T$  such that for all  $\omega \in \Omega, \mathbf{q}_i^\tau(\cdot; \omega) = \mathbf{q}_i^{\tau+1}(\cdot; \omega) = \mathbf{q}_i^{\tau+2}(\cdot; \omega) = \dots$ . Hence we can write  $\mathbf{q}_i^\tau$  by  $\mathbf{q}_i^\infty$  and  $\varphi_i^\tau$  by  $\varphi_i^\infty$ .

### 151.3 The Result

We can now state the main theorem, and we will give an example to illustrate it.

**Theorem 1** *Let  $G$  be a strategic form game. Suppose that the players in  $G$  have the knowledge structure with  $\mu$  a common prior. Let  $S$  be a coalition in a game  $G$ . In a coalition  $S$ -communication process  $\pi_S(G)$  according to an  $S$ -protocol  $\text{Pr}_S$  among all members in  $S$ , the  $|S|$ -tuple of their conjectures  $(\varphi_i^t)_{(i,t) \in S \times T}$  converges to a coalition  $S$ -Nash equilibrium of the game in finitely many rounds.*

*Proof* will be omitted, and it will be appeared elsewhere.

*Example 1* Let us consider the game  $G = \langle N, (A_i)_{i \in N}, (g_i)_{i \in N} \rangle$  as follows:

- The set of players  $N = \{ 1, 2, 3 \}$ ;
- The action sets  $A_1 = \{H, T\}, A_2 = \{H, T\}, A_3 = \{W, E\}$ ;
- The payoff functions  $g_1, g_2, g_3$  are given in Table 151.1:

The game  $G$  has the unique Nash equilibrium  $(\frac{1}{2}H + \frac{1}{2}T, \frac{1}{2}h + \frac{1}{2}t, W)$ , and let us reconsider the situation as follows: *Each player knows his/her own actions, but he/she cannot know the other players' action.* To model the situation we introduce the game  $G$  as a Bayesian game equipped with the below information partition  $(\Pi_i)_{i = 1, 2, 3}$ :

- $\Omega = \{\omega_1, \omega_2, \dots, \omega_8\}$  with the equal probability  $\mu$ ; i.e.,  $\mu(\omega) = \frac{1}{8}$ ;
- The partitions  $(\Pi_i)_{i = 1, 2, 3}$  on  $\Omega$ :
  - The partition  $\Pi_1: \Pi_1(\omega) = \{\omega_1, \omega_2, \omega_5, \omega_6\}, \Pi_1(\omega) = \{\omega_3, \omega_4, \omega_7, \omega_8\}$ ,
  - The partition  $\Pi_2: \Pi_2(\omega) = \{\omega_1, \omega_3, \omega_5, \omega_7\}, \Pi_2(\omega) = \{\omega_2, \omega_4, \omega_6, \omega_8\}$ .
  - The partition  $\Pi_3: \Pi_3(\omega) = \{\omega_1, \omega_2, \omega_3, \omega_4\}, \Pi_3(\omega) = \{\omega_5, \omega_6, \omega_7, \omega_8\}$ .
- $\mathbf{a}_i$  is defined by

**Table 151.1** Game G in Example 1

W	h	t	E	h	t
H	1, 0, 2	0, 1, 2	H	1, 0, 3	0, 1, 0
T	0, 1, 2	1, 0, 2	T	0, 1, 0	1, 0, 3

$$\begin{aligned} \mathbf{a}_1(\omega) &= \mathbf{H} \text{ for } \omega = \omega_i(i = 1, 2, 5, 6), & \mathbf{a}_1(\omega) &= \mathbf{T} \text{ for } \omega = \omega_i(i = 3, 4, 7, 8). \\ \mathbf{a}_2(\omega) &= \mathbf{h} \text{ for } \omega = \omega_i(i = 1, 3, 5, 7), & \mathbf{a}_2(\omega) &= \mathbf{t} \text{ for } \omega = \omega_i(i = 2, 4, 6, 8). \\ \mathbf{a}_3(\omega) &= \mathbf{W} \text{ for } \omega = \omega_i(i = 1, 2, 3, 4), & \mathbf{a}_3(\omega) &= \mathbf{E} \text{ for } \omega = \omega_i(i = 5, 6, 7, 8). \end{aligned}$$

We can observe that the conjectures  $\varphi_i(a_j) = \mathbf{q}_i(a_j; \omega_5)$  at  $\omega_5$  are:  $\varphi_2(a_1) = \varphi_3(a_1) = \frac{1}{2}\mathbf{H} + \frac{1}{2}\mathbf{T}$ ,  $\varphi_1(a_2) = \varphi_3(a_2) = \frac{1}{2}\mathbf{h} + \frac{1}{2}\mathbf{t}$ ,  $\varphi_1(a_3) = \varphi_2(a_3) = \frac{1}{2}\mathbf{W} + \frac{1}{2}\mathbf{E}$ . This shows that for each player  $i$ , any other players than  $i$  must agree on every  $i$ 's actions, but these distributions  $(\varphi_3(a_1), \varphi_1(a_2), \varphi_2(a_3))$  cannot form the Nash equilibrium for  $G$ , but  $(\varphi_2(a_1), \varphi_1(a_2)) = (\varphi_3(a_1), \varphi_1(a_2))$  forms a coalition  $\{1, 2\}$ -Nash equilibrium. However  $(\varphi_3(a_1), \varphi_1(a_3))$  is not a  $\{1, 3\}$ -Nash equilibrium. It should be noted that  $(\varphi_3(a_1), \varphi_1(a_3))$  is commonly known among  $\{1, 3\}$ , and so the notion of common-knowledge cannot always yield a coalition Nash equilibrium.

Let us consider the coalition  $\{1, 3\}$ -communication process for game  $G$  with the protocol  $\text{Pr} : T \rightarrow S = \{1, 3\}$ . After two rounds in the communication, the below information partition can be obtained:

- The partitions  $(\Pi_i^\infty)_{i \in N}$  on  $\Omega$ :
  - The partition  $\Pi_1^\infty$  on  $\Omega$ :  $\Pi_1^\infty(\omega) = \{\omega_1, \omega_2\}, \Pi_1^\infty(\omega) = \{\omega_5, \omega_6\}, \Pi_1^\infty(\omega) = \{\omega_3, \omega_4\}, \Pi_1^\infty(\omega) = \{\omega_7, \omega_8\}$ .
  - The partition  $\Pi_2^\infty$  on  $\Omega$  is the same as the initial partition  $\Pi_2$ :
  - The partition  $\Pi_3^\infty$  on  $\Omega$  is the same as the initial partition  $\Pi_3$ :

We can observe that the conjectures  $\varphi_i^\infty(a_j) = \mathbf{q}_i^\infty(a_j; \omega)$  at  $\omega_5$  are:  $\varphi_2^\infty(a_1) = \varphi_3^\infty(a_1) = \frac{1}{2}\mathbf{H} + \frac{1}{2}\mathbf{T}$ ,  $\varphi_1^\infty(a_2) = \varphi_3^\infty(a_2) = \frac{1}{2}\mathbf{h} + \frac{1}{2}\mathbf{t}$ ,  $\varphi_1^\infty(a_3) = \varphi_2^\infty(a_3) = \mathbf{W}$ , and these distributions  $(\varphi_3^\infty(a_1), \varphi_1^\infty(a_3))$  are a  $\{1, 3\}$ -Nash equilibrium. Furthermore,  $(\varphi_2^\infty(a_1), \varphi_1^\infty(a_2), \varphi_1^\infty(a_3))$  forms the Nash equilibrium for  $G$ .

**Remark 1** When a coalition  $S$  is the ground coalition  $N$ , the above theorem shows that the conjectures of the players lead to a Nash equilibrium through communication (Matsuhisa [3]). The notion of common-knowledge controls to form a Nash equilibrium for a game (Aumann and Brandenburger [1]), but it no longer play a crucial role to form a coalition Nash equilibrium.

### 151.4 Conclusion

This paper points out that for coalition Nash equilibrium, common-knowledge cannot play such role as for mixed strategy Nash equilibrium. In fact, the profile of conjectures of a coalition may not yield a coalition Nash equilibrium even when the profile is commonly known among all the members of the coalition. To improve the situation we adopt the knowledge revisions model as a communication model. The main theorem shows that communication instead of common-knowledge plays an essential role to form a coalition Nash equilibrium. We have observed that in a

communication process with revisions of players' beliefs about the other actions among all the members in a coalition, their predictions induce a coalition Nash equilibrium of the game in the long run.

## References

1. Aumann, R. J., & Brandenburger, A. (1995). Epistemic conditions for mixed strategy Nash equilibrium. *Econometrica*, *63*, 1161–1180.
2. Binmore, K. (1992). *Fun and games*. xxx+642pp. Lexington, MA: D. C. Heath and Company.
3. Matsuhsa, T. (2000). Communication leading to mixed strategy Nash equilibrium I. In T. Maruyama (Ed.) *Mathematical economics*, Suri-Kaiseki-Kenkyusyo Kokyuroku, vol. 1165, (pp. 245–256).
4. Parikh, R., & Krasucki, P. (1990) Communication, consensus, and knowledge. *Journal of Economic Theory* *52*, 178–189.

# Chapter 152

## An Optimization Model of the Layout of Public Bike Rental Stations Based on B+R Mode

Liu He, Xuhong Li, and Dawei Chen

**Abstract** In order to find out the optimal layout of bike rental stations for B+R mode, a bi-level programming model combined of genetic algorithm and the joint model with mode split and traffic assignment model is built. The optimal layout plan can minimize the total travelling cost and facility cost. A case is used to test and verify the practicability of the model. The result shows that the model can effectively solve the layout problem of bike rental stations for B+R mode and can offer suggestions for related planning.

### 152.1 Introduction

The connection modes of metro stations are mainly walking, bus, and car. As the trip distance of walking is short, it can be regarded as direct passenger flow. People prefer to reach the trip destination directly considering refueling and parking. For bus, the door-to-door service is restricted by lines and much time is wasted walking to the bus stop and waiting there. For B+R mode, public bus expands the covering area that metro stations radiate on the one hand and offer a second connection for bus on the other hand. As a result, the attraction is improved along metro line and more indirect passenger flow is brought in. So the study of the layout of bike rental stations has an essential meaning for metro passenger flows.

Current researches into the layout of bike rental stations are focused on the analysis, forecast, and principles. Li-hui Li et al. put forward the principles as “Control total quantity, divide and sort, balance the scale, and adjust flexibly” and divide rental stations into five types: bus stops, public buildings, blocks, resting places, and schools. They point out that the scale of blocks should be the same as the others [1]. Zhenghao Li works out the capacities that each bike rental station needs

---

L. He (✉) • X. Li • D. Chen  
School of Transportation, Southeast University, Nanjing 210096, Jiangsu, China  
e-mail: [124675282@qq.com](mailto:124675282@qq.com)

with Markov chain model [2]. Xue Geng et al. get the number and size of bike rental station in Paris after calculating the average number of trip [3]. Jenn-Rong Lin and Ta-Hui Yang choose level of service as the factor that affects the layout of public bike rental station [4]. Zhonghua Wei et al. analyze the road density that R+B model needs [5]. Karel takes three European cities in his study and shows how public bike's trip distance, purpose, and car affect each other [6].

The chapter considers what influence public bike has on the passenger flow attraction of metro stations. Based on bi-level planning, an optimization model of the layout of bike rental stations is built and solved to offer support for making such layout decisions.

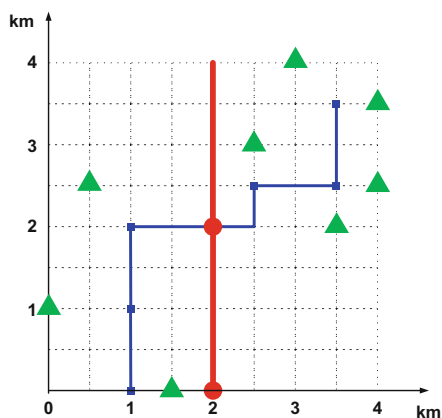
## 152.2 Modeling

Imagining there exists a road network illustrated in Fig. 152.1, bold lines are metro and bus lines and the rest are roads; dots are metro stations, diamonds are bus stops, and triangles are residential areas. As is illustrated in Fig. 152.2, public bike raises passenger attractions from the original two to five, with direct two and indirect one.

When planning public bike rental stations in the region, residential areas and bus stops are both alternative, and public bike is only allowed within the region. Starting from residential areas, travellers choose to get to metro station by car, by bus, or on foot before reaching the destination by metro. The trip modes to be chosen are illustrated in Fig. 152.3.

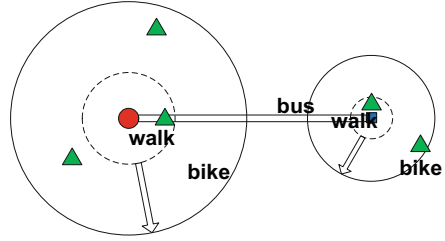
Based on the above assumptions and analysis, the question is brought forward: How to determine the layout of public bike rental stations so as to maximize the passenger flow volume of metro to the destination?

The layout optimization of public bike rental stations is a discrete network design problem (DNDP). Based on the DNDP theory, the chapter builds a planning model that is suitable for discrete transportation network design problem (DTNDP).

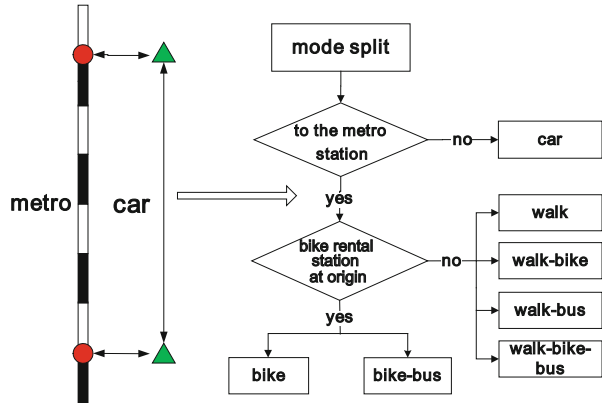


**Fig. 152.1** A sketch map of a certain region

**Fig. 152.2** The effect public bike has on the radiation



**Fig. 152.3** Trip modes within the area



The optimization model of the layout of public bike rental stations based on B+R mode is a two-layer model, in which the upper model is the generation and screening of the plan set, and the lower model is the evaluation of every single plan. With the feedback between the two models, the best layout plan that satisfies both layers' targets can be found.

The building of the lower level should fully describe users' trip modes and route selection on the existing network, that is, mode split and traffic assignment.

In the whole system, every traveller needs to decide the trip mode and route that cost the least. Considering the IIA (independence from irrelevant) among different trip modes [7], selections among all modes use mixed logit model [8, 9]:

$$P_{in} = \int \frac{\exp(\lambda V_{in})}{\sum_j \exp(\lambda V_{jn})} g(\beta/\theta) d\beta \tag{152.1}$$

$g(\beta/\theta)$  uses the normal distribution function [10] and the utility function uses linear form whose limbs including metro, walk, bicycle, public bicycle, car, and bus. The variables of trip mode are fare and travelling time while the variables of travellers are whether to have a car, whether to have a private bicycle, age, and trip purpose.

In the assignment model, every traveller selects the path whose travelling cost is the lowest. When the network reaches equilibrium, travellers cannot reduce their

costs by changing the selected mode and path. Though user equilibrium model achieves theoretical optimization, users actually make the decision according to their expected cost. In that case, stochastic user equilibrium (SUE) model is more suitable.

When planning the layout of bike rental stations, the government should consider the characteristics of regional travel demand. Then the optimal layout can be chosen to minimize the total travelling cost and facility cost.

At the same time, according to the principles for the layout of public bike rental stations, four constraints should be satisfied. The first is the distance between bike rental stations should be reasonable considering travelling distance of bicycle; the second is the density of rental stations should be within a certain range; the third is the proportion of public bike travelling should reach a certain standard; the fourth is each rental station should have enough space. Then constraints (152.3), (152.4), and (152.5) of the model can be found and listed below.

From all the above, the model of the upper level can be expressed as follows:

$$\max z(x) = \sum_{i \in R} Q_i - \sum_{i \in R} Q_i P_i^c \tag{152.2}$$

$$s.t. \quad d_{\min} \leq d_{ij} \leq d_{\max} \quad i, j \in T, \zeta_i = 1, \zeta_j = 1 \tag{152.3}$$

$$\rho_{\min} \leq \sum_{i \in T} \zeta_i / S \leq \rho_{\max} \tag{152.4}$$

$$\min(C_i) \geq Q^b_{\min} \quad \zeta_i = 1 \tag{152.5}$$

In constraint (152.3),  $d$  stands for distance and is the set of alternative public bike rental stations;  $\zeta$  means whether to build a rental station there and 1 means yes, while 0 no. In constraint (152.4),  $\rho$  is the density and  $S$  is the area of the region.  $C$  in constraint (152.5) means the capacity of the rental station.

The upper level is a system optimization problem whose objective function can get the best solution of the lower level.

### 152.3 Model Solving

The bi-level planning model is combined of the upper level from the aspect of the government and the lower level from the users. The government hopes that public bike can help improve the proportion of metro travelling while users want to minimize their broad travelling expenses, the average of time, and narrow travelling expenses. The final result is a best layout plan found to optimize the two objective functions.

When solving the lower level, a super network should be created to do mode spilt and traffic assignment in consideration of multi-modes.

When solving the upper level, an effective plan should be built for the lower layer according to the constraints. Genetic algorithm has an advantage in solving the upper level.

When doing mode spilt with mixed logit model, the difficulty lies in building utility matrix of different modes, whose solving depends on creating a super network. In addition, traffic assignment is needed to get each rental station's size. So the lower level is combined of two steps: mode spilt and traffic assignment. The detailed steps are listed below:

**Step 1: Creating the network.** A super network is created according to the layout of public bike rental stations.

**Step 2: Calculating utility.** The shortest paths for each OD of different modes are found and then used to calculate the utility matrix. Dijkstra algorithm can be used here. Penalty fees increase when transferring.

**Step 3: Mode spilt.** The travelling proportion of each mode is calculated with (152.1).

**Step 4: Traffic assignment.** Trace back with Dijkstra algorithm in Step 2 and distribute OD volume onto the super network. The scale of each rental station is obtained at the same time.

The location of public bike rental stations needs the support of modeling and quantitative method, and the bi-level planning model has the character of NP (nondeterministic polynomial time complete, NP-complete). The layout plan of the rental stations can be described with a series of discrete genetic data and then explored by crossover and mutation. Due to its outstanding advantage in solving such problems, genetic algorithm is applied to the upper level [11].

**One: Construction of Solutions** A  $n$ -piece gene cluster is used to stand for a set of layout plans of public bike rental stations. Each gene is composed an  $n$ -unit 0–1 variable. Each unit means whether to build a rental station in the location it stands for, and 0 means no, 1 means yes.

**Two: Generating Original Cluster**  $L$  individuals are generated randomly to compose the original cluster on condition that the coding plan is satisfied. The cluster is named  $G_0 = \{g_1, g_2, \dots, g_L\}$ , and its feasibility should be verified.

**Three: Fitness Function** Fitness function is one that measures the degree individuals approach the optimal solution. In the problem here, it measures each plan. The higher the value is, the closer it is to the optimal solution.

$$Fit\tilde{z}(x) = f - (f)_{\min}^k + \xi^k \quad (152.6)$$

where  $\tilde{z}(x)$  is the objective function of SO problem;  $f$  is the objective function value of the lower level;  $f_{\min}^k$  is the minimized objective function value of individuals in the  $k$ th generation; and  $\xi^k$  is the adjusted value of selection pressure, which decreases as  $k$  increases. It's set as follows:



$$\begin{cases} \xi^0 = M \\ \xi^k = c \cdot \xi^{k-1} \end{cases} \quad (152.7)$$

where  $f_{\min}^k$  is the smallest individual of the  $k$ th generation and  $M, c$  are constants,  $c \in [0.9, 0.999]$ .

**Four: Genetic Manipulation** There are three simple genetic manipulations of genetic algorithm: selection, crossover, and mutation. Selection is a process during which individuals with strong vitality produce new clusters; crossover means two homologous chromosomes recombine through mating to form a new chromosome during the evolution; and mutation means new chromosome is produced due to gene mutation and shows some new biological traits.

### 152.4 Case Study

Imagine the travel demand of all settlements in a region during a certain period as illustrated in Fig. 152.4. The travelling distance obeys uniform distribution between 4 and 12 km. The bi-level model is used to solve the optimization problem.

Constraints in the upper level are processed with penalty function methods. Those that don't satisfy constraints (152.3), (152.4), and (152.5) are punished to one-third of the original value.

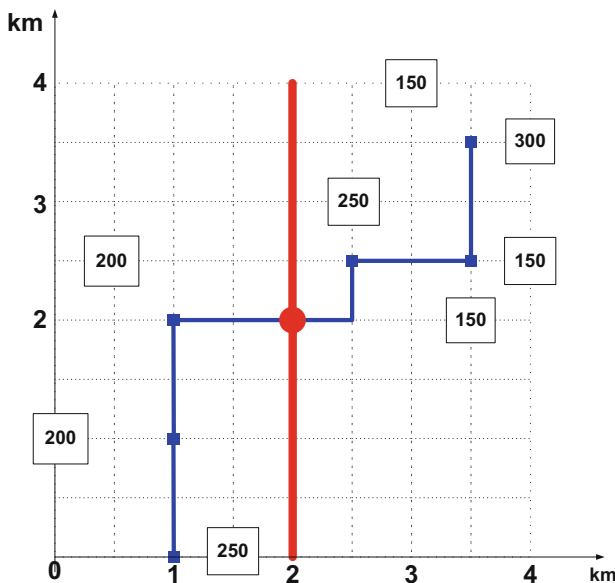
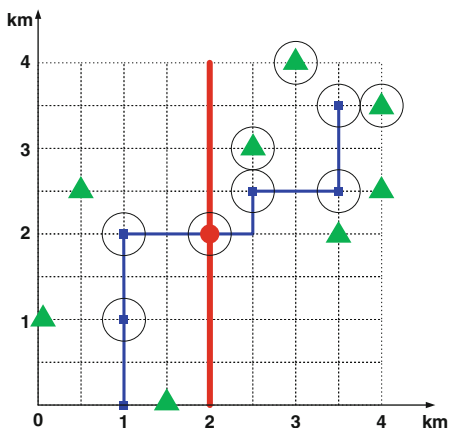


Fig. 152.4 The production volume in settlements

**Fig. 152.5** The optimal layout plan



**Table 152.1** Scale of public bike rental stations

No.	Volume of picking up	Volume of dropping off	Recommended capacity
21	45	28	50
39	64	22	70
41	0	116	120
51	0	39	40
53	55	39	60
60	36	0	40
71	0	76	80
72	83	0	90
79	37	0	40
Total	320	320	590

The selection of genes uses roulette strategy. That is to say, the selection probability is connected with fitness and individuals are selected at random probability to the next generation. The crossover probability is 0.9 and mutation probability is 0.04.

Matlab programming can help realize the two levels' algorithm of bi-level planning and be applied to the case. After 60 generations' evolution, the optimization plan is found. The optimal layout plan of station is illustrated in Fig. 152.5.

The final recommended capacity of each rental station is listed in Table 152.1.

As is shown in Table 152.1, 9 public bike rental stations need to be built in the region with a total capacity of 320; thus, the recommended capacity is 590.

## 152.5 Conclusion

The optimal layout of public bike rental stations based on B+R mode is an essential way to increase the attraction of metro's passenger flow. Considering the characteristics of the layout of public bike rental stations, discrete transportation network design method and bi-level planning model are chosen to describe and solve the problem. For the upper level, to maximize the metro trip volume, from the aspect of travel planners, the model designs a reasonable layout plan under all constraints. And the lower level uses a combined model of mode split and traffic assignment to describe users' travelling habits. To validate the bi-level model, a case is taken as an example. The result shows bi-level model can not only well describe the optimal layout problem of public bike but also offer technical and scientific support for decisions of layout of public bike rental stations around metro stations.

## References

1. Li, L.-h., Chen, H., & Sun, X.-l. (2009). Bike rental station deployment planning in Wuhan [J]. *Urban Transport of China*, 7(4), 39–44 (In Chinese).
2. Li, Z. (2010). Analysis of the size of public bike rental stations development [J]. *Jiaotong Jieneng Yu Huanbao*, 2, 44–46 (In Chinese).
3. Geng, X., Kai Tian, Y., & Zhang, Q. L. (2009). Bike rental station planning and design in Paris [J]. *Urban Transport of China*, 7(4), 21–29 (In Chinese).
4. Lin, J.-R., & Yang, T.-H. (2010). Strategic design of public bicycle sharing systems with service level constraints [J]. *Transportation Research Part E*, 47, 284–294.
5. Wei, Z., Huabing, D., & Ren, F. (2005). A research into R\_B tripping in large cities of China [J]. *Road Traffic & Safety*, 5(4), 1–4 (In Chinese).
6. Martens, K. (2004). The bicycle as a feeder mode: Experiences from three European countries [J]. *Transportation Research Part D*, 9, 281–294.
7. Huapu, L. (2006). *Theory and method in transportation planning (second edition) [M]* (pp. 24–28). Beijing: Qsinghua University Press (In Chinese).
8. Jia, W. (2011). *Mixed logit model and its application research [D]*. Ningbo: Ningbo University (In Chinese).
9. Wang, S.-s., Huang, W., & Zhen-bo, L. (2006). Study on mixed logit model and its application in traffic mode split [J]. *Journal of Highway and Transportation Research and Development*, 23(5), 88–91 (In Chinese).
10. Li, H.-m., Huang, H.-j., & Liu, J.-f. (2009). Parameter estimation of the mixed logit model and its application [J]. *Journal of Transportation Systems Engineering and Information Technology*, 7(4), 39–46 (In Chinese).
11. Chen, D.-w., Zhong, X., & Li, X.-h. (2010). Layout optimizing model of alternative interchange in highway network [J]. *Journal of Traffic and Transportation Engineering*, 10(3), 72–76 (In Chinese).

# Chapter 153

## Modeling of Train Control Systems Using Formal Techniques

Bingqing Xu and Lichen Zhang

**Abstract** Train control systems must guarantee a very high level of safety because their incorrect functioning may have very serious consequences such as loss of human life, large-scale environmental damages, or considerable economical penalties. The software reliability is related to several factors, such as completeness, consistency, and lack of ambiguity. Formal methods are widely recognized as fault avoidance techniques that can increase dependability by removing errors during the specification of requirements and during the design stages of development. In this chapter, a brief overview of existing results on formal specification of train control systems is first presented. Then we propose an integrated formal approach to specify train control systems; this integrated approach combines CSP and Object-Z with Clock theory to specify the Railway Control System concerning both the linear track and crossing area, especially the time delay between any two aspects of the railway system.

### 153.1 Introduction

Train control systems must guarantee a very high level of safety because their incorrect functioning may have very serious consequences such as loss of human life, large-scale environmental damages, or considerable economical penalties [1]. To meet safety and reliability requirements, the relative international standards recommend the application of formal methods in specifying development specifications and design for train control systems, the approaches using formal methods which can eliminate ambiguities of specification, and specify and prove system specification in a rigorous mathematical way, were highly recommended for the railway systems which belong to software safety integrity level 4 [2]. Although the

---

B. Xu • L. Zhang (✉)  
Shanghai Key Laboratory of Trustworthy Computing, East China Normal University,  
Shanghai 200062, China  
e-mail: [zhanglichen1962@163.com](mailto:zhanglichen1962@163.com)

train control system safety and reliability depend not only on software but also on hardware, without a proper software system support, the system cannot work perfectly. In addition, compared with hardware faults which are mostly physical, the detecting and correcting of faults in software systems are usually more abstract and more troublesome. The software reliability is related to several factors, such as completeness, consistency, and lack of ambiguity [3]. Formal methods are widely recognized as fault avoidance techniques that can increase dependability by removing errors during the specification of requirements and during the design stages of development. Formal methods can be used to increase the safety of systems by formally verifying that certain safety properties hold on a model of system. In addition most formal method approaches are well suitable to be mechanized and a great variety of tools for automatic validation are nowadays available. Complex system such as train control system is a system with many different aspects, and the mechanism of communication between different aspects is hard to define. With the help of formal methods, we can find a way to construct a detailed specification of each aspect and the link mechanism among various aspects, while a communication mechanism is not enough to describe the state change and data change in the system. Above all, the author tends to use Communicating Sequential Processes (CSP) [4] to specify the communication part of the train control system. Concerning the time characteristics in the system, clock [5] specifies the system time requirements better [6]. For the state and data changes, Object-Z [7] is ideal for analysis in data change in a schema box form.

In this chapter, we propose an integrated formal approach to specify train control systems; this integrated approach combines CSP and Object-Z with Clock theory to specify the Railway Control System concerning both the linear track and crossing area, especially the time delay between any two aspects of the railway system.

## 153.2 Related Works

Formal methods are approaches, based on the use of mathematical techniques and notations, for describing and analyzing properties of software systems. That is, descriptions of a system are written using notations which are based on mathematical expressions rather than informal notations. These mathematical notations are typically drawn from areas of discrete mathematics, such as logic, set theory, or graph theory. There exists a lot of work on applications of formal methods in train control system.

Modeling the controller of the railway network, having resource sharing based on mutual exclusion constraints, is an important problem. Ahmad, Farooq and Khan, and Sher Afzal firstly address the specification of safety properties for the model of a complex railway crossing. The operations, i.e., occupied, free, and block, are formalized to describe the safety properties along railway crossing. Second, to develop the control model of the crossing system, they construct the subnet representing the train flow along the tracks in the crossing region and the set

of monitors or supervisors are also modeled as subnets. Arc-constant colored Petri net (ac-CPN) is used to construct the train flow subnet while the monitors are modeled using the place/transition net. Arc-constant colored Petri net enforces the specification of not to shift the train from a track to another one. Bottom-up approach is adopted to model the control for railway crossing as a synchronous synthesis of the subnets is applied to build the final model. Finally, to verify the safety properties in the developed controller, the coverability tree method is used for the analysis of the final model [8].

Defects in requirements specification of train control system may have fatal consequences. Zhao Lin, Tang Tao, Cheng Ruijun, and He Liyun present a property-based requirements analysis approach for train control system [9], which provides support for constructing precise, complete, and consistent requirements specification and analyzing them with formal techniques.

Level crossings (LCs) are considered to be a safety black spot for railway transportation since LC accidents/incidents dominate the railway accident landscape in Europe, thus considerably damaging the reputation of railway transportation. LC accidents cause more than 300 fatalities every year throughout Europe, which represents up to 50 % of all deaths for railway. That is why LC safety is a major concern for railway stakeholders in particular and transportation authorities in general. LCs with an important traffic moment <sup>1</sup> are generally equipped with automatic protection systems (APSs). Here, they focus on two main risky situations, which have caused several accidents at LCs. The first is the short opening duration between successive closure cycles relative to trains passing in opposite directions. The second is the long LC closure duration relative to slow trains. Mekki Ahmed, Ghaze Mohamed, and Toguyeni Armand suggest a new APS architecture that prevents these kinds of scenarios and therefore increases the global safety of LCs [10]. To validate the new architecture, a method based on well-formalized means has been developed, allowing one to obtain sound and trustworthy results. Their method uses a formal notation, i.e., timed automata (TA), for the specification phase and the model-checking formal technique for the verification process. All the steps are progressively discussed and illustrated.

Railway interlocking system is a distributed, safety, monetary, and environmentally critical system, and its failure may cause the loss of human life, severe injuries, loss of money, and environmental damages. The complexity of this system requires formal modeling and step-by-step refinement for its construction and development. The formal specification-based languages, such as VDM, Z-notation, and RAISE, have been used for its modeling using crisp (two-valued logic) theory. However, due to the continuous and inexact features, like speed, weight, and moving block (breaking distance including length of a train), fuzzy distributed multi-agent approaches are required to capture the inexactness and uncertainty present in the existing system.

---

<sup>1</sup> LCs with an important traffic moment are generally equipped with automatic protection systems (APSs).

The RBC (radio block center) handover is an important part of European Train Control System (ETCS) level 2 which is a typical safety-critical hybrid system. Liu et al. [11] build a formal model of RBC handover procedure using Differential Dynamic Logic, which is a first-order dynamic logic for specifying and verifying hybrid systems, and identify some constraints that are necessary for ensuring safety of train control, including collision avoidance as well as derailment avoidance. Moreover, they formally verify the safety-related properties of their model with deductive verification tool KeYmaera. The experimental results show the validity and feasibility of the method. Meanwhile, the safety constraints and safety-related properties verified in the chapter can be helpful to the practical application of train control.

Flaws in requirements may have unacceptable consequences in the development of safety-critical applications. Formal approaches may help with a deep analysis that takes care of the precise semantics of the requirements. However, the proposed solutions often disregard the problem of integrating the formalization with the analysis, and the underlying logical framework lacks either expressive power, or automation. We propose a new, comprehensive approach for the validation of functional requirements of hybrid systems, where discrete components and continuous components are tightly intertwined. The proposed solution allows to tackle problems of conversion from informal to formal, traceability, automation, user acceptance, and scalability.

Jochen Hoenicke uses a combination of three techniques for the specification of processes, data, and time: CSP, Object-Z, and Duration Calculus [12–16]. The basic building block in our combined formalism CSP-OZ-DC is a class. First, the communication channels of the class are declared. Every channel has a type which restricts the values that it can communicate. There are also local channels that are visible only inside the class and that are used by the CSP, Z, and DC parts for interaction. Second, the CSP part follows; it is given by a system of (recursive) process equations. Third, the Z part is given which itself consists of the state space, the Init schema and communication schemas. For each communication event a corresponding communication schema specifies in which way the state should be changed when the event occurs. Finally, below a horizontal line the DC part is stated. The combination is used to specify parts of a novel case study on radio-controlled railway crossings. Johannes Faber formally specifies a part of the ETCS with the specification language CSP-OZ-DC treating the handling of emergency messages.

### 153.3 Formal Specification of Train Control Systems

Train control system is composed of two parts [3]: onboard system and control center. The onboard system is composed of the following: (1) a basic state detection subsystem, including position and speed of train; (2) an ATP subsystem, which is used to monitor the speed and will produce appropriate actions if certain situations

happen (unresponsive train operator, earthquake, disconnected rail, overrun of the authority, etc.) to prevent accidents from happening; (3) a communication subsystem, which is used to send and receive messages and commands to and from the control center when it is necessary, such as approaching the station and departing from the station; and (4) a record subsystem. The control center is mainly composed of the following: (1) a communication system, which is used to send and receive commands and messages to and from trains. The control center knows the position and station block where every train is, which is the basis of authorization generation; (2) an interlocking system, which is responsible for the set of points, and route set based on information received from the trains; and (3) a traffic operation management, which creates train timetables. All trains share their information with the control center periodically. In the event of a natural disaster, once the train loses communication with the control center for a certain time, it will stop automatically.

In order to keep the description focused, we concentrate on some particular points in train control systems rather than the detailed descriptions of all development process. The specification is made by integrating Object-Z, CSP, and Clock theory. The Object-Z is used to specify the data aspect and the operation on the data of the system; CSP is used to specify communication aspect of the system, and clock is used to specify the time aspect of the system.

Clock theory [5] puts forward the possibility to describe the event in physical world by using a clock and can analyze and record the event by clock. To use clock to specify cyber physical systems, the time description is clearer to every event and can link continuous world with discrete world better [6].

Let  $e$  be an event. Clock ( $e$ ) denotes the clock that records the time instants when the event  $e$  occurs. And we use clock (event( $c$ )) to denote the event that take place at every time instant  $c[i]$ .

The controller should take care of trains running over the track. It should control the safety of the configuration, i.e., no two trains may enter the critical section. When one critical section is occupied, some others, which share some part of section with this one, should be locked. The controlled can control the status, speed, and position of trains as shown in Fig. 153.1.

We integrate CSP and Object-Z with Clock theory to specify real-time aspects as shown in Fig. 153.2.

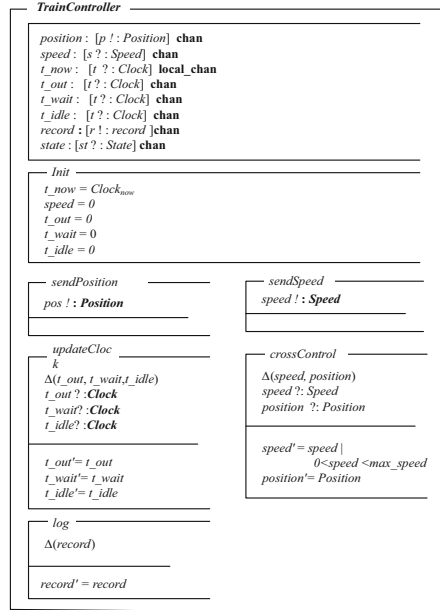
Finally, woven model is shown as Fig. 153.3.

## 153.4 Conclusion

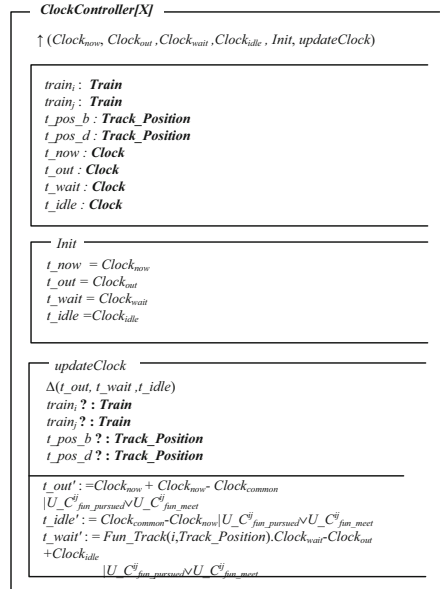
Train control systems must guarantee a very high level of safety because their incorrect functioning may have very serious consequences such as loss of human life, large-scale environmental damages, or considerable economical penalties. The software reliability is related to several factors, such as completeness, consistency, and lack of ambiguity. Formal methods are widely recognized as fault



**Fig. 153.1** Train control system

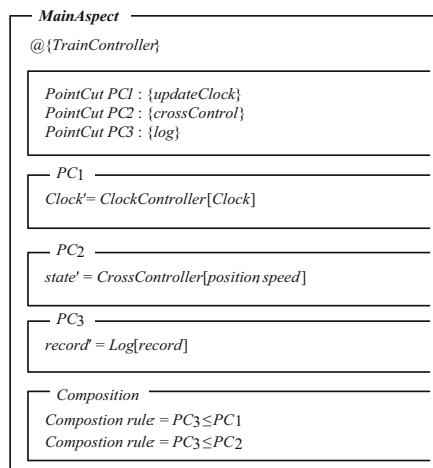


**Fig. 153.2** Model of clock



avoidance techniques that can increase dependability by removing errors during the specification of requirements and during the design stages of development. In this chapter, we propose an integrated formal approach to specify train control system; this integrated approach combines CSP and Object-Z with Clock theory to specify

**Fig. 153.3** Woven aspects of diagram



the train control system. In this chapter, we have applied the proposed approach to the specification of train control systems; the proposed approach is expressive enough to represent the functional requirements, behavior requirements, and real-time aspects of train control system, yet simple enough to allow for the use by nonexperts in formal methods.

Future work focuses on the verification tool development of our proposed method.

**Acknowledgments** This work is supported by Shanghai Knowledge Service Platform Project (No. ZF1213); National High Technology Research and Development Program of China (No. 2011AA010101); National Basic Research Program of China (No. 2011CB302904); the National Science Foundation of China under Grant Nos. 61173046, 61021004, 61061130541, and 91118008; Doctoral Program Foundation of Institutions of Higher Education of China (No. 20120076130003); and National Science Foundation of Guangdong Province under Grant No. S2011010004905.

## References

1. Jo, H.-J., Yoon, Y.-K., & Hwang, J.-G. (2009). Analysis of the formal specification application for train control systems. *Journal of Electrical Engineering & Technology*, 4(1), 87–92.
2. IEC62278:2002. *Railway applications: Specification and demonstration of reliability, availability, maintainability and safety (RAMS)*.
3. Xie, G., Hei, X., Mochizuki, H., Takahashi, S., & Nakamura, H. (2013). Safety and Reliability Estimation of Automatic Train Protection and Block System. Quality and Reliability Engineering International. © John Wiley & Sons, Ltd.
4. Reed, G. M., & Roseoe, A. W. (1988). A timed model for communicating sequential processes. *Theoretical Computer Science*, 58, 249–261.
5. Xu, B. Q., He, J., & Zhang, L. C. (2013). Specification of cyber physical systems based on clock theory. *International Journal of Hybrid Information Technology*, 6(3), 45–54.

6. Xu, B. Q., et al. (2013). Specification of cyber physical systems by clock. In *AST2013, SERSC* (Vol. 20, pp. 111–114). SERSC (Science & Engineering Research Support Society) Korea
7. Najafi, M., & Haghghi, H. (2013). An integration of UML-B and object-Z in software development process. In K. Elleithy & T. Sobh (Eds.), *Innovations and advances in computer, information, systems sciences, and engineering* (pp. 633–648). New York: Springer.
8. Ahmad, F., & Khan, S. A. (2013). Specification and verification of safety properties along a crossing region in a railway network control. *Applied Mathematical Modelling*, 37(7), 5162–5170.
9. Zhao, L., Tang, T., Cheng, R., & He, L. (2013). Property based requirements analysis for train control system. *Journal of Computational Information Systems*, 9(3), 915–922.
10. Mekki, A., Ghaze, M., & Toguyeni, A. (2012). Validation of a new functional design of automatic protection systems at level crossings with model-checking techniques. *IEEE Transactions on Intelligent Transportation Systems*, 13(2), 714–723.
11. Dewang Chen, Rong Chen, & Yidong Li. Formal modeling and verification of RBC handover of ETCS using differential dynamic logic. In *Proceedings of 2011 10th international symposium on autonomous decentralized systems* (pp. 67–72).
12. Hoenicke, J. (1999). Specification of radio based railway crossings with the combination of CSP. In G. Smith & I. Hayes (Eds.), *Towards real-time object-Z, Lecture notes in computer science* (pp. 49–65). Berlin: Springer.
13. Hoenicke, J. (2006). *Combination of processes, data, and time*. PhD thesis, University of Oldenburg.
14. Hoenicke, J., & Maier, P. (2005). Model-checking of specifications integrating processes, data and time. In J. S. Fitzgerald, I. J. Hayes, & A. Tarlecki (Eds.), *FM 2005: Formal methods* (Vol. 3582, pp. 465–480). Berlin: Springer.
15. Hoenicke, J., & Olderog, E.-R. (2002). CSP-OZ-DC: A combination of specification techniques for processes, data and time. *Nordic Journal of Computing*, 9(4), 301–334.
16. Hoenicke, J., & Olderog, E.-R. (2002). Combining specification techniques for processes data and time. In M. Butler, L. Petre, & K. Sere (Eds.), *Integrated formal methods. Lecture notes in computer science* (Vol. 2335, pp. 245–266). Berlin: Springer.

# Chapter 154

## A Clock-Based Specification of Cyber-Physical Systems

Bingqing Xu and Lichen Zhang

**Abstract** In cyber-physical systems, the elapse of time becomes the most important property of system behavior, and time is central to predicting, measuring, and controlling properties of the physical world. A cyber-physical system is composed of two interacting subsystems: a cyber system and a physical system. The behavior of the cyber system is controlled by the execution of programs on a distributed digital computer system, while the laws of physics control the behavior of the physical system. The different models of time—continuous physical time in the physical system versus discrete execution time in the cyber system and the impossibility of perfect synchronization of the physical clocks of the nodes of a distributed computer system, lead to interesting phenomena concerning the joint behavior of these two subsystems. The chapter describes the case studies in applying clock theory to the production cell. The clock theory described is very simple, in that it models clocks as potentially infinite lists of reals. Xenon's paradox and similar problems are avoided by specifying limits on clock rates, which effectively means that the model sits somewhere between a discrete synchronous model and a fully dense continuous-time model as assumed by some other formalisms. The case study of the specification of the production cell shows that using clock theory to specify cyber-physical systems can give a more detailed description of the every subsystem and give a much more considerate observation of the time line and sequence of every event.

---

B. Xu • L. Zhang (✉)  
Shanghai Key Laboratory of Trustworthy Computing, East China Normal University,  
Shanghai 200062, China  
e-mail: [zhanglichen1962@163.com](mailto:zhanglichen1962@163.com)

## 154.1 Introduction

Cyber-physical systems consist of the class of large-scale infrastructures that have significant cyber and physical components and have wide-ranging impact on society in their deployment [1]. Time is central to predicting, measuring, and controlling properties of the physical world: given a physical model, the initial state, the inputs, and the amount of time elapsed, one can compute the current state of the plant. A cyber-physical system consists of two interacting subsystems: a cyber system and a physical system. The behavior of the cyber system is controlled by the execution of programs on a distributed digital computer system, while the laws of physics control the behavior of the physical system. The different models of time—dense physical time in the physical system versus discrete execution time in the cyber system and the impossibility of perfect synchronization of the physical clocks of the nodes of a distributed computer system, lead to interesting phenomena concerning the joint behavior of these two subsystems.

Cyber-physical systems-related research is based on two, originally different worldviews: on the one hand, the dynamics and control (DC) worldview and on the other hand the computer science (CS) worldview. The DC worldview is that of a predominantly continuous-time system, which is modeled by means of differential (algebraic) equations or by means of a set of trajectories. The CS worldview is that of a predominantly discrete-event system. A well-known model is a (hybrid) automaton, which modeling of discrete-event systems is also based on. As new CPS applications start to interact with the physical world using sensors and actuators, there is a great need for ensuring that the actions initiated by the CPS is timely. Cyber-physical systems are characterized by their stringent requirements for time constraints such as predictable end-to-end latencies and timeliness.

Since cyber-physical systems are dynamic systems that exhibit both continuous and discrete dynamic behavior, the continuously bilateral interaction between discrete events and continuous time flow makes it hard to know the dynamic feature of the system. So specifying the timing issues is a really vital work in the early stage. This chapter describes the case study in applying clock theory to the production cell. The clock theory described is very simple, in that it models clocks as potentially infinite lists of reals. Xeno's paradox and similar problems are avoided by specifying limits on clock rates, which effectively means that the model sits somewhere between a discrete synchronous model and a fully dense continuous-time model as assumed by some other formalisms.

## 154.2 Related Works

Duggirala et al. present an algorithm for checking global predicates from distributed traces of cyber-physical systems for an individual agent [2]. Each observation has a possibly inaccurate timestamp from the agent's local clock. The challenge is to symbolically over-approximate the reachable states of the entire system from the

unsynchronized traces of the individual agents. The presented algorithm first approximates the time of occurrence of each event, based on the synchronization errors of the local clock, and then over-approximates the reach sets of the continuous variables between consecutive observations. The algorithm is shown to be sound; it is also complete for a class of agents with restricted continuous dynamics and when the traces have precise information about timing synchronization inaccuracies. Experimental results illustrate that interesting properties like safe separation, correct geocast delivery, and distributed deadlocks can be checked for up to 20 agents in minutes.

The specification of modeling and analysis of real-time and embedded (MARTE) systems is an extension of the unified modeling language (UML) in the domain of real-time and embedded systems. Even though MARTE time model offers a support to describe both discrete and dense clocks, the biggest effort has been put so far on the specification and analysis of discrete MARTE models. To address hybrid real-time and embedded systems, Liu et al. propose to extend statecharts using both MARTE and the theory of hybrid automata [3]. As a case study, they model the behavior of a train control system with hybrid MARTE statecharts to demonstrate the benefit.

The Object Management Group (OMG) UML profile for MARTE aims at using the general-purpose modeling language UML in the domain of real-time and embedded (RTE) systems. We have also defined a non-normative concrete syntax called the CCSL to demonstrate what can be done based on this structure. Mallet Frédéric gives a brief overview of this syntax and its formal semantics and shows how existing UML model elements can be used to apply this syntax in a graphical way and benefit from the semantics [4].

The UML Profile for MARTE has been recently adopted. The CCSL allows the specification of causal, chronological, and timed properties of MARTE models. The IEEE Property Specification Language (PSL) provides a formal notation for expressing temporal logic properties that can be automatically verified on electronic system models [5]. We identify and restrict the CCSL constructs that cannot be expressed in temporal logics so that CCSL become tractable in temporal logics.

Event clock automata (ECA) are a model for timed languages that has been introduced by Alur, Fix, and Henzinger as an alternative to timed automata, with better theoretical properties (for instance, ECA are determinizable while timed automata are not). Gilles et al. revisit and extend the theory of ECA [6].

In the chapter of Lamport [7], the concept of “happening before” defines an invariant partial ordering of the events in a distributed multiprocess system. The representation of a closed finitary real-time system as a graph annotated with clock constraints is called a timed automaton [8], since clocks range over the nonnegative reals, every nontrivial timed automaton has infinitely many states. If the clocks of a finitary real-time system are permitted to drift with constant, rational drift bounds, one obtains a finitary drifting-clock system. The representation of a closed finitary drifting-clock system as a graph annotated with constraints on drifting clocks is called an initialized rectangular automaton [9]. Two popular specification languages for the algorithmic verification of untimed systems are finite automata and

propositional temporal logics. In order to specify timing constraints, these languages can be extended by adding clock variables. If we judiciously add clocks to finite automata, one obtains the timed automata (TA); from propositional linear temporal logic, one obtains the real-time logic TPTL [10]; and from the propositional, branching-time logic CTL, one obtains the real-time logic TCTL [11].

Bujorianu et al. presented a multiclock model for real-time abstractions of hybrid systems [12]. They call hybrid time systems the resulting model, which is constructed using category theory. They define a timed (or clock) system as a functor from a category of states to a category of time values. They further define concurrent composition operators and bisimulation.

### 154.3 Clock Theory

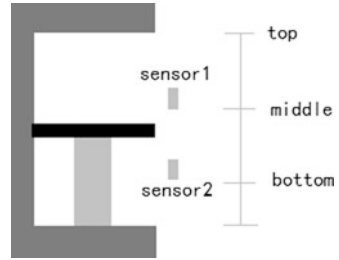
Clock theory [13] puts forward the possibility to describe the event in physical world by using a clock, and can analyze and record the event by clock. To use clock to specify cyber-physical systems, the time description is clearer to every event and can link continuous world with discrete world better [14].

### 154.4 Case Study: Specification of the Production Cell Steam Based on Clock Theory

The production cell [15] is composed of two conveyor belts, a positioning table, a two-armed robot, a press, and a travelling crane. Metal plates inserted in the cell via the feed belt are moved to the press [16]. There, they are forged and then brought out of the cell via the other belt and the crane [17]. As is known to all, the production cell is widely used in factories, and it has brought a large amount of benefit to the manufacture [18]. Since its multifunction and high efficiency, metal blank can be made into various shapes to fit different requirements [19]. Figure 154.1 shows the typical type of the production cell; there are sensors named sensor1 and sensor2 which conveys information when the press is loaded or unloaded separately. First of all, the lower part of the press is raised to middle, and the metal blank is fed. When the press is loaded, sensor1 conveys the information [20]. When the lower part is raised to the top, the press moves down to the bottom position to place the metal blank to the lower convey belt. Now sensor 2 conveys the information when the press is unloaded. Then the lower part is raised to the middle position and waits for the press to be loaded again. The production cell is a kind of cyclical control system [21]. All the necessary parameters are listed in Table 154.1.

To guarantee the safety of the press, the two basic rules are as follows. The current height of press is always lower than or equal to the top position, and

**Fig. 154.1** The production cell



**Table 154.1** Parameter list for the production cell

Parameter	Value
v	Current speed of press
h	Height of press
c	Height change of press
Bottom, middle, top	Position of press
unload, load, pressing	Operations of press
mov_unload, mov_load	Press moves to load and unload

it is always higher than or equal to the bottom position. Bottom position is always lower than middle position, and so is the middle position to top position.

$$\text{Bottom} \leq h \leq \text{Top}$$

$$\text{Bottom} < \text{Middle} < \text{Top}$$

First of all, we consider the process of press; h denotes the height of press; c denotes the height change of press; v denotes the current speed of press; and load and unload are events which control the convey belt to feed metal blanks. During the process, the press moves down and up. Suppose the initial situation is that the press is at the bottom position and is ready to be raised to the middle position.

$$\begin{aligned} \text{climb}(h, \text{Middle} - c) &\leq \text{clock}(\text{load}) \leq \text{clock}(\text{pressing}) \\ &\leq \text{climb}(h, \text{Top} - c) \\ &\leq \text{drop}(h, \text{Bottom} + c) \leq \text{clock}(\text{unload}) \end{aligned}$$

Between different phases, there exists time latency.

$$\begin{aligned} \rho(\text{climb}(h, \text{Middle} - c), \text{clock}(\text{load})) &\leq c/v \\ \rho(\text{clock}(\text{load}), \text{clock}(\text{pressing})) &\leq c/v \\ \rho(\text{clock}(\text{pressing}), \text{climb}(h, \text{Top} - c)) &\leq c/v \\ \rho(\text{climb}(h, \text{Top} - c), \text{drop}(h, \text{Bottom} + c)) &\leq c/v \\ (\text{drop}(h, \text{Bottom} + c), \text{clock}(\text{unload})) &\leq c/v \end{aligned}$$

From above, it is obvious that some couple of events in the equations above have noninterference, since mov\_load and mov\_unload cannot happen at the same time, and so do load and unload.

$$\begin{aligned} &\text{clock}(\text{mov\_load}) [1]) > 0 \\ \text{clock}(\text{mov\_load}) &\leq \text{clock}(\text{mov\_unload}) \leq \text{clock}(\text{mov\_load}) ' \end{aligned}$$



$$\begin{aligned} \text{clock}(\text{mov\_unload}) \wedge \text{clock}(\text{mov\_load}) &= \emptyset \\ \text{clock}(\text{load})[1] &> 0 \\ \text{clock}(\text{load}) &\leq \text{clock}(\text{unload}) \leq \text{clock}(\text{load}) ' \\ \text{clock}(\text{load}) \wedge \text{clock}(\text{unload}) &= \emptyset \end{aligned}$$

In the equations,  $\mathbf{V}$  denotes continuous speed change, and  $\mathbf{v}$  denotes discrete speed at each clock unit. The continuous variable and discrete variable can be linked as below:

$$\begin{aligned} \mathbf{V} &= \mathbf{v} \text{ init } \mathbf{v}_0 \\ \mathbf{v} &= \mathbf{V} \text{ every } c \text{ init } \mathbf{v}_0 \end{aligned}$$

## 154.5 Conclusion

In cyber-physical systems, the elapse of time becomes the most important property of system behavior and time is central to predicting, measuring, and controlling properties of the physical world. A cyber-physical system is composed of two interacting subsystems: a cyber system and a physical system. The behavior of the cyber system is controlled by the execution of programs on a distributed digital computer system, while the laws of physics control the behavior of the physical system. The different models of time—continuous physical time in the physical system versus discrete execution time in the cyber system and the impossibility of perfect synchronization of the physical clocks of the nodes of a distributed computer system, lead to interesting phenomena concerning the joint behavior of these two subsystems. This chapter presented the case study in applying clock theory to the production cell. Case study shows that using clock theory to specify cyber-physical systems can give a more detailed description of every subsystem and give a much more considerate observation of the time line and sequence of every event.

It is brilliant to connect the event with clock, while it is so difficult to handle so many local clocks with the global clock. In my point of view, it is always hard work to make local clocks keeping consistent with the global clock, and the verification of the security and accuracy of the synchronization is very complicated, and we need more ideas to do this work.

**Acknowledgements** This work is supported by Shanghai Knowledge Service Platform Project (No. ZF1213), National High Technology Research and Development Program of China (No. 2011AA010101), National Basic Research Program of China (No.2011CB302904), the National Science Foundation of China under grants (No.61173046, No. 61021004, No. 61061130541, No.91118008), Doctoral Program Foundation of Institutions of Higher Education of China (No. 20120076130003), National Science Foundation of Guangdong Province under grant (No. S2011010004905).

## References

1. Eidson, J., Lee, E. A., Matic, S., Seshia, S. A., & Zou, J. (2012). Distributed real-time software for cyber-physical systems. *Proceedings of the IEEE (Special Issue on CPS)*, 100(1), 45–59.
2. Duggirala, P. S. Johnson, T. T, Zimmerman, A., et al. (2012). Static and dynamic analysis of timed distributed traces. *Proceedings of the 2012 I.E. 33rd Real-Time Systems Symposium* (pp. 173–182). IEEE.
3. Liu, J., Liu, Z., He, J., Frédéric, M., & Ding, Z. (2013). Hybrid MARTE statecharts. *Frontiers of Computer Science*, 7(1), 95–108.
4. Frédéric, M. (2008). Clock constraint specification language: Specifying clock constraints with UML/MARTE. *Innovations in Systems and Software Engineering*, 4(3), 309–314.
5. Regis, G., Frederic, M., & Julien, D. (2011). Logical time and temporal logics: Comparing UML MARTE/CCSL and PSL. In *Proceedings of the International Workshop on Temporal Representation and Reasoning* (pp. 141–148). IEEE.
6. Gilles, G., Jean-François, R., & Nathalie, S. (2011). Event clock automata: From theory to practice. In U. Fahrenberg & S. Tripakis (Eds.), *Formal modeling and analysis of timed systems*: Vol. 6919. *Lecture notes in computer science* (pp. 209–224). Berlin: Springer.
7. Lamport, L. (1978). Time, clocks, and the ordering of events in a distributed system. *Communications of the ACM*, 21(7), 558–565.
8. Rajeev, A., & Dill, D. L. (1994). A theory of timed automata. *Theoretical Computer Science*, 126(2), 183–235.
9. Henzinger, T.A., Kopke, P.W., Puri, A. & Varaiya, P. (1995). What’s decidable about hybrid automata? In *Proceedings of the 27th Annual Symposium on Theory of Computing* (pp. 373–382). ACM Press.
10. Alur, R., & Henzinger, T. A. (1994). A really temporal logic. *Journal of the ACM*, 41(1), 181–204.
11. Alur, R., Courcoubetis, C., & Dill, D. L. (1993). Model checking in dense real time. *Information and Computation*, 104(1), 2–34.
12. Bujorianu, M. C. Bujorianu, L. M., & Langerak, R. (2008). An interpretation of concurrent hybrid time systems over multi-clock systems. In *Proceedings of the 17th IFAC World Congress* (Vol. 17, Part 1, pp. 3635–3640). IFAC
13. Xu, B., He, J., & Zhang, L. (2013). Specification of cyber physical systems based on clock theory. *International Journal of Hybrid Information Technology*, 6(3), 45–54.
14. Xu, B., et al. (2013). Specification of cyber physical systems by clock. In *AST 2013, ASTL 20:111–114*, 2013 © SERSC 2013.
15. Back, R. J., Petre, L., & Porres, I. (2000). Generalizing action systems to hybrid systems. *Lecture Notes in Computer Science*, 1926, 202–213.
16. Abouttab, M. S., Brockway, M., Counsell, S., & Hierons, R. M. (2013). Testing real-time embedded systems using timed automata based approaches. *Journal of Systems and Software*, 86(5), 1209–1223.
17. Budde, R. (1995). ESTEREL applied to the case study production cell. In: C. Lewerentz and T. Lindner eds. *Formal Development of Reactive Systems-Case Study Production Cell*, Springer, Berlin, 75–100
18. Lewerentz, C., & Lindner, T. (1995). Case study ‘production cell’: A comparative study in formal specification and verification. In: Manfred Broy, Stefan Jähnichen. Berlin. *Lecture Notes in Computer Science*, 1009, Springer Berlin Heidelberg, 388–416.
19. Burns, A. (2003). How to verify a safe real-time system: The application of model checking and timed automata to the production cell case study. *Real-Time Systems*, 24(2), 135–151.
20. Benghazi Akhlaki, K., Capel Tuñón, M. I., Holgado Terriza, J. A., & Mendoza Morales, L. E. (2007). A methodological approach to the formal specification of real-time systems by transformation of UML-RT design models. *Science of Computer Programming*, 65(1), 41–56.
21. El-Maddah Islam, A. M. (2005). Component-based development of process control systems. In *3rd ACS/IEEE International Conference on Computer Systems and Applications* (pp. 797–804).

# Chapter 155

## Polymorphic Worm Detection Using Position-Relation Signature

Huihui Liang, Jiwen Chai, and Yong Tang

**Abstract** This chapter proposes a novel worm signature that is appropriate for the polymorphic worm detection. Most of the recent worm signatures are constructed based on worm bytes themselves or relationships between worm bytes. In this case, most of these signatures cannot detect the polymorphic worms successfully. Our worm signature takes the worm bytes themselves and the relationships between worm bytes into consideration. So, it is called position-relation signature (PRS). The new signature is capable of handling certain polymorphic worms. The experiments show that the algorithm could be used as a basis to implement a worm detection system.

### 155.1 Introduction

In recent years, Internet worms have proliferated with the development of computer hardware and software. When a host is found to be worm infected, a large number of hosts may have been infected. So, fast spreading worms have presented a huge threat to the security of the Internet. Whatever in academia or in industry, there are two major problems that must be solved, the method [1] to detect worms in time and how to curb worm propagation effectively [2, 3]. In this chapter, we mainly focus on extracting the signatures of the worms to do a basis for detection.

With the development of polymorphism technology, the next generation of Internet worms is likely to be polymorphic. Polymorphic worms can be changed in their spreading process. As a consequence, copies of a polymorphic worm might no longer share a common invariant substring of sufficient length, and the existing systems will not recognize the network streams containing the worm copies as the manifestation of a worm outbreak. Therefore, the polymorphic worms that have

---

H. Liang (✉) • J. Chai • Y. Tang  
Sichuan Electric Power Research Institute, Chengdu 610072, China  
e-mail: [liang\\_huihui@sohu.com](mailto:liang_huihui@sohu.com)

evolved to the network have brought great threat. Defense of the damage effectiveness mainly depends on the quality of extracted signatures of worms. Most of the recent worm signatures are constructed based on worm bytes themselves or relationships between worm bytes. Polymorphic worms can automatically change its content, so to detect them is a challenging task. To detect polymorphic worm, extracting the invariant signatures of the same polymorphic worm in different situations is the key to success.

This chapter presents a novel worm signature based on signature fusion. This worm signature combines the worm bytes themselves and the relationships between worm bytes-neighborhood-relation signature. It is called position-relation signature (PRS). And this chapter is structured as follows. Section 155.2 discusses related work. Section 155.3 presents the description of the new signature. In Sect. 155.4, we present the experimental results using different signatures to evaluate our signature. Section 155.5 briefly concludes and points out limitations of the current signatures.

## 155.2 Related Works

Signature is the most fundamental factors in worm detection system that is based on the signatures. The effectiveness of the polymorphic worm defense system depends on signature describing the ability of worms. The complexity of worm technology that is increasing puts forward higher request for the ability of detection and defense about worms. A lot of research is dedicated to produce worm signatures.

In recent years, most worm signatures mainly divided into two categories are as follows:

### 155.2.1 Signatures Based on Worm Bytes Themselves

In such signatures, the longest common substring (LCS) is a typical signature [4, 5]. Many systems use LCS. However, it cannot detect polymorphic worms. For example, when the samples of polymorphic worms added instructions, LCS cannot detect the samples after deformation. Newsome, Karp, and Song [6] put forward a polymorphic worm signature generation system—polygraph—and present a kind of suitable system for matching the characteristics of the polymorphic worms. Polygraph system extracts many tokens from suspicious traffic pool; these tokens referring to independent substring have emerged at least  $k$  times in the suspicious pool  $n$  series.

Position-aware distribution signature (PADS) is proposed by Tang and Chen [7]. It is a collection of position-aware byte-frequency distribution function. Compared with tokens signature, PADS is more flexible. When the polymorphic worm uses encryption technology and uses limited decryption routines, the decryption

routines can be identified if we get enough samples of worms. However, we find it difficult to get enough polymorphic worm samples in a short period of time, so suspicious traffic pool cannot contain all polymorphic worm samples produced by each routine. Therefore, the PADS extracted from suspicious traffic pool cannot detect polymorphic worm which includes many decryption routines successfully.

### ***155.2.2 Signatures Based on Relationships Between Worm Bytes***

Most of recent worm signatures are constructed based on worm bytes themselves. They can be used to detect one pattern of worms successfully, but are not appropriate when treating on polymorphic worms since these worms can change their patterns dynamically. A class of neighborhood-relation signatures (NRSs) [8] is proposed based on neighborhood relationship between worm bytes. The NRS can exhibit characteristics of polymorphic worm and can be used to detect polymorphic worm with no noise efficiently. However, NRS is difficult to have a good performance when the polymorphic worm has noise.

Each of these two category signatures has its own advantages and disadvantages. In this chapter, we combine these two kinds of signatures, combine these advantages, and gain a better performance.

## **155.3 Position-Relation Signature Description**

Before introducing the signatures of worms, we give an introduction on how to determine the significant regions of the worms. Given a set  $S$  of worm variants, at first, we know either the significant regions of the worm variants or the signature of this worm. If we know the signature of a worm category, we can compute the significant region of a worm variant; in the same way, given the significant region, we can compute the signature of the worm category. This is a “missing data problem” in statistics; in this chapter, we solve it using expectation-maximization (EM) algorithm.

Expectation-maximization (EM) algorithm can estimate the maximum-likelihood parameter by using an iterative procedure. In this chapter, we use it to determine the signature and the significant regions. Given a set  $S = S_1, S_2, \dots, S_n$  of collected worm byte sequence, let  $(a_1, a_2, \dots, a_n)$  represent the start positions of the significant regions and parameter  $F$  represent the signature.

At first, we initialize the start positions  $(a_1, a_2, \dots, a_n)$  of the significant regions for the worm byte sequences  $(S_1, S_2, \dots, S_n)$  and compute the maximum-likelihood estimate of the signature  $F$ . Then, we calculate the new locations of the significant regions based on the computed signature  $F$ . In this algorithm, the new start location

$a_j(j = 1, 2, \dots, n)$  of the significant region of the worm variant  $S_j(j = 1, 2, \dots, n)$  is the position that the worm byte sequence has the maximum match score with the current signature  $F$ . The formulation of the start position of the significant region is shown as follows:

$$a_i = \arg \max_{a_i} \text{Score}(F, S_i, a_i), \quad (155.1)$$

where  $\text{Score}(F, S_i, a_i)$  is the matching score function; it is formulated in Function 1.5. This is the expectation step. And then, based on the current locations of the significant regions, we compute the maximum-likelihood estimate of signature; this is called a maximization step.

The EM algorithm iterates between the expectation step and the maximization step. It terminates while the difference of the current average matching score and the score of the previous iteration is less than  $\epsilon$ . In this chapter, the signature width is denoted by  $w$ . The EM algorithm decides the significant regions and the signature. In the following, we introduce three kinds of signatures.

### 155.3.1 Position-Aware Distribution Signature

The PADS signature is a frequency distribution in different byte positions. We use  $(f_1, f_2, \dots, f_w)$  to denote the worm signature or the byte-frequency distribution of the worm, and let  $f_0$  represent the byte-frequency distribution of the normal traffic or the normal signature. The PADS signature is  $F_{pads} = (f_0, f_1, f_2, \dots, f_w)$ .

Then we give detailed information on how to compute the signature while the significant regions of the worm sequence are known. At first, we compute the byte-frequency distribution for each significant region position. For position  $p(p = 1, 2, \dots, w)$ , the maximum-likelihood estimation of frequency  $f_p(x)$  is defined as follows:

$$f_p(x) = \frac{c_{p,x}}{n} \quad (155.2)$$

where  $c_{p,x}$  is the number of times  $x$  appears at position  $p$  in the significant regions  $x \in [0, \dots, 255]$ . The above function is constrained by  $\sum_{b=0}^{255} f_p(b) = 1$ .

### 155.3.2 Neighborhood-Relation Signature

NRS takes the frequency of the neighbor distance of the bytes of the significant regions as the feature. The worm variant of each kind of worm contains at least one

significant region to complete the function of the worm. Most of the worm detection method detects worms directly based on the worm payload bytes or subsequences. Because the load contents of the polymorphic worm are often changed, such methods cannot effectively detect polymorphic worms. However, NRS can be used to detect polymorphic worms effectively. This signature combines the characteristics of the polymorphic worms and takes the relationship of the payload bytes and the subsequence into important consideration. NRS has a good performance on polymorphic worm detection.

For convenience, at first, we give the definition of neighbor distance. Given a sequence  $S_j = c_1c_2 \cdots c_n$ , assume the byte distance between  $c_i$  and  $c_{i+1}$  is  $d_{i,i+1}$ , and  $d_{i,i+1}$  is referred to as the neighbor distance of location  $i$  of the sequence  $S_j$ . The formulation of the neighbor distance  $d_{i,i+1}$  is as follows:

$$d_{i,i+1} = |c_{i+1} - c_i| (i = 1, 2, \dots, n - 1). \tag{155.3}$$

The sequence set  $S = S_1, S_2, \dots, S_n$  is a kind of collected worm variants. Assume the length of the significant regions of such kind of worm is  $w$ . The start locations of the significant regions of the worm sequences are  $a_i (i = 1, 2, \dots, n)$ . Then we introduce how to compute NRS of this worm.

Given a value  $p (p = 1, 2, \dots, w - 1)$ , the neighbor distance of the location index  $a_i + p$  of sequence  $S_i$  is  $d_{a_i+p, a_i+p+1}$ ;  $count(p, d)$  represents the occurrence number of value  $d$  of  $d_{a_i+p, a_i+p+1}$ ,  $(i = 1, 2, \dots, n)$  for the sequences of the sequence set  $S$ . Define the distribution function of the neighbor distance as follows:

$$f'_p(d) = \frac{count(p, d)}{n} \tag{155.4}$$

with the constraint,  $\sum_{d=0}^{255} f'_p(d) = 1, \quad p = 1, 2, \dots, w - 1.$

$(f'_1, f'_2, \dots, f'_{w-1})$  is the NRS of the  $n$  sequences of the sequence set  $S$ . The signature length is  $w - 1$ .  $f'_0$  denotes the normal signature. The NRS signature of a worm category is denoted as  $F_{nrs} = (f'_0, f'_1, f'_2, \dots, f'_{w-1})$ .

### 155.3.3 Position-Relation Signature

PRS is a novel signature that combines the byte-frequency information and the neighbor distance information. The PRS  $F$  is represented as  $(F_{pads}, F_{nrs})$ . This signature combines the merits of PADS and NRS. In the following experiments, it has a good performance on worm detection.

For an incoming connection  $S_i$ , we want to decide whether  $S_i$  is a worm variant. Assume the length of  $S_i$  is  $l$ .  $S_{i,j} (j = 1, 2, \dots, l)$  is the byte of  $S_i$  at position  $j$ .

$seg(S_i, a_i)$  denotes the  $w$ -byte segment of  $S_i$  with the start position  $a_i$ . The matching score of  $seg(S_i, a_i)$  with the PRS is defined as follows:

$$Score(F, S_i, a_i) = \prod_{p=1}^w \frac{f_p(S_i, a_i+p-1)}{f_0(S_i, a_i+p-1)} + \prod_{p=1}^{w-1} \frac{f'_p(d_{a_i+p, a_i+p+1})}{f'_0(d_{a_i+p, a_i+p+1})}. \quad (155.5)$$

We define the maximum of  $Score(F, S_i, a_i)$  among all possible position  $a_i$  as the matching score of the byte sequence  $S_i$  with the PRS. The formulation is shown as follows:

$$Score(F, S_i) = \max_{a_i=1}^{l-w+1} Score(F, S_i, a_i). \quad (155.6)$$

Alternatively, we reformulate the score function as the logarithmic form. The final matching score of  $S_i$  with the PRS  $F$  as shown is

$$\Theta(F, S_i) = \max_{a_i=1}^{l-w+1} \left( \sum_{p=1}^w \frac{1}{w} \log \left( \frac{f_p(S_i, a_i+p-1)}{f_0(S_i, a_i+p-1)} \right) + \sum_{p=1}^{w-1} \frac{1}{w-1} \frac{f'_p(d_{a_i+p, a_i+p+1})}{f'_0(d_{a_i+p, a_i+p+1})} \right). \quad (155.7)$$

The  $w$ -byte segment that maximizes  $\Theta(F, S_i)$  is called the significant region of  $S_i$ . The matching score of the whole byte sequence is the matching score of the significant region.

For the incoming byte sequence  $S_i$ , if  $\Theta(F, S_i)$  is greater than a threshold (here, the threshold is set as 0), we take  $S_i$  as a worm byte sequence. If the value of  $\Theta(F, S_i)$  is above 0, it means  $S_i$  is closer to the worm signature; else if the value is below 0, it means  $S_i$  is closer to the normal signature.

## 155.4 Experiments

This experiment adopted Blaster worm and SQL Slammer worm as the test cases. Blaster worm spreads through the leak of Windows RPC DCOM. Blaster worm will get a copy of the infected files in the system, when the attack after successful execution. SQL Slammer worm uses a leak of SQL Server to attack during buffer. In the experiment, we use the polymorphic technology to generate Blaster worm sample and SQL Slammer worm sample.

*Generate data sets.* First of all, to generate 1,000 Blaster worm sequences, the number of samples in the pool is 1,000, and then, replace worm sequences with noise sequence. In this experiment, we consider five kinds of situations. The details are in Table 155.1 to extract PADS [8], NRS, and PRS, respectively, and to use them to detect worms. The results of detection are expressed in the missing report rate (MRR) and error report rate (ERR) in Table 155.1.



**Table 155.1** Missing report rate and error report rate about Blaster worm

Number of noise	MMR (%)			ERR (%)		
	NRS	PADS	PRS	NRS	PADS	PRS
0	0.0	0.0	0.0	0.0	0.0	0.0
100	67.2	0.0	0.0	0.0	80.8	0.0
200	0.0	0.0	0.0	0.0	92.4	0.0
300	64.8	0.0	19.2	0.0	96.8	49.2
400	63.7	0.0	0.0	0.0	97.9	0.0
500	0.0	0.0	0.0	0.0	99.9	0.0

**Table 155.2** Missing report rate and error report rate about SQL Slammer worm

Signature length	MMR (%)			ERR (%)		
	NRS	PADS	PRS	PADS	PADS	PRS
10	0.0	12.5	0.0	0.0	0.0	0.0
20	0.0	30.6	0.0	0.0	0.0	0.0
30	5.5	30.2	3.2	0.0	0.0	0.0
40	5.7	30.1	4.8	0.0	0.0	0.0
50	14.5	30.4	14.5	0.0	0.0	0.0
60	17.6	30.9	16.8	0.0	0.0	0.0
70	19.3	31.2	10.3	0.0	0.0	0.0
80	26.2	31.3	20.3	0.0	0.0	0.0
90	27.3	31.7	28.9	0.0	0.0	0.0
100	29.8	31.3	25.6	0.0	0.0	0.0

*Experiment results.* As seen in Table 155.1, when suspicious pool contained in article 200 and 500 of the noise sequences, the MRR of NRS is 0; however, it has higher MRR in other situations. The MRR of PADS and PRS is 0. That means the NRS have unstable performance. In this case, the performance of PRS is similar with PADS. In Table 155.1, PADS perform higher ERR than NRS and PRS mainly because PADS does not exclude noise jamming. The PADS not only include worm features but also include features about noise sequences.

Table 155.2 shows the MRR and ERR about SQL Slammer worm signatures. PADS, NRS, and PRS have different performance when the signature has different length. From the result, the PRS has better performance than others in MRR and ERR.

## 155.5 Conclusion

To detect polymorphic worm, this chapter presented a novel worm signature: PRS which is based on signature fusion. This worm signature combines the worm bytes themselves—PADS and the relationships between worm bytes, NRS. In this chapter, a large number of experiments are completed. We get the following conclusion:

PRS is a combination of two kinds of signatures. It is better than PADS and NRS. It is more suitable for the polymorphic worm detection, which is complicated and changing.

## References

1. Korczyński, M. (2012). Classifying application flows and intrusion detection in the internet traffic. PhD thesis, UNIVERSITÉ DE GRENOBLE.
2. Fan, W. K. G. (2012). An adaptive anomaly detection of WEB-based attacks. *International Conference on Computer Science & Education (ICCSE)* (pp. 690–694). Melbourne, VIC: IEEE.
3. Magkos, E., Avlonitis, M., Kotzanikolaou, P., & Stefanidakis, M. (2013). Toward early warning against Internet worms based on critical-sized networks. *Security and Communication Networks*, 6(1), 78–88.
4. Cai, M., Hwang, K., & Pan, J. (2007). WormShield: Fast worm signature generation with distributed fingerprint aggregation. *IEEE Transactions on Dependable and Secure Computing*, 4(2), 88–104.
5. Portokalidis, G., & Bos, H. (2007). SweetBait: Zero-hour worm detection and containment using low- and high-interaction honeypots. *Computer Networks*, 51(11), 1256–1274.
6. Newsome, J, Karp, B, Song, D. (2005). Polygraph: Automatically generating signatures for polymorphic worms. *IEEE symposium on Security and Privacy Symposium* (pp. 226–241). Washington, DC: IEEE Computer Society.
7. Tang, Y., & Chen, S. (2007). An automated signature-based approach against polymorphic internet worms. *IEEE Transactions on Parallel and Distributed Systems*, 18(7), 879–892.
8. Wang, J, Wang, J, Sheng, Y, Chen, J. (2009). Polymorphic worm detection using signatures based on neighbourhood relation. *IEEE International Conference on High Performance Computing and Communications* (pp. 347–353). Seoul: IEEE.

# Chapter 156

## Application of the Wavelet-ANFIS Model

Rijun Zhang, Caishui Hou, Hui Lin, Meiyan Zhuo, Meixin Zhang,  
Zhongsheng Li, Liwu Sun, and Fengqin Lin

**Abstract** Since many predicting methods, such as CS and PP are not very precise, Wavelet-ANFIS with high estimation precision is always used to model the decomposed series recently. This chapter uses wavelet analysis to decompose water level series and then uses ANFIS to model the decomposed series; in the end, it combined these series and predicted Lingxi Reservoir's runoff. The runoff forecast of reservoir is essential for its flood control safety. The forecast result shows that the prediction accuracy of Wavelet-ANFIS is very high and the model is quite fit to use in daily runoff and water level prediction.

### 156.1 Introduction

Lingxi Reservoir was built and put into operation in June 1956, whose total capacity is 30.6 million m<sup>3</sup>. It is a comprehensive utilization of medium-sized reservoir, which is combined with functions of water supply, irrigation, and flood protection. Currently, the reservoir is responsible for industries and domestic water supply of Fujian Refinery and Hui'an county, and its downstream is an important transport area. Therefore, the runoff and water level forecast of the reservoir is essential for its flood control safety [1].

This chapter predicts Lingxi Reservoir's runoff; it is the basic of influence research of the dam project. Once there were many predicting methods, such as CS and PP [2], but these methods are not very precise. This chapter uses wavelet analysis to decompose water level series and then it uses ANFIS to model the decomposed series; in the end, it combined these series.

Wavelet analysis is a good multi-resolution frequency method; by the multi-scale sequence analysis, it can effectively recognize the main frequency

---

R. Zhang (✉) • C. Hou • H. Lin • M. Zhuo • M. Zhang • Z. Li • L. Sun • F. Lin  
Fujian College of Water Conservancy and Electric Power, Yong'an 366000, China  
e-mail: [Zhangrj\\_vip@163.com](mailto:Zhangrj_vip@163.com)

components and local information. ANFIS has strong nonlinear approximation functions and self-learning, adaptive characteristics. So they can be combined and give full play to the advantage of both [3].

## 156.2 Wavelet Analysis

Wavelet analysis is a window fixed but the shape can variable (variable bandwidth and when wide) changing in the time-frequency analysis method. It has adaptive time-frequency window: high frequency, frequency-domain window increased, the time window reduced, the time window expands, and the frequency domain window reduced. Wavelet analysis is the key to satisfy certain conditions, the introduction of the basic wavelet function  $\psi(t)$  to replace the Fourier transform of basic functions  $e^{-i\omega t}$ . The next is stretching and translation functions:

$$\psi_{a,b}(t) = |a|^{-1/2} \psi\left(\frac{t-b}{a}\right) \quad a, b \in \mathbb{R}, a \neq 0 \quad (156.1)$$

in which  $\psi_{a,b}$  is called the analysis wavelet or continuous wavelet;  $a$  is measurements (telescopic) factor, in a sense, is corresponding to frequency  $\omega$ ;  $b$  is the time (translation) factor, and it reacts to time's translation.

## 156.3 ANFIS

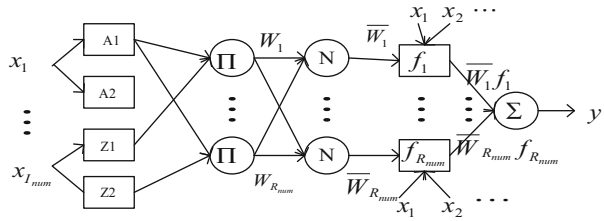
Artificial Fuzzy Neural Network (ANFIS) [4, 5] is combined by Artificial Neural Network and fuzzy theory. It uses ANN to construct the fuzzy system. According to input and output sample, it can automatically design and adjust the design parameters of the fuzzy system, and then it can achieve fuzzy system's self-learning and adaptive function. It can fit to complex input and output's linear and nonlinear mapping relations. So it is especially applicable to complex nonlinear hydrological system [6].

The network structure of Artificial Fuzzy Neural Network (ANFIS) is shown in Fig. 156.1.

In Fig. 156.1, the connecting line between the nodes only shows the flow of signal; it doesn't associate with weight. Square nodes represent nodes which have adjustable parameters; circular nodes represent nodes which don't have adjustable parameters. ANFIS's structure can be divided into five levels:

Level 1: Membership  $z_{ij} = \mu_i(x_j, \theta_i)$ ,  $i = 1, 2, L, M_{\text{num}}$ ,  $j = 1, 2, L, I_{\text{num}}$ ,  $M_{\text{num}}$  is the number of membership functions;  $I_{\text{num}}$  is the number of input variable;  $\mu(\cdot)$  is the generalized membership function; and the commonly used membership

**Fig. 156.1** Structure of ANFIS



functions are triangle membership function, trapezoidal membership function, Gaussian membership function, and bell-shaped membership function. The form of triangle and trapezoidal membership functions is simple, and their computing efficiency is high. However, because their membership function is constructed by linear line, the corner points of some specified parameters are not smooth enough. Gaussian function and bell-shaped membership function have smooth and simple representation, so they are the most commonly used form for definite fuzzy sets. Formula 1 is the bell-shaped membership function:

$$\mu_i(x, \theta_i) = 1 / \left[ 1 + \left[ \frac{x - c_i}{a_i} \right]^{2b_i} \right] \tag{156.2}$$

$x$  and  $\theta_i = [a_i \ b_i \ c_i]$  are respectively I the membership function's input and original reasoning parameter set.

Level 2:  $k$  the incentive intensity, in which  $R_{num}$  is the number of fuzzy rules.

$$W_k = \prod^{I_{num}} Z_{ij}, \quad i \in [1, 2, \dots, M_{num}], \quad i = 1, 2, \dots, R_{num} \tag{156.3}$$

Level 3: Normalized incentive intensity is the ratio of the rule's incentive intensity and the sum of all of the rule's incentive intensity.

$$\bar{W}_k = \frac{W_k}{\sum_{j=1}^{R_{num}} W_j}, \quad k = 1, 2, \dots, R_{num} \tag{156.4}$$

Its vector form is:

$$\bar{W} = [\bar{W}_1, \bar{W}_2, \dots, \bar{W}_{R_{num}}]^T \tag{156.5}$$

Level 4: Fuzzy rule's conclusion, which is accurate output.

$$f_i = p_{i1}x_1 + p_{i2}x_2 + \dots + p_{ij}x_i + r_i \quad i = 1, 2, \dots, R_{num}, \quad j = I_{num} \tag{156.6}$$

The parameter set, which is composed by all of  $\{p_{ij}, r_i\}$ , is called consequent parameter set.

Level 5: After weight-average, the overall output of the net can get:

$$y = \sum_{i=1}^{R_{num}} \overline{W}^T F \quad (156.7)$$

in which  $F = [f_1, f_2, L, f_{R_{num}}]$ .

This chapter uses hybrid learning algorithm to optimally select the parameter of ANFIS; this gets the smallest sum of square error between the final output result and the goal. The core idea of this algorithm is that in the forward calculation, it keeps the value of all of the original reasoning parameter unchanged and improves the value of consequent parameter by recursive least squares; then, it keeps the value of all of the improved consequent parameter unchanged and improves the value of original reasoning parameter by error back-propagation.

## 156.4 Wavelet-ANFIS Rainoff Forecast Model

Since wavelet analysis is a good multi-resolution frequency method and ANFIS has strong nonlinear approximation functions and self-learning, adaptive characteristics, the result will work better and give full play to the advantage of both if they are combined. It is named Wavelet-ANFIS rainfall forecast model.

In the model, first, data sequence is decomposed and reconstructed  $J$  times by Mallat wavelet, by which a low frequency signal  $C_J$  and the high frequency signals ( $D_1, D_2, \dots, D_J$ ) in  $J$  scale can be obtained; then ANFIS models are established for each decomposed signal ( $C_J, D_1, D_2, \dots, D_J$ ); after that it is time to determine the model parameters and to start prediction. Finally, the ultimate forecast result is obtained by synthesis of forecast results of the decomposed signals, and accuracy of the results is analyzed. The combined model flowchart is shown in Fig. 156.2.

## 156.5 Calculation Example

This chapter uses Wavelet-ANFIS model to find the suitable level of the runoff series; by some trying, it finds level 2 as the best level. The decomposed result is delayed as shown in Fig. 156.3.

In the figure we can clearly see that wavelet analysis can well separate the trend item and the wave item.

Next we use ANFIS to separately predict the trend item and the wave item.

We use runoff historical data from 1954 to 1993 as the modeling series and 1994–1996 as the predict series. In order to combine them in the end, we select the

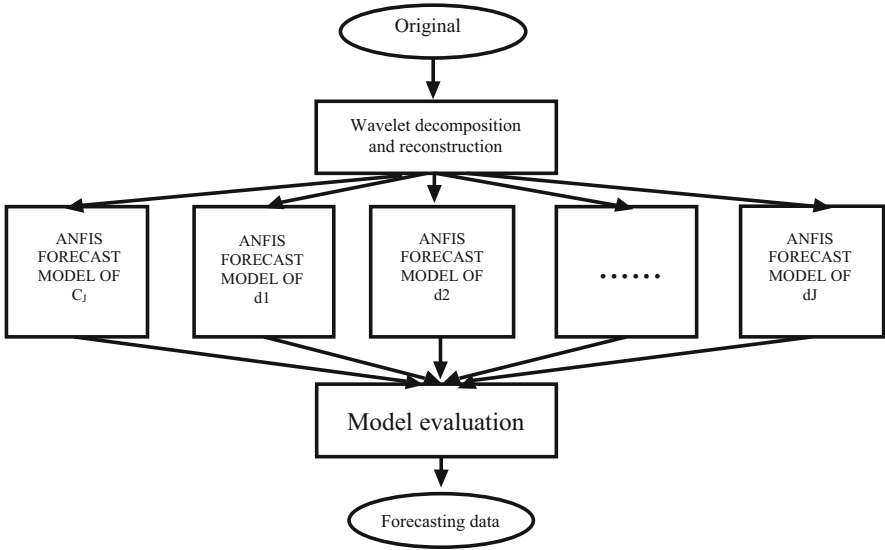


Fig. 156.2 Wavelet-ANFIS model flow diagram

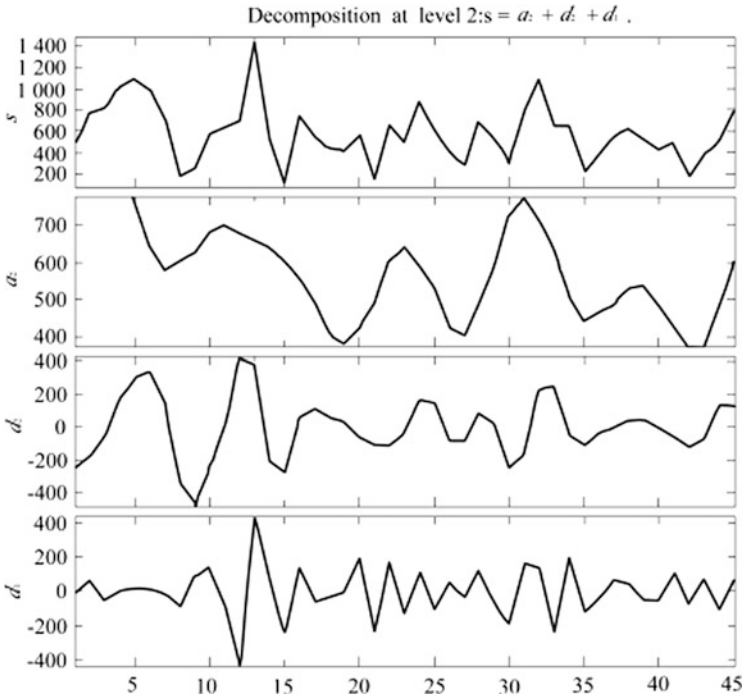
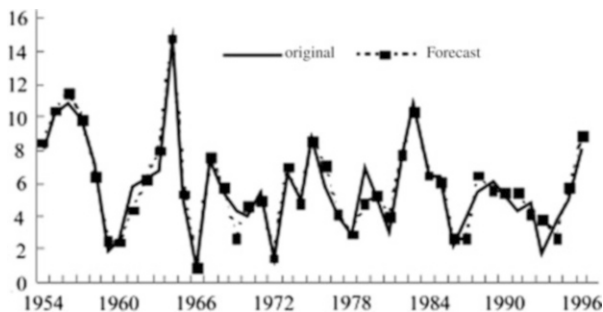


Fig. 156.3 The decomposed result of wavelet analysis

**Fig. 156.4** Comparison of original sequence and the final result



**Table 156.1** Compared result

Relative error (%)	<5	5–20	>20
Proportion (%)	94.97	5.03	0.00

same function for these series. This chapter adopts ANFIS function of MATLAB’s toolbox to predict these series; the type of function is bell-shaped membership function. After the 20th iteration, we can get a good simulating result. Then this chapter combined the result of this series to make the final result. The contrast result of the forecast result and the original sequence is delayed as shown in Fig. 156.4.

The evaluation of this model is delayed as shown in Table 156.1.

From Fig. 156.3 and Table 156.1, we can see that the result is very good. The relative error is nearly all less than 5 %. So we can use Wavelet-ANFIS model to forecast Lingxi Reservoir’s runoff. The precision is very high. This is also the basic research of other water resources’ researches.

### 156.6 Conclusion

This chapter combined wavelet analysis and ANFIS to predict Lingxi Reservoir’s runoff, which is called Wavelet-ANFIS model with the advantage of both; through series of analysis, we can find that Wavelet-ANFIS model is very fit to predict water level. The forecast result is very close to the original sequence and can be used to evaluate the influence of Lingxi Reservoir.

### References

1. Ma, X., Mu, X., & Guo, H. (2012). Reservoir monthly runoff forecast model based on wavelet-ANFIS analysis. *Hydroelectric Energy*, 1(1), 12.
2. Ma, X., He, X., & Zhao, D. (2012). BP network hidden layer on the water quality impact analysis results of the evaluation. *Hydroelectric Energy*., 20(3), 121.



3. Wang, W., Ding, J., & Li, Y. (2005). *Hydrological wavelet analysis* (pp. 32–36). Beijing: Chemical Industry Publishing House.
4. Jang, J. S. R. (2011). ANFIS: Adaptive-network-based fuzzy inference system. *IEEE Transactions on System, Man and Cybernetics*, 2, 235–238.
5. Zhang, Z., & Sunetc, C. (2010). *Neuron fuzzy and soft computing* (pp. 86–92). Xi'an: Xi'an Communication University Press.
6. Zhang, B., et al. (2011). *Research on Poyang Lake* (pp. 22–28). Shanghai: Shanghai Technology Press.

# Chapter 157

## Visualization of Clustered Network Graphs Based on Constrained Optimization Partition Layout

Fang Huang, Wenjie Xiao, and Hao Zhang

**Abstract** Hybrid layout is a common visualization technique for clustered network graphs. Since most previous hybrid layout methods do not consider a reasonable balance between screen utilization and layout aesthetics of the network graphs, the inappropriate partition of the display region may result in unpleasant display effect of network graphs. This chapter proposes to address this problem with nonlinear constrained optimization techniques. This chapter analyzes why the circular algorithm would fail in region partition. To ensure that every subgroup of network nodes can be assigned to a rectangular region, the maximal utilization of the display area is taken as an objective function and the rectangular ratio is taken as constraints. The constrained optimization layout model leads to efficient balance between regional utilization and layout aesthetic. Experimental results show that the constrained optimal partition layout generates more balanced relation network graphs with better visual effects.

### 157.1 Introduction

Visualization of clustered network graphs is an important research topic in social relations modeling and analysis based on network information technology. To clearly display relation networks, it is necessary to construct a clustered graph by dividing all nodes into clusters according to network closeness degrees. In the process of clustered graphs visualization, it is crucial to decide how to display all the nodes in various clusters in the screen. Besides, in order to achieve a balanced aesthetic effect, we need to spread nodes to the whole area. Integrating the two issues together, we can obtain efficient and coordinated picture by optimizing

---

F. Huang (✉) • W. Xiao • H. Zhang  
School of Information Science and Engineering, Central South University,  
Changsha 410083, China  
e-mail: [hfang@mail.csu.edu.cn](mailto:hfang@mail.csu.edu.cn)

layout of nodes for clustered network graphs. The optimal layout can present the nodes and associations of network diagram more clearly and evenly.

In 2008, Huang proposed a hybrid layout method [1] which is commonly used in visualization of clustered graphs. This algorithm firstly divides all the nodes into subgroups according to the tightness of contact in relation network structure, then partitions the screen into several rectangular regions in accordance with the number of subgroups, and finally layouts nodes in the rectangular regions. In the hybrid layout, it is supposed that the display area is fully utilized, and a reasonable length-to-width ratio is required so that the layout is aesthetic. If only one of the two constraints are not satisfied, the program executing would be terminated, which will cause the subgroups to be not incompletely distributed to the rectangular region and sequentially impacts on the effect of network layout. Therefore, this chapter proposes an optimization region partition method for the hybrid layout, which merges utilization rate of display area, aesthetic proportion of rectangle as well as uniform density of nodes. We employ the constrained nonlinear optimization model to conduct region partition, which can ensure that every subgroup can occupy a rectangular region by dividing the screen completely and meet the conditions of aesthetic layout. As a result, the effect of visualization of relation network graphs can be improved.

So far, the visualization of clustered graphs is still an open topic in social relations modeling and analysis. In 2006, Eades proposed that the planar clustered graphs can be converted into convex polygons and can be drawn by a straight line, which is named as straight-line algorithm [2]. Since confusion visualizations are usually caused by many cross-edges in the clustered graphs, in 2007, Omote provided a novel force-directed algorithm, which reduces the number of the cross-edges with clearer graphs [3]. In 2008, Huang made a synthesis process for the network graphs visualization, which integrates partition of clusters and layout of nodes [1]. In the scheme, the hierarchy-clustering algorithm is employed to build a node-clustered tree, then utilize the region partition algorithm to assign the display area for every cluster and, lastly, take force-directed algorithm to layout the nodes in each area. In 2009, Battisia raised a directly visualized method that organizes all nodes by clustered trees and the network graph shows the shape of the tree [4]. In 2012, Liu achieved more effective clustering while considering the relationship strength between nodes and more attributes of nodes [5]. Since the region partition algorithm plays an important role in clustered graphs visualization, it has been paid much attention. In 2005, Nguyen investigated the region partition algorithms, such as circular region partition, Squarified treemaps, and Sunburst, and compared the balance of nodes and orderliness of the edges [6]. The conclusion is that the circular region partition is better in effect and aesthetic. So, in 2008, Huang adopted circular region partition as a part of his hybrid layout process [1]. However, in the method, the unbalanced node of clustered trees easily leads to an incomplete partition process. The purpose of the study is to explore an optimal region partition to improve visual effect of clustered graphs.

## 157.2 Analysis of Circular Region Partition

Hybrid layout is composed of nodes clustering [7], circular region partition [6], and force-directed layout [8]. Firstly, nodes closely related to each other are classified as various subgroups by nodes clustering. Then, the circular region partition is adopted to divide the display screen into several rectangular regions roughly so that each subgroup is positioned in their respective rectangle. Finally, the force-directed layout is applied to spread all the nodes of the subgroup in the rectangle in order to achieve balanced effect. In the hybrid layout, the circular region partition plays a decisive role for the overall effect of the relation networks visualization. The circular region partition assigns a rectangular region to each subgroup one by one clockwise along the four edges of the display screen, which is a key factor for aesthetics and integrity of the circular region partition.

- Calculating the Size of Rectangles

In Fig. 157.1, the width of display area in the screen is denoted by  $W$ , and the height is denoted by  $H$ . Assume there are  $n$  subgroups to be assigned, and we denote the number of nodes in subgroup  $i$  by  $wg_i$  to serve as the subgroup weight. For example,  $m$  subgroups are divided into rectangles in the left part of the display area as shown in Fig. 157.1. Since the area of rectangle  $R_i$  is  $S_i$ , the width  $w_i$ , height  $h_i$ , and  $S_i$  of each rectangle can be calculated, respectively, by the following:

$$Wp_i = wg_i / \sum_{i=1}^n wg_i \quad (157.1)$$

$$S_i = W * H * Wp_i. \quad (157.2)$$

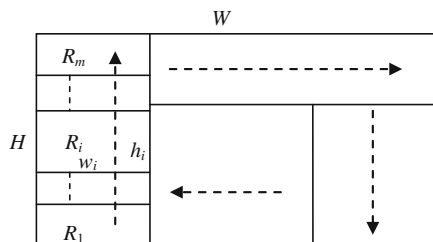
In formulas (157.1) and (157.2),  $Wp_i$  represents the percentage of the weight of subgroup  $i$  in all subgroups. Expression (157.2) means the area  $S_i$  of the rectangle  $R_i$  is a percentage of the total area, which corresponds to the ratio between the number of subgroup nodes and the total number of all nodes in the network.

Expression (157.3) shows that  $h_i$  is the percentage of the vertical height of the display screen, which is the ratio between the number of subgroup nodes and the total number of nodes in the  $m$  subgroups along the vertical edge.

$$h_i = H \left( wg_i / \sum_{i=1}^m wg_i \right) \quad (157.3)$$

$$w_i = W \left( \sum_{i=1}^m wg_i / \sum_{i=1}^n wg_i \right) \quad (157.4)$$

**Fig. 157.1** Circular region partition



From expression (157.4),  $w_i$  is decided by the ratio of the number of  $m$  subgroup nodes to the total number of nodes. According to the above formulas, it is sure that the sum of all rectangular areas is equal to the display area. In other words, the circular partition algorithm takes full utilization of display area as a precondition.

- Constraining the Proportion of Width and Height

In addition to the width and height of a rectangle, the ratio between them must also meet the aesthetic criteria of a rectangle in the circular partition. Otherwise, the algorithm will be terminated. That is to say, not only the ratio of rectangles decides the aesthetics but also the probability of partition success. Here, we use  $\lambda_i = w_i/h_i$  to represent the ratio between width and height. The ratio is set to be larger than 0.67 by trial and error and it is manually set in order to achieve the best rectangular layout [6].

- Process of Circular Region Partition

In Fig. 157.1, the partition divides the first rectangle on the bottom-left corner along the direction of the vertical dashed arrow. First, the area, height and width of the rectangle is calculated according to expression (157.2) (157.3) and (157.4). Next, the width/height ratio of the rectangle is checked. If it meets the constraints, try to assign two subgroups in the direction and calculate the height, area, and width of the two rectangles successively. Similarly, check whether the ratio of each rectangle meets the requirements. Trying and computing repeatedly, an extra subgroup is added in each iteration until one of these rectangles doesn't meet the proportional constraints. At this point, the partition in the current direction will be ended. The partition continues in the next direction clockwise. If the attempts in all direction are finished but the remaining subgroups cannot be assigned to any suitable region, the circular region partition fails. The number of rectangles arranged on one edge of the display screen is  $m$ , which is in accordance with the ratio constraints [6].

In the circular region partition, dividing rectangles is based on assuming rectangular seamless arrangement. If the premise does not hold, the rectangular formula is not valid, and the segmentation process cannot be carried out. Thus, the first rule of the circular partition algorithm is the full utilization of display area. Secondly, the aspects ratios of the rectangle should satisfy the constraint condition to ensure

the aesthetics of networks. Regardless of whether the other subgroups are partitioned to the region, the algorithm will terminate if the constraint cannot be established. In other words, the circular partition has the effectiveness and aesthetics as two absolutely independent conditions. However, in applications, the aesthetics should be on the basis of the validity. It is first to assure that all subgroups can be assigned to a rectangular region, then to consider the aesthetics. Meanwhile, the aesthetic criteria can be flexible, if only a good visual effect is kept. Therefore, to take the above factors into consideration, we propose a nonlinear constraint optimization for the region partition layout.

### 157.3 Constrained Optimization Partition Layout

Constrained optimization layout model includes an objective function and constraints [9]. For the region partition of subgroups after nodes clustering, it is the main goal to guarantee that all subgroups can get the right rectangle from the display area. On this basis, adjusting the size and position of the rectangle leads to maximum utilization of the display screen. We designed the optimization model which aims to maximize the sum area of all rectangles and takes rectangular aesthetics and reasonable position as constraints. A set of the rectangles with optimal size and position are found by solving the model so as to achieve the optimal layout.

#### 157.3.1 *The Effect of Rectangular Aspects Ratio on Aesthetics*

The aspects ratio decides the shape and size of rectangles in the process of region partition. Its purpose is to layout nodes of the subgroup in the region, namely, the regional division of the subgroups. We should conduct the quantitative evaluation for the aesthetic effect of display area objectively. Since visualization should meet the aesthetic standards of people, the feasible solution of constrained optimization model must be limited in the acceptable range. Therefore, the quantitative range of aesthetics can be derived to be the model constraints.

Battista proposed a method to measure the effect of visualization according to aesthetic criteria which include the overall edge lengths  $d_i$ , variance  $e_i$  of edge lengths, the number of cross-edges  $c_i$ , and variance  $a_i$  of an included angle between associate edges [10]. If the overall edge lengths are longer, the layout of nodes is more dispersed. Minimizing the variance of edge lengths can result in more balanced layout of nodes to a certain extent. The less cross-edges show more reasonable layout of nodes and smaller variance of angles makes network diagrams clearer. We build a sample dataset of relationship networks to evaluate the effect of

rectangle aspects ratios on aesthetics. For  $k$  relationship networks, the comprehensive assessment for the above four rules can be defined as following formulas:

$$D_{ave} = \frac{1}{k} \sum_{i=1}^k d_i, E_{ave} = \frac{1}{k} \sum_{i=1}^k e_i, C_{ave} = \frac{1}{k} \sum_{i=1}^k c_i, A_{ave} = \frac{1}{k} \sum_{i=1}^k a_i. \quad (157.5)$$

Here  $D_{ave}$ ,  $E_{ave}$ ,  $C_{ave}$ , and  $A_{ave}$  are the average values of total edge lengths, variance of edge lengths, the number of cross-edges, and variance of angle, respectively. In addition,  $DL_i$ ,  $DE_i$ ,  $CE_i$ , and  $AN_i$  defined below are the proportion of deviation from  $i$  network graph to the average value of all networks on the above four rules, respectively.

$$DL_i = \frac{d_i - D_{ave}}{D_{ave}}, DE_i = \frac{e_i - E_{ave}}{E_{ave}}, CE_i = \frac{c_i - C_{ave}}{C_{ave}}, AN_i = \frac{a_i - A_{ave}}{A_{ave}} \quad (157.6)$$

We also define  $NET_i$  as a comprehensive estimation for network  $i$  in formula (157.7).

$$NET_i = DL_i - DE_i - CE_i - AN_i \quad (157.7)$$

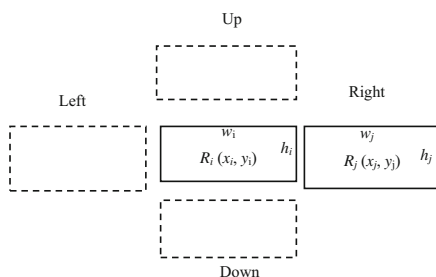
In the current relation network, when  $d_i$  gets larger, the layout of nodes is more dispersed. If  $d_i - D_{ave}$  is positive, it shows that the overall edge lengths in the current network will be larger than the average value. That is,  $DL_i$  contributes to the comprehensive assessment, so it is a positive sign. For smaller  $e_i$ , the layout of nodes is more uniform. If  $e_i - E_{ave}$  is positive, the variance of edge lengths will be greater than the average; at the same time, the  $DE_i$  has a negative impact on the comprehensive assessment, so we can use minus sign for the item.  $CE_i$  and  $AN_i$  are similar to  $DE_i$ . In other words, when  $DL_i$  is positive and  $DE_i$ ,  $CE_i$ , and  $AN_i$  are negative, the  $NET_i$  is positive. That means that the aesthetics of current network is better than the average effect of all networks. The comprehensive assessment reflects the relative value of visualization between the current network and all networks.

In order to determine the effect of rectangular aspects ratios on the aesthetics of the display area, take the network containing 21 nodes as inputs, set the rectangular aspects ratio  $\lambda$  from 0.1 to 0.7 for constraint layout, and then get the quantitative evaluations, respectively.

In Table 157.1, when the rectangular aspects ratio equals 0.1, 0.2, 0.3, 0.4, and 0.5, respectively, the comprehensive evaluations are negative. That indicates the aesthetics of these networks is poorer than the average effect. When the ratio is 0.6 or 0.7 and the assessments are positive, it means a better effect. So we set the rectangular aspects ratio greater than 0.6 as the constraint of rectangular aesthetics.

**Table 157.1** Quantitative evaluations of visual networks

Ratio $\lambda$	Various aesthetics criteria				
	$d_i$	$e_i$	$c_i$	$a_i$	$NET_i$
0.1	1,235	14.08	58	11.08	-0.4789
0.2	1,220	12.9	44	15	-0.15
0.3	1,204	16.65	38	18	-0.34
0.4	1,180	16.21	32	16	-0.03
0.5	1,157	12.35	26	23.71	-0.13
0.6	1,162	13.73	12	16	0.64
0.7	844	12.81	8	17	1.1



**Fig. 157.2** Nonoverlapping between rectangles

### 157.3.2 Nonoverlapping Criteria

All rectangles in the display area should not be overlapping, which is a rigorous constraint in the process of region partition. In Fig. 157.3, there are two rectangles:  $R_i$  and  $R_j$ . Supposing  $R_i$  is fixed, there are four situations that  $R_j$  is not overlapping with  $R_i$ : the  $R_j$  is on the right, left, up, and down of the  $R_i$ , respectively.

In Fig. 157.2,  $x_i$  and  $y_i$  are the center coordinate of  $R_i$ ,  $w_i$  and  $h_i$  is the width and height, and  $x_j$  and  $y_j$  are the center of the  $R_j$ . For example, when the  $R_j$  locates on the right of the  $R_i$ , the nonoverlap rule is that the abscissa of  $R_j$  left edge must be greater than the abscissa of  $R_i$  right edge. When the  $R_j$  is on the left of  $R_i$ , the abscissa of the  $R_j$  right edge is smaller than the abscissa of the  $R_i$  left edge. Similarly,  $R_j$  stands above  $R_i$ ; the ordinate of  $R_j$  bottom must be greater than the ordinate of  $R_i$  top, and  $R_j$  is below  $R_i$ ; the ordinate of  $R_j$  top is smaller than the ordinate of  $R_i$  bottom.

### 157.3.3 Nonlinear Constrained Optimization Layout Model

In optimization layout, the objective function and constraints are defined by the following.

- Objective Function for the Maximizing Utilization of Display Area



In layout partitioning, every subgroup must correspond with a rectangle in display area and minimize the remaining area. In other words, we maximize the utilization to reduce center-of-gravity shift and unbalance of networks in the display area, which can cause too much blank areas. Therefore, the utilization of display screen is taken as the objective function:

$$\max \sum_{i=1}^n w_i h_i \quad (157.8)$$

Here,  $w_i$  and  $h_i$  are the width and height of  $R_i$ ;  $n$  is the number of all rectangles.

- Constraints of Rectangular Aesthetics

$$0.6 \leq \lambda_i = \frac{w_i}{h_i} \leq 1.67 \quad i = 1, 2, \dots, n \quad (157.9)$$

Here  $\lambda_i$  is the ratio between width and height of  $R_i$ .

- Nonoverlapping Constraints

$$\left(x_i + \frac{w_i}{2}\right) - \left(x_j - \frac{w_j}{2}\right) \leq 0, \quad i = 1, 2, \dots, n, j = 1, 2, \dots, n, i \neq j \quad (157.10)$$

or

$$\left(x_i - \frac{w_i}{2}\right) - \left(x_j + \frac{w_j}{2}\right) \geq 0 \quad i = 1, 2, \dots, n, j = 1, 2, \dots, n, i \neq j \quad (157.11)$$

or

$$\left(y_i + \frac{h_i}{2}\right) - \left(y_j - \frac{h_j}{2}\right) \leq 0 \quad i = 1, 2, \dots, n, j = 1, 2, \dots, n, i \neq j \quad (157.12)$$

or

$$\left(y_i - \frac{h_i}{2}\right) - \left(y_j + \frac{h_j}{2}\right) \geq 0 \quad i = 1, 2, \dots, n, j = 1, 2, \dots, n, i \neq j \quad (157.13)$$

Formulas (157.10)–(157.13) show the constraints of rectangular center coordinate to ensure nonoverlapping and just located on right, left, up, and down.

- Constraints for Exceeding Boundary

$$h_i \leq Hi = 1, 2, \dots, n \quad (157.14)$$

$$w_i \leq Wi = 1, 2, \dots, n \quad (157.15)$$

$$\frac{w_i}{2} \leq x_i \leq W - \frac{w_i}{2} \quad i = 1, 2, \dots, n \quad (157.16)$$

$$\frac{h_i}{2} \leq y_i \leq H - \frac{h_i}{2} \quad i = 1, 2, \dots, n \quad (157.17)$$

The constraints limit of the width and height of rectangles should be smaller than the width and height of display area, respectively.

- Relaxed Constraint for the Rectangular Area

$$H \times W \times Wp_i \times (1 - 5\%) \leq S_i \leq H \times W \times Wp_i \times (1 + 5\%) \quad (157.18)$$

The constraint of rectangular areas, formula (157.18), is a relaxation condition based on the rectangular area accounted for the proportion of total area which is equal to the subgroup weight to total weight. Since small changes of the rectangular area bring about little impact on the visual aesthetics, the relaxation can help to expand the range of feasible solutions. The relaxed range of rectangular area is 5%.

The best partition layout is determined by calculating the width and height of rectangles and the abscissa and ordinate of centers in constrained optimization model.

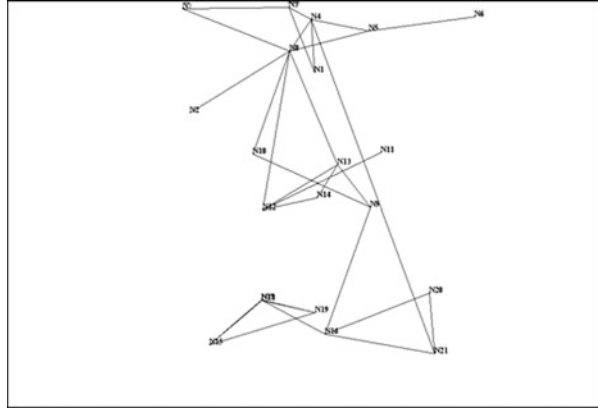
## 157.4 Comparison and Analysis of Visualization

Since the objective of the proposed optimal region partition is to avoid layout failures of the circular region partition, it is necessary to compare the visual effect for circular region partition and optimal region partition. Four relationship networks containing 21, 40, 80, and 160 nodes, respectively, are chosen to implement the experiments.

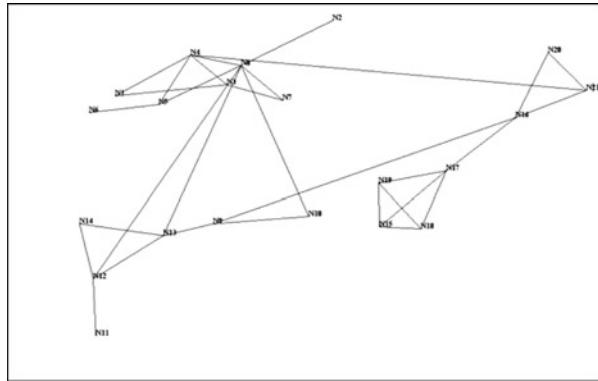
### 157.4.1 Visual Comparison

Figure 157.3 is for the circular region partition; since the ratio of the rectangle in the middle of display area does not satisfy the aesthetic constraint, it leads to partition failure. However, the process of layout of nodes has to be implemented by the rectangle with inappropriate proportion so that it is more crowding and has many cross-edges in the middle. For optimal region partition, Fig. 157.4 shows more scattered nodes, fewer cross-edges, and clearer visualization.

**Fig. 157.3** 21 nodes by circular partition layout



**Fig. 157.4** 21 nodes by optimizing partition layout

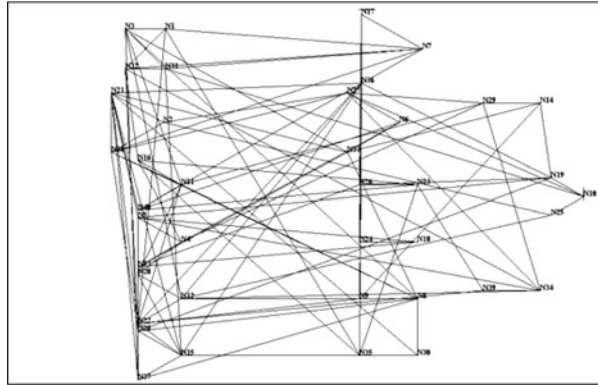


In Fig. 157.5, since the rectangle of inappropriate ratio stands on the left space of display area, it results in nonuniform distribution of nodes on the left. After optimizing the partition layout, nodes in Fig. 157.6 are distributed in a more uniform way.

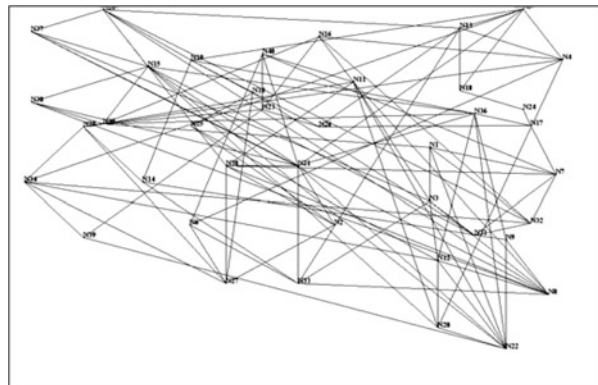
Figure 157.7 is a network diagram by the circular region partition for 80 nodes, and all the nodes are concentrated at the middle of the display area. The visualization result is poor because of small distances among the nodes. After optimal region partition, the nodes in Fig. 157.8 are more uniformly and clearly distributed.

Figures 157.9 and 157.10 show the visualization results for a network with 160 nodes using circular and optimization partition layout, respectively. Comparing with Fig. 157.9, Fig. 157.10 presents the network distribution with better visual effect. In particular, we choose the network with fewer connections, in order to better contrast on vision.

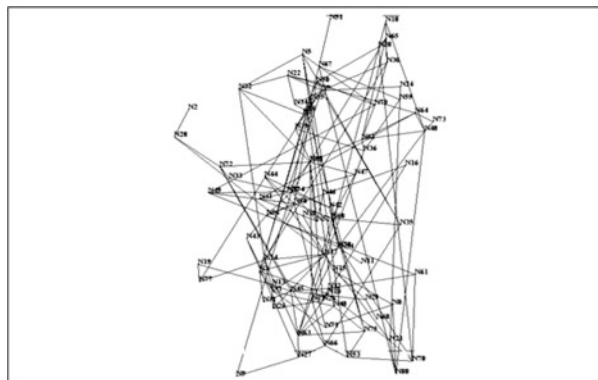
**Fig. 157.5** 40 nodes by circular partition layout



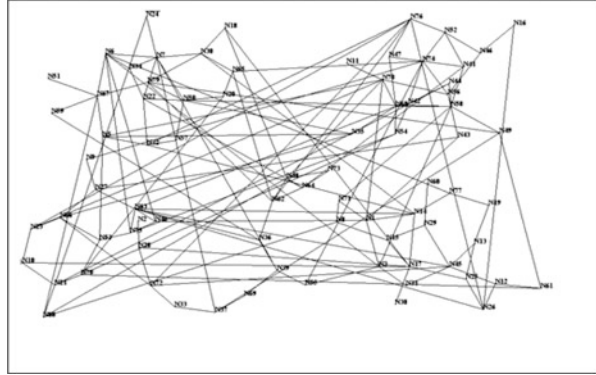
**Fig. 157.6** 40 nodes by optimizing region layout



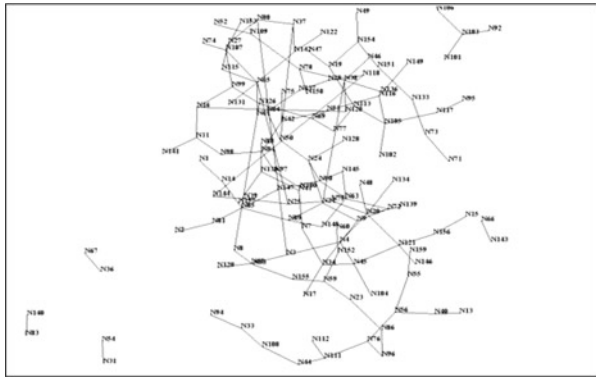
**Fig. 157.7** 80 nodes by circular partition layout



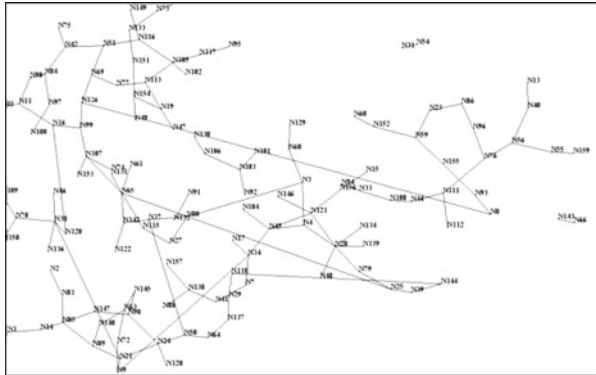
**Fig. 157.8** 80 nodes by optimizing partition layout



**Fig. 157.9** 160 nodes by circular partition layout



**Fig. 157.10** 160 nodes by optimizing partition layout



**Table 157.2** Aesthetics evaluations of 21 nodes

Aesthetics criteria	$d$	$e$	$c$	$a$
Circular partition	2,030	125.75	26	21.3
Optimizing partition	2,046	162	20	18.6

**Table 157.3** Aesthetics evaluations of 40 nodes

Aesthetics criteria	$d$	$e$	$c$	$a$
Circular partition	11,232	115	985	24.2
Optimizing partition	16,287	210	754	29.3

**Table 157.4** Aesthetics evaluations of 80 nodes

Aesthetics criteria	$d$	$e$	$c$	$a$
Circular partition	22,500	48.5	3,048	29.1
Optimizing partition	34,596	62.7	2,169	25.3

**Table 157.5** Aesthetics evaluations of 160 nodes

Aesthetics criteria	$d$	$e$	$c$	$a$
Circular partition	2,076	65.5	385	101
Optimizing partition	2,112	76.4	283	95

### 157.4.2 Assessment on Aesthetics of Visualization

We elaborate the quantitative evaluation of visualization for the above four networks including 21, 40, 80, and 160 nodes, respectively. Aesthetics criteria include the total edge lengths  $d$ , edge length variance  $e$ , the number of cross-edges  $c$ , and variance of angle  $a$ . The statistical results are shown in Tables 157.2, 157.3, 157.4, and 157.5 separately.

From the above four tables, it can be concluded that optimization partition layout can produce greater total length of edges, fewer cross-edges, and smaller angle variance, which result in wider distribution of nodes and more uniform and clearer network. The most critical point is the cross-edge by using optimization layout less than the number generated by circular region partition. It makes network connections clearer. In the four criteria, only the edge length variance is larger by circular region partition, which indicates that the distance between nodes should not be more uniform, but a smaller variance of edge length cannot absolutely guarantee a uniform distribution of nodes. The optimal partition performs better results on the three aesthetic criteria. Synthetically, the relation network graphs produced from optimization layout is superior to circular partition layout.

## 157.5 Conclusion

In the optimization region partition, the nonlinear constrained optimization model is employed to avoid the failure partition in circular region layout. It effectively balances the contradictions of the aesthetics and utilization by taking the utilization of the display area as the objective function and putting the rectangular ratio as constraints. It ensures that all subgroups of nodes can be assigned to a rectangular region. We randomly generated 60 relationship networks as the experimental inputs, then ran programs with the following three steps: nodes clustering, optimizing region partition, and force-directed node layout. Since the optimization region partition is implemented on LINGO 13 and other two parts are programmed by Java, the total running time is the sum of the three parts. The experiment results showed that the average running time of circular region partition and optimal region partition are 1.23 and 1.8 s, respectively. However, the latter has a 92 % average utilization of the display screen in all experiment networks and effectively avoids the layout failures. The results showed that the optimization partition layout not only achieved better quantitative evaluations on aesthetic criteria but also made the visual effects of relationship networks more balanced and clearer. In future work, we will further improve the algorithm in order to display scalable large-scale relation networks.

**Acknowledgements** This work was supported by Project 61073105 of National Natural Science Foundation of China.

## References

1. Huang, M. L., & Nguyen, Q. V. (2008). Large graph visualization by hierarchical clustering. *Journal of Software, China*, 9(8), 1933–1946.
2. Eades, P., Feng, Q. W., Lin, X. M., & Nagamochi, H. (2006). Straight-line drawing algorithms for hierarchical graphs and clustered graphs. *Algorithmica*, 44(1), 1–32.
3. Omote, H., & Sugiyama, K. (2007). Force-directed drawing method for intersecting clustered graphs. *Proceedings of Asia-Pacific Symposium on Visualization 2007 (APVIS2007)* (pp. 85–92). Sydney: IEEE.
4. Battista, G. D., Drovandi, G., & Frati, F. (2009). How to draw a clustered tree. *Journal of Discrete Algorithms*, 7(4), 479–499.
5. Liu, X., Glänzel, W., & Moor, B. D. (2012). Optimal and hierarchical clustering of large-scale hybrid networks for scientific mapping. *Scientometrics*, 91(2), 473–493.
6. Nguyen, Q. V., & Huang, M. L. (2005). EncCon: An approach to constructing interactive visualization of large hierarchical data. *Information Visualization*, 4(1), 1–21.
7. Huang, M. L., & Nguyen, Q. V. (2007). A fast algorithm for balanced graph clustering. *Proceedings of IEEE 11th International Conference Information Visualization (IV'07)* (pp. 46–52). Zurich: IEEE.
8. Eades, P. (1984). A heuristic for graph drawing. *Utilitas Mathematica*, 42(11), 149–160.
9. Yuan, Y. X., & Sun, W. Y. (1999). *Optimization theory and methods* (pp. 35–75). Beijing, China: Science Press.
10. Battista, G. D., Eades, P., Tamassia, R., & Tollis, I. G. (1999). *Graph drawing: Algorithms for the visualization of graphs* (pp. 187–234). Upper Saddle River, NJ: Prentice-Hall.

# Chapter 158

## An Ultra-Wideband Cooperative Communication Method Based on Transmitted Cooperative Reference

Tiefeng Li, Ou Li, and Zewen Zhou

**Abstract** In order to decrease the power waste of relay node, the paper presents a novel ultra-wideband cooperative communication method that uses two relay nodes to transmit reference impulses and data impulses separately. A transmitted cooperative-reference UWB cooperative communication model is developed in this paper. Based on the model and sampling expansion approach, a closed-form SER expression was deduced for delay-hopped transmitted-reference UWB systems which use cooperation strategy of decode and forward relaying and equal-gain combining. Simulation results show that the transmitted cooperative-reference method can obtain multi-order diversity gains.

### 158.1 Introduction

It becomes the current research hotspot to involve cooperative diversity technique in the ultra-wideband communication system because of the additional spatial diversity. By doing this, both transmission reliability and communication coverage will be substantially enhanced. At present, studies of cooperative UWB technology are mainly focused on multiband orthogonal frequency division multiplexing ultra-wideband (MB-OFDM-UWB). There are a few essays which refer to the issue that combines impulse radio ultra-wideband with cooperative communication technique and few researches on cooperative communication that are based on the transmitted reference (TR) [1]. The paper just simply combines two techniques and chooses a relay node with best SNR to forward. In fact, it had developed a relay-forward

---

T. Li (✉) • O. Li • Z. Zhou  
China National Digital Switching System Engineering and Technological Research Center,  
Zhengzhou 450002, China  
e-mail: [13838146019@126.com](mailto:13838146019@126.com)



channel without diversity gain [1]. As well known, TR techniques can implement reliable communication with low complexity in the random or unknown channel [2]. Therefore, it is necessary to research the performance of cooperative TR-UWB system.

The transmitted-reference UWB can be classified into three main categories: time-domain transmitted reference [3], frequency-domain transmitted reference [4], and code-domain transmitted reference [5]. In the time-domain transmitted-reference system, since the nodes not only transmit data impulse but also send reference impulses that independent of the data, so it wastes half of the total energy. If all cooperative nodes are using this modulation to participate in cooperative, apparently the relay nodes have to pay half power as much as the source node. This cooperative approach can be named “transmitted-reference UWB cooperative.” This paper presents a new method that is named “transmitted cooperative-reference UWB cooperative.” It uses two relay cooperative nodes to send cooperative-reference impulses and cooperative-data impulse separately. Obviously, it is effective to save transmitting power by using this cooperation strategy.

The paper is organized as follows: In Sect. 158.2, we build transmitted cooperative-reference UWB communication system model. Next, in Sect. 158.3, based on the sampling expansion approach, a closed-form SER formulation was deduced from the delay-hopped transmitted-reference UWB systems which use cooperation strategy of decode and forward relaying, equal-gain combining. Then, Sect. 158.4 details the experiment results and discussions.

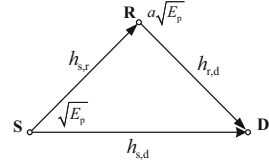
## 158.2 Transmit Cooperative Reference UWB Cooperative Communication System Model

We consider a cooperation strategy with two phases and three nodes in the cooperative model, which can be extended to the multi-node. Figure 158.1 shows the specific three-node cooperative model:

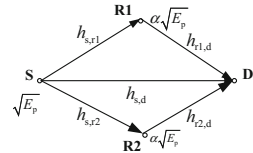
Compared to the three-node model, the transmitted cooperative-reference UWB cooperative communication model added an additional relay node as shown in Fig. 158.2. The model includes a source node ( $S$ ), two relay nodes ( $R1$  and  $R2$ ), and a destination node ( $D$ ). Firstly,  $S$  sends a signal;  $R1$ ,  $R2$ , and  $D$  receive it. Secondly,  $R1$  and  $R2$  forward the decoded reference impulses and data impulses separately, and then  $D$  receives them.

Assume that the system can always find two relay nodes  $R1$  and  $R2$ , which can establish a good time-synchronous accordance with the destination node through the upper layer cooperate protocol. Assume  $R1$  only sends the cooperative-reference impulse, and  $R2$  only sends cooperative-data impulse. Good synchronization means that when the impulse transmitted from each relay nodes arrived to the destination node, the time delay between the reference impulse and data impulse sent separately by  $R1$  and  $R2$  has the same effect as  $R$  in Fig. 158.1. So, the model is

**Fig. 158.1** Transmitted-reference UWB cooperative model



**Fig. 158.2** Transmitted cooperative-reference UWB cooperative model



equivalent to the classic three-node model. R in Fig. 158.1 can replace the role of R1 and R2 in Fig. 158.2. In cooperative communication network with multi-nodes, it is possible to satisfy such conditions.

Figure 158.2 depicts the transmitted cooperative-reference UWB cooperative model; in stage one, the source node broadcasts to the destination node and the relay nodes. The signals received by destination and relay, respectively, are defined as:

$$\begin{aligned}
 y_{s,d} &= \sqrt{E_p}h_{s,d}s + n_{s,d} \\
 y_{s,r1} &= \sqrt{E_p}h_{s,r1}s + n_{s,r1} \\
 y_{s,r2} &= \sqrt{E_p}h_{s,r2}s + n_{s,r2}
 \end{aligned}
 \tag{158.1}$$

where  $E_p$  is transmission energy carried by single impulse and  $S$  is the symbol signal calculated as:

$$s(t) = \sum_{n=0}^{\frac{N_s}{2}-1} (p(t - 2nT_f) + d_0p(t - nT_f - T_r))
 \tag{158.2}$$

$h_{s,d}, h_{s,r1}, h_{s,r2}, h_{r2,d}$ , and  $h_{r1,d}$  are the channel coefficients correspond to  $S$  to  $D$  or  $S$  to  $R$ , and  $R$  to  $D$ , respectively, are modeled as zero mean, complex Gaussian random variables with variances  $\delta_{s,d}^2, \delta_{s,r1}^2, \delta_{s,r2}^2$ , etc.  $n_{s,d}, n_{s,r1}$ , and  $n_{s,r2}$  are the additive white Gaussian noise.  $T_f$  is the delay time between the reference impulse and data impulse. A frame consisted of a pair of impulse, that is, reference impulse and data impulse.  $T_r$  is the time delay between the reference impulse and data impulse.  $N_s$  is the number of impulse in one symbol signal.  $d_0$  is the binary data bit.

In stage two, the relay nodes separately decode and only forward reference impulse or data impulse. For simplicity to derivat and analyze, the role of R1 and R2 in Fig. 158.2 can be replaced by R in Fig. 158.1.

After two periods of this stage, the destination received three copies of signals via source and relay channel. The received SNR of the destination is compounded with the SNR of  $S$  and  $R$ . The optimal method of maximizing the total SNR is maximum-ratio combining (MRC). The output SNR via MRC is equal to the summation of SNR of each branch. However, it needs to know the real-time SNR of each branch. It implies that we need to obtain the real-time channel estimation. On the contrary, the advantages of transmitting-reference scheme can implement reliable receive without channel estimation. It is more suited for transmitted-reference scheme with the gain-equal-combination strategy. Based on the gain-equal combination, the output  $r_{rec}$  at destination is given by

$$\begin{aligned}
 r_{rec} &= y_{s,d} + y_{r,d} \\
 &= (\sqrt{E_p}h_{s,d} + a\sqrt{E_p}h_{s,r}h_{r,d})s + (n_{s,d} + ah_{r,d}n_{s,r} + n_{r,d}) \quad (158.3) \\
 &= s' + n'
 \end{aligned}$$

where

$$\begin{aligned}
 s' &= (\sqrt{E_p}h_{s,d} + a\sqrt{E_p}h_{s,r}h_{r,d})s = \sqrt{E'_p}\alpha_1s \\
 E'_p &= E_p(1 + ah_{s,r})^2, n' = (n_{s,d} + ah_{r,d}n_{s,r} + n_{r,d}).
 \end{aligned}$$

We assume that three transmission channels are independent and identically obey the Rayleigh distribution. The fading coefficient here is represented by the variable  $h$ . Since all signals transmitted in the channel of  $S$  to  $R$  have been processed by relay, we employed a constant  $\bar{h}_{s,r}$  to represent the statistical average of the fading coefficient of the channel. Assuming that the components of channels of  $S$  to  $R$  have been combined by relay, we extend the above model to the multipath case. Let

$$c = \sum_{l=1}^{L_{CAP}} h^{(l)} = \sum_{l=1}^{L_{CAP}} \alpha_l \quad (158.4)$$

$$\begin{aligned}
 r_{rec\_multi} &= \sqrt{E'_p}cs + n'' \\
 &= \sum_{m=0}^{\frac{N_s}{2}-1} b'_r(t - 2mT_f) + d_0b'_d(t - 2mT_f - T_r) + n \quad (158.5)
 \end{aligned}$$

where  $n'' = (n_{s,d} + ah_{r,d}n_{s,r} + n_{r,d})$ ;  $b'_r(t) = b'_d(t) \triangleq \sum_{l=1}^{L_{CAP}} \sqrt{E'_p}\alpha_l p(t - 2mT_f)$ .  $L_{CAP}$  is the number of maximum multipath that can be captured by the receiver.

### 158.3 Performance Analysis of DHTR-UWB System

In this paper, we focus on delay-hopped transmitted-reference ultra-wideband (DHTR-UWB). The structure of DHTR-UWB receiver is shown in Fig. 158.3.

The decision signal  $Z_{TR}$  can be given as

$$\begin{aligned}
 Z_{TR} &\triangleq \sum_{m=0}^{\frac{N_s}{2}-1} \int_0^T \left[ \tilde{b}'_r(t + 2mT_f) + n''(t + 2mT_f) \right] \\
 &\quad \cdot \left[ d_0 \tilde{b}'_d(t + 2mT_f + T_r) + n''(t + 2mT_f + T_r) \right] dt \\
 &= \sum_{m=0}^{\frac{N_s}{2}-1} \int_0^T \left( w'_m(t) + \eta'_m(t) \right) \left( d_0 w'_m(t) + \xi'_m(t) \right) dt \\
 &= \sum_{m=0}^{\frac{N_s}{2}-1} U_m \tag{158.6}
 \end{aligned}$$

where  $U_m \triangleq \int_0^T (w'_m(t) + \eta'_m(t))(d_0 w'_m(t) + \xi'_m(t))dt$ ,  $\tilde{b}'_r(t + 2mT_f)$ , and  $\tilde{b}'_d(t + 2mT_f + T_r)$  are the impulse response of the BPZF, denoted as reference impulse and data impulse, respectively.  $\eta'_m(t) = n''(t + mT_f)$ ;  $\xi'_m(t) = n''(t + 2mT_f + T_r)$ .  $m$  is the frame number in a symbol.

We assumed that all impulses experience the same channel, implying that

$$w'_m(t) \triangleq \tilde{b}'_r(t + 2mT_f) = \tilde{b}'_d(t + 2mT_f + T_r). \tag{158.7}$$

Though it is easy to solve SER with traditional Gaussian approximation approach, some restrictions must be taken into account [6]. In addition, the analytical solution cannot be obtained by using Gaussian approximation approach, thus making it difficult to carry out further research, such as power allocation and relay selection. An analytical method is introduced that combines sampling expansion approach with time transmitted reference and solved for the analytical solution of the system [6–8]. In this paper, we extend it to DHTR scheme to analyze the performance of cooperative DHTR-UWB system.

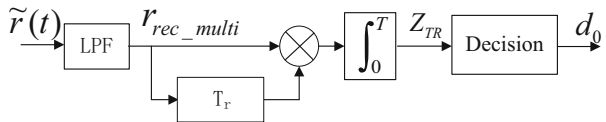


Fig. 158.3 DHTR-UWB receiving model

According to the sampling theorem, we consider to use 2WT-dimensional discrete signal to project the received continuous waveform  $U_m$  without losing:

$$\begin{aligned} U'_m &= \frac{1}{2W} \sum_{l=1}^{2WT} (w'_{m,l} + \eta'_{m,l}) (d_0 w'_{m,l} + \xi'_{m,l}) \\ &= \frac{1}{2W} \sum_{l=1}^{2WT} (d_0 w'^2_{m,l} + w'_{m,l} \xi'_{m,l} + d_0 w'_{m,l} \eta'_{m,l} + \eta'_{m,l} \xi'_{m,l}) \end{aligned} \quad (158.8)$$

where the  $l$ th sample of  $w'_m(t)$ ,  $d_m(t)$ ,  $\eta'_m(t)$  are  $w'_{m,l}(t)$ ,  $d_{m,l}(t)$ ,  $\eta'_{m,l}(t)$ , respectively. In the condition on  $d_0$ , we can rewrite  $U'_m$  as follows:

$$U'_{m|d_0=+1} = \frac{1}{2W} \sum_{l=1}^{2WT} (w'^2_{m,l} + w'_{m,l} \xi'_{m,l} + w'_{m,l} \eta'_{m,l} + \eta'_{m,l} \xi'_{m,l}) \quad (158.9)$$

$$U'_{m|d_0=-1} = \frac{1}{2W} \sum_{l=1}^{2WT} (-w'^2_{m,l} + w'_{m,l} \xi'_{m,l} - w'_{m,l} \eta'_{m,l} + \eta'_{m,l} \xi'_{m,l}) \quad (158.10)$$

We obtain the following simplification:

$$U'_{m|d_0=+1} = \sum_{l=1}^{2WT} \left[ \left( \frac{1}{\sqrt{2W}} w'_{m,l} + \chi_{1,m,l} \right)^2 - \chi_{2,m,l}^2 \right] \quad (158.11)$$

$$U'_{m|d_0=-1} = \sum_{l=1}^{2WT} \left[ - \left( \frac{1}{\sqrt{2W}} w'_{m,l} - \chi_{2,m,l} \right)^2 + \chi_{1,m,l}^2 \right] \quad (158.12)$$

where  $\chi_{1,m,l}$  and  $\chi_{2,m,l}$  are given by Eqs. 158.13 and 158.14

$$\begin{aligned} \chi_{1,m,l} &= \frac{1}{2\sqrt{2W}} (\xi'_{m,l} + \eta'_{m,l}) \\ &= \frac{1}{2\sqrt{2W}} \left[ \left( n_{1,s,d}^{(l)} + ah_{1,r,d}^{(l)} n_{1,s,r} + n_{1,r,d}^{(l)} \right) + \left( n_{2,s,d}^{(l)} + ah_{2,r,d}^{(l)} n_{2,s,r} + n_{2,r,d}^{(l)} \right) \right] \end{aligned} \quad (158.13)$$

$$\begin{aligned} \chi_{2,m,l} &= \frac{1}{2\sqrt{2W}} (\xi'_{m,l} - \eta'_{m,l}) \\ &= \frac{1}{2\sqrt{2W}} \left[ \left( n_{1,s,d}^{(l)} + ah_{1,r,d}^{(l)} n_{1,s,r} + n_{1,r,d}^{(l)} \right) - \left( n_{2,s,d}^{(l)} + ah_{2,r,d}^{(l)} n_{2,s,r} + n_{2,r,d}^{(l)} \right) \right] \end{aligned} \quad (158.14)$$

$\xi'_{m,l}$  and  $\eta'_{m,l}$  are white Gaussian noises, with the following mean and variance:

$$E\{\xi'_{m,l}\} = E\{\eta'_{m,l}\} = 0. \quad (158.15)$$

The variance of DHTR-UWB can be defined as

$$\sigma'^2_{TR} = E\left\{\left[\xi'_{m,l} \pm \eta'_{m,l}\right]^2\right\} = \frac{N_0}{4} (2 + a^2 \delta_{r,d}^2) \quad (158.16)$$

We can define the four normalized random variables as shown below:

$$Y'_1 \triangleq \frac{1}{2\sigma'^2_{TR}} \sum_{m=0}^{\frac{N_s}{2}-1} \sum_{l=1}^{2WT} \left( \frac{1}{\sqrt{2W}} w'_{m,l} + \chi_{1,m,l} \right)^2 \quad (158.17)$$

$$Y'_2 \triangleq \frac{1}{2\sigma'^2_{TR}} \sum_{m=0}^{\frac{N_s}{2}-1} \sum_{l=1}^{2WT} \chi_{2,m,l}^2 \quad (158.18)$$

$$Y'_3 \triangleq \frac{1}{2\sigma'^2_{TR}} \sum_{m=0}^{\frac{N_s}{2}-1} \sum_{l=1}^{2WT} \left( \frac{1}{\sqrt{2W}} w'_{m,l} - \chi_{2,m,l} \right)^2 \quad (158.19)$$

$$Y'_4 \triangleq \frac{1}{2\sigma'^2_{TR}} \sum_{m=0}^{\frac{N_s}{2}-1} \sum_{l=1}^{2WT} \chi_{1,m,l}^2 \quad (158.20)$$

where  $Y'_2$  and  $Y'_4$  are central chi-squared random variables with  $N_s WT$  degrees.  $Y'_1$  and  $Y'_3$  are noncentral chi-squared random variables with same degrees. The noncentrality parameter is given by

$$\mu'_{TR} \triangleq \frac{1}{2\sigma'^2_{TR}} \sum_{m=0}^{\frac{N_s}{2}-1} \sum_{l=1}^{2WT} \frac{1}{2W} w'^2_{m,l}. \quad (158.21)$$

So, we obtain as

$$\mu'_{TR} = \frac{1}{2\sigma'^2_{TR}} \sum_{m=0}^{\frac{N_s}{2}-1} \int_0^T w_m'^2(t) dt = \frac{(1 + a\bar{h}_{s,r})^2}{(2 + a^2 \delta_{r,d}^2)} \frac{E_s}{N_0} \sum_{l=1}^{LCAP} \alpha_l^2 \quad (158.22)$$

where  $E'_s = E'_p N_s = (1 + a\bar{h}_{s,r})^2 E_p N_s$ ,  $\alpha_l^2 = h^{(l)2}$ .

We assume that  $q_{TR} = N_s WT/2$ ,  $\gamma'_{TR} = \mu'_{TR}/2$ , and the conditional SER of cooperative DHTR-UWB system are as follows [6]:

$$\begin{aligned}
 P\{e|\gamma'_{TR}\} &= P\{Y_2 > Y_1 | d_0 = +1\} \\
 &= \frac{e^{-\gamma'_{TR}}}{2^{q_{TR}}} \sum_{i=0}^{q_{TR}-1} \frac{(\gamma'_{TR})^i}{i!} \sum_{k=i}^{q_{TR}-1} \frac{1}{2^k} \frac{(k + q_{TR} - 1)!}{(k - i)!(q_{TR} + i - 1)!}
 \end{aligned} \tag{158.23}$$

Considering the uncertainty of channel coefficient  $h^2$ , we define the expectation of  $P\{e|\gamma'_{TR}\}$  as follows [7, 9]:

$$\begin{aligned}
 P_{e,TR} &= E\{P\{e|\gamma'_{TR}\}\} \\
 &= \frac{1}{2^{q_{TR}}} \sum_{i=0}^{q_{TR}-1} \frac{E\{(\gamma'_{TR})^i e^{-\gamma'_{TR}}\}}{i!} \sum_{k=i}^{q_{TR}-1} \frac{1}{2^k} \frac{(k + q_{TR} - 1)!}{(k - i)!(q_{TR} + i - 1)!} \\
 &= \frac{1}{2^{q_{TR}}} \sum_{i=0}^{q_{TR}-1} \frac{(-j)^i}{i!} \frac{d^i}{dv^i} \psi'_{\gamma'_{TR}}(jv) \Big|_{jv=-1} \cdot \sum_{k=i}^{q_{TR}-1} \frac{1}{2^k} \frac{(k + q_{TR} - 1)!}{(k - i)!(q_{TR} + i - 1)!} \\
 &\triangleq P_e(\psi'_{\gamma'_{TR}}(jv), q_{TR})
 \end{aligned} \tag{158.24}$$

where  $\psi'_{\gamma'_{TR}}(jv)$  is the CF of  $\gamma'_{TR}$  [9]:

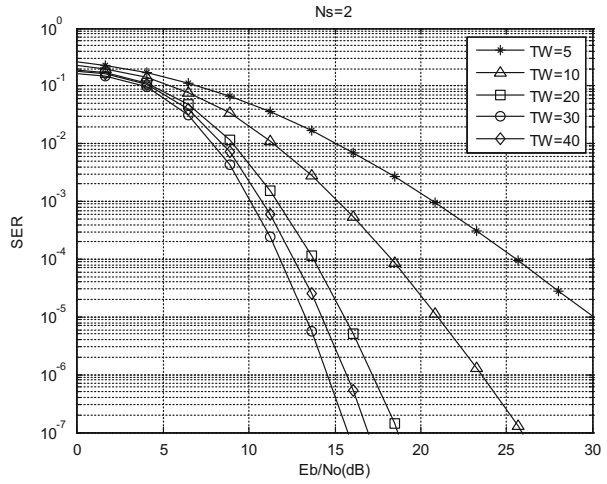
$$\psi'_{\gamma'_{TR}}(jv) = \prod_{l=1}^{L_{CAP}} \psi'_{l,\gamma'_{TR}}(jv) = \prod_{l=1}^{L_{CAP}} \frac{1}{K(1 - jv\bar{\gamma}'_{TR})} \tag{158.25}$$

where  $\bar{\gamma}'_{TR} = \left[ \frac{(1 + a\bar{h}_{s,r})^2}{2(2 + a^2\delta_{r,d}^2)} \right] \frac{\Omega E_s}{N_0}$ ,  $\Omega = \frac{1}{L}$ .

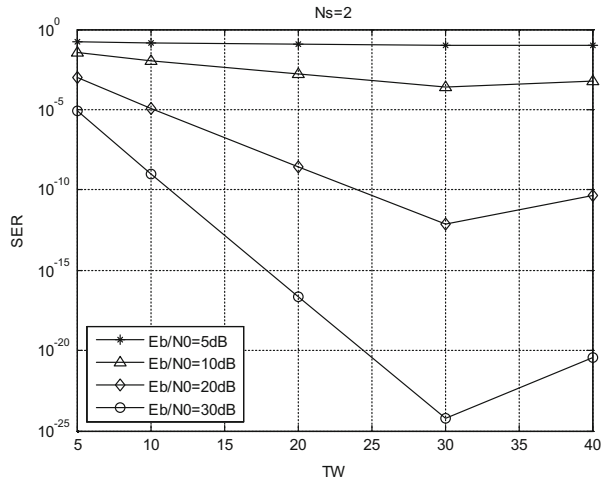
### 158.4 Results and Discussion

In this section, we evaluate the performance of DHTR-UWB system. We consider a dense multipath channel, where each path's signal experiences flat Rayleigh fading.  $h_{s,d}$ ,  $h_{s,r}$ ,  $h_{r,d}$  are the channel fading coefficient which are defined as complex Gaussian random variables with zero mean and normalized square deviation  $\delta_{s,d}^2 = \delta_{s,r}^2 = \delta_{r,d}^2 = 1$ , respectively. The actual number of multipath components is  $L = 30$ , and the number of multipath components captured by the autocorrelation receiver (AcR) is denoted by  $L_{CAP=TW}$ . We also discuss the effect of time-bandwidth product, number of impulse, and instantaneously received SNR on the cooperative DHTR-UWB system as shown in Figs. 158.4, 158.5 and 158.6.

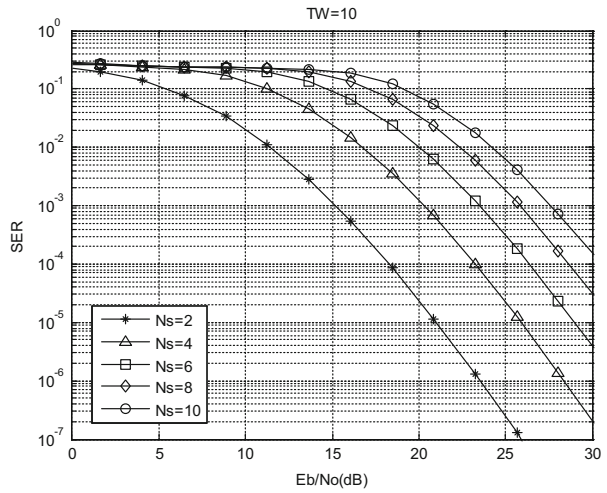
**Fig. 158.4** DHTR-UWB's SER performance with  $N_s = 2$



**Fig. 158.5** Effect of TW on the performance of DHTR-UWB



**Fig. 158.6** SER performance of DHTR-UWB with  $N_s$





In Fig. 158.4, we note that the tendency of *SER* performance is from bad to good then turns into bad again as  $TW$  increases. This is due to the fact that time-bandwidth product  $TW$  represents for the integral duration.  $L < TW$  implies that multipath components cannot be captured completely, thus severely degrades the *SNR* at the receiver termination and drops the *SER* performance of the cooperative system. On the contrary, increasing  $TW$  beyond  $L$  will only accumulate more noise energy, and it degrades performance as well.

We assumed that the actual number of multipath components is  $L = 30$  in this simulation. For the case  $TW = 30$ , at which point integral interval equals to the multipath delay, the energy of multipath components is fully captured and the noise is maximally restrained. As we can see from Fig. 158.3, the cooperative DHTR-UWB system gains five-order diversity under the condition of optimum integration interval. This embodies the advantage of impulse scheme in dense multipath channel transmission.

Figure 158.5 details the effect of time-bandwidth product  $TW$  on the *SER* performance of DHTR-UWB cooperative system. It can be observed that the *SER* decreases with  $TW$  until it rises to a value of 30, which is equal to  $L$ . We obtain the best *SER* performance at the inflection point  $TW = 30$  for each fixed  $E_b/N_0$ . Due to the fixed bandwidth  $W$  at the transmitting terminal, we can easily deduce that the optimum integral interval is  $30/W$ .

In order to analyze the number of impulse  $N_s$  impact on *SER* performance, we fixed time-bandwidth product and specify the optimum integral interval  $T = 10/W$ . We observe that *SER* decreases with  $N_s$  that correspond to Fig. 158.5. Increasing  $N_s$  implies less energy per impulse and leads to more noise energy accumulation. This can be seen as the reason why performance degradation as  $N_s$  increases in Fig. 158.6.

Obviously, the transmitted cooperative-reference UWB cooperative approach is not suitable for simply amplify and forward (AF) because in AF the relay node do not decode. They cannot separate and regenerate the reference impulses or data impulses to forward further alternatively. Only decode and forward supports transmitted cooperative-reference cooperative approach.

Further, the transmitted cooperative-reference UWB cooperative approach is only applicable to the transmitted-reference system and applies only to time-domain transmitted-reference system. For the frequency-domain and code-domain transmitted-reference system, the node finally transmits the composite impulse of the reference impulse and the data impulse [4, 5]. It is not possible that the relay node decodes and forwards only one of the reference impulse and the data impulse.

Furthermore, in the TR-UWB cooperative communication network, it can be the degradation or evolution between of three cooperative communication modes: two-relay transmitted cooperative-reference mode, single-relay transmitted cooperative-reference mode, and direct transmission mode. System will convert to two-relay transmitted cooperative-reference mode when we got two ideal relays, and the power of each relay will reduce by half. Similarly, single-relay mode will be chosen if there is only one proper relay in the system. Otherwise, the system will degenerate into direct transmission mode.

## 158.5 Conclusion

The paper proposes a novel transmitted cooperative-reference UWB cooperative communication method that uses two relay nodes to forward reference impulses and data impulses separately. For each relay node is concerned, it can save half of the energy and help to encourage nodes to participate in cooperation. It also provides a new relay strategy for cooperative communication. We established a cooperation model based on this method and analyzes the performance of it. Simulation result shows that the DHTR-UWB cooperative system has favorable SER performance. System can obtain  $SER = 10^{-3}$  performance under optimal integration time with receiving SNR is 10 dB. Cooperative system can gain approximately five-order diversity in dense multipath channels. Future research will focus on the effect of distance and [power distribution](#) on DHTR-UWB system.

## References

1. Shen, Q., Wu, X., Lin, D., & Qiu, X. (2010). Performance analysis of cooperative ultra-wideband communication system. *2010 International Conference on Communications and Mobile Computing (CMC), IEEE* (pp. 217–220).
2. Chao, Y. L., & Scholtz, R. (2003). Optimal and suboptimal receivers for ultra-wideband transmitted reference systems. *IEEE Global Telecommunications Conference, 2003. GLOBECOM'03* (pp. 759–763).
3. Hoor, R., & Tomlinson, H. (2002). Delay-hopped transmitted-reference RF communications. *Digest of Papers, 2002 I.E. Conference on Ultra Wideband Systems and Technologies* (pp. 265–269).
4. Goecke, D. L., & Zhang, Q. (2007). Slightly frequency-shifted reference ultra-wideband (UWB) radio. *IEEE Transactions on Communications, 55*(3), 508–519.
5. Jian, Z., Han-Ying, H., Luo-Kun, L., & Tie-Feng, L. (2007). Code orthogonalized transmitted reference ultra-wideband wireless communication system. *International Conference on Wireless Communications, Networking and Mobile Computing, 2007. WiCom 2007* (pp. 528–532).
6. Quek, T. Q. S., & Win, M. Z. (2005). Analysis of UWB transmitted-reference communication systems in dense multipath channels. *IEEE Journal on Selected Areas in Communications, 23* (9), 1863–1874.
7. Quek, T. Q. S., Win, M. Z., & Dardari, D. (2007). Unified analysis of UWB transmitted-reference schemes in the presence of narrowband interference. *IEEE Transactions on Wireless Communications, 6*(6), 2126–2139.
8. Ngo, H. Q., Quek, T. Q. S., & Shin, H. (2010). Amplify-and-forward two-way relay networks: error exponents and resource allocation. *IEEE Transactions on Communications, 58*(9), 2653–2666.
9. Simon, M. K., & Alouini, M. S. (1998). A unified approach to the performance analysis of digital communication over generalized fading channels. *Proceedings of the IEEE, 86*(9), 1860–1877.

# Author Index for Volume 1

## A

Ao, X., 421  
Arshad, M.R.H.M., 355

## B

Bai, C., 385  
Bai, J., 337, 577  
Bao, W., 577

## C

Cai, W., 107  
Cao, G., 671  
Cao, L., 635  
Cao, S., 671  
Cao, Y., 549  
Chai, J., 643  
Chen, C., 559, 609  
Chen, G., 219  
Chen, J., 365  
Chen, K., 671  
Chen, L., 481  
Chen, M., 163  
Chen, Q., 55  
Chen, X., 243, 515, 635  
Chen, Y., 47  
Chen, Z., 171, 307  
Cheng, W., 481  
Cheng, Y., 541

## D

Dai, M., 577  
Dong, J., 99  
Duerling, B., 329

## F

Fang, P., 153  
Fang, X., 9  
Fard, S.P., 125, 355  
Fei, S., 209  
Feng, W., 577  
Feng, X., 265

## G

Gai, S., 625  
Gao, J., 171  
Gao, S., 609  
Gao, Z., 31  
Geng, X., 395  
Gu, C., 91, 145  
Gu, H., 429  
Gu, Y., 375  
Guan, H., 345  
Guan, Y., 653  
Guo, C., 21  
Guo, J., 291

## H

Han, W., 255  
Hao, X., 577  
He, H., 489  
Hong, L., 541  
Hong, Z., 499  
Hou, C., 321  
Hou, L., 489  
Hu, H., 437  
Hu, S., 345  
Huang, X., 329, 661  
Huang, Y., 507  
Huang, Z., 91, 375

**J**

Jiang, C., 345  
 Jiang, M., 593  
 Jiao, Z., 365  
 Jin, Z., 209

**K**

Kang, Y., 73, 83

**L**

Li, D., 411, 635  
 Li, G., 65, 193  
 Li, H., 177, 465, 593, 679  
 Li, J., 47  
 Li, K., 177, 679  
 Li, L., 201, 321, 585  
 Li, M., 643  
 Li, Q., 3  
 Li, T., 39  
 Li, X., 55  
 Li, Y., 115  
 Li, Z., 193, 201  
 Liang, H., 643  
 Liang, T., 533  
 Lin, A., 525  
 Lin, M., 617  
 Lin, Y., 73, 83  
 Liu, J., 107  
 Liu, Q., 617  
 Liu, W., 65  
 Liu, X., 115, 255  
 Liu, Y., 185  
 Lu, H., 73, 83  
 Luo, P., 153  
 Luo, S., 177, 679  
 Luo, X., 533  
 Luo, Y., 617

**M**

Ma, J., 291  
 Ma, L., 403  
 Ma, R., 307  
 Ma, Y., 337, 585  
 Meng, Q., 177, 679  
 Ming, J., 695  
 Mohamed, H.H., 355  
 Mousavi, S.A., 355  
 Mu, F., 201  
 Mu, S., 337

**P**

Peng, X., 489

**Q**

Qiu, L., 445  
 Qiu, Q., 153

**S**

Sedaghat, L., 329  
 Shen, H., 559  
 Shen, J., 55  
 Shen, Z., 403  
 Shi, M., 255  
 Shu, S., 281  
 Sumari, P., 355  
 Sun, J., 365  
 Sun, S., 3  
 Sun, Z., 299

**T**

Tan, X., 163  
 Tang, J., 243  
 Tang, Z., 329  
 Tao, Q., 91, 145

**W**

Wang, B., 219  
 Wang, F., 47, 201, 411  
 Wang, L., 585  
 Wang, P., 429  
 Wang, Q., 177, 403  
 Wang, R., 695  
 Wang, S., 653  
 Wang, W., 533  
 Wang, X., 273, 661  
 Wang, Y., 365, 395, 455  
 Wang, Z., 411  
 Wen, D., 437  
 Wen, Q., 209  
 Wu, C., 671  
 Wu, Y., 321

**X**

Xia, C., 99  
 Xia, X., 625  
 Xiao, B., 525  
 Xiao, Y., 185, 601

Xing, Y., 567  
Xu, D., 465  
Xu, H.-Y., 9  
Xu, J., 201  
Xu, R., 653  
Xu, Z., 201, 307

**Y**

Yan, X., 365  
Yang, B., 411, 679  
Yang, F., 567  
Yang, G., 473  
Yang, H., 227  
Yang, J., 265, 687  
Yang, T., 515  
Yang, Y., 201, 445  
Yang, Z., 385  
Yao, C., 541  
Yin, L., 39  
Yin, S., 403  
Ying, S., 345  
Yu, B., 567  
Yu, F., 65  
Yu, P., 465  
Yu, Q., 617  
Yu, X., 429  
Yu, Z., 385  
Yuan, F., 39  
Yuan, H., 235, 421  
Yuan, J., 671

**Z**

Zainuddin, Z., 125  
Zeng, H., 55  
Zeng, W., 695  
Zhang, B., 609  
Zhang, C., 91, 473, 577  
Zhang, H., 219  
Zhang, J., 235, 291  
Zhang, M., 107  
Zhang, S., 307  
Zhang, T., 671  
Zhang, W., 99  
Zhang, X., 193, 473  
Zhang, Y., 73, 83, 635  
Zhang, Z., 39, 135  
Zhao, D., 375  
Zhao, J., 3  
Zhao, L., 507  
Zhao, T., 421  
Zhao, Y., 429  
Zhao, Z., 321  
Zheng, S., 489  
Zhou, D., 21  
Zhou, H., 465  
Zhou, L., 107  
Zhou, W., 153  
Zhu, S., 687  
Zhu, X.-f., 687  
Zhu, Y., 525  
Zou, G., 337

# Subject Index for Volume 1

## A

Active decoy, 437–440, 442, 443  
Adaptive, 3–8, 65, 107–114, 171–176, 292,  
379, 386, 457, 478, 507–510, 512,  
525–532, 580, 631, 701  
filter, 171–176  
sampling, 107–114  
AF. *See* Array factor (AF)  
Affinity propagation, 177–183  
Android, 154, 155, 161, 371  
ANNs. *See* Artificial neural networks (ANNs)  
Ant colony algorithm, 227–233  
Ant colony system, 73–80, 83–90  
Anti-radiation missile (ARM), 437–443  
Apache, 210, 265  
Apriori algorithm, 281–285  
Architectural space, 549–558  
ARM. *See* Anti-radiation missile (ARM)  
Array factor (AF), 40, 41, 43  
Artificial neural networks (ANNs), 22, 48,  
50–52, 100, 102, 125, 126, 220,  
467, 672  
Association rule, 281–288, 458

## B

Bayesian models, 385–392  
Bayesian networks (BNs), 386, 391, 392,  
525–532, 649, 650  
BCI. *See* Brain–computer interface (BCI)  
Beam position, 5, 6, 8  
Binary image, 478, 479, 483, 688, 691, 692  
Binary phase coded, 445–451  
Biometrics, 465, 466, 469  
BNs. *See* Bayesian networks (BNs)

BP neural network (BPNN), 21–30, 243–250,  
671–677  
Brain–computer interface (BCI), 355–362  
Business process execution language (BPEL)  
process, 345–352

## C

Camera calibration, 484–486  
Centroid, 122, 499, 515–522, 561, 562, 657  
Certainty-based active learning, 235–242  
CH. *See* Cluster head (CH)  
Chaos perturbation, 230, 233  
Classification  
accuracy, 51, 321, 322, 326, 327, 500, 503,  
504, 578, 583, 584, 586, 591, 661,  
666–668, 676, 698, 701  
and recognition, 672–674, 676  
Classify, 50, 322, 326, 357, 468, 500, 502,  
594, 596, 598, 600, 602, 644–648,  
650, 661, 662  
Cloud computing, 202, 207, 209, 210, 425, 429  
Cluster head (CH), 65–71, 110, 111, 163–170  
Clustering, 48, 65, 66, 69, 85, 108, 110,  
163–170, 177–183, 221–223, 301, 322,  
339, 455–462, 501, 635–641, 680, 697  
Collection sites, 395–402  
Color management, 617–619  
Component docking, 473–480  
Compression matrix, 281–288  
Computer technology, 48, 300, 549–558  
Computer vision (CV), 223, 224, 301, 507,  
533, 609, 654  
Conditional independence tests, 526  
Connected-element interferometry, 411–418

Contour features, 597, 599  
 Convergence, 22–26, 28–30, 43, 44, 65, 80, 97,  
 104, 150, 172, 174, 175, 185–187, 193,  
 194, 229, 230, 233, 526, 531, 544, 546,  
 585–592, 664  
 Copy-move forgery, 680, 684  
 Corner extraction, 482, 485, 486  
 Curve on surface, 10, 11, 14, 16–19  
 CV. *See* Computer vision (CV)  
 Cycle analysis, 56, 69

## D

Data center network, 429–435  
 Data intensive, 375–383  
 Data mining, 209, 210, 281, 300, 458, 499, 644,  
 645, 649  
 DBN. *See* Deep belief network (DBN)  
 Decision degree, 277, 278  
 Decision-theoretic rough set model (DTRS),  
 273–279  
 Decoupling interrupts, 291–296  
 Deep belief network (DBN), 661–668  
 Dense matching, 625–633  
 Density-sensitive, 177–183  
 Descriptive, 339, 508  
 Device driver, 329–336  
 Dictionary learning, 541, 578, 580–581  
 Difference received signal strength (DRSS),  
 115–123  
 Differential geometric characteristic, 13  
 Dijkstra algorithm, 153, 155–157, 227, 233  
 Distributed computing, 228  
 Distributed system, 31, 74, 75, 80, 83–90, 404  
 Document object model (DOM), 368–370  
 Domain ontology learning, 489–496  
 Double approximate identity functions, 126, 127  
 Double flexible approximate identity functions,  
 126, 132  
 Double flexible approximate identity neural  
 networks, 125–132  
 DRSS. *See* Difference received signal strength  
 (DRSS)  
 DTRS. *See* Decision-theoretic rough set model  
 (DTRS)  
 Dynamic route guidance system (DRGS),  
 227–233  
 Dynamic voltage scaling (DVS), 73–80

## E

ECFG. *See* Exception control flow graph  
 (ECFG)  
 Electroencephalographic (EEG) signal  
 classification, 355–362

Energy consumption, 65, 66, 69, 71, 74–77,  
 79, 80, 108, 110, 112, 113, 164–166,  
 168–170, 558  
 Energy effective, 107–114  
 ENN. *See* Extension neural network (ENN)  
 Exception control flow graph (ECFG),  
 345–352  
 Exception handling, 345, 346, 349, 351  
 Extended Kalman filter, 610, 611,  
 613–614  
 Extension neural network (ENN), 219–225

## F

Factorization, 255, 258, 263, 321–327  
 Fall detection, 653–659  
 FAS algorithm. *See* Frequency based sampling  
 (FAS) algorithm  
 FCA. *See* Formal concept analysis (FCA)  
 Feature extraction, 48, 49, 243, 245, 322, 355,  
 356, 358, 362, 466, 469, 471, 476, 516,  
 594, 597, 600, 662  
 Feature selection, 299–306, 361, 635–641,  
 645–646  
 Flexible approximate identity functions,  
 126, 132  
 Forecasting model, 101–102, 104, 248  
 Forgery detection, 680, 683  
 Formal concept analysis (FCA), 490–493,  
 495, 496  
 Formal method, 307  
 Forward-error correction, 291  
 Fragments, 507–512, 534  
 Frequency based sampling (FAS) algorithm,  
 107–114  
 Frequency identification, 243–250  
 2FSK. *See* Pseudorandom binary frequency  
 shift keying (2FSK)  
 Fuzzy system, 99–106

## G

GA. *See* Genetic algorithm (GA)  
 Gabor wavelet, 662, 663, 668  
 Gait feature, 515–522  
 Gait recognition, 516  
 Gauss-Jacobi iterative, 542, 543, 546  
 Genetic algorithm (GA), 21–30, 39–46, 79,  
 88, 89, 92, 146, 457  
 Geographic information system (GIS), 338,  
 339, 342–344  
 Gesture manipulation, 135–142  
 GLCM. *See* Gray-Level Co-occurrence  
 Matrix (GLCM)  
 Global orientation, 653–659

Graph cut, 538  
 Graphical user interface (GUI) testing, 385–392  
 Gray-Level Co-occurrence Matrix (GLCM), 662, 663, 665, 666, 668  
 Green supply chain management (GSCM), 395, 396, 402  
 Grey model, 220, 224, 225  
 GUI testing. *See* Graphical user interface (GUI) testing

**H**

Hadoop, 209–216, 265–269, 300, 376, 425, 591  
 Hand tracking, 135–142  
 Harmonic function (HF), 321–327  
 HBase, 265–271, 376  
 HCI. *See* Human–computer interaction (HCI)  
 Heterogeneous, 73–80, 83–90, 164, 186, 191, 375, 379  
 Heterogeneous system, 73–80, 84, 85  
 Heuristic, 75, 77–80, 84–89, 111, 229, 230, 258, 291, 293, 294, 395–402, 526, 676, 700  
 HF. *See* Harmonic function (HF)  
 Hough transformation, 571–573, 675  
 HowNet, 635–641  
 Human–computer interaction (HCI), 135–142, 559, 675  
 Human shape, 653–659  
 Hybrid particle swarm optimization, 185–191  
 Hyperbola model, 609–616

**I**

ICC profile. *See* International Color Consortium (ICC) profile  
 Identifying, 343, 490, 538, 567–575, 581  
 Image  
   forensics, 679, 680  
   monitoring, 475  
   processing, 322, 475, 476, 481–486, 568, 594, 668  
   restoration, 541–546  
   skeleton, 687–691  
   softproofing, 617–624  
   tampering, 601  
 Implicit information, 255, 260–262  
 Improved tolerance relation, 274, 279  
 IMRT. *See* Intensity modulated radiation therapy (IMRT)  
 Incomplete decision table, 274–279  
 Indoor three-dimensional positioning, 115–123

Information extraction, 202, 366, 489, 579  
 Information service, 455–462  
 Infrared (IR), 695–697  
 Integral sliding mode controller, 31–37  
 Intelligent visual surveillance, 653  
 Intense illumination, 567–575  
 Intensity modulated radiation therapy (IMRT), 193–198  
 Interlocking, 307–319  
 International Color Consortium (ICC) profile, 617, 622  
 IR. *See* Infrared (IR)

**J**

Jsoup, 368, 369

**K**

Kernel-nearest neighbor (kNN), 48, 52, 326, 499–505, 680, 683  
 K-optimal path, 231  
 K–T transform, 578, 579, 582, 583

**L**

Lane detection, 609–616  
 Lane mark, 567–575, 615  
 Lattice reduction, 262  
 LEACH-SC, 66–71  
 Least mean square error (LMS) algorithm, 171–176  
 Least squares method, 615  
 Lifetime, 65, 70, 71, 169, 170  
 Linear antenna array, 39–46  
 Linear matrix inequality (LMI), 34–37  
 Link16, 55–63  
 LLL algorithm, 256, 258  
 LMS algorithm. *See* Least mean square error (LMS) algorithm  
 Loading data, 265–271

**M**

Machine learning, 240, 241, 299–301, 305, 321, 500, 580, 585, 608  
 MapReduce, 210, 214, 215, 265–269, 299–306, 376, 429, 585–592  
 MapReduce-support vector machine (MR-SVM), 585–592  
 Markov models, 291–296  
 Mass data, 403–409  
 Maude, 307–319



Mean comparison, 412, 416, 417  
 Mean shift, 136–142  
 Mel Frequency Cepstrum Coefficient (MFCC),  
 244–247, 249, 250  
 Mellin approximate identity functions, 126  
 Mismatch eliminating, 629  
 Missile-borne phased array radar, 3–8  
 MongoDB, 376, 377, 379, 403–409  
 MR-SVM. *See* MapReduce-support vector  
 machine (MR-SVM)  
 Multihop communication, 164  
 Multi-objective optimization, 193  
 Multi-sensor fusion, 559–565  
 Multispectral remote sensing images, 578,  
 581, 584  
 Mutual information, 300–301, 303–305, 526,  
 527, 529

## N

NAR. *See* Nonparametric auto-regression  
 analysis (NAR)  
 Navigation software, 153–161  
 NDVI. *See* Normalized Difference Vegetation  
 Index (NDVI)  
 Network control system (NCS), 31–37  
 NMF. *See* Non-negative matrix factorization  
 (NMF)  
 Non-dominated neighbor-based immune  
 algorithm (NNIA), 193–198  
 Non-local, 541–546  
 Non-negative matrix factorization (NMF),  
 321–327  
 Nonparametric auto-regression analysis  
 (NAR), 219–225  
 Non-relational databases, 404  
 Normalized Difference Vegetation Index  
 (NDVI), 578–579, 582, 583  
 Not Only SQL (NoSQL), 266, 376,  
 377, 404

## O

On-line, 108, 109, 136, 201, 204, 376, 456,  
 458–460, 462, 507–512, 559–565, 637  
 Online gesture recognition, 559–565  
 Optimization, 6, 22, 27–28, 30, 40, 41, 43, 48,  
 50, 75, 76, 84–88, 91–98, 103, 108, 138,  
 145–151, 183, 185–191, 193–198, 228,  
 230, 233, 248, 281, 324, 343, 422, 457,  
 538, 580, 586, 589, 590, 592, 595, 604,  
 612, 647, 697

## P

PageRank algorithm, 204–206  
 Pairwise constraints, 178, 180–182  
 Palmprint, 465–471  
 Paper submission, 201–208  
 Parallel  
   processing framework, 375–383  
   projection, 9–19  
 Parallel computing, 186–189, 191, 210,  
 376, 377  
 Parallelize, 188  
 Parameter optimization, 50, 592  
 Partial code replicated, 451  
 Particle swarm optimization (PSO), 39, 50,  
 91–98, 185–191  
 PCA. *See* Principal components analysis (PCA)  
 PCA-SIFT, 679–685  
 PC radar. *See* Pulse compression (PC) radar  
 P300 detection, 355–362  
 Performance analysis, 57–62, 294, 421, 422  
 Performance evaluation, 112–113, 376, 422,  
 470, 640–641  
 Piano, 243–250  
 Polar space, 533–540  
 Predict interaction between proteins, 48  
 Prediction model, 100, 219–225, 426, 697  
 Principal components analysis (PCA), 466,  
 467, 471, 682–684, 695–701  
 Processing time, 21–30, 75, 76, 84, 88,  
 379–381  
 Process quality monitoring, 473–480  
 Pseudorandom binary frequency shift keying  
 (2FSK), 171–176  
 PSO. *See* Particle swarm optimization (PSO)  
 Pulse compression (PC) radar, 445–451

## Q

Quantum-inspired genetic algorithm (QGA),  
 145–151

## R

Range deception, 449  
 Range-gate, 450, 451  
 RBF neural network, 99–106, 126  
 Recommendation, 201–208, 455–462  
 Region growing, 625–633  
 Regularization, 103, 541–546, 647  
 Remote sensing image, 577–584, 671–677  
 Rendering intent, 617–624  
 Repositioning, 505

Restricted searching area, 153–159, 161  
 Reverse engineering, 329–336  
 Reverse logistics, 395–402  
 Rewriting logic, 307–309, 318  
 Road impedance factor, 229–233  
 Routing protocol, 66, 69, 163

## S

SA. *See* Simulated annealing (SA)  
 Sag measurement, 481–486  
 Sample selection, 236–241, 674–676  
 Scheduling algorithm, 74–76, 84, 85, 89  
 SDA. *See* Subclass discriminant analysis (SDA)  
 Search, 3–7, 23, 40, 41, 50, 76, 77, 80, 86,  
   95, 153, 157, 161, 172, 181, 186, 188,  
   194, 195, 201–208, 210, 228, 291,  
   307, 378, 455, 457, 460, 474, 511,  
   526, 612, 613, 626, 630, 631, 633,  
   636, 680, 683, 700  
 Search engine, 201–208, 636  
 Segmentation, 204, 208, 240, 248, 250,  
   477–479, 490, 491, 520, 533–540, 638  
 Semantic dependency, 489–496  
 Semantics, 206, 309, 348, 376–378, 382, 459,  
   489–496, 635–641  
 Semi-supervised clustering, 177–183  
 Semi-supervised learning, 178, 179, 321–327  
 Server-centric architecture, 433, 435  
 Shortest path, 153–161, 228  
 Short-term load forecasting, 99–106  
 Short texts clustering, 635–641  
 Side lobe level, 39–46  
 Similarity, 177–183, 202, 206, 207, 222, 239,  
   303, 322, 372, 460, 461, 466, 469, 502,  
   508–510, 512, 541–544, 560–563, 580,  
   635–640, 682, 683  
 Simulated annealing (SA), 186, 525–532  
 Simulation, 3–8, 36, 37, 48, 69–70, 74, 104,  
   112–114, 167–176, 194, 196, 232, 233,  
   244, 292, 300, 307, 412, 416–418, 433,  
   434, 437–443, 445–451, 474, 479,  
   550–557, 563–564, 614, 615  
 Simulation analysis, 168–170, 418,  
   550–557  
 Slot allocation, 55–63  
 Sparse principal component analysis (SPCA),  
   695–701  
 Sparse representations, 577–584  
 Spatial analysis, 337–344  
 Spatial data, 343, 375–383  
 Spectral analysis, 243, 518–520

Speeded-up robust features (SURF), 594,  
   597–599  
 Spontaneous speech summarization, 235  
 SPRINT, 209–216  
 Standard processing time table (SPTT),  
   22–30  
 State feedback, 33, 37  
 Statistics, 21, 154, 159, 300, 301, 338, 339,  
   341, 342, 490, 529, 593, 602, 667, 681  
 Structure learning, 525–532  
 Subclass discriminant analysis (SDA),  
   465–471  
 Supply chain network, 91–98, 145–151  
 Support vector machine (SVM), 49, 52, 236,  
   240–242, 300, 304, 305, 467, 499, 500,  
   502–505, 582–600, 602–605, 607, 608,  
   643–650, 695–701  
 SURF. *See* Speeded-up robust features (SURF)  
 Survival time, 168, 170  
 Switching function, 32–34, 37  
 Synthetic aperture radar (SAR) oil spill image,  
   661–668

## T

Taylor approximation, 12, 19  
 TB. *See* Tuberculosis (TB)  
 TDMA, 55, 57–62  
 Template matching method, 559  
 Text classification, 222, 499–505  
 Thinning procedure, 687–692  
 Time delay, 411–418  
 TOLU, 291–296  
 Topological properties, 430, 431, 435  
 Topology, 125, 220, 232, 375, 387–389, 392,  
   430, 586, 689, 691, 692  
 Tracing, 10, 11, 534, 551, 555, 655  
 Tracking, 3, 135–142, 411, 412, 439, 440,  
   448–450, 474, 507–512, 567–575, 610,  
   611, 613, 614, 635, 654  
 Transforming, 145–151, 179, 536–537  
 Transforming operator, 92–94, 96, 98, 148,  
   149, 151  
 Transmission delay, 58–61  
 Tuberculosis (TB), 337–344, 380, 381, 383

## U

Uncertainty, 31–37, 102, 237, 300, 357, 502,  
   626, 629, 676  
 Uneven clustering, 170  
 Universal approximation, 125–132

USB, 295, 296, 329–336  
User model, 456–459, 462

**V**

Variable precision rough set (VPRS),  
500–503, 505  
Vehicle classification, 593–600  
Vertical search, 201–208  
Visual design, 549–558

**W**

WarpingLCSS, 559–565  
Wavelet analysis, 47–52  
Wavelet transform, 48–50, 99–106  
Web sentiment, 219–225  
WiFi, 115–123  
Wind turbine, 403–409  
Wireless sensor networks (WSN),  
65–71, 107–114, 163–170  
Worm detection, 643–650

# Author Index for Volume 2

## A

Alahmadi, A., 731

## C

Cao, B., 799  
Chai, J., 1365  
Chen, B., 723  
Chen, D., 1341  
Chen, G., 755  
Chen, K.-L., 1045  
Chen, Q., 783, 825  
Chen, X., 1323  
Chen, Y., 1209  
Chen, Y.-W., 713  
Chen, Z., 799  
Chi, T., 739  
Chou, W., 955  
Chu, Y.-Y., 713  
Collados, K., 1191  
Cui, Y., 905

## D

Di, Y., 1099  
Dong, M., 955  
Dong, W., 1201  
Dong, Y., 783  
Dou, L., 1233  
Du, J., 755  
Du, Y., 755

## F

Feng, S., 1117, 1125  
Fu, T., 747

## G

Gao, L., 1289  
Gao, W., 971  
Gao, Y., 747  
Gorricho, J.-L., 1191  
Guan, H., 1173  
Guo, W., 971  
Guo, X., 1245

## H

He, L., 1341  
He, Q., 963  
He, Y., 913  
He, Z., 997  
Hou, C., 979, 1373  
Hu, D., 809  
Hu, P., 1109  
Hu, Q., 883  
Hu, W., 1017  
Hu, X., 855  
Huang, F., 1381  
Huang, J.-F., 1035, 1045  
Huang, T., 705  
Huang, Y., 883

## J

Jiang, C., 1173  
Jiang, D., 1007  
Jiang, F., 855  
Jiang, Z., 971  
Jiangn, Y.-J., 1035  
Jiao, J., 1091

**L**

Li, B., 791, 865  
 Li, F., 1073, 1083  
 Li, G., 835  
 Li, J., 1225  
 Li, M., 971  
 Li, O., 1395  
 Li, T., 1395  
 Li, W., 905, 1073, 1083, 1143  
 Li, X., 929, 939, 1341  
 Li, Y., 1217, 1263  
 Li, Z., 963, 1373  
 Liang, H., 1365  
 Lin, F., 1373  
 Lin, H., 1373  
 Liu, D., 835  
 Liu, F., 1181  
 Liu, H., 1165  
 Liu, J., 997  
 Liu, L., 835, 1063  
 Liu, T., 739  
 Liu, Y., 799, 955, 1055,  
 1253, 1281  
 Lu, L., 847, 1233  
 Luan, D., 921  
 Lv, S., 791, 865

**M**

Man, Q., 763  
 Mao, T., 775  
 Matsuhisa, T., 1331  
 Miao, Z., 1263

**N**

Nie, W., 809

**P**

Peng, I.-H., 713  
 Peng, Y., 893

**Q**

Qian, K., 1307  
 Qin, P., 755  
 Qin, Z., 1099

**R**

Ren, Y., 1055  
 Rui, P., 763

**S**

Serrat, J., 1191  
 Shao, Y., 929, 939  
 Shi, B., 1271  
 Shi, F., 929, 939  
 Shuo, C., 1201  
 Soh, B., 731  
 Song, B., 971  
 Su, A.Y.S., 713  
 Su, D., 1073, 1083  
 Su, Q., 921  
 Sun, J., 963  
 Sun, L., 1373  
 Suo, Y., 739

**T**

Tang, Y., 1365  
 Tang, Y.H., 873  
 Tao, H., 979  
 Tian, M., 791, 865  
 Tian, Y., 913

**W**

Wang, H., 963  
 Wang, K., 783  
 Wang, L., 947, 1109  
 Wang, P., 971, 1297  
 Wang, S., 1263  
 Wang, Y., 723, 893,  
 1055, 1209  
 Wang, Z., 723, 847, 1217  
 Wei, S., 835, 1063  
 Wen, J.-Y., 1035  
 Wu, C., 989  
 Wu, G., 1181  
 Wu, H., 1181  
 Wu, J., 997  
 Wu, L., 1165, 1209  
 Wu, R., 783  
 Wu, Y., 809  
 Wu, Z., 1289

**X**

Xi, J., 1281  
 Xia, M., 783  
 Xiao, J., 1143, 1263  
 Xiao, T., 947  
 Xiao, W., 1281, 1381  
 Xie, L., 913  
 Xu, B., 1349, 1357

Xu, C., 1253  
Xu, K., 1253  
Xu, M., 1201, 1245  
Xu, P., 1217  
Xu, W., 1263  
Xu, Y., 775  
Xu, Z., 775  
Xue, Y., 825  
Xue, Z., 1017

**Y**

Yang, B., 775  
Yang, F., 1027  
Yang, J., 747  
Yang, J.-L., 1045  
Yang, L., 1165  
Yang, T., 913  
Yang, Y., 763  
Yang, Z., 1099, 1233, 1245  
Yao, J., 1181  
Yi, C.-F., 713  
Yi, D., 979  
Yin, B., 1027  
Yin, S., 835, 1063  
Yin, Y., 1007  
Ying, S., 1173  
Yu, G., 747  
Yu, Q., 1133  
Yuan, H., 817, 1315

**Z**

Zha, F., 971  
Zhai, J., 1073, 1083, 1323  
Zhang, C.N., 1133  
Zhang, F., 1297  
Zhang, G., 1073, 1083  
Zhang, H., 1381  
Zhang, L., 755, 1117, 1125, 1349, 1357  
Zhang, M., 1217, 1373  
Zhang, P., 921  
Zhang, Q., 1099, 1263  
Zhang, R., 1373  
Zhang, X., 847, 905, 1099  
Zhang, Y., 791, 865, 1017, 1165  
Zhang, Z., 817, 1099, 1315  
Zhao, H., 1271  
Zhao, S., 847  
Zhao, T., 817, 1315  
Zhao, W., 1297  
Zhao, X., 963  
Zheng, H., 1191  
Zheng, Z., 847  
Zhou, F., 1155  
Zhou, X., 1225  
Zhou, Y., 1017  
Zhou, Z., 1395  
Zhu, M., 1063  
Zhuo, M., 1373  
Zou, Y., 1063  
Zuo, Y., 1233

# Subject Index for Volume 2

## A

AADL. *See* Architecture Analysis and Design Language (AADL)  
Access category (AC), 1165, 1171  
Access control, 784, 791–797, 1203–1206, 1294  
Advanced persistent threat (APT), 1297–1304  
Advertising delivery strategy, 856  
Aesthetic criterion, 1384, 1385, 1393, 1394  
Android, 740, 749, 788, 837, 1099–1106, 1112–1116  
ANFIS, 1373–1378  
APT. *See* Advanced persistent threat (APT)  
Architecture Analysis and Design Language (AADL), 1117–1123, 1125–1131  
Area efficiency, 1064, 1069, 1071  
Aspect-oriented, 1125–1131  
Asynchronous mode, 1001, 1002, 1290

## B

Basic Detection Strategy, 802–805, 807  
Bike rental station, 1341–1348  
Bi-level model, 1346, 1348  
Bipartite graph, 826, 829, 830, 832  
Bounded retransmission protocol (BRP), 1245–1251  
B+R mode, 1341–1348

## C

Cache, 764–767, 786, 900, 1064, 1143–1151  
Carrier aggregation, 714, 715, 786, 990–992, 1212, 1291  
Chain network, 1209–1216  
Channel reservation, 1166, 1169, 1199

Churn, 1263, 1264, 1267–1269  
Clock, 787, 840, 843, 906, 907, 910, 989, 993, 1050, 1290, 1350, 1353–1355, 1357–1362  
Clock theory, 1350, 1353, 1354, 1358, 1360–1362  
Cloud computing, 723–727, 747–754, 791–797, 799–801, 805, 807, 817–823, 865, 866, 893–902, 1112  
Cloud GIS, 893, 894, 896–902  
Cloud security, 759, 865–871, 894  
Cloud storage, 847–853, 895, 898  
Cluster-based, 1253–1261  
CMOS, 843, 906, 909, 912, 1036, 1040, 1042, 1043, 1046, 1047, 1051, 1091–1095, 1097  
Collaborative filtering, 1017–1024  
Communicating Sequential Processes (CSP), 1350, 1352–1354  
Communication, 724, 775–782, 784–787, 790, 813, 835, 836, 843, 844, 868, 875, 886, 905, 963, 997, 1005, 1007, 1008, 1012, 1035–1043, 1056, 1063, 1073, 1091, 1109–1116, 1118, 1120, 1128, 1165, 1182, 1183, 1191, 1202, 1204, 1206, 1209, 1223, 1246, 1264, 1271, 1294, 1302, 1308, 1309, 1319, 1331–1339, 1350, 1352, 1353, 1395–1405  
Complex network, 1226, 1228–1231, 1281, 1282  
Computer graphics, 947  
Computer power management system, 913–920  
Configurability, 998, 999, 1005  
Conjecture, 1332, 1334–1338  
Constrained optimizing layout, 1381–1394

Context-triggered piecewise hash (CTPH),  
1100, 1102, 1103, 1105, 1106  
Continuous and discrete, 1128  
Control mode, 929–936  
Cooperative communication, 1395–1405  
Coq, 1233–1243  
Cost optimization, 1343, 1344  
Co-training, 979–986  
CPS. *See* Cyber physical systems (CPS)  
Cross-domain, 791–797  
CSP. *See* Communicating Sequential  
Processes (CSP)  
Cyber counterwork, 1155–1162  
Cyber physical systems (CPS), 1117–1119,  
1125–1130, 1353, 1357–1362

## D

D/A converter, 1091–1097  
Data analysis, 810, 815, 870, 1254  
Decision-making, 767, 855–862, 885, 886, 1157  
Decision support system, 825–832  
Dedicated short-range communication  
(DSRC), 1035–1043  
Delay-hopped transmitted reference UWB,  
1396, 1399  
Denotational semantics, 1240  
Desktop virtualization, 751–752, 754  
Digital library, 747–754  
Digital publishing, 888, 889  
Discrete-time Markov chain (DTMC),  
1246–1248  
Dispatching, 723–730  
Distributed file system, 817–823, 847, 898  
Distribution network, 1077, 1083–1089,  
1323–1330  
Distribution system, 1073–1080, 1084,  
1086, 1088, 1324  
Domain specific languages, 1217–1224  
Drowsy, 1143–1151  
DSRC. *See* Dedicated short-range  
communication (DSRC)  
DTMC. *See* Discrete-time Markov chain  
(DTMC)  
DVB-RCS, 776–777, 782  
Dynamic adaptive, 1156  
Dynamic routing switching, 1317  
Dynamic voltage scaling, 989–996

## E

Earliest deadline first, 989–996  
E-commerce application level, 873–881  
EDCA. *See* Enhanced distributor channel  
access (EDCA)

Edge marking algorithm, 947–952  
Educational Technology, 871  
e-Health monitoring, 731–738  
Electronic product code (EPC), 1271–1278  
Electronic vehicle (EV), 1055–1060  
Embedded Linux, 956–959, 993  
Embedded processor, 1143, 1150, 1151  
Embedded system, 947, 990, 1007–1015, 1117,  
1130, 1217, 1218, 1359  
Energy conservation, 913, 914  
Energy consumption, 920, 963–969, 990,  
995, 996, 1144, 1191, 1210, 1211,  
1214, 1215, 1307–1309, 1311, 1313,  
1316, 1319  
Enhanced distributor channel access (EDCA),  
1166, 1169–1171  
EPC. *See* Electronic Product Code (EPC)  
Exception handling framework, 1173–1179  
Expert system, 809–816  
Extended Petri net, 1297–1304

## F

Face detection, 913–920  
Fault detection, 911, 1083–1089, 1327  
Fault diagnosis, 737, 825–832, 1323–1330  
Fault-tolerance design, 731, 732, 735–738  
Fault-tolerant requirements, 732  
Field authority, 1017–1024  
Filtering, 803, 804, 971–977, 1017–1024,  
1103, 1203, 1204, 1240  
Fire-fighting, 741–743  
Forecasting model, 1377  
Formal expression, 1301–1304  
Formal specification, 1128, 1242, 1351–1353  
Format-compliant, 778–781  
Forwarding probability, 1308–1310  
Frequency synthesizer, 1045–1053  
Fuzzy logic, 1192, 1193, 1199

## G

GA. *See* Genetic algorithm (GA)  
General interface system, 998  
Genetic algorithm (GA), 1060, 1345, 1346  
Genre, 1018–1021, 1024  
GTIN-716, 1272–1276  
Guest OS, 1202, 1204–1206

## H

Hadoop, 817–823, 847, 848, 851, 898  
Hadoop Distributed File System (HDFS),  
817–823, 847–850, 853, 898  
Handheld controller, 955–961



HDFS. *See* Hadoop Distributed File System (HDFS)

Healthcare, 733, 1126

Hierarchical local-interconnection, 1063–1071

Hierarchical mobile IPv6 network, 763–773

Hierarchical scene analysis, 1298, 1299, 1301, 1302

High-dimensional data, 980, 981, 986

High impedance fault, 1073–1080, 1083, 1323

Host OS, 1203

Human role, 1155–1162

Hybrid layout, 1382, 1383

## I

IaaS. *See* Infrastructure as a Service (IaaS)

Identity federation, 756, 759–761

IEEE802.15.4, 1210, 1213–1216

IEEE 802.11p, 1165–1172

Incipient fault, 1083–1087, 1089, 1323–1330

Indoor GIS, 739–745

Indoor map, 739–745

InfiniBand switch, 963–969

Influence diagram, 855–862

Information dissemination, 886, 1225–1231

Infrastructure as a Service (IaaS), 895, 896, 899, 900

Instruction, 815, 1101, 1102, 1143–1151, 1205, 1275, 1366

Integration, 744, 750, 751, 896, 897, 905, 1005, 1056, 1119, 1123, 1130, 1404, 1405

Integrity measurements, 869

Intelligent transportation system (ITS), 1035, 1165

Interface bridge, 835–844

Internal resistance, 1028–1030, 1033

Inter-turn short circuit, 940

IPv6, 763–773, 1007–1015

Isolated neutral system, 1087, 1324

ITS. *See* Intelligent transportation system (ITS)

IVI, 1008, 1009

## K

Key management, 1115, 1134, 1138, 1141

Knowledge, 710, 810–814, 816, 856, 857, 860, 874, 983, 1002, 1056, 1134, 1158, 1166, 1177, 1178, 1298, 1317, 1321, 1331–1339

## L

Laser bending, 921–927

Learning system, 1198

Light protocol, 1209–1216

Load control scheme, 763–773

Localization, 1067, 1253–1261

Long-Term Evolution Advanced (LTE-A), 713–720

## M

Mac protocol, 784–786, 790

Magnetic resonance, 1056

Map updating, 744

MDA. *See* Model-driven architecture (MDA)

MDS-MAP, 1253–1261

Measurement indicators, 875

Media selection, 855–862

Memory controller, 837, 905–912

Message, 715, 724, 764–767, 770, 771,

801, 1110, 1111, 1113, 1133,

1135–1137, 1140, 1166, 1168,

1173–1179, 1205, 1219, 1225,

1234, 1241, 1245, 1255, 1256, 1275,

1290–1292, 1309–1311, 1332, 1335,

1336, 1352, 1353

Method of moments (MoM), 1056–1058, 1060

Method of Poynting vector, 1028, 1033

Mirror-role, 797

Mobile payment, 706–711

Model-based development, 1130, 1243

Model-driven architecture (MDA), 1118

Modelica, 1117–1123

Model transformation, 1118, 1123, 1127

MoM. *See* Method of moments (MoM)

Multi-bank, 705, 710, 711

Multi-loop theory, 939–945

Multiple-target, 835–844

Multiple view, 979–986

## N

NAND flash, 836, 841–844

Nash equilibrium, 1332, 1334, 1335, 1337–1339

NAT-PT, 1008, 1009

Nearest neighbor, 1017, 1020–1022

Near field communication (NFC), 1109–1116

Network design, 783–790, 1264, 1342, 1348

NFC. *See* Near field communication (NFC)

Nonlinear frequency analysis, 1073–1080

## O

OAuth, 756–759, 761

Object Linking and Embedding for Process Control (OPC), 997–1006

- Object-Z, 1350, 1352–1354  
 Off-line payment, 706, 709  
 One-way hash chain, 1133–1141  
 On-line payment, 708–709  
 OODA loop, 1156–1157, 1162  
 OPC. *See* Object Linking and Embedding for Process Control (OPC)  
 Operating system (OS), 749, 750, 752, 821, 837, 851, 867, 869–871, 894, 897, 994, 1013, 1112, 1156, 1158, 1202–1206, 1223, 1299, 1300, 1302, 1303  
 OPNET, 1210, 1213, 1214, 1294  
 OS. *See* Operating system (OS)
- P**  
 PaaS. *See* Platform as a Service (PaaS)  
 Packet delivery, 764–768, 770, 772  
 Packing algorithm, 847–853  
 Parking guidance, 783–790  
 p-CSMA, Persistent carrier sense multiple access (p-CSMA)  
 Performance, 714, 723, 732, 750, 764, 776, 805, 810, 817, 836, 847, 870, 874, 886, 907, 919, 922, 963, 976, 980, 990, 1000, 1018, 1028, 1041, 1048, 1056, 1089, 1091, 1112, 1127, 1140, 1143, 1162, 1169, 1182, 1193, 1215, 1247, 1263, 1290, 1308, 1315, 1327, 1367, 1396  
 Peripheral environment simulation, 1217–1224  
 Persistent carrier sense multiple access (p-CSMA), 1290, 1291, 1293, 1294, 1296  
 Phase-locked loop, 1048  
 Plasma column antenna, 1027–1029, 1032  
 Platform as a Service (PaaS), 895, 896, 899, 900  
 PLL, 1046, 1048, 1049, 1052, 1053  
 Poisson distribution, 726, 1283, 1290, 1292  
 Polygon filling, 947, 951  
 Polymorphic, 1365–1372  
 Position-relation signature (PRS), 1365–1372  
 Power, 711, 714, 737, 785, 810, 844, 858, 912, 924, 929, 939, 955, 963, 989, 1010, 1028, 1035, 1045, 1055, 1073, 1083, 1095, 1109, 1118, 1126, 1134, 1143, 1192, 1229, 1235, 1271, 1282, 1297, 1308, 1319, 1323, 1352, 1396  
 Power amplifier, 1039  
 Power-law distribution, 1282, 1283, 1285–1287  
 Power saving, 713–720, 914–916, 919, 920  
 Prediction, 875, 974, 980, 1017, 1018, 1020–1023, 1084, 1143–1151, 1159, 1324, 1332, 1336, 1337, 1339, 1376  
 PRISM, 1246, 1247, 1250, 1251  
 Probabilistic model checking, 1246, 1247, 1251  
 Probability, 727, 728, 770, 803, 825–832, 856, 858, 860–862, 870, 906, 934, 942, 1166, 1168–1171, 1225–1231, 1247, 1249–1251, 1290, 1291, 1295, 1296, 1307–1313, 1333, 1334, 1337, 1347, 1384  
 Process control, 898, 997–1005  
 Propagation delay, 1168, 1169, 1171, 1291, 1296  
 Protocol, 708, 756, 763, 775, 784, 836, 956, 963, 997, 1008, 1035, 1110, 1120, 1134, 1158, 1166, 1174, 1203, 1209–1245, 1263, 1290, 1307, 1315, 1332, 1396  
 PRS. *See* Position-relation signature (PRS)  
 Public bike, 1341–1348
- Q**  
 QMR. *See* Quadruple Modular Redundant (QMR)  
 QoS. *See* Quality of services (QoS)  
 Qt GUI, 959, 960  
 Quadrupe robot, 972, 975–977  
 Quadruple Modular Redundant (QMR), 905–912  
 Quality of services (QoS), 714–716, 732, 1125–1131, 1166, 1171, 1192–1198  
 Query processing, 1315, 1316, 1321
- R**  
 Radiation resistance, 1028, 1031–1033  
 Radio-frequency identification (RFID), 733, 1110, 1271  
 Radio resource management, 1191–1199  
 Radio resource scheduling, 713–720  
 Range-free, 1253–1261  
 Rating prediction, 1020–1023  
 RCAA. *See* Resource congestion avoidance algorithm (RCAA)  
 RC4 Based Hash Function (RC4 BHF), 1133–1141  
 RC4 BHF. *See* RC4 Based Hash Function (RC4 BHF)  
 RC4 stream cipher, 1134–1136  
 Real-time, 741, 743, 776, 782, 826, 898, 955–961, 971–977, 989–991, 995, 996, 1126–1128, 1157, 1173, 1175, 1181, 1215, 1218, 1246, 1302, 1326, 1353, 1354, 1359, 1360, 1398  
 Recommender system, 1017

- Reconfigurable processing unit (RPU), 1063–1071
- Regional registration, 764, 766–768, 770, 772
- Region partition, 1382–1385, 1387, 1389, 1390, 1393, 1394
- Reinforcement learning, 1191–1199
- Reliability, 710, 731, 735, 736, 751, 755, 870, 871, 905–912, 1005, 1089, 1118, 1159, 1210, 1211, 1234, 1243, 1245–1247, 1308, 1330, 1349, 1350, 1353, 1395
- Repackage, 1099, 1106
- Resonance, 972, 1032, 1033
- Resource allocation, 714–720, 724, 902, 1128, 1181–1189
- Resource congestion avoidance algorithm (RCAA), 1183–1189
- Resource discovery, 1263, 1264
- Resource information organization, 1263–1270
- Resource management, 723–725, 728, 729, 1191–1199
- RFID. *See* Radio-frequency identification (RFID)
- Routing protocol, 1172, 1307–1313, 1315–1321
- RPUs. *See* Reconfigurable processing unit (RPU)
- R-2R ladder, 1091–1097
- Ruby, 1218, 1220–1222
- Runoff forecast, 1373
- S**
- SaaS. *See* Software as a Service (SaaS)
- Sampling expansion, 1396, 1399
- Satellite communication, 775–782
- Scale-free network, 1230, 1281–1287
- Scoring system, 809–816
- SD. *See* Security Digital (SD)
- SDH network, 1181–1189
- SEC. *See* Spectral embedded clustering (SEC)
- Security, 709–710, 724, 740, 745, 748, 753, 756, 758–761, 775, 792–797, 836, 865–871, 894, 897, 898, 914, 1109–1112, 1115, 1118, 1119, 1123, 1126–1128, 1130, 1133–1136, 1138–1141, 1159, 1201–1206, 1278, 1362, 1365
- Security Digital (SD), 835–844
- Selective content encryption, 775–782
- Server virtualization, 749–751, 754
- Service channel, 1165–1172, 1196
- SET. *See* Single event transient (SET)
- SEU. *See* Single event upset (SEU)
- SGTIN-798, 1272–1277
- Sheet, 921–927
- Short circuit fault, 932, 1084
- Similarity, 984, 1020, 1021, 1100, 1101, 1103–1106, 1121, 1282
- Simulation testing environment, 769–770, 827, 1217–1224
- Simulation training platform, 809–816
- Single event transient (SET), 906, 907
- Single event upset (SEU), 905–909, 912
- Single sign-on, 755–761
- SIR model, 1226–1228, 1231
- Small file, 847–853
- S4n-logic, 1333
- Software as a Service (SaaS), 895, 896, 899–900
- Software engineering, 1119, 1233
- Software testing, 1217
- Spectral embedded clustering (SEC), 979–986
- Stateful translation, 1007–1015
- Stateless translation, 1007–1015
- Stator, 931, 939–945
- Storage efficiency, 847, 848
- Storage system, 725, 817, 818, 847, 848, 898, 963, 964, 969
- S-transform, 1323–1330
- Structure, 715, 744, 747, 776, 788, 810, 826, 837, 868, 894, 906, 914, 929, 939, 948, 957, 979, 990, 999, 1029, 1036, 1056, 1063, 1121, 1128, 1134, 1145, 1156, 1176, 1209, 1219, 1225, 1234, 1263, 1273, 1283, 1294, 1299, 1307, 1331, 1359, 1366, 1374, 1382, 1399
- Structured P2P, 1263, 1264, 1266, 1269
- Synchronous generator, 939, 940, 942
- T**
- Tag, 1001, 1004, 1005, 1018, 1019, 1112, 1143–1151, 1174, 1271–1273, 1276, 1277
- TCP/IP, 776, 777, 821, 836, 955–961, 1007, 1010–1013, 1166–1168, 1170, 1171, 1203–1206
- Technical support, 901
- Thermal stress, 921, 922
- Threshold Strategy, 804, 805
- Time analysis, 1325
- Titanium, 921–927
- Tracking, 976, 977, 998–1002, 1005, 1159, 1350, 1351, 1353
- Traffic, 717, 719, 727, 763, 770, 772, 773, 964, 1014, 1035, 1105, 1112, 1166, 1171, 1172, 1183, 1264, 1267–1269, 1291, 1307, 1343–1345, 1348, 1351, 1353, 1366–1368
- Train control systems, 1349–1355

Transceiver front-end, 1035–1043  
 Transmitted cooperative reference, 1395–1405  
 Trust, 760, 791–797, 865–871, 1138, 1139  
 Trusted cloud, 865–871  
 Trusted computing, 865–871  
 T-type network, 1091–1097  
 Twist, 921–927

## U

Ubiquitous computing (UBICOMP), 835, 836, 844  
 UML sequence diagrams, 1233–1243  
 Unhealthy cloud system status, 799–807

## V

VANET. *See* Vehicular ad hoc network (VANET)  
 VCO. *See* Voltage-Controlled Oscillator (VCO)  
 VDL2, 1294–1295  
 Vehicular ad hoc network (VANET), 1130, 1165–1172  
 Verification, 745, 842, 843, 959–961, 1130, 1206, 1233–1243, 1245–1251, 1281–1287, 1351, 1352, 1355, 1360, 1362  
 Vertex-degree sequence, 1281–1287  
 Video transmission, 955–961

Virtualization, 747–754, 820, 897, 898, 1201  
 Virtual machine monitor, 820, 1202  
 Visualization of clustered network graph, 1381–1394  
 VMware vSphere, 748, 749, 751, 752  
 Voltage-Controlled Oscillator (VCO), 1046–1052  
 Voltage transient stability, 929–936

## W

Wavelet analysis, 1087, 1373, 1374, 1376–1378  
 Wavelet transform, 1083–1085, 1323–1325, 1327  
 Wavelet-ANFIS, 1373–1378  
 Web Service, 899, 900, 1173–1179  
 Weighted current, 1091–1097  
 Wind generator, 929–936  
 Wireless Local Area Network (WLAN), 1109–1116  
 Wireless Power Transmission (WPT), 1055–1058, 1060  
 Wireless sensor network (WSN), 784, 843, 1209, 1254, 1307–1313, 1315–1321  
 WLAN. *See* Wireless Local Area Network (WLAN)  
 Worm detection, 1365–1372  
 WPT. *See* Wireless Power Transmission (WPT)  
 WSN. *See* Wireless sensor network (WSN)