

Omer H. Abdelrahman

Erol Gelenbe

Gokce Gorbil

Ricardo Lent

*Editors*

# Information Sciences and Systems 2015

30th International Symposium on  
Computer and Information Sciences  
(ISCIS 2015)

# Lecture Notes in Electrical Engineering

Volume 363

## Board of Series editors

Leopoldo Angrisani, Napoli, Italy  
Marco Arteaga, Coyoacán, México  
Samarjit Chakraborty, München, Germany  
Jiming Chen, Hangzhou, P.R. China  
Tan Kay Chen, Singapore, Singapore  
Rüdiger Dillmann, Karlsruhe, Germany  
Haibin Duan, Beijing, China  
Gianluigi Ferrari, Parma, Italy  
Manuel Ferre, Madrid, Spain  
Sandra Hirche, München, Germany  
Faryar Jabbari, Irvine, USA  
Janusz Kacprzyk, Warsaw, Poland  
Alaa Khamis, New Cairo City, Egypt  
Torsten Kroeger, Stanford, USA  
Tan Cher Ming, Singapore, Singapore  
Wolfgang Minker, Ulm, Germany  
Pradeep Misra, Dayton, USA  
Sebastian Möller, Berlin, Germany  
Subhas Mukhopadhyay, Palmerston, New Zealand  
Cun-Zheng Ning, Tempe, USA  
Toyoaki Nishida, Sakyo-ku, Japan  
Bijaya Ketan Panigrahi, New Delhi, India  
Federica Pascucci, Roma, Italy  
Tariq Samad, Minneapolis, USA  
Gan Woon Seng, Nanyang Avenue, Singapore  
Germano Veiga, Porto, Portugal  
Haitao Wu, Beijing, China  
Junjie James Zhang, Charlotte, USA

### *About this Series*

“Lecture Notes in Electrical Engineering (LNEE)” is a book series which reports the latest research and developments in Electrical Engineering, namely:

- Communication, Networks, and Information Theory
- Computer Engineering
- Signal, Image, Speech and Information Processing
- Circuits and Systems
- Bioengineering

LNEE publishes authored monographs and contributed volumes which present cutting edge research information as well as new perspectives on classical fields, while maintaining Springer’s high standards of academic excellence. Also considered for publication are lecture materials, proceedings, and other related materials of exceptionally high quality and interest. The subject matter should be original and timely, reporting the latest research and developments in all areas of electrical engineering.

The audience for the books in LNEE consists of advanced level students, researchers, and industry professionals working at the forefront of their fields. Much like Springer’s other Lecture Notes series, LNEE will be distributed through Springer’s print and electronic publishing channels.

More information about this series at <http://www.springer.com/series/7818>

Omer H. Abdelrahman · Erol Gelenbe  
Gokce Gorbil · Ricardo Lent  
Editors

# Information Sciences and Systems 2015

30th International Symposium on Computer  
and Information Sciences (ISCIS 2015)

 Springer

*Editors*

Omer H. Abdelrahman  
Department of Electrical and Electronic  
Engineering  
Imperial College  
London  
UK

Gokce Gorbil  
Department of Electrical and Electronic  
Engineering  
Imperial College  
London  
UK

Erol Gelenbe  
Department of Electrical and Electronic  
Engineering  
Imperial College  
London  
UK

Ricardo Lent  
Department of Engineering Technology  
University of Houston  
Houston, TX  
USA

ISSN 1876-1100                      ISSN 1876-1119 (electronic)  
Lecture Notes in Electrical Engineering  
ISBN 978-3-319-22634-7            ISBN 978-3-319-22635-4 (eBook)  
DOI 10.1007/978-3-319-22635-4

Library of Congress Control Number: 2015946093

Springer Cham Heidelberg New York Dordrecht London  
© Springer International Publishing Switzerland 2016

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.

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.

The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

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

# Preface

The 30th anniversary of the *International Symposium on Computer and Information Sciences* (ISCIS) series, which have regularly over this long period of time brought together Computer Scientists from around the world and from Turkey, was held at Imperial College London, UK during 21–24 September 2015.

This year the conference attracted 82 submissions from 21 countries—with contributions coming mostly from Europe, America and the Far East—out of which 39 carefully refereed proposals were selected, along with two Invited Papers, for inclusion in the proceedings. Several other Invited Papers were presented orally.

In addition to the papers that are contained in these proceedings, the symposium was preceded by a conference comprising some 50 keynote and invited presentations in honour of Prof. Erol Gelenbe, who started ISCIS back in 1986, and kept it going constantly with the help of several colleagues from Europe and Turkey.

This volume provides a compact yet broad view of recent developments in the Computer and Information sciences, and covers exciting research areas in the field including Green and Cloud computing, Performance Modelling, Cybersecurity, Big Data, and Smart Algorithm design for computer, biological and chemical systems. This symposium also highlights recent results from the EU FP7 NEMESYS Project.

We are very grateful to the authors of all of the submitted papers, and to the authors of accepted papers, for their contributions, and to the technical programme committee members who each evaluated several papers, without whom this exciting programme would not have been possible.

The ISCIS 2015 Chairs

London  
Houston  
September 2015

Omer H. Abdelrahman  
Erol Gelenbe  
Gokce Gorbil  
Ricardo Lent

# Programme Committee

Omer Abdelrahman  
Ethem Alpaydin  
Cevdet Aykanat  
Selim Balcisoy  
Madalina Baltatu  
Javier Barria  
Olivier Beaumont  
Leszek Borzemski  
Jeremy Bradley  
Manfred Broy  
Fazli Can  
Sophie Chabridon  
Ilyas Cicekli  
Nihan Cicekli  
Tadeusz Czachorski  
Gokhan Dalkilic  
Mariangiola Dezani  
Engin Erzin  
Taner Eskil  
Jean-Michel Fourneau  
Erol Gelenbe  
Stephen Gilmore  
Mariusz Glabowski  
Gokce Gorbil  
Krzysztof Grochla  
Adam Grzech  
Lin Guan  
Ugur Gudukbay  
Attila Gursoy  
Ugur Halici  
Peter Harrison

Yorgo I Stefanopoulos  
Alain Jean-Marie  
Sylwester Kaczmarek  
Jacek Kitowski  
Stefanos Kollias  
Jerzy Konorski  
Ibrahim Koreoglu  
Stanislaw Kozielski  
Olcay Kursun  
Ricardo Lent  
Albert Levi  
Aristidis Likas  
George Limperopoulos  
Peixiang Liu  
Chris Mitchell  
Marek Natkaniec  
Ender Ozcan  
Oznur Ozkasap  
Zdzislaw Papier  
Ferhan Pekergin  
Nihal Pekergin  
Yves Robert  
Alexandre Romariz  
Georgia Sakellari  
Ercan Solak  
Andreas Stafylopatis  
Halina Tarasiuk  
Eleni Theodoropoulou  
Dao Thi  
Nigel Thomas  
Hakki Toroslu  
Salvatore Tucci  
Dimitrios Tzovaras  
Ozgur Ulusoy  
Ozlem Uzuner  
Krzysztof Walkowiak  
Wei Wei  
Jozef Wozniak  
Zhiguang Xu  
Adnan Yazici  
Cemal Yilmaz  
Thomas Zeugmann  
Qi Zhu



## **Additional Reviewers**

Hasan Ates  
Esma Fatima Bilgin Tasdemir  
Elif Bozkurt  
Cem Bozsahin  
Tib Chis  
Tugrul Dayar  
M. Taner Eskil  
Erol Gelenbe  
Ira Goldstein  
Krzysztof Grochla  
Cigdem Gunduz-Demir  
Przemysław Głomb  
Ali Inan  
Altynbek Isabekov  
Elmar Juergens  
Rafal Krenz  
Olcay Kursun  
Adrian Langowski  
Amir Ligata  
Andrea Marin  
Rasha Osman  
Mustafa Ozdal  
Tuan Phung-Duc  
Kostas Rapantzikos  
Piotr Remlein  
Djamé Seddah  
R. Oguz Selvitopi  
Ercan Solak  
Pawel Sroka  
Alexander Trofimovsky  
Krzysztof Turowski

Ayşegül Tüysüz Erman  
Juan Manuel Vara  
Michał Wąrowski  
Olca Taner Yıldız  
Piotr Zwierzykowski

# Contents

## Part I Invited Papers

<b>ISCIS and Erol Gelenbe’s Contributions</b> . . . . .	3
Mehmet Ufuk Çağlayan	
<b>Interoperability and Semantics: A Never-Ending Story</b> . . . . .	19
Erich Neuhold and Elmar Kiesling	

## Part II Green Computing and Networking

<b>Performance of an Autonomous Energy Harvesting Wireless Sensor</b> . . . . .	35
Erol Gelenbe and Yasin Murat Kadioglu	
<b>Towards Assessment of Energy Consumption and Latency of LTE UEs During Signaling Storms</b> . . . . .	45
Frederic Francois, Omer H. Abdelrahman and Erol Gelenbe	
<b>A State-Dependent Control for Green Computing</b> . . . . .	57
Evsey Morozov and Alexander Romyantsev	
<b>Environment Friendly Energy Efficient Distributed Data Centers</b> . . . .	69
Ahsan Ali and Oznur Ozkasap	
<b>Modeling Power Consumption in Multicore CPUs with Multithreading and Frequency Scaling</b> . . . . .	81
D. Cerotti, M. Gribaudo, P. Piazzolla, R. Pincirolì and G. Serazzi	

**Part III Network Security**

**A Role and Activity Based Access Control for Secure Healthcare Systems . . . . .** 93  
 Naim Alperen Pulur, Duygu Karaođlan Altop and Albert Levi

**Bandwidth Usage—Based Detection of Signaling Attacks. . . . .** 105  
 Mihajlo Pavloski, Gökçe Görbil and Erol Gelenbe

**A BRPCA Based Approach for Anomaly Detection in Mobile Networks . . . . .** 115  
 Stavros Papadopoulos, Anastasios Drosou, Nikos Dimitriou, Omer H. Abdelrahman, Gokce Gorbil and Dimitrios Tzouvaras

**LBP-DCT Based Copy Move Forgery Detection Algorithm . . . . .** 127  
 Beste Ustubioglu, Guzin Ulutas, Mustafa Ulutas, Vasif Nabiyeve and Arda Ustubioglu

**Smart Mobile Ecosystem Security: Existing Solutions, MNO Requirements and Business Model . . . . .** 137  
 George Lyberopoulos, Helen Theodoropoulou, Konstantinos Filis and Ioanna Mesogiti

**Undermining Isolation Through Covert Channels in the Fiasco.OC Microkernel . . . . .** 147  
 M. Peter, M. Petschick, J. Vetter, J. Nordholz, J. Danisevskis and J.-P. Seifert

**Part IV Smart Algorithms**

**A Hybrid Movie Recommender Using Dynamic Fuzzy Clustering . . . . .** 159  
 Fatih Gurcan and Aysenur Akyuz Birturk

**Generating Minimum Height ADSs for Partially Specified Finite State Machines . . . . .** 171  
 Robert M. Hierons and Uraz Cengiz Türker

**Line-Search Aided Non-negative Least-Square Learning for Random Neural Network . . . . .** 181  
 Yonghua Yin

**A Novel Concise Specification and Efficient F-Logic Based Matching of Semantic Web Services in Flora-2** . . . . . 191  
 Shahin Mehdipour Ataee and Zeki Bayram

**Fast Frequent Episode Mining Based on Finite-State Machines** . . . . . 199  
 Stavros Papadopoulos, Anastasios Drosou and Dimitrios Tzovaras

**Hybrid Heuristic Algorithms for the Multiobjective Load Balancing of 2D Bin Packing Problems** . . . . . 209  
 Muhammed Beyaz, Tansel Dokeroglu and Ahmet Cosar

**Part V Stochastic Modelling and Computer Networks**

**Smoothing the Input Process in a Batch Queue**. . . . . 223  
 F. Ait Salaht, H. Castel Taleb, J.M. Fourneau, T. Mautor and N. Pekergin

**Network-Based Job Dispatching in the Cloud** . . . . . 233  
 Byungseok Kang

**Discrete Time Stochastic Automata Network with Steady-State Product Form Distribution**. . . . . 241  
 J.M. Fourneau

**Modelling Dynamics of TCP Flows in Very Large Network Topologies**. . . . . 251  
 Monika Nycz, Tomasz Nycz and Tadeusz Czachórski

**Numerically Efficient Analysis of a One-Dimensional Stochastic Lac Operon Model** . . . . . 261  
 Neslihan Avcu, Nihal Pekergin, Ferhan Pekergin and Cüneyt Güzeliş

**Part VI Image Processing and Computer Vision**

**Brain MR Image Denoising for Rician Noise Using Intrinsic Geometrical Information** . . . . . 275  
 Hamit Soyel, Kamil Yurtkan, Hasan Demirel and Peter W. McOwan

**Image Analysis in a Parameter-Free Setting** . . . . . 285  
 Yu Zhu and Thomas Zeugmann

<b>Age Estimation Based on Hybrid Features of Facial Images . . . . .</b>	<b>295</b>
Asuman Günay and Vasif V. Nahiye	

## **Part VII Algorithm Design for Biological and Chemical Systems**

<b>Proposal of a New Method for de Novo DNA Sequence Assembly Using de Bruijn Graphs. . . . .</b>	<b>307</b>
Adriano Donato Couto, Fabio Ribeiro Cerqueira, Ricardo dos Santos Ferreira and Alcione de Paiva Oliveira	

<b>A New Graph Algorithm for the Analysis of Conformational Dynamics of Molecules. . . . .</b>	<b>319</b>
D. Barth, S. Bougueroua, M.-P. Gaigeot, F. Quessette, R. Spezia and S. Vial	

## **Part VIII Natural Language Processing and Language Design**

<b>Two-Stage Feature Selection for Text Classification . . . . .</b>	<b>329</b>
Levent Özgür and Tunga Güngör	

<b>Constructing a Turkish Constituency Parse TreeBank. . . . .</b>	<b>339</b>
Olca Taner Yıldız, Ercan Solak, Şemsinur Çandır, Razieh Ehsani and Onur Görgün	

<b>A TV Content Augmentation System Exploiting Rule Based Named Entity Recognition Method . . . . .</b>	<b>349</b>
Yunus Emre Işıklar and Nihan Çiçekli	

<b>A Comparison Study on Ensemble Strategies and Feature Sets for Sentiment Analysis. . . . .</b>	<b>359</b>
Deniz Aldogan and Yusuf Yaslan	

<b>Feature Selection for Enhanced Author Identification of Turkish Text . . . . .</b>	<b>371</b>
Yasemin Bay and Erbuğ Çelebi	

<b>Noun Phrase Chunking for Turkish Using a Dependency Parser . . . . .</b>	<b>381</b>
Mucahit Kutlu and Ilyas Cicekli	

**Enabling Secure and Collaborative Document Sharing in BIM Processes** . . . . . 393  
 Carlo Argiolas, Nicoletta Dessi, Maria Grazia Fugini and Barbara Pes

**Tactical Graphics Description Language** . . . . . 403  
 İsmail Kiliñ, Hüseyin Ateş, Bülent Özhorasan and Hüseyin Korkmaz

**Part IX Wireless Networks**

**EASER: Energy Aware Scalable and Reactive Replication Protocol for MANETs** . . . . . 417  
 Saeed Nourizadeh Azar, Kaan Karaagacli and Oznur Ozkasap

**Fractional Frequency Reuse Based Adaptive Power Control Scheme for Interference Mitigation in LTE-Advanced Cellular Network with Device-to-Device Communication** . . . . . 429  
 Sok Chhorn, Si-o Seo, Ji-eun Song, Suk-ho Yoon, Seung-yeon Kim and Choong-ho Cho

**Influence of the Management Protocols on the LTE Self-configuration Procedures' Performance** . . . . . 439  
 Mariusz Slabicki and Krzysztof Grochla

**Subcarrier Allocation for LTE Soft Frequency Reuse Based on Graph Colouring** . . . . . 447  
 Krzysztof Grochla and Konrad Połys

**Erratum to: Information Sciences and Systems 2015** . . . . . E1  
 Omer H. Abdelrahman, Erol Gelenbe, Gokce Gorbil and Ricardo Lent

**Author Index** . . . . . 455

**Part I**  
**Invited Papers**



# ISCIS and Erol Gelenbe's Contributions

Mehmet Ufuk Çağlayan

**Abstract** We celebrate the 30th annual ISCIS (International Symposium on Computer and Information Sciences) that Erol Gelenbe started in 1986 and pursued each year since then uninterruptedly as a service to the Turkish Computer Science and Engineering Community. We also outline his scientific contributions and wide collaborations over the last fifteen years covering half of the life-time of ISCIS. These include his innovative work on a new representation of Intermittent or Renewable Energy Sources, and *Energy Packet Networks* which are a convenient representation for the flow, storage and consumption of electrical energy, both at the microscopic level (in electronic chips) and at the macroscopic level (e.g. in buildings or data centres), his work on ICT systems that parsimoniously use energy in order to achieve quality of service (QoS). Pioneering work on Autonomic Communications and the Cognitive Packet Network is also reviewed, followed by network security, Emergency Management Systems, Gene Regulatory Networks, and analytic models of computer systems and networks.

## 1 Introduction

The International Symposia on Computer and Information Sciences (ISCIS) was started in 1986 by Erol Gelenbe to provide the Computer Engineering and Computer Science community in Turkey with a venue for their academic research, interacting with their peers from abroad, creating a tradition of refereed international quality publications in this new community. Since 1986, ISCIS has been held annually without interruption, with the 30th held at Imperial College in September 2015. In recent years, the Proceedings have been published by Springer.

The first ISCIS was held at Bilkent University, Ankara, with the active support of Professor Ali Dođramacı. Bilkent University was newly formed, and emphasised Computer Engineering, Electrical and Electronic Engineering and Industrial Engineering, as key areas for Turkey's future. Very quickly, several Turkish universities,

---

M.U. Çağlayan (✉)

Department of Computer Engineering, Bođaziçi University, Bebek, Istanbul, Turkey  
e-mail: caglayan@boun.edu.tr

including the Middle East Technical University, Boğaziçi University in Istanbul, Ege University in Izmir, Istanbul Technical and Yıldız Technical Universities. Sabancı University in Istanbul participated in organising this annual event. In the early years ISCIS was actively supported by the Department Chairs Council for Computer Engineering, itself chaired by the late Professor Oğuz Manas of Ege University. Active colleagues included Professors Erol Arkun and Cevdet Aykanat of Bilkent University, Professors Ugur Halıcı, Hakkı Toroslu and Adnan Yazıcı of the Middle East Technical University, Professors Ali Rıza Kaylan and myself at Bogaziçi University.

Erol each year selected the venues and teams that would run the conference, “transferred” the know how and procedures to make it all work, training his colleagues in the “art” of organising scientific events without relying on subsidies, including the organisation, venues, publicity, refereeing and paper selection, identifying and attracting invited speakers from abroad, and remaining within budget. There were indeed difficulties for lack of funds, speakers from across Turkey with limited travel funds, quality control issues, third party attempts to obstruct the publication of proceedings due to foreign political pressure, as in North Cyprus in 2009. Since 1998, the conference was held in Europe, Turkey and America:

- 29. ISCIS 2014: Krakow, Poland [20],
- 28. ISCIS 2013: Paris, France [85],
- 27. ISCIS 2012: Paris, France [83],
- 26. ISCIS 2011: London, UK [89],
- 25. ISCIS 2010: London, UK [90],
- 24. ISCIS 2009: North Cyprus [1],
- 23. ISCIS 2008: Istanbul, Turkey,
- 22. ISCIS 2007: Ankara, Turkey,
- 21. ISCIS 2006: Istanbul, Turkey [146],
- 20. ISCIS 2005: Istanbul, Turkey [161],
- 19. ISCIS 2004: Antalya, Turkey [11],
- 18. ISCIS 2003: Antalya, Turkey [160],
- 17. ISCIS 2002: Orlando, Florida, USA [39],
- 15. ISCIS 2000: Istanbul, Turkey,
- 14. ISCIS 1999: Ege University, Izmir, Turkey,
- 13. ISCIS 1998: Ankara, Turkey.

Turning to Erol himself, it is hard to describe the scientific content of an ongoing career starting in the early 1970s for a highly productive researcher whose curiosity ranges widely in Computer Science, Applied Probability, Operational Research, Electrical Engineering and even Theoretical Biology. The variety of Erol’s work over the past fifteen years is unknown to those who know him through some specific research area. Thus our brief review points to his work in several areas since the year 2000, such as energy efficiency of ICT [14] and autonomic communications [24] where Erol has authored influential papers.

## 2 Energy Packet Networks

Erol's first work published in 2015 analyses the link between the random nature of harvested energy, and the random nature of the data collection activities of a wireless sensor [59], leading to an original analysis of "synchronisation" between the two resources that enable wireless communications: the data packets and the energy packets, first studied in a paper published in 2014 [58]. However in earlier, he had introduced of a novel way to view energy as a "packet-based" resource that can be modelled in discrete units which he called Energy Packets [53, 54]. While Ohm's Law is a good way to analyse the flow of electricity in RLC networks, there are areas where other models are needed:

- At a nano-scopic level, say at the level of the flow of individual electrons, both the stochastic nature of the sources and the physical non-homogeneities which govern the medium (e.g. metal) imply that different models may be needed; thus Erol recently proposed a stochastic flow model that addresses the conveyance of energy and information by individual particles [30, 57].
- At a more macroscopic level, when one deals with intermittent sources of energy so that energy must be stored in batteries or other storage units (such as compressed air cylinders) that can include conversion losses to and from the electrical storage, and energy usage itself is intermittent, models descended from G-Networks [35, 38] become useful [5, 55, 56].
- This approach has raised interesting questions about how such large networks may be analysed in the presence of flow of energy and flow of work [58, 62] and some recent interesting results regarding "product form solutions" for such multi-hop networks have also been obtained [100].

## 3 Energy in ICT and Its Optimisation

However, Erol's concern for energy consumption for communications actually started a decade earlier [40, 80] in the context of Wireless Ad-Hoc Networks, contributing a technique to extend overall life of a multi-hop network by using paths that have the most energy in reserve, i.e. the most full batteries. This work was pursued in papers related to network routing and admission control based on energy considerations [96, 98, 102, 103, 115, 116, 123, 157] and this resulted in a practical design for an energy aware routing protocol.

His research group's involvement with energy consumption in information technology was also developed through their participation in EU Fit4Green Project which resulted in a widely cited paper [14] regarding the energy optimisation of Cloud Computing servers and of software systems [152].

Although energy consumption by ICT is an important issue, it must be viewed as a compromise between the two aspects, where a reduction in energy consumption in the manner a specific system is being operated, for instance as a function of workload

or of workload distribution, is “paid for” by a loss in performance or an increase in the response times experienced by users. This issue has been studied in several of Erol’s recent papers [81, 82, 86, 152].

Similar problems arise in wireless communications, but of course at far lower levels of energy consumption. Here the purpose is to minimise the amount of energy consumed per correctly received packet or bit. Indeed, in the wireless case, increasing the transmission power is often possible. This will overcome noise, but it has the opposite (negative) effect if *all* cooperating transceivers raise their power level, resulting in greater wireless signal *interference* and hence larger error probabilities for all parties. This in turn will *lengthen* the time needed to correctly receive a data unit, and hence will also increase the net energy consumed per correctly received bit or packet [69, 84, 109, 150].

## 4 Autonomic Communications and CPN

Erol has been long intrigued with the adaptive control of computer systems and networks since the 1970s [12, 13, 77, 78, 110, 154], where the challenge is to deal both with the very large size of the systems encountered in computer science, the imperfection of the dynamic models that describe them, and the very large size of these dynamic models.

His most recent foray into this area, starting with early papers [66, 79, 108, 127] that describe the Cognitive Packet Network (CPN) routing algorithm both for wired and wireless networks that uses reinforcement learning to provide network Quality of Service (QoS) in an automatic manner, is a pioneering initiative in the field of Autonomic Communications [41]. He is also the co-author of a paper that made this field popular a few years later [24].

CPN related was a clear break with Erol’s traditional research which has relied essentially on mathematical modeling and simulation [31–33, 104, 117]. Related work [114, 151] suggests that decisions that have a “natural” appearance could also be incorporated in a similar manner, simulation systems where complex agent interactions occur, and agents take decisions based on their collective best interest. Similar questions have also been discussed with other methods in the context of search algorithms in dangerous environments [61].

The basic idea of CPN, amply tested in many experiments [87, 88, 91, 92, 94] is to use probe or “smart” Cognitive Packets (CPs) to search for paths and to measure QoS while the network is in operation. The search for paths is run via Reinforcement Learning using a Random Neural Network, based on the QoS objective of goal pursued by the end user. The CPs furnish information to the end user about the QoS offered by different paths, and in particular those actually being used by the end user, but in CPN it is the end user, which may be a representative decision maker for a QoS Class, that actually decides to switch to a new path or select a given path [50, 80, 93]. An extension to CPN that uses genetic algorithms to construct hitherto untested paths based on predicted QoS was also proposed [147].

More recent work has considered CPN for specific applications. For instance in [73] the issue of dealing with web access applications where the uplink may require short response times for Web requests, while the download may require high bandwidth and low packet loss for video downloads, is considered, and a specific system that supports these needs is designed, implemented and tested. Similarly, other recent work deals with the QoS needs of Voice, and a VoCPN system is designed and evaluated [158] on an experimental test-bed.

One of the interesting developments of CPN relates to novel energy aware routing algorithms [96, 97] that link to Erol's concern with energy savings. Other useful application concerns admission control [111] and denial of service defense [65, 95]. Other work that is unrelated to CPN but that proposes adaptive techniques for the management of wireless sensor networks in order to achieve better QoS are discussed in [105–107, 148], while adaptivity for the management of secondary memory systems is discussed in [163].

Of course, the Random Neural Network was reported in Erol's earlier work [34, 36, 37, 63, 64, 71, 99]. Some of its other applications can be found in [2–4, 8–10, 18, 19, 67, 74–76, 118, 133] and several papers reviewing this subject can be found in the papers of the special issue in [38].

## 4.1 Network Security

Erol's work on network security came through some work on the impact of Distributed Denial of Service (DDoS) Attacks on network QoS, and a proposal to use CPN as a way to detect DDoS, counter-attack by tracing the attacking traffic upstream and use CPN's ACK packets as a tool to give "drop orders" to upstream routers that are conveying the attacking traffic [95, 149]. This approach was also evaluated on a large network test-bed as a means to detect worm attacks and react to them by sending the users' traffic on routes that avoid the infected nodes [155, 156]. His work on security continued with more algorithmic issues [162], but recently moved to the analysis of signalling storms in mobile networks [7, 128, 129] and is currently one of the main centres of his attention.

## 5 Emergency Management Systems

Like many people around the world who have a personal experience of large scale disasters, Erol was deeply struck by what he saw in Turkey in 1999, right after the major earthquake that took place near Istanbul, in the areas around Izmit and Yalova. Although he was not present during the earthquake, he went to the sites just *after* the earthquakes with family members in order to try to find two family members who were missing and who has in fact perished during the event. Finding, identifying, and transporting the bodies was a traumatic experience. He was impressed by the very rudimentary nature of technology that was being used to seek, locate and try

to evacuate the victims. The destruction of roads, bridges, electrical power lines and telecommunication networks, meant that only rudimentary construction technologies could be used by the numerous family members and professional rescuers who flocked to the area in the days that followed the earthquake.

Since that time, Erol has devoted a part of his research effort to better understanding the ongoing research of emergency management technologies [124, 126], and developing a deeper understanding of the appropriate models and algorithms which are specific to this domain of research [21, 22, 68, 125]. Unfortunately, many aspects of emergency management have a significant overlap with tactical planning of military operations [72, 121, 122] whose purpose is to destroy the capabilities of an adversary but also to save the lives of friendly forces and evacuate the injured of both sides. In particular, his team has investigated different simulation approaches [25, 70, 72] that could be used to represent the extremely dynamic and fast changing “transient” events in an emergency, and then developed a novel agent based simulator named the Distributed Building Evacuation Simulator (DBES) [23] for evacuation planning and simulation. A constant concern of this work has been to develop decentralised techniques that do not require large and expensive infrastructures [28], and his team has organised a series of annual workshops related to the Pervasive Communications conferences of the ACM [26, 112, 113].

His team studied fast decision algorithms based on learning a wide range of problem instances and their optimal rescuer and rescue vehicle allocations [29, 119, 120], and selecting in real-time the allocation that best matches the current observed emergency situation. They also studied low-cost, light-weight and disruption tolerant techniques that can offer robust communications in emergency environments [132], and many of these methods actually span both the military and the civilian domain [27, 130, 131, 145, 159]. More recent work has been devoted to autonomic routing techniques based on ideas derived from CPN or from directional techniques so that the management of evacuees can be carried out without the intervention of any centralised decision making agent [15–17, 60, 144]

## 6 Gene Regulatory Networks

Erol’s interest in Gene Regulatory Networks [45], a very important topic in health sciences, started fortuitously during a visit in the mid 2000s to the well known French gene splicing centre, the Genopole in Evry, near Paris. He had been invited there by a friend, Gabriel Mergui who was then the Genopole’s industrialisation and marketing director and who knew about Erol’s interests in the interface between biology and computer science [52], to give a seminar (on his work in general) and to meet other colleagues.

At the Genopole, he met Professor Gilles Bernot who had known Erol when Gilles was an Assistant at the Laboratoire de Recherche en Informatique, co-founded by Erol at the Université de Paris Sud in Orsay, south of Paris, towards 1978–1979. As we say in Turkish “nereden nereye”—as an expression of surprise about the

fortuitous chains of events... “from where to where”. At the Genopole, Gilles (whose background is in Formal Methods in Computer Science) was heading a group specialised on the formal specification and simulation of Gene Regulatory Networks ... and he handed to Erol some of the early papers on graph models of gene regulatory networks (GRNs). On his return to London, Erol jumped into the subject and came up with the basic model that was published in the previously cited paper [45] and various conferences such as [42]. This initial paper took some time to attract interest from biologists, but it did attract interest in two directions.

It gave rise to an interesting development regarding the use of Erol's GRN model to detect anomalies in genetic data that can help detect or point to predisposition to certain diseases [135–137, 139, 141, 142]. Typically, this line of work involved using Erol's GRN model to represent a set of gene interactions, including some measure of the time scale of these interactions through appropriate time constants, and then estimating the parameters from measured micro-array from known “normal” (i.e. non-disease) data [138, 140], for instance using a learning algorithm similar to some earlier work [36, 71]. Once this phase of model identification is complete, the model can be used for comparison with other micro-array data, to determine whether this other data shows an anomaly or a propensity for some disease such as cancer [134, 143]. Another line of collaboration also emerged with biologists [153] regarding difficult problems of protein interaction networks. Interestingly enough, Erol also investigated how some of the underlying chemistry could be modelled [48].

## 7 How About Analytical Models?

In fact, in all of the areas we have listed, Erol's work has been driven by the use, or invention, of appropriate probability models [101]. However, an interesting development has been Erol's incursion into the literature on Statistical Physics with his theoretical papers on analytical solutions for chemical master equations [47], and the related issue of stochastic modeling in gene regulatory networks [46]. Another multi-disciplinary link he has been able to make links the behaviour of adversarial biological populations (such as germs and cells) to viruses in software or in computer networks [43]. His recent work has also linked particle motion in non-homogenous media [6, 51], where he has studied both the time it takes for a set of particles to attain a target, and the energy that is consumed in the process, which is related to the random motion of packets in a very large multi-hop sensor network [44].

Finally in my list of unusual analytic results, let me mention his work where he considers an (economic) market composed of  $N$  English auctions [49] where customers arrive according to a random process, select some auction and bid for a product with a probability that may depend on the current value that has been attained by that product, leave the marketplace if they are successful in purchasing the product, and may go to some other auction (or may leave the market) if they are unsuccessful. This analysis leads to a closed form expression for the equilibrium prices of all the products in the market.

## References

1. The 24th International Symposium on Computer and Information Sciences, ISCIS 2009, 14–16 September 2009, North Cyprus. IEEE (2009)
2. Abdelbaki, H., Gelenbe, E., El-Khamy, S.E.: Random neural network decoder for error correcting codes. In: International Joint Conference on Neural Networks, 1999. IJCNN'99, vol. 5, pp. 3241–3245. IEEE (1999)
3. Abdelbaki, H., Gelenbe, E., Kocak, T.: Matched neural filters for emi based mine detection. In: International Joint Conference on Neural Networks, 1999. IJCNN'99, vol. 5, pp. 3236–3240. IEEE (1999)
4. Abdelbaki, H., Gelenbe, E., Koçak, T., El-Khamy, S.E.: Random neural network filter for land mine detection. In: Proceedings of the Sixteenth National Radio Science Conference, 1999. NRSC'99, pp. C43–1. IEEE (1999)
5. Abdelrahman, O.H., Gelenbe, E.: Packet delay and energy consumption in non-homogeneous networks. *Comput. J.* **55**(8), 950–964 (2012)
6. Abdelrahman, O.H., Gelenbe, E.: Time and energy in team-based search. *Phys. Rev. E* **87**(3), 032125 (2013)
7. Abdelrahman, O.H., Gelenbe, E.: Signalling storms in 3g mobile networks. In: IEEE International Conference on Communications, ICC 2014, Sydney, Australia, June 10–14, 2014, pp. 1017–1022. IEEE (2014). doi:[10.1109/ICC.2014.6883453](https://doi.org/10.1109/ICC.2014.6883453). <http://dx.doi.org/10.1109/ICC.2014.6883453>
8. Aguilar, J., Gelenbe, E.: Task assignment and transaction clustering heuristics for distributed systems. *Inf. Sci.* **97**(1), 199–219 (1997)
9. Atalay, V., Gelenbe, E.: Parallel algorithm for colour texture generation using the random neural network model. *Int. J. Pattern Recognit. Artif. Intell.* **6**(02n03), 437–446 (1992)
10. Atalay, V., Gelenbe, E., Yalabik, N.: The random neural network model for texture generation. *Int. J. Pattern Recognit. Artif. Intell.* **6**(01), 131–141 (1992)
11. Aykanat, C., Dayar, T., Korpeoglu, I. (eds.): In: Proceedings of Computer and Information Sciences—ISCIS 2004, 19th International Symposium, Kemer-Antalya, Turkey, October 27–29, 2004. Lecture Notes in Computer Science, vol. 3280. Springer (2004)
12. Badel, M., Gelenbe, E., Leroudier, J., Potier, D.: Adaptive optimization of a time-sharing system's performance. *Proc. IEEE* **63**(6), 958–965 (1975)
13. Badel, M., Gelenbe, E., Leroudier, J., Potier, D., Lenfant, J.: Adaptive optimization of the performance of a virtual memory computer. *ACM SIGMETRICS Perform. Eval. Rev.* **3**(4), 188 (1974)
14. Berl, A., Gelenbe, E., Di Girolamo, M., Giuliani, G., De Meer, H., Dang, M.Q., Pentikousis, K.: Energy-efficient cloud computing. *Comput. J.* **53**(7), 1045–1051 (2010)
15. Bi, H., Desmet, A., Gelenbe, E.: Routing emergency evacuees with cognitive packet networks. In: Information Sciences and Systems 2013, pp. 295–303. Springer International Publishing (2013)
16. Bi, H., Gelenbe, E.: Routing diverse evacuees with cognitive packets. [arXiv:1311.4818](https://arxiv.org/abs/1311.4818) (2013)
17. Bi, H., Gelenbe, E.: A cooperative emergency navigation framework using mobile cloud computing. In: Information Sciences and Systems 2014, pp. 41–48. Springer International Publishing (2014)
18. Cramer, C., Gelenbe, E., Bakircioglu, H.: Video compression with random neural networks. In: Proceedings of International Workshop on Neural Networks for Identification, Control, Robotics, and Signal/Image Processing, 1996, pp. 476–484. IEEE (1996)
19. Cramer, C., Gelenbe, E., Bakircioglu, H.: Low bit-rate video compression with neural networks and temporal subsampling. *Proc. IEEE* **84**(10), 1529–1543 (1996)
20. Czachórski, T., Gelenbe, E., Lent, R. (eds.): In: Information Sciences and Systems 2014—Proceedings of the 29th International Symposium on Computer and Information Sciences, ISCIS 2014, Krakow, Poland, October 27–28, 2014. Springer (2014). doi:[10.1007/978-3-319-09465-6](https://doi.org/10.1007/978-3-319-09465-6). <http://dx.doi.org/10.1007/978-3-319-09465-6>



21. Desmet, A., Gelenbe, E.: Graph and analytical models for emergency evacuation. *Future Internet* **5**(1), 46–55 (2013)
22. Desmet, A., Gelenbe, E.: Interoperating infrastructures in emergencies. In: *Computer and Information Sciences III*, pp. 123–130. Springer, London (2013)
23. Dimakis, N., Filippopolitis, A., Gelenbe, E.: Distributed building evacuation simulator for smart emergency management. *Comput. J.* bxq012 (2010)
24. Dobson, S., Denazis, S., Fernández, A., Gañi, D., Gelenbe, E., Massacci, F., Nixon, P., Saffre, F., Schmidt, N., Zambonelli, F.: A survey of autonomic communications. *ACM Trans. Auton. Adapt. Syst. (TAAS)* **1**(2), 223–259 (2006)
25. Filippopolitis, A., Gelenbe, E.: A distributed simulation platform for urban security. In: *IEEE International Conference on Green Computing and Communications (GreenCom)*, 2012, pp. 434–441. IEEE (2012)
26. Filippopolitis, A., Gelenbe, E., Abdelrahman, O.H., Ardito, L., Bulut, E., Desmet, A., Falcarin, P., Gorbil, G., Islam, S., Loukas, G. et al.: Pernem'14: the fourth international workshop on pervasive networks for emergency management, 2014-welcome and committees welcome message from the pernem'14 co-chairs. In: *Proceedings PerCOM 2014* (2014)
27. Filippopolitis, A., Gorbil, G., Gelenbe, E.: Spatial computers for emergency management. In: *Fifth IEEE Conference on Self-Adaptive and Self-Organizing Systems Workshops (SASOW)*, 2011, pp. 61–66. IEEE (2011)
28. Filippopolitis, A., Gorbil, G., Gelenbe, E.: Pervasive emergency support systems for building evacuation. In: *IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, 2012, pp. 525–527. IEEE (2012)
29. Filippopolitis, A., Hey, L., Loukas, G., Gelenbe, E., Timotheou, S.: Emergency response simulation using wireless sensor networks. In: *Proceedings of the 1st International Conference on Ambient Media and Systems*, p. 21. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering) (2008)
30. Gelenbe, E.: Errors and power when communicating with spins. *IEEE Trans. Emerg. Top. Comput*
31. Gelenbe, E.: A unified approach to the evaluation of a class of replacement algorithms. *IEEE Trans. Comput.* **100**(6), 611–618 (1973)
32. Gelenbe, E.: On approximate computer system models. *J. ACM (JACM)* **22**(2), 261–269 (1975)
33. Gelenbe, E.: Probabilistic models of computer systems. *Acta Informatica* **12**(4), 285–303 (1979)
34. Gelenbe, E.: Stability of the random neural network model. *Neural Comput.* **2**(2), 239–247 (1990)
35. Gelenbe, E.: G-nets and learning recurrent random networks. In: *Proceedings of International Conference on Artificial Neural Networks*, Brighton, England (1992)
36. Gelenbe, E.: Learning in the recurrent random neural network. *Neural Comput.* **5**(1), 154–164 (1993)
37. Gelenbe, E.: G-networks: a unifying model for neural and queueing networks. *Ann. Oper. Res.* **48**(5), 433–461 (1994)
38. Gelenbe, E.: The first decade of g-networks. *Eur. J. Oper. Res.* **126**(2), 231–232 (2000)
39. Gelenbe, E. (ed.): *International Symposium on Computer and Information Sciences*. CRC Press, Boca Raton (2002)
40. Gelenbe, E.: Quality of service in ad hoc networks. *Ad Hoc Netw.* **2**(3), 203 (2004)
41. Gelenbe, E.: Users and services in intelligent networks. *IEE Proc. Intell. Transp. Syst.* **153**(3), 213–220 (2006)
42. Gelenbe, E.: Analytical solution of gene regulatory networks. In: *IEEE International Fuzzy Systems Conference, 2007. FUZZ-IEEE 2007*, pp. 1–6. IEEE (2007)
43. Gelenbe, E.: Dealing with software viruses: a biological paradigm. *Inf. Secur. Tech. Rep.* **12**(4), 242–250 (2007)
44. Gelenbe, E.: A diffusion model for packet travel time in a random multi-hop medium. *ACM Trans. Sens. Netw.* **3**(2) (2007)

45. Gelenbe, E.: Steady-state solution of probabilistic gene regulatory networks. *Phys. Rev. E* **76**(3), 031903 (2007)
46. Gelenbe, E.: Steady-state solution of probabilistic gene regulatory networks. *Phys. Rev. E* **76**(1), 031903 (2007)
47. Gelenbe, E.: Network of interacting synthetic molecules in equilibrium. *Proc. Roy. Soc. A* **464**, 22192228 (2008)
48. Gelenbe, E.: Network of interacting synthetic molecules in steady state. *Proc. R. Soc. A: Math. Phys. Eng. Sci.* **464**(2096), 2219–2228 (2008)
49. Gelenbe, E.: Analysis of single and networked auctions. *ACM Trans. Internet Technol. (TOIT)* **9**(2), 8 (2009)
50. Gelenbe, E.: Steps toward self-aware networks. *Commun. ACM* **52**(7), 66–75 (2009)
51. Gelenbe, E.: Search in unknown random environments. *Phys. Rev. E* **82**, 061112 (2010)
52. Gelenbe, E.: Ubiquity symposium: natural computation. *Ubiquity* **2011**, 1 (2011)
53. Gelenbe, E.: Energy packet networks: adaptive energy management for the cloud. In: *Proceedings of the 2nd International Workshop on Cloud Computing Platforms*, p. 1. ACM (2012)
54. Gelenbe, E.: Energy packet networks: Ict based energy allocation and storage. In: *Green Communications and Networking*, pp. 186–195. Springer, Berlin (2012)
55. Gelenbe, E.: Energy packet networks: smart electricity storage to meet surges in demand. In: *Proceedings of the 5th International ICST Conference on Simulation Tools and Techniques*, pp. 1–7. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering) (2012)
56. Gelenbe, E.: Adaptive management of energy packets. In: *IEEE 38th International Computer Software and Applications Conference Workshops (COMPSACW)*, 2014, pp. 1–6. IEEE (2014)
57. Gelenbe, E.: Error and energy when communicating with spins. In: *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2014, pp. 784–787. IEEE (2014). doi:[10.1109/GlobalSIP.2014.7032226](https://doi.org/10.1109/GlobalSIP.2014.7032226)
58. Gelenbe, E.: A sensor node with energy harvesting. *ACM SIGMETRICS Perform. Eval. Rev.* **42**(2), 37–39 (2014)
59. Gelenbe, E.: Synchronising energy harvesting and data packets in a wireless sensor. *Energies* **8**(1), 356–369 (2015). doi:[10.3390/en8010356](https://doi.org/10.3390/en8010356)
60. Gelenbe, E., Bi, H.: Emergency navigation without an infrastructure. *Sensors* **14**(8), 15142–15162 (2014)
61. Gelenbe, E., Cao, Y.: Autonomous search for mines. *Eur. J. Oper. Res.* **108**(2), 319–333 (1998)
62. Gelenbe, E., Ceran, E.: Central or distributed energy storage for processors with energy harvesting. In: *The Fourth International Conference on Sustainable Internet and ICT for Sustainability*. IEEE (2015)
63. Gelenbe, E., Feng, Y., Krishnan, K.R.R.: Neural network methods for volumetric magnetic resonance imaging of the human brain. *Proc. IEEE* **84**(10), 1488–1496 (1996)
64. Gelenbe, E., Feng, Y., Ranga, K., Krishnan, R.: Neural networks for volumetric mr imaging of the brain. In: *International Workshop on Neural Networks for Identification, Control, Robotics, and Signal/Image Processing*, 1996. pp. 194–202. IEEE (1996)
65. Gelenbe, E., Gellman, M., Loukas, G.: An autonomic approach to denial of service defence. In: *Sixth IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks*, 2005. *WoWMoM 2005*, pp. 537–541. IEEE (2005)
66. Gelenbe, E., Gellman, M., Su, P.: Self-awareness and adaptivity for quality of service. In: *Proceedings of Eighth IEEE International Symposium on Computers and Communication*, 2003. (ISCC 2003), pp. 3–9. IEEE (2003)
67. Gelenbe, E., Ghanwani, A., Srinivasan, V.: Improved neural heuristics for multicast routing. *IEEE J. Sel. Areas Commun.* **15**(2), 147–155 (1997)
68. Gelenbe, E., Gorbil, G., Wu, F.J.: Emergency cyber-physical-human systems. In: *21st International Conference on Computer Communications and Networks (ICCCN)*, 2012, pp. 1–7. IEEE (2012)

69. Gelenbe, E., Gunduz, D.: Optimum power level for communications with interference. In: 24th Thyrrenian International Workshop on Digital Communications–Green ICT (TIWDC), 2013, pp. 1–6. IEEE (2013)
70. Gelenbe, E., Hussain, K., Kaptan, V.: Simulating autonomous agents in augmented reality. *J. Syst. Softw.* **74**(3), 255–268 (2005)
71. Gelenbe, E., Hussain, K.F.: Learning in the multiple class random neural network. *IEEE Trans. Neural Netw.* **13**(6), 1257–1267 (2002)
72. Gelenbe, E., Kaptan, V., Wang, Y.: Biological metaphors for agent behavior. In: *Computer and Information Sciences-ISCIS 2004*, pp. 667–675. Springer, Berlin (2004)
73. Gelenbe, E., Kazhmaganbetova, Z.: Cognitive packet network for bilateral asymmetric connections. *IEEE Trans. Industr. Inf.* **10**(3), 1717–1725 (2014). doi:[10.1109/TII.2014.2321740](https://doi.org/10.1109/TII.2014.2321740). <http://dx.doi.org/10.1109/TII.2014.2321740>
74. Gelenbe, E., Koçak, T.: Area-based results for mine detection. *IEEE Trans. Geosci. Remote Sens.* **38**(1), 12–24 (2000)
75. Gelenbe, E., Koçak, T., Wang, R.: Wafer surface reconstruction from top-down scanning electron microscope images. *Microelectron. Eng.* **75**(2), 216–233 (2004)
76. Gelenbe, E., Koubi, V., Pekergin, F.: Dynamical random neural network approach to the traveling salesman problem. In: *International Conference on Systems, Man and Cybernetics, 1993. Systems Engineering in the Service of Humans, Conference Proceedings*, pp. 630–635. IEEE (1993)
77. Gelenbe, E., Kurinckx, A.: Random injection control of multiprogramming in virtual memory. *IEEE Trans. SE Softw. Eng.* **4**(1), 2–17 (1978)
78. Gelenbe, E., Labed, A.: Esprit Itr project 8144 lydia load balancing and g-networks: design, implementation and evaluation. Technical Report, IHEI, University of René Descartes, Paris V (1996)
79. Gelenbe, E., Lent, R.: Mobile ad-hoc cognitive packet networks. In: *Proceedings of IEEE ASWN*, pp. 2–4 (2002)
80. Gelenbe, E., Lent, R.: Power-aware ad hoc cognitive packet networks. *Ad Hoc Netw.* **2**(3), 205–216 (2004)
81. Gelenbe, E., Lent, R.: Optimising server energy consumption and response time. *Theor. Appl. Informatics* **24**(4), 257–270 (2012)
82. Gelenbe, E., Lent, R.: Trade-offs between energy and quality of service. In: *Sustainable Internet and ICT for Sustainability (SustainIT), 2012*, pp. 1–5. IEEE (2012)
83. Gelenbe, E., Lent, R. (eds.): In: *Computer and Information Sciences III—27th International Symposium on Computer and Information Sciences, Paris, France, October 3–4, 2012*. Springer (2013). doi:[10.1007/978-1-4471-4594-3](https://doi.org/10.1007/978-1-4471-4594-3). <http://dx.doi.org/10.1007/978-1-4471-4594-3>
84. Gelenbe, E., Lent, R.: Energy-qos trade-offs in mobile service selection. *Future Internet* **5**(2), 128–139 (2013)
85. Gelenbe, E., Lent, R. (eds.): In: *Information Sciences and Systems 2013—Proceedings of the 28th International Symposium on Computer and Information Sciences, ISCIS 2013, Paris, France, October 28–29, 2013, Lecture Notes in Electrical Engineering*, vol. 264. Springer (2013). doi:[10.1007/978-3-319-01604-7](https://doi.org/10.1007/978-3-319-01604-7). <http://dx.doi.org/10.1007/978-3-319-01604-7>
86. Gelenbe, E., Lent, R., Douratsos, M.: Choosing a local or remote cloud. In: *Second Symposium on Network Cloud Computing and Applications (NCCA), 2012*, pp. 25–30. IEEE (2012)
87. Gelenbe, E., Lent, R., Montuori, A., Xu, Z.: Cognitive packet networks: Qos and performance. In: *10th IEEE International Symposium on Modeling, Analysis and Simulation of Computer and Telecommunications Systems, 2002. MASCOTS 2002. Proceedings*, pp. 3–9. IEEE (2002)
88. Gelenbe, E., Lent, R., Nunez, A.: Self-aware networks and qos. *Proc. IEEE* **92**(9), 1478–1489 (2004)
89. Gelenbe, E., Lent, R., Sakellari, G. (eds.): *Computer and Information Sciences II—26th International Symposium on Computer and Information Sciences, London, UK, 26–28 September 2011*. Springer (2011). doi:[10.1007/978-1-4471-2155-8](https://doi.org/10.1007/978-1-4471-2155-8). <http://dx.doi.org/10.1007/978-1-4471-2155-8>

90. Gelenbe, E., Lent, R., Sakellari, G., Sacan, A., Toroslu, I.H., Yazici, A. (eds.): In: *Computer and Information Sciences—Proceedings of the 25th International Symposium on Computer and Information Sciences*, London, UK, September 22–24, 2010, *Lecture Notes in Electrical Engineering*, vol. 62. Springer (2010). doi:10.1007/978-90-481-9794-1. <http://dx.doi.org/10.1007/978-90-481-9794-1>
91. Gelenbe, E., Lent, R., Xu, Z.: Design and performance of cognitive packet networks. *Perform. Eval.* **46**(2), 155–176 (2001)
92. Gelenbe, E., Lent, R., Xu, Z.: Measurement and performance of a cognitive packet network. *Computer Networks* **37**(6), 691–701 (2001)
93. Gelenbe, E., Lent, R., Xu, Z.: Towards networks with cognitive packets. In: *Performance and QoS of next generation networking*, pp. 3–17. Springer, London (2001)
94. Gelenbe, E., Liu, P.: Qos and routing in the cognitive packet network. In: *Sixth IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks*, 2005. *WoW-MoM 2005*, pp. 517–521. IEEE (2005)
95. Gelenbe, E., Loukas, G.: A self-aware approach to denial of service defence. *Comput. Netw.* **51**(5), 1299–1314 (2007)
96. Gelenbe, E., Mahmoodi, T.: Energy-aware routing in the cognitive packet network. In: *ENERGY 2011, The First International Conference on Smart Grids, Green Communications and IT Energy-aware Technologies*, pp. 7–12 (2011)
97. Gelenbe, E., Mahmoodi, T.: Distributed energy-aware routing protocol. In: *Computer and Information Sciences II*, pp. 149–154. Springer, London (2012)
98. Gelenbe, E., Mahmoodi, T., Morfopoulou, C.: Energy aware routing in packet networks. *E-Energy* (2010)
99. Gelenbe, E., Mao, Z.W., Li, Y.D.: Function approximation with spiked random networks. *IEEE Trans. Neural Netw.* **10**(1), 3–9 (1999)
100. Gelenbe, E., Marin, A.: Interconnected wireless sensors with energy harvesting. In: *Proceedings of Analytical and Stochastic Modelling Techniques and Applications*, ASMTA 2015. Springer International Publishing (2015)
101. Gelenbe, E., Mitrani, I.: *Analysis and Synthesis of Computer Systems*. Imperial College Press, World Scientific (2010)
102. Gelenbe, E., Morfopoulou, C.: A framework for energy-aware routing in packet networks. *Comput. J.* **54**(6), 850–859 (2011)
103. Gelenbe, E., Morfopoulou, C.: Routing and g-networks to optimise energy and quality of service in packet networks. In: *Energy-Efficient Computing and Networking*, pp. 163–173. Springer, Berlin (2011)
104. Gelenbe, E., Muntz, R.R.: Probabilistic models of computer systems part i (exact results). *Acta Informatica* **7**(1), 35–60 (1976)
105. Gelenbe, E., Ngai, E.C.: Adaptive qos routing for significant events in wireless sensor networks. In: *5th IEEE International Conference on Mobile Ad Hoc and Sensor Systems*, 2008. *MASS 2008*, pp. 410–415. IEEE (2008)
106. Gelenbe, E., Ngai, E.C.: Adaptive random re-routing in sensor networks. In: *Proceedings of the Annual Conference of ITA (ACITA08) September 16–18*, 348–349 (2008)
107. Gelenbe, E., Ngai, E.C., Yadav, P.: Routing of high-priority packets in wireless sensor networks. In: *IEEE Second International Conference on Computer and Network Technology*, IEEE (2010)
108. Gelenbe, E., Núñez, A.: Self-aware networks and quality of service. In: *Artificial Neural Networks and Neural Information Processing ICANN/ICONIP 2003*, pp. 901–908. Springer, Berlin (2003)
109. Gelenbe, E., Oklander, B.: Cognitive users with useful vacations. In: *IEEE International Conference on Communications Workshops (ICC)*, 2013, pp. 370–374. IEEE (2013)
110. Gelenbe, E., Potier, D., Brandwajn, A., Lenfant, J.: Gestion optimale d'un ordinateur multi-programmé à mémoire virtuelle. In: *5th Conference on Optimization Techniques Part II*, pp. 132–143. Springer, Berlin (1973)

111. Gelenbe, E., Sakellari, G., D'ariento, M.: Admission of qos aware users in a smart network. *ACM Trans. Auton. Adapt. Syst. (TAAS)* **3**(1), 4 (2008)
112. Gelenbe, E., Sakellari, G., Filippoupolitis, A.: Pernem 2012: Second international workshop on pervasive networks for emergency management 2012, committees and welcome. In: *Proceedings PerCOM 2012* (2012)
113. Gelenbe, E., Sakellari, G., Filippoupolitis, A.: Pernem 2013: Third international workshop on pervasive networks for emergency management 2013-committees and welcome. In: *Proceedings PerCOM 2013* (2013)
114. Gelenbe, E., Seref, E., Xu, Z.: Simulation with learning agents. *Proc. IEEE* **89**(2), 148–157 (2001)
115. Gelenbe, E., Silvestri, S.: Optimisation of power consumption in wired packet networks. In: *Quality of Service in Heterogeneous Networks*, pp. 717–729. Springer, Berlin (2009)
116. Gelenbe, E., Silvestri, S.: Reducing power consumption in wired networks. In: *Computer and Information Sciences, 2009. ISCIS 2009*, pp. 292–297. IEEE (2009)
117. Gelenbe, E., Stafylopatis, A.: Global behavior of homogeneous random neural systems. *Appl. Math. Model.* **15**(10), 534–541 (1991)
118. Gelenbe, E., Sungur, M., Cramer, C., Gelenbe, P.: Traffic and video quality with adaptive neural compression. *Multimedia Syst.* **4**(6), 357–369 (1996)
119. Gelenbe, E., Timotheou, S.: Random neural networks with synchronized interactions. *Neural Comput.* **20**(9), 2308–2324 (2008)
120. Gelenbe, E., Timotheou, S.: Synchronized interactions in spiked neuronal networks. *Comput. J.* **51**(6), 723–730 (2008)
121. Gelenbe, E., Wang, Y.: A mathematical approach for mission planning and rehearsal. In: *Defense and Security Symposium*, pp. 62,490Q–62,490Q. International Society for Optics and Photonics (2006)
122. Gelenbe, E., Wang, Y.: Modelling large scale autonomous systems. In: *9th International Conference on Information Fusion, 2006*, pp. 1–7. IEEE (2006)
123. Gelenbe, E., Wu, F.J.: Distributed networked emergency evacuation and rescue. In: *IEEE International Conference on Communications (ICC), 2012*, pp. 6334–6338. IEEE (2012)
124. Gelenbe, E., Wu, F.J.: Large scale simulation for human evacuation and rescue. *Comput. Math. Appl.* **64**(12), 3869–3880 (2012)
125. Gelenbe, E., Wu, F.J.: Sensors in cyber-physical emergency systems. In: *IET Conference on Wireless Sensor Systems (WSS 2012)*. IEEE (2012). doi:[10.1049/cp.2012.0578](https://doi.org/10.1049/cp.2012.0578)
126. Gelenbe, E., Wu, F.J.: Future research on cyber-physical emergency management systems. *Future Internet* **5**(3), 336–354 (2013)
127. Gelenbe, E., Xu, Z., Seref, E.: Cognitive packet networks. In: *Proceedings of 11th IEEE International Conference on Tools with Artificial Intelligence, 1999*, pp. 47–54. IEEE (1999)
128. Görbil, G., Abdelrahman, O.H., Gelenbe, E.: Storms in mobile networks. In: P. Mueller, L. Foschini, R. Yu (eds.) *Q2SWinet'14, Proceedings of the 10th ACM Symposium on QoS and Security for Wireless and Mobile Networks*, Montreal, QC, Canada, September 21–22, 2014, pp. 119–126. ACM (2014). doi:[10.1145/2642687.2642688](https://doi.org/10.1145/2642687.2642688). <http://doi.acm.org/10.1145/2642687.2642688>
129. Gorbil, G., Abdelrahman, O.H., Pavloski, M., Gelenbe, E.: Storms in mobile networks. [arXiv:1411.1280](https://arxiv.org/abs/1411.1280) (2014)
130. Gorbil, G., Filippoupolitis, A., Gelenbe, E.: Intelligent navigation systems for building evacuation. In: *Computer and Information Sciences II*, pp. 339–345. Springer, London (2012)
131. Görbil, G., Gelenbe, E.: Design of a mobile agent-based adaptive communication middleware for federations of critical infrastructure simulations. In: *Critical Information Infrastructures Security*, pp. 34–49. Springer, Berlin (2010)
132. Gorbil, G., Gelenbe, E.: Opportunistic communications for emergency support systems. *Procedia Comput. Sci.* **5**, 39–47 (2011)
133. Hocaoglu, A.K., Gader, P.D., Gelenbe, E., Kocak, T.: Optimal linear combination of order statistics filters and their relationship to the delta-operator. In: *AeroSense'99*, pp. 1323–1329. International Society for Optics and Photonics (1999)

134. Kim, H., Atalay, R., Gelenbe, E.: G-network modelling based abnormal pathway detection in gene regulatory networks. In: *Computer and Information Sciences: 26th International Symposium on Computer and Information Sciences*, p. 257. Springer (2011)
135. Kim, H., Atalay, R., Gelenbe, E.: G-network modelling based abnormal pathway detection in gene regulatory networks. In: *Computer and Information Sciences II*, pp. 257–263. Springer, London (2012)
136. Kim, H., Gelenbe, E.: Stochastic gene expression model base gene regulatory networks. In: *EKC 2009 Proceedings of the EU-Korea Conference on Science and Technology*, pp. 235–244. Springer, Berlin (2010)
137. Kim, H., Gelenbe, E.: Reconstruction of large-scale gene regulatory networks using bayesian model averaging. In: *IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, 2011, pp. 202–207. IEEE (2011)
138. Kim, H., Gelenbe, E.: Reconstruction of large-scale gene regulatory networks using bayesian model averaging. *IEEE Trans. NanoBiosci.* **11**(3), 259–265 (2012). doi:[10.1109/TNB.2012.2214233](https://doi.org/10.1109/TNB.2012.2214233)
139. Kim, H., Gelenbe, E.: Stochastic gene expression modeling with hill function for switch-like gene responses. *IEEE/ACM Trans. Comput. Biol. Bioinf.* **9**(4), 973–979 (2012)
140. Kim, H., Gelenbe, E.: Stochastic gene expression modeling with hill function for switch-like gene responses. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **9**(4), 973–979 (2012). doi:[10.1109/TCBB.2011.153](https://doi.org/10.1109/TCBB.2011.153)
141. Kim, H., Gelenbe, E., et al.: G-networks based two layer stochastic modeling of gene regulatory networks with post-translational processes. *Interdisc. Bio Central* **3**(1), 8 (2011)
142. Kim, H., Park, T., Gelenbe, E.: Identifying disease candidate genes via large-scale gene network analysis. *Int. J. Data Min. Bioinf.* **10**(2), 175–188 (2014)
143. Kim, H., Park, T., Gelenbe, E.: Identifying disease candidate genes via large-scale gene network analysis. *Int. J. Data Min. Bioinf.* **10**(2), 175–188 (2014). doi:[10.1504/IJDMB.2014.064014](https://doi.org/10.1504/IJDMB.2014.064014)
144. Kokuti, A., Gelenbe, E.: Directional navigation improves opportunistic communication for emergencies. *Sensors* **14**(8), 15387–15399 (2014)
145. Lent, R., Abdelrahman, O.H., Gorbil, G., Gelenbe, E.: Fast message dissemination for emergency communications. In: *8th IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM Workshops)*, 2010, pp. 370–375. IEEE (2010)
146. Levi, A., Savas, E., Yenigün, H., Balcisoy, S., Saygin, Y. (eds.): In: *Computer and Information Sciences - ISCIS 2006, 21th International Symposium, Istanbul, Turkey, November 1–3, 2006, Proceedings, Lecture Notes in Computer Science*, vol. 4263. Springer (2006)
147. Liu, P., Gelenbe, E.: Recursive routing in the cognitive packet network. In: *3rd International Conference on Testbeds and Research Infrastructure for the Development of Networks and Communities*, 2007. TridentCom 2007, pp. 1–6. IEEE (2007)
148. Ngai, E.C., Gelenbe, E., Humber, G.: Information-aware traffic reduction for wireless sensor networks. In: *IEEE 34th Conference on Local Computer Networks*, 2009. LCN 2009, pp. 451–458. IEEE (2009)
149. Oke, G., Loukas, G., Gelenbe, E.: Detecting denial of service attacks with bayesian classifiers and the random neural network. In: *IEEE International Fuzzy Systems Conference*, 2007. FUZZ-IEEE 2007, pp. 1–6. IEEE (2007)
150. Oklander, B., Gelenbe, E.: Optimal behaviour of smart wireless users. In: *Information Sciences and Systems 2013*, pp. 87–95. Springer International Publishing (2013)
151. Ören, T.I., Numrich, S.K., Uhrmacher, A.M., Wilson, L.F., Gelenbe, E.: Agent-directed simulation: challenges to meet defense and civilian requirements. In: *Proceedings of the 32nd conference on Winter simulation*, pp. 1757–1762. Society for Computer Simulation International (2000)
152. Pernici, B., Aiello, M., vom Brocke, J., Donnellan, B., Gelenbe, E., Kretsis, M.: What is can do for environmental sustainability: a report from caise2011 panel on green and sustainable is. *Commun. Assoc. Inf. Syst.* **30**(1), 18 (2012)

153. Phan, H.T., Stemberg, M.J., Gelenbe, E.: Aligning protein-protein interaction networks using random neural networks. In: IEEE International Conference on Bioinformatics and Biomedicine (BIBM), 2012, pp. 1–6. IEEE (2012)
154. Potier, D., Gelenbe, E., Lenfant, J.: Adaptive allocation of central processing unit quanta. *J. ACM (JACM)* **23**(1), 97–102 (1976)
155. Sakellari, G., Gelenbe, E.: Adaptive resilience of the cognitive packet network in the presence of network worms. In: Proceedings of the NATO Symposium on C3I for Crisis, Emergency and Consequence Management, pp. 11–12 (2009)
156. Sakellari, G., Hey, L., Gelenbe, E.: Adaptability and failure resilience of the cognitive packet network. In: DemoSession of the 27th IEEE Conference on Computer Communications (INFOCOM2008), Phoenix, Arizona, USA (2008)
157. Sakellari, G., Morfopoulou, C., Mahmoodi, T., Gelenbe, E.: Using energy criteria to admit flows in a wired network. In: Computer and Information Sciences III, pp. 63–72. Springer, London (2013)
158. Wang, L., Gelenbe, E.: An implementation of voice over ip in the cognitive packet network. In: Information Sciences and Systems 2014, pp. 33–40. Springer International Publishing (2014)
159. Witkowski, M., White, G., Louvieris, P., Gorbil, G., Gelenbe, E., Dodd, L.: High-level information fusion and mission planning in highly anisotropic threat spaces. In: 11th International Conference on Information Fusion, 2008, pp. 1–8. IEEE (2008)
160. Yazici, A., Sener, C. (eds.): In: Computer and information sciences—ISCIS 2003, 18th International Symposium, Antalya, Turkey, November 3–5, 2003, Proceedings, Lecture Notes in Computer Science, vol. 2869. Springer (2003)
161. Yolum, P., Güngör, T., Gürgeç, F.S., Özturan, C.C. (eds.): Computer and Information Sciences—ISCIS 2005, 20th International Symposium, Istanbul, Turkey, October 26–28, 2005, Proceedings, Lecture Notes in Computer Science, vol. 3733. Springer (2005)
162. Yu, C.M., Ni, G.K., Chen, I.Y., Gelenbe, E., Kuo, S.Y.: Top-k query result completeness verification in tiered sensor networks. *IEEE Trans. Inf. Forensics Secur.* **9**(1), 109–124 (2014)
163. Zhu, Q., Gelenbe, E., Qiao, Y.: Adaptive prefetching algorithm in disk controllers. *Perform. Eval.* **65**(5), 382–395 (2008)

# Interoperability and Semantics: A Never-Ending Story

Erich Neuhold and Elmar Kiesling

**Abstract** Sending and receiving data is an important requirement of all computing infrastructures with components that need to interact. This interoperability requirement has been the focus of research since the beginning of the computer age. In this paper we will illustrate that in order to achieve interoperability, computer hardware and software have to be able to interpret data, i.e., understand their structure and their meaning. That understanding has to be conveyed somehow from the real world by the creators of the computing components and brought into a machine-processable form. In most cases, that means that formal models of those real world aspects have to be created and transferred into the computing environment. Here we will show from a data-centric point of view how those models developed over time from early relational database models, via document models and the (Semantic) Web to Linked Data and the Internet of Things.

## 1 Introduction

Over the last decades, the complexity of information systems has increased considerably. These systems have grown tremendously from single (large) computers to the virtually unlimited network hosting the World Wide Web. The need to exchange data whenever multiple system components interact has existed from the early days of computing and motivated interoperability research for decades. Data interoperability is required to facilitate communication (networks), to interface with the real world (sensors and actuators) or to enable cooperation between (application) processes. In all cases, the meaning, i.e., semantics of those data has to be interpreted by humans as system developers or programmers and as a consequence made processable by

---

E. Neuhold  
University of Vienna, Wien, Austria  
e-mail: erich.neuhold@univie.ac.at

E. Kiesling (✉)  
Vienna University of Technology, Wien, Austria  
e-mail: elmar.kiesling@tuwien.ac.at



machines. When meaning is attached to data, they become information objects or entities and we will use those terms to distinguish between raw and interpreted data. In this paper, we focus on interoperability and integration issues that arise in application systems rather than problems to be handled in the network layers of complex systems. Hence, networks, network standards, network protocols and optimization issues remain outside the scope of this paper.

The size of complex systems makes it mandatory that the understanding of data can be derived (semi-) automatically from the environment and the context in which those systems are operating. The modeling of semantics therefore has played a fundamental role in all interoperability and data integration efforts since the beginning of the computer age.

In this paper, we trace the progress that has been made in coping with these problems and argue that few issues have been resolved definitely. As new interoperability challenges arise constantly, much remains to be done. The current discussions of Linked Open Data and Big Data are just intermediary steps with more to come.

Heterogeneities in data and processes and the problem of ‘semantics’ have been investigated in the context of interoperability and integration in various domains (e.g., [1–4]). These heterogeneities have roughly been categorized into (i) technical heterogeneities, (ii) structural and syntactic heterogeneities, and (iii) semantic heterogeneities. Those heterogeneities give rise to three parallel types of interoperability and integration challenges: technical, structural and syntactic, and semantic. In the following, when we refer to interoperability, we shall refer to the latter two notions. Because of space restrictions, we will concentrate on the ‘data’ aspects of interoperability and integration and will mostly disregard related research and solution approaches developed within the Artificial Intelligence, Data/Information Mining and Service Science communities.

The remainder of this paper is structured as follows. In Sect. 2, we will start with early approaches to represent the meaning of data that were developed to exchange data between application programs, to access data from different sources, and to integrate them into conceptually consistent aggregates. In Sect. 3, we discuss semi-structured and document models, based on the notion that large amounts of data are textual with embedded structural elements. In Sect. 4, we analyze in some detail the largest document and data store in existence—the Web and the Semantic Web as its machine-interpretable extension. In Sect. 5, we move to interoperability and semantic interpretation problems of the Linked Data world. We conclude in Sect. 6 with a summary and some ideas on what issues future research in interoperability and integration may have to tackle.

## 2 Database Models Then and Now

It has been nearly 50 years since data interoperability problems first arose, originally in the context of concurrent access to data. As it became necessary for multiple programs to access and process the same data, it became evident that this could

not be solved satisfactorily with the then widely used file systems and file system descriptions. Hence, more powerful tools for pulling all those files together and integrating them in some fashion would be needed. To this end, various organizations and corporations developed what was to become known as (model-driven) database management systems. Early such models were the *CODASYL standard*, that was developed by the database task group at CODASYL and published in 1971 (cf. Taylor and Frank [5]), and the *relational model* published by Codd in 1970 [6]. Arguably due to its structural clarity and relative simplicity, the relational model found widespread adoption in early database systems and today still remains the most widely used model for structured data.

## 2.1 Relational Model

With its strong mathematical foundation, its table representations, and with SQL as a simple query language, the relational model had a seminal influence on data integration and application interoperability research and practice. Through the concept of normal forms, the meaning of data as a model of the real world can be expressed simply and explicitly. The first normal form still allows for redundancy and interdependency, which may cause anomalies. These problems can be reduced by functional dependency rules and the foreign key concept of the third normal form.

Around the time the relational model established itself as *the* dominant database paradigm, the rise of networks and communication systems fueled the need for integrated use of multiple remote databases in a single application. As a consequence, research turned towards *distributed databases* and mechanisms to handle those multiple databases. In particular, two approaches towards distributed database design emerged: (i) *homogeneous modeling* was grounded in a top-down design paradigm starting from a single (global) schema that got split into multiple sub-databases; (ii) *heterogeneous modeling*, by contrast, followed a bottom-up design approach starting from multiple separate databases to form a homogenized global schema.

In the first case, interoperability of the constituent databases was assured. Prototype homogeneous systems that were developed include Distributed INGRES [7] and POREL [8]. In the second case, the problem of interoperability required a more general approach and explicit mappings from the individual databases to the global schema. Research prototypes of heterogeneous distributed databases include MULTIBASE [9] and SIRIUS-DELTA [10].

Over time, the homogeneous approach became a feature of all major database management system products, whereas the heterogeneous approach remains difficult to handle. The main problem lies in the scarcity of semantics in the relational model, which is not expressive enough to describe the various mappings from local to global scheme and then on to the applications.

## 2.2 Entity-Relationship Model

As the limitations of the relational model in terms of semantic expressiveness became increasingly apparent, programmers and system developers required a more explicit language to express the meaning of data. Such a conceptual modeling language became available in 1975 when Peter Chen published the Entity-Relationship Model [11], which is based on the concept of modeling entities, attributes and relationships between them. Over the years, a conference dedicated to the ER-model has explored the model and expanded it with additional concepts such as *is-a* inheritance, which was added in 1985.

The ER-model, together with the Object-Role Model [12], formed one of the foundations of what later became the Unified Modeling Language (UML), which is today widely used in software and system design (see the 2090 ISO standard 19505).

## 2.3 The Object-Oriented Data Model

In the 1980ies, object-oriented programming (Smalltalk, C++, Java) came to the center of software engineering research. The object-oriented paradigm conceives the real world as a collection of objects, each described in terms of their properties and associated behavior. The goal was to provide a ‘natural’ way to model meaning, i.e., to represent semantics as it is to be processed by computers. With respect to data, the ER-model was considered a sound basis for object-oriented abstractions. In addition, the data was extended by behavior to the entities, attributes, and relationships of the model. Other important concepts that were introduced by the object-oriented model include object identity (today typically implemented by means of URIs), type-extensibility, multi-instance objects, and property inheritance.

The need for persistence of such objects resulted in what became known as *object-oriented databases* (OODBs). A number of commercial products were developed based on this paradigm, but the expected success did not materialize and “pure” OODBs remain a niche product until today. However, all of the major database products today incorporate selected object-oriented aspects.

## 3 Semantics in Semi-structured and Document Models

Databases provide persistent storage for well-structured and well-understood data. Driven by developments such as the rise of the web less well-structured data such as documents, gained in importance. Interoperability problems on the semantic level spun off an array of different approaches to represent semi-structured data of all kinds in machine-processable form, including textual, visual, and other forms of media. To this end, appropriate modeling concepts had to be developed.

One such approach arose from the need of the publishing industry to use new means for electronic document exchange, which required standardized computer processable information about documents and their structure. This *meta-information* mostly reflected structural information through formatting tags, indexing tags, author and publisher information etc. The need to represent such information led—in various steps—to the standardization of the Standard General Markup Language (SGML) as ISO 8879 in 1986. In effect, SGML is not a markup language per se, but a meta-language for describing mark-up languages. To this end, it introduced document type definitions to declare the syntax for the mark-up and the allowed structural and textual elements. Despite initial success, SGML's complexity soon led to the development of HTML (Hypertext Markup Language) as a more specialized, presentation-oriented mark-up language [13]. Derived from SGML, it allowed users to easily create structured electronic documents and, with the addition of the Universal Resource Locator (URL), the distribution of the documents via the rapidly developing communication networks. Those networks, characterized by the 7-layer Open System Interconnection (OSI) Model standardized in 1984 as OSI 7498, introduced a Hypertext Transfer Protocol (HTTP) in the application (7th) layer. It could be used to distribute and access HTML documents all over the network—the World Wide Web was born.

However, it soon turned out that despite the success of the web and the flexibility of the HTTP/HTML concept, the representation of meta-information was not defined well enough to allow for automated processing of the information by machines. As a consequence, a stricter approach was developed in the form of XML (Extensible Modeling Language) and XML Schema as the corresponding meta-language for describing the allowed concepts for an XML model. XML is a somewhat simplified derivative of SGML/DDT and became widely used. However, it did not replace HTML in the web context, but rather enhanced it.

The various models that have been developed in the database, information, and document systems communities have facilitated persistent storage and exchange of widely distributed data. However, when processing or integrating those data, the lack of semantic information as part of the meta-information increasingly became a problem.

As the web grew in size and complexity, the necessity to integrate the wealth of available disparate information increased. In this context, it became apparent that this was not just a structural or technical problem. XML allows users to impose arbitrary structure on their documents, but it does not define the meaning of this structure. This flexible approach has many benefits and has arguably spurred the growth of the early web, but it largely precluded automatic integration and interoperation. In order to satisfy information needs, web users need to search for relevant web documents, access them with their browser, and integrate the information (often implicitly) in a tedious manual process. This unsatisfactory situation gave rise to the vision of a machine-interpretable semantic web “where software agents roaming from page to page can readily carry out sophisticated tasks for users” [14].

## 4 Interoperation and the Semantic Web

With the advent of the Web, the focus of interoperability research shifted from integrating well-structured databases to also integrating semi- and unstructured data in large-scale distributed document systems. This shift was associated with new conceptual challenges. Distributed databases could, at least in principle, be integrated by mapping local schemas to a single global schema. On the scale of the web, agreeing on such a centralized global schema—which requires users to subscribe to a single conceptualization of the world—is neither feasible nor desirable. Furthermore, a database schema would not provide sufficient explicit semantics to interpret data consistently and unambiguously [2]. Object-oriented integration approaches can support structural homogenization through abstraction, classification, and taxonomies, but they do not resolve the central issue of large-scale semantic integration of distributed content.

To tackle these issues and release the value locked in relational databases [15] Berners Lee et al. proposed the idea of a semantic web that should emerge from “a new form of Web content that is meaningful to computers” [14]. The semantic web would be decentralized, i.e., it would not rely on a single semantic model of the world, but rather allow users to define and use their own context-specific conceptualizations. For the sake of scalability, flexibility and versatility, the choice of this decentralized paradigm accepted that the different conceptualizations used cannot always be matched completely and that inconsistencies will be introduced. The vision was that once semantics become encoded into web pages, autonomous programs called agents would be able to interpret semantic documents, use rules to make inferences, exchange results with other programs, choose courses of action, and answer users’ questions [14].

To accomplish this ambitious goal, a stack of semantic languages was developed to express data and rules for reasoning about the data. RDF (Resource Description Framework, originally proposed in 1997) constitutes the basis of this stack on top of the XML structural description. It is influenced by knowledge representation formalisms and designed as an infrastructure that enables the encoding, exchange, and reuse of structured metadata [16].

RDF encodes meaning in sets of subject-verb-object expressions called triples, which are used to make statements about resources, in a similar vein as earlier conceptual approaches, such as E-R or object-oriented modeling, make statements about entities or objects. A defining characteristic of RDF, however, is its use of URIs (Uniform Resource Identifier) or IRI (Internationalized Resource Identifier) as identifiers to refer to resources. Whereas identifiers in a typical database have no meaning outside the database, URIs/IRIs are globally unique and referable. Associating a URI with a resource means that anyone can link to it, refer to it, or retrieve a representation of it [15]. Multiple data sets can refer to the same identifier, which facilitates data integration. According to the semantic web vision, URI/IRIs should also ensure that concepts are tied to a unique definition that everyone can find on the Web, rather than being just words in a document [14]. In particular, the subject and the predicate of

an RDF statement must always be resources identified by a URI (or IRI); the object can either be a resource or a literal node (cf. the example in Fig. 5, where subjects are represented as ellipses, predicates as arrows, and literals as boxes).

A collection of RDF statements expressed as triples form a labeled, directed multi-graph. This multi-graph does not necessarily have a predefined schema and can grow arbitrarily as additional statements are added. Compared to the relational model, RDF hence provides a much more flexible schema-less and distributed data model to express arbitrary statements. Without any external means to describe the vocabularies used to make such statements, however, their precise meaning remains unclear. Therefore, additional stack layers are necessary to incorporate explicit semantics. In this context, RDF Schema (RDFS) was developed to provide basic constructs for the description of vocabulary terms used in an RDF model. It took the basic RDF specification and extended it to provide a minimal ontology representation language. RDFS allows the definition of classes and properties and to arrange them in simple subsumption hierarchies. Its overall expressiveness, however, is limited.

The Web Ontology Language (OWL) [17] was developed to address these limitations. The term *ontology* has developed a distinct meaning within information science, with limited grounding in philosophy, the domain from which it was originally adopted. In the context of the semantic web, “an ontology is an explicit specification of a conceptualization” [18]. OWL ontologies provide formal descriptions of concepts and their relationships that exist in a certain universe of discourse, together with a shared vocabulary to refer to these concepts [19]. OWL is based on attribute-like and relationship-like properties and also includes several other expressive modeling primitives (e.g., class union and intersections, cardinality restrictions on properties, etc.). Today, OWL is not a single language but a family of sublanguages with varying levels of expressiveness (in increasing order): OWL Lite, OWL DL, and OWL Full.

Ontologies were considered the key to solve the problem of semantic heterogeneity by allowing users to explicitly define the structure of their knowledge of a domain. The vision was that this explicit semantic specification would allow communities to reach a shared understanding and thereby reduce semantic heterogeneity. Despite varying vocabularies, agents would then be able to discover common meaning by referring to concepts in (interlinked) OWL ontologies.

Many of the standards developed within the semantic web framework, including RDF, RDFS and the expressive query language SPARQL are in widespread use today. The amount of data published in RDF has grown tremendously and OWL has also seen significant adoption as a representation in particular knowledge domains and for particular applications (particularly in scientific communities, e.g., the life sciences). Interlinked ontologies (through RDF), which allow ontologies to be distributed across systems and refer to each other’s terms, by contrast, remain relatively uncommon today. Overall, OWL has not played the envisaged central role as “semantic glue” that brings representation to the open web. Hence, it is not surprising that the additional layers proposed in the original semantic web stack such as unifying logic, proof, and trust (as well as cryptography as an orthogonal aspect) have not been realized. As a consequence, the vision of a semantic web in which agents automatically perform tasks for their users has so far largely failed to materialize. The

focus of interoperability research has therefore shifted again, concentrating mostly on more successful aspects of the semantic web vision, which gave rise to the concept of Linked (Open) Data.

## 5 Linked (Open) Data and Micro Formats

Linked (Open) Data can be seen as a more pragmatic approach towards a “web of data”, which leaves the vision of a completely semantic web and intelligent agents aside. Tim-Berners Lee postulated a set of Linked Data principles [20] that accentuate the data-centric aspects of existing semantic web technologies. These principles are (i) use URIs to identify things; (ii) use HTTP URIs so that people can look up those names; (iii) when someone looks up a URI, provide useful information, using the standards (RDF, SPARQL); and (iv) include links to other URIs, so that one can discover more things.

All datasets published as Linked Data use RDF as a common data model. The encoding used, however, is not uniform. The use of RDF/XML proposed in the initial RDF standard [21] turned out to be an impractical compromise between readability for humans and machines. Therefore, various other syntax notations and serialization formats for RDF have emerged and are in widespread use (e.g., N-Triples, Turtle, N3, JSON-LD). Based on the common data model, syntactic integration of different Linked Data sets is straightforward, even if they are encoded in different formats.

The use of URIs as a resource identification mechanism remains a key principle. When an RDF triple contains URIs from different namespaces in subject and object position, this triple establishes a link between the entity identified by the subject (which may be described in a source dataset using namespace *A*) with the entity identified by the object (described in another target dataset using namespace *B*). Through the typed RDF links, data items are effectively interlinked [22]. De-referencability means that URIs are not just used for identifying entities, but also enable locating and retrieving resources describing and representing these entities on the Web.

The use of URIs as identifiers does not, however, entail standardization. Even though reuse of existing URIs is encouraged, data providers are free to publish their data in RDF using arbitrary URIs. Identity resolution therefore remains a major issue. Pre-defined concepts (e.g., *rdfs:seeAlso*, *owl:sameAs*, *skos:closeMatch*, etc.) can be used to establish explicit links, which facilitates vocabulary mapping, i.e., translating terms from different vocabularies into a single target schema. Complex semantic integration requires schema and instance matching techniques and expressing found matches as alignments of RDF vocabularies and ontologies in terms of additional triple facts [22]. It also requires the resolution of conflicts between data sources. Such complex integration workflows are a far cry from the original semantic web vision of fully automated semantic integration and reasoning, but they do make it possible to reconcile, reuse, and find new applications for previously isolated data with moderate effort. In that sense, the Linked Data approach has been highly successful in bringing



Fig. 1 Linked open data example visualization ([http://en.lodlive.it/?http://dbpedia.org/resource/The\\_Library\\_of\\_Babel](http://en.lodlive.it/?http://dbpedia.org/resource/The_Library_of_Babel), accessed May 14, 2015)

us closer to a web of data. The number of open datasets published based on these principles has grown tremendously in recent years. As of 2015, the global Linked Data Graph spanned by these datasets has grown to approximately 90 billion triples from almost 4,000 working datasets.<sup>1</sup> Figure 1 shows a small example of an excerpt of this global graph. It visualizes the DBpedia entry of an example book using LODlive. Amongst others, the book is linked to the Wikipedia article the information was derived from, to author, country, and language information, and the literary genre (for space reasons, only selected links are shown).

A related development that has emerged from web design practice and that has seen increasing adoption in recent years is the use of HTML markup tags to attach semantics to structured data embedded within web pages. Typical use cases include calendar entries, contact information, blog posts, products, reviews, or cooking recipes. Like earlier attempts to make the web of documents machine-readable, these practices aim to make the content accessible to applications. Compared to more “heavyweight” approaches, however, they tend to value ease of authoring and human-readability

<sup>1</sup><http://stats.lod2.eu>, accessed April 30, 2015.



above machine-interpretability, convention over formal standardization, and simplicity over explicitly codified semantics. Therefore, these approaches do not use ontologies to formally standardize vocabularies, but normalize conventions derived from web-publishing behavior. Whereas the semantic web approach is based on a complex stack of technologies and languages, these approaches embed semantics in HTML directly following commonly accepted markup formats. Thereby, they aim to lower the barrier of adoption among web developers to avoid the deadlock that has plagued most semantic web technologies. Major web companies, including Google, Facebook, Yahoo!, and Microsoft, have started to make use of the data embedded into web pages in microformats e.g., to enrich search results. The currently most prevalent formats used (cf. [23]) are Microformats,<sup>2</sup> which annotate html with terms from a fixed set of vocabularies; RDFa [24], which can be used to embed any kind of RDF data into html pages; and Microdata [25], a recent format developed in the context of HTML 5.

*Microformats* make structured use of the *class* and *rel* attributes to associate web content with a particular meaning, e.g., through vCards, iCal events, and friendship relations. Microformats result in very compact syntax with little explicit semantics, which are rather implicitly defined by codified convention and embodied in parser and application code. Development and extension of the format is based on a community process and no mechanisms for custom extensions exist.

*RDFa* [24], developed and supported by the W3C, is a set of rules that can be used as a module for XHTML2. It reuses the *meta* and *link* attributes from standard XHTML and makes it easy to extract RDF triples from an RDFa annotated document. Compared to Microdata and Microformats, RDFa is more versatile and allows publishers to use arbitrary vocabularies and modular schemas.

Finally, *Microdata* [25] is a recent approach that is driven by the major web search companies. Standardization through the W3C has been initiated [26], but the standard has not advanced to the state of a Candidate Recommendation. Microdata uses a supporting vocabulary to describe an item and name-value pairs to assign values to its properties. Search engines publish markup vocabularies via *schema.org*. Publishers are encouraged to re-use such existing vocabularies, but ad-hoc vocabularies are also possible.

All three formats have found considerable adoption and led to a growth of semantics on the Web. The structural richness of the published data, however, is limited, as most websites only use a small set of rather generic properties to describe entities [23]. As a consequence, interoperability issues remain largely unresolved.

## 6 Conclusions and an Outlook on the Future

Looking retrospectively at the long history of interoperability research against the backdrop of changing tides of technological development, we find that few issues have been ultimately solved. Early on, it became clear that in order to facilitate inter-

---

<sup>2</sup><http://microformats.org>, accessed April 30, 2015.

operability, it was necessary to make data machine-interpretable. This has motivated various attempts to add semantics to raw data in order to facilitate meaningful interactions between machines.

In the database community, this led to the development of conceptual approaches such as Entity-Relationship modeling. In the programming domain, it was reflected in the development of the object-oriented paradigm as a way to represent reality more naturally and hence facilitate interpretation. In the context of documents, it first led to the development of semi-structured document models that are machine-interpretable at least to the extent necessary to present them to the user. Subsequently, it led to the vision of a web of documents that could be navigated and interpreted by machines. This triggered the development of various semantic web languages, technologies, and concepts, but the vision of interoperable intelligent agents that can reason and draw conclusions autonomously, while automatically resolving semantic issues using logical connections of terms, failed to materialize.

Linked Data took a step back from the ambitious vision of a completely semantic web and followed a more pragmatic approach that centers on data publishing. Finally, microformats, which have seen increasing adoption, are even more pragmatic in that they instill meaning into arbitrary web documents only selectively where it provides tangible benefits. This approach is much more in line with actual incentives for developers to invest into interoperability.

More generally, a lesson we can draw from the retrospective in this paper is that the lack of interoperability that still exists is largely not caused by technological limitations, but is rather driven by economic and social factors. The semantic web, for instance, is plagued with some technical issues (such as computational complexity, limitations of reasoning engines and triple stores, limited network bandwidth etc.), but the more fundamental issues are arguably a lack of concrete incentives for adoption on the one hand, and the social nature of semantics on the other hand. The idea that semantic models can be specified independently by users in a decentralized manner to accommodate their individual needs, contexts and respective worldview is necessary. However, the assumption that this semantic heterogeneity could be tackled technologically through additional mechanisms, if only the same language were used to express semantics explicitly, proved unfounded. It did not hold up to reality because communities need to arrive at agreed-upon semantics in a social process. Semantics is also not absolute, but subject to personal interpretation and frequently a moving target. This is critical, because inconsistencies on the semantic web can be introduced easily by anyone, breaking its logic and hence its interoperability.

More successful recent approaches that foster interoperability, such as Linked Data and Microformats, are focused more on the actual needs of data publishers and consumers. They are successful in facilitating limited interoperability because they require less codified semantics. However, Linked Data tends to shift the problem of interpretation to data publishers, which may lead to highly fragmented and isolated datasets. To achieve interoperability, Linked Data requires strong links between datasets and strong vocabularies that anchor the data expressed. Microformats shift interpretation even further from explicit models back into the parser and application implementations. Programmers must hence hardcode a consensus on accepted

practices into their applications. That consensus must either emerge socially, or be promoted by actors with sufficient leverage, such as search engines. In practice, topics, formats, and vocabularies used to represent data are therefore largely determined by the major consumers the data is targeted at [23].

Overall, recent pragmatic approaches, which are based around the idea that “a little semantics goes a long way”, have improved interoperability, but they have not solved the actual problem of semantic interoperability on a large scale. With the onset of even larger issues in the context of “Big Data” and the “Web of Things” [27], new challenges for integration will surely arise.

## References

1. Ouksel, A.M., Sheth, A.: Semantic interoperability in global information systems. *SIGMOD Rec.* **28**(1), 5–12 (1999)
2. Sheth, A.P., Larson, J.A.: Federated database systems for managing distributed, Heterogeneous, and autonomous databases. *ACM Comput. Surv.* **22**(3), 183–236 (1990)
3. Visser, P.R.S., Jones, D.M., Bench-Capon, T.J.M., Shave, M.J.R.: An analysis of ontological mismatches: Heterogeneity versus interoperability. In *AAAI 1997 spring symposium on ontological engineering*. Stanford, USA (1997)
4. Wache, H.: *Semantische mediation für heterogene Informationsquellen*. University of Bremen, Ph.D. diss (2003)
5. Taylor, R.W., Frank, R.L.: CODASYL data-base management systems. *ACM Comput. Surv.* **8**(1), 67–103 (1976)
6. Codd, E.F.: A relational model of data for large shared data banks. *Commun. ACM* **13**(6), 377–387 (1970)
7. Stonebraker, M., Neuhold, E.J.: A distributed database version of INGRES. In *Berkeley workshop*, pp. 19–36 (2002)
8. Neuhold, E.J., Biller, H.: POREL: A distributed data base on an Inhomogeneous computer network. In *VLDB*, pp. 380–395 (2002)
9. Landers, T., Rosenberg, R.L.: An overview of multibase. In: Chu, W.W. (ed.) *Distributed Data Base Systems*, pp. 391–421. Artech House Inc., Norwood (1986)
10. Litwin, W., Boudenat, J., Esculier, C., Ferrier, A., Glorieux, A.M., La Chima, J., Kabbaj, K., Moulinoux, C., Rolin, P., Stangret, C.: Sirius system for distributed data management. In *Distributed data base systems*, pp. 311–366 (1982)
11. Chen, P.P.-S.: The entity-relationship model—toward a unified view of data. *ACM Trans. database Syst.* **1**(1), 9–36 (1976)
12. Halpin, T.: Object-role modeling (ORM/NIAM). In *Handbook on architectures of information systems*, pp. 81–102 (1998)
13. Berners-Lee, T., Conolly, D.: RFC 1866 - hypertext markup language—2.0. <https://tools.ietf.org/html/rfc1866> (1995). Accessed 20 Apr 2015
14. Berners-Lee, T., Hendler, J., Lassila, O.: The semantic web. *Sci. Am.* **284**(5), 28–37 (2001)
15. Shadbolt, N., Berners-Lee, T., Hall, W.: The semantic web revisited. *IEEE Intell. Syst.* **21**(3), 96–101 (2006)
16. Fensel, D.: *Ontologies—a silver bullet for knowledge management and electronic commerce*. Springer, Berlin (2004)
17. W3C: *OWL 2 web ontology language structural specification and functional-style syntax (Second Edition)*
18. Gruber, T.R.: A translation approach to portable ontology specifications. *Knowl. Acquisition* **5**, 199–220 (1993)

19. Ziegler, P., Dittrich, K.R.: Three decades of data integration – all problems solved? In: Jacquart, R. (ed.) *Building the Information Society*, pp. 3–12. Springer, New York (2004)
20. Berners-Lee, T.: Linked data—design issues. <http://www.w3.org/DesignIssues/LinkedData.html> (2006). Accessed 21 Apr 2015
21. Beckett, D., McBride, B.: *RDF/XML syntax specification (Revised)*. W3C (2004)
22. Auer, S., Bryl, V., Tramp, S. (eds.): *Linked open data—creating knowledge out of interlinked data*, vol. 8661. Springer International Publishing, Cham (2014)
23. Bizer, C., Eckert, K., Meusel, R., Mühleisen, H., Schuhmacher, M., Völker, J.: Deployment of RDFa, microdata, and microformats on the web—a quantitative analysis. In *the semantic Web-ISWC*, pp. 17–32. Springer (2013)
24. Herman, I., Adida, B., Sporny, M., Birbeck, M.: *RDFa 1.1 primer—third edition*. <http://www.w3.org/TR/xhtml-rdfa-primer/> (2015). Accessed 29 Apr 2015
25. Haas, K., Mika, P., Tarjan, P., Blanco, R.: Enhanced results for web search. In: *Proceedings of the 34th International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 725–734. New York (2011)
26. Hickson, I.: *HTML microdata—W3C working group note 29 October 2013*. <http://www.w3.org/TR/microdata/> (2013). Accessed 29 Apr 2015
27. Baccelli, E., Ragget, D.: The promise of the internet of things and the web of things—introduction to the special Theme. *ERCIM News* **101**, (2015)

**Part II**  
**Green Computing and Networking**

# Performance of an Autonomous Energy Harvesting Wireless Sensor

Erol Gelenbe and Yasin Murat Kadioglu

**Abstract** While complex autonomic self-organising systems have received much attention, simple autonomous systems are also needed for remote sensing applications, as well as for the Internet of Things. Such autonomous stand-alone unattended devices may not have access to reliable sources of mains power, and will have to harvest energy locally from ambient sources such as vibrations, heat or light. However energy leakage will also be a problem. This paper proposes a mathematical model to analyse the performance of such systems in the presence of a random source of energy, as well as a random source of data. The equilibrium between random energy, random data and random leakage results in an interesting performance analysis of these small but ubiquitous systems as a whole. A discussion is also provided about the effect of transmission errors.

**Keywords** Autonomous energy harvesting · Performance evaluation · Autonomous wireless sensor

## 1 Introduction

In recent years, much work has been conducted on self-organised or “autonomic” communication systems [1] and biologically inspired [2] or adaptive control methods [3] have been suggested for their management. However many stand-alone autonomous systems require a very *simple* organisation for unattended long term operation. One such area of application is in stand-alone wireless sensors which need to operate remotely without a change of batteries. Such systems are also motivated

---

E. Gelenbe · Y.M. Kadioglu (✉)  
Department of Electrical and Electronic Engineering, Intelligent Systems  
and Networks Group, Imperial College, Sw7 2bt, London  
e-mail: y.kadioglu14@imperial.ac.uk

E. Gelenbe  
e-mail: e.gelenbe@imperial.ac.uk

© Springer International Publishing Switzerland 2016  
O.H. Abdelrahman et al. (eds.), *Information Sciences and Systems 2015*,  
Lecture Notes in Electrical Engineering 363,  
DOI 10.1007/978-3-319-22635-4\_3

by the cost or difficulty of access, and also by the potential environmental damage when one discards batteries, and by the need to save energy in ICT [4, 5].

Thus energy harvesting from solar, thermal, vibrational, or ambient electromagnetic radiation including light, are of particular interest [6, 7], especially in remote sensing and security applications [8–10], and recent research has addressed such technologies for communications [11]. However much work still needs to be done to understand the performance of such systems which need to operate autonomously for very long periods of time.

Earlier work [12] studied the performance of an autonomous energy harvesting communication node as a function of the random flow of harvested energy using an “energy packet” model which discretises both the data flow and the energy flow in the sensor node [13] based on queueing networks [14]. Here we extend the work to study the ability of such a system to operate in an unattended and autonomous manner, in the presence of leakage from energy storage units such as batteries or capacitors.

### 1.1 The Mathematical Model

An energy harvesting wireless sensor communication node collects data to transmit at rate  $\lambda$  data packets (DPs) per second from sensing activities, and harvests energy at rate  $\Lambda$  energy packets (EPs) per second, where one energy packet is the unit of electrical energy, e.g. micro-joules, collected from light, heat or vibrations, that is needed to transmit one data packet. The energy harvesting rate  $\Lambda$  and data gathering rate  $\lambda$  are assumed to be small (i.e. very slow) as compared to the time it takes to create and transmit a packet via wireless, which may be in the nano or picoseconds. The node stores energy in a capacitor or battery, and energy leaks in a random manner at rate of  $\mu$  energy packets per second.

The state of the system is represented by  $K(t)$ , the number of data packets stored at the node, and by  $M(t)$  the number of energy packets that are stored, at time  $t \geq 0$ . Since the transmission time at the node is very short, whenever energy is available and there are data packets waiting they will be instantaneously transmitted till the energy is depleted: from a state  $K(t) > 0$ ,  $M(t) > 0$  the system instantaneously moves to either state  $(0, M(t) - K(t))$  if  $M(t) \geq K(t)$ , or to  $(K(t) - M(t), 0)$  if  $K(t) \geq M(t)$ . Writing  $p(n, m, t) = \text{Prob}[K(t) = n, M(t) = m]$ , we therefore only consider  $p(n, m, t) \in S$  with  $S = \{(0, 0), (n, 0), (0, m) : n > 0, m > 0\}$ .

## 2 Finite Capacity Data and Energy Buffers with Energy Leakage

First note that if both the data buffer and the energy storage capacity are finite, the system can be modelled as a finite Markov chain whose set of states are given in Sect. 1.1 with  $0 \leq n \leq B$ ,  $0 \leq m \leq E$ . We note that the process  $[K(t), M(t), t \geq 0]$

is an irreducible and aperiodic Markov chain so that the stationary probabilities  $p(n, m) = \lim_{t \rightarrow \infty} Pr[K(t) = n, M(t) = m]$  exist and are computed from the equations:

$$p(0, 0)[\lambda + \Lambda] = \Lambda p(1, 0) + \lambda p(0, 1) + \mu p(0, 1), \quad (1)$$

since state  $(0, 0)$  is reached if either there was just one data packet and it was transmitted as soon as an energy packet arrived, or there was one energy packet and it was consumed as soon as a data packet arrived, or one energy packet was lost due to leakage. When  $0 < n < B$  we have:

$$p(n, 0)[\lambda + \Lambda] = \Lambda p(n + 1, 0) + \lambda p(n - 1, 0), \quad (2)$$

while:

$$p(B, 0)\Lambda = p(B - 1, 0)\lambda. \quad (3)$$

We note that these equations have a solution of the form:

$$p(n, 0) = \alpha^n C_d, \quad \alpha = \frac{\lambda}{\Lambda}, \quad (4)$$

where  $C_d$  is a constant. For the energy storage system, when  $0 < m < E$  we have:

$$p(0, m)[\lambda + \Lambda + \mu] = \Lambda p(0, m - 1) + \lambda p(0, m + 1) + \mu p(0, m + 1), \quad (5)$$

while:

$$p(0, E)[\lambda + \mu] = p(0, E - 1)\Lambda. \quad (6)$$

We note that these equations have a solution of the form:

$$p(0, m) = \theta^m C_e, \quad \theta = \frac{\Lambda}{\lambda + \mu}, \quad (7)$$

where  $C_e$  is a constant. Since the unique stationary probability distribution exists, we must have  $C_d = C_e = p(0, 0)$  by considering Eq. (1):

$$p(0, 0)(\lambda + \Lambda) = \Lambda \left(\frac{\lambda}{\Lambda}\right) C_d + (\lambda + \mu) \left(\frac{\Lambda}{\lambda + \mu}\right) C_e, \quad (8)$$

$$0 = (p(0, 0) - C_d)\lambda + (p(0, 0) - C_e)\Lambda. \quad (9)$$



Using the fact that the probabilities sum to one we have:

$$1 = p(0, 0) + \sum_{n=1}^B p(n, 0) + \sum_{m=1}^E p(0, m), \quad (10)$$

$$= p(0, 0)[1 + \sum_{n=1}^B \alpha^n + \sum_{m=1}^E \theta^m], \quad (11)$$

$$= p(0, 0)[1 + (\frac{\alpha(\alpha^B - 1)}{\alpha - 1}) + (\frac{\theta(\theta^E - 1)}{\theta - 1})]. \quad (12)$$

Hence:

$$p(0, 0) = \frac{1 - \alpha - \theta + \alpha\theta}{\alpha^{B+1}(\theta - 1) + \theta^{E+1}(\alpha - 1) + 1 - \alpha\theta}, \quad (13)$$

$$p(n, 0) = \alpha^n \frac{1 - \alpha - \theta + \alpha\theta}{\alpha^{B+1}(\theta - 1) + \theta^{E+1}(\alpha - 1) + 1 - \alpha\theta}, \quad 0 \leq n \leq B, \quad (14)$$

$$p(0, m) = \theta^m \frac{1 - \alpha - \theta + \alpha\theta}{\alpha^{B+1}(\theta - 1) + \theta^{E+1}(\alpha - 1) + 1 - \alpha\theta}, \quad 0 \leq m \leq E. \quad (15)$$

Also, we can express the marginal probabilities for the queue length of DPs and EPs as:

$$p_d(n) = \sum_{m=0}^{\infty} p(n, m) = p(n, 0), \quad n > 0, \quad (16)$$

$$p_d(0) = \sum_{m=0}^{\infty} p(0, m) = \sum_{m=0}^{\infty} \theta^m p(0, 0) = \frac{1 - \theta^{E+1}}{1 - \theta} p(0, 0). \quad (17)$$

Similarly,

$$p_e(m) = \sum_{n=0}^{\infty} p(n, m) = p(0, m), \quad m > 0, \quad (18)$$

$$p_e(0) = \sum_{n=0}^{\infty} p(n, 0) = \sum_{n=0}^{\infty} \alpha^n p(0, 0) = \frac{1 - \alpha^{B+1}}{1 - \alpha} p(0, 0). \quad (19)$$

Hence:

$$p_d(n) = \alpha^n \frac{1 - \alpha - \theta + \alpha\theta}{\alpha^{B+1}(\theta - 1) + \theta^{E+1}(\alpha - 1) + 1 - \alpha\theta}, \quad 0 < n \leq B, \quad (20)$$

$$p_e(m) = \theta^m \frac{1 - \alpha - \theta + \alpha\theta}{\alpha^{B+1}(\theta - 1) + \theta^{E+1}(\alpha - 1) + 1 - \alpha\theta}, \quad 0 < m \leq E. \quad (21)$$

### 3 Energy and Data Packet Loss Due to Finite Energy and Data Buffers

When the energy storage capacity is finite, or the data packet buffer is finite, we are bound to have some level of energy loss or data packet loss. These loss rates  $L_e$ ,  $L_d$  in energy and data packets per second, respectively, can easily be computed as:

$$L_e = \Lambda \sum_{n=0}^{\infty} p(n, E) = \Lambda \theta^E \frac{1 - \alpha - \theta + \alpha\theta}{\alpha^{B+1}(\theta - 1) + \theta^{E+1}(\alpha - 1) + 1 - \alpha\theta}. \quad (22)$$

$$L_d = \lambda \sum_{m=0}^{\infty} p(B, m) = \lambda \alpha^B \frac{1 - \alpha - \theta + \alpha\theta}{\alpha^{B+1}(\theta - 1) + \theta^{E+1}(\alpha - 1) + 1 - \alpha\theta}. \quad (23)$$

For the assumption of very large buffer sizes i.e., both B and E tend to infinity, the following cases can be considered:

**Case 1** If  $\alpha > 1$  and hence  $\theta < 1$  or equivalently  $\Lambda < \lambda$ , so that the energy is not sufficient enough for the data and  $L_e \rightarrow 0$  and  $L_d \rightarrow \lambda - \Lambda$ , as would be expected.

**Case 2** If  $\alpha = 1$  and hence  $\theta < 1$  or equivalently  $\Lambda - \mu < \lambda$ , the expressions for  $L_e$  and  $L_d$  are in indeterminate form. However, after some algebra we get  $L_e \rightarrow 0$  and  $L_d \rightarrow 0$ .

**Case 3** If  $\alpha < 1$  and  $\theta < 1$  or equivalently  $\lambda < \Lambda < \lambda + \mu$ , in this case there is no leakage for both buffer, and  $L_e \rightarrow 0$  and  $L_d \rightarrow 0$ .

**Case 4** If  $\alpha < 1$  and  $\theta > 1$  or equivalently  $\Lambda > \lambda + \mu$ , so that the energy is more plentiful than it is needed, and  $L_e \rightarrow \Lambda - \lambda - \mu$  and  $L_d \rightarrow 0$ , as would be expected.

**Case 5** If  $\alpha < 1$  and  $\theta = 1$  or equivalently  $\Lambda - \lambda = \mu$ , the expressions for  $L_e$  and  $L_d$  are in indeterminate form. However, after some algebra we get  $L_e \rightarrow 0$  and  $L_d \rightarrow 0$ .

### 4 Sensors with Transmission Errors

When a transmission error is detected, the same DP will be retransmitted until an error-free transmission occurs or until energy is depleted. Thus if a DP is waiting in queue and an EP arrives, the DP is transmitted and a transmission error may occur with probability  $\pi$ , so that the DP is not deleted from the queue. Similarly, if a DP arrives to the node when one or more EPs are waiting, then a transmission error may occur with probability  $p$  and the transmission will be repeated independently with the probability of error  $p$ , and this will be repeated until success occurs or until all the EPs are depleted.

When the transmission errors are considered, we still have the same leakage rate ( $\mu$ ), DP arrival rate ( $\lambda$ ) and EP arrival rate ( $\Lambda$ ) and these rates are responsible for the state transitions  $(0, m) \rightarrow (0, m - 1)$ ,  $(n, 0) \rightarrow (n + 1, 0)$  and  $(0, m) \rightarrow (0, m + 1)$ , respectively. On the other hand, due to existence of  $\pi$  and  $p$ , we need to define some new state transition rates.

The rate  $\Lambda\pi$  is for the transition  $(n, 0) \rightarrow (n, 0)$  when  $n \geq 1$ . For this transition, an EP arrives to an empty energy buffer and since upon arrival of an EP, if another DP transmission is requested immediately just after a DP transmission with probability  $\pi$ , then due to lack of an EP, the new DP transmission will not be successful and will replace the previous one. Similarly,  $\Lambda(1 - \pi)$  is the rate for the transition  $(n, 0) \rightarrow (n - 1, 0)$  when  $n \geq 1$ , for which an EP arrives to an empty energy buffer and another DP transmission is not requested after the DP transmission caused by the arrival of an energy packet. The rate  $\lambda p$  is for the transition  $(0, 1) \rightarrow (1, 0)$  which occurs when a DP arrives to an empty data buffer and another DP transmission is requested after the DP transmission caused by the EP already in the queue. In this case, the DP will have to wait for arrival of another EP. Furthermore, the transition  $(0, m) \rightarrow (0, m - 1)$  when  $m > 0$  occurs with rate  $\lambda(1 - p)$ . We have this transition when a DP arrives to an empty data buffer and the transition is served by an EP already in the queue so that the number of EPs is reduced by 1. Likewise, arrival of a DP results in a sequence of  $k$  successive repetitions of DPs arriving to the node so that there will be a set of transitions starting from states of the form  $(0, m)$ ,  $m > 0$ , with probability  $p^k(1 - p)$  where  $m \geq k \geq 0$ . Therefore, the rate  $\lambda p^{k-1}(1 - p)$  is responsible for the transition  $(0, m) \rightarrow (0, m - k)$  when  $m \geq k > 1$ . Finally, when we consider arrival of a DP, the number of EP reduces by 1 in energy storage and it may be followed by another DP transmission request. The number of EPs decreases with the each DP transmission request so that the  $m$ th and final transmission request cannot be satisfied due to the fact that EPs are depleted. The system moves into state  $(1, 0)$  having depleted all its EPs and having one final DP waiting to be transmitted. Thus, this transition is  $(0, m) \rightarrow (1, 0)$  with rate  $\lambda p^m$ . Notice that for any  $m > 0$  the sum of these probabilities is one:

$$\sum_{k=0}^{m-1} p^k(1 - p) + p^m = 1. \quad (24)$$

These transitions lead to the equilibrium equations:

$$p(0, 0)[\lambda + \Lambda] = \lambda \sum_{l=1}^{\infty} p^{l-1}(1 - p)p(0, l) + \Lambda(1 - \pi)p(1, 0) + \mu p(0, 1), \quad (25)$$

$$p(1, 0)[\lambda + \Lambda(1 - \pi)] = \lambda \sum_{l=0}^{\infty} p^l p(0, l) + \Lambda(1 - \pi)p(2, 0), \quad (26)$$

$$p(n, 0)[\lambda + \Lambda(1 - \pi)] = \lambda p(n - 1, 0) + \Lambda(1 - \pi)p(n + 1, 0), \quad n > 1, \quad (27)$$

$$p(0, m)[\lambda + \Lambda + \mu] = \lambda \sum_{l=1}^{\infty} p^{l-1} (1-p)p(0, m+l) \quad (28)$$

$$+ \Lambda p(0, m-1) + \mu p(0, m+1), \quad m > 0. \quad (29)$$

**Theorem** If  $(\Lambda - \mu)(1-p) < \lambda < \Lambda(1-\pi)$ , the stationary distribution exists and is given by:

$$p(0, m) = p(0, 0)Q^m, \quad m \geq 0, \quad (30)$$

$$p(n, 0) = p(1, 0)q^{n-1}, \quad n \geq 1, \quad (31)$$

where

$$q = \frac{\lambda}{\Lambda(1-\pi)}, \quad Q = \frac{\lambda + \mu + \Lambda p - \sqrt{(\lambda + \mu + \Lambda p)^2 - 4\mu\Lambda p}}{2\mu p}, \quad (32)$$

and

$$\frac{p(1, 0)}{p(0, 0)} = \frac{q}{(1-pQ)} = \frac{\lambda 2\mu}{\Lambda(1-\pi)[\mu - \lambda - \Lambda p + \sqrt{(\lambda + \Lambda p + \mu)^2 - 4\mu\Lambda p}]}, \quad (33)$$

with

$$\begin{aligned} p(0, 0) &= \frac{(1-q)(1-Q)(1-pQ)}{q(1-Q) + (1-q)(1-pQ)} \\ &= \left( \frac{\frac{2\mu\lambda}{\Lambda(1-\pi)-\lambda}}{[\mu - \lambda - \Lambda p + \sqrt{(\lambda + \Lambda p + \mu)^2 - 4\mu\Lambda p}]} \right. \\ &\quad \left. + \frac{2\mu p}{2\mu p - (\lambda + \Lambda p + \mu) + \sqrt{(\lambda + \mu + \Lambda p)^2 - 4\mu\Lambda p}} \right)^{-1}. \end{aligned} \quad (34)$$

*Proof* To proceed with the proof, we substitute (30) in (28), which after some algebra becomes:

$$Q^m[\lambda + \Lambda + \mu] = \lambda(1-p)Q^{m+1} \frac{1}{1-pQ} + \Lambda Q^{m-1} + \mu Q^{m+1}, \quad (35)$$

$$0 = (Q-1)[Q^2(\mu p) + Q(-\Lambda p - \lambda - \mu) + \Lambda], \quad (36)$$

$$Q_{1,2} = \frac{\lambda + \mu + \Lambda p \pm \sqrt{(\lambda + \mu + \Lambda p)^2 - 4\mu\Lambda p}}{2\mu p}, \quad (37)$$

note that  $Q$  has to be smaller than 1,

$$\frac{\lambda + \mu + \Lambda p + \sqrt{(\lambda + \mu + \Lambda p)^2 - 4\mu\Lambda p}}{2\mu p} \geq \frac{\lambda + \mu + \Lambda p}{2\mu p} > \frac{1}{2} \left( \frac{1}{p} + \frac{\Lambda}{\mu} \right) > 1, \quad (38)$$

since  $p < 1$  and  $\mu < \Lambda$ . The root  $Q_1$  should not be considered, and  $Q_2$  is the only viable root which leads  $(\Lambda - \mu)(1 - p) < \lambda$ .

Also we must have  $q < 1$  and we note that the equation in (31) has a solution of the form:

$$p(n, 0) = q^{n-1} C_d, \quad q = \frac{\lambda}{\Lambda(1-\pi)} \text{ and } C_d = p(1, 0) \text{ with } \lambda < \Lambda(1-\pi). \quad (39)$$

After further analysis using the fact that the probabilities must sum to one, and  $q < 1$ ,  $Q < 1$ , we obtain  $p(0, 0)$ :

$$\left( \frac{\frac{2\mu\lambda}{(\Lambda(1-\pi)-\lambda)}}{[\mu - \lambda - \Lambda p + \sqrt{(\lambda + \Lambda p + \mu)^2 - 4\mu\Lambda p}] + \frac{2\mu p}{2\mu p - (\lambda + \Lambda p + \mu) + \sqrt{(\lambda + \mu + \Lambda p)^2 - 4\mu\Lambda p}}} \right)^{-1}. \quad \square$$

## 5 Conclusions

Complex autonomic self-organising systems have been at the centre of attention over the last decade and have included several EU research projects and have met with wide interest in the literature. On the other hand, simple autonomous systems which have to operate in unattended remote environments have met with less research interest. Yet important areas such as remote sensing applications, as well as for the the Internet of Things, require that unattended autonomous systems operate reliably over long periods of time.

One important area of application is when one cannot change batteries of simple devices which may also be difficult to connect to the mains for their power needs. In such cases, the autonomous systems will have to harvest energy locally from ambient sources such as vibrations, heat or light, and energy leakage from temporary storage units, such as rechargeable batteries or capacitors, will also be a problem.

This paper analyses the performance of such systems that are devoted to gathering data and transmitting it, and which use energy harvesting for their operation. We have proposed a mathematical model to analyse the performance of such systems in the presence of a random source of energy, a random source of data, and random energy leakage. The equilibrium between random energy, random data and random leakage results in an interesting performance analysis of these small but ubiquitous systems. A preliminary discussion has also been given about how transmission errors can be included in the model.

Future work will have to incorporate the actual transmission scheme, including noise and interference, to better understand the optimum autonomous performance of such systems.

**Acknowledgments** The authors gratefully acknowledge the support of the ERA-NET ECROPS Project under grant No. Grant No. EP/K017330/1 from the UK Engineering and Physical Science Research Council (EPSRC) to Imperial College.

## References

1. Dobson, S., et al.: A survey of autonomic communications. *ACM Trans. Auton. Adapt. Syst.* **1**(2), 223–259 (2006)
2. Gelenbe, E.: Steps towards self-aware networks, *Comm. ACM* **52**, 66–75 (2009)
3. Gelenbe, E.: Sensible decisions based on qos. *Comput. Manage. Sci.* **1**(1), 1–14 (2003)
4. Pettey, C.: Gartner estimates ict industry accounts for 2 percent of global CO<sub>2</sub> emissions, vol. 14, p. 2013 (2007). <https://www.gartner.com/newsroom/id/503867>
5. Uysal-Biyikoglu, E., Prabhakar, B., El Gamal, A.: Energy-efficient packet transmission over a wireless link. *IEEE/ACM Trans. Networking* **10**(4), 489–749 (2002)
6. Rodoplu, V., Meng, T.H.: Bits-per-joule capacity of energy-limited wireless networks. *IEEE Trans. Wireless Commun.* **6**(3), 857–865 (2007)
7. Meshkati, F., Poor, H.V., Schwartz, S.C., Mandayam, N.B.: An energy-efficient approach to power control and receiver design in wireless data networks. *IEEE Trans. Commun.* **53**(11), 1885–1894 (2005)
8. Gelenbe, E., Hussain, K., Kaptan, V.: Simulating autonomous agents in augmented reality, *J. Syst. Softw.* **74**, 255–268 (2005)
9. Alippi, C., Galperti, C.: An adaptive system for optimal solar energy harvesting in wireless sensor network nodes. *IEEE Trans. Circuits Syst. I: Regul. Pap.* **55**(6), 1742–1750 (2008)
10. Seah, W.K., Eu, Z.A., Tan, H.-P.: Wireless sensor networks powered by ambient energy harvesting (wsn-heap)-survey and challenges. In: *IEEE, 1st International Conference on Wireless Communication, Vehicular Technology, Information Theory and Aerospace and Electronic Systems Technology, Wireless VITAE*, p. 15 (2009)
11. Gelenbe, E., Gesbert, D., Gunduz, D., Kulah, H., Uysal-Biyikoglu, E.: Energy harvesting communication networks, optimization and demonstration: The e-crops project. In: *24th TIWDC, Tyrrhenian International Workshop 2013 on Digital Communications: Green ICT, IEEE Xplore* (2013)
12. Gelenbe, E.: A sensor node with energy harvesting. *SIGMETRICS Perform. Eval. Rev.* **42**(2), 37–39 (2014)
13. Gelenbe, E.: Energy packet networks: ict based energy allocation and storage. *Green Commun. Network.* 186–195 (2012)
14. Gelenbe, E.: The first decade of g-networks. *Eur. J. Oper. Res.* **126**(2), 231–232 (2000)

# Towards Assessment of Energy Consumption and Latency of LTE UEs During Signaling Storms

Frederic Francois, Omer H. Abdelrahman and Erol Gelenbe

**Abstract** Signaling storms in mobile networks, which congest the control plane, are becoming more frequent and severe because misbehaving applications can nowadays spread more rapidly due to the popularity of application marketplaces for smartphones. This paper aims to investigate how signaling storms affect both the energy consumption and bandwidth allocation of normal and misbehaving LTE User Equipments (UEs) by constructing a mathematical model which captures the interaction between the UE traffic and the Radio Resource Control state machine and bandwidth allocation mechanism at the eNodeB. Our results show that even if only a small proportion of the UE population is misbehaving, the energy consumption of the radio subsystem of the normal UEs can increase significantly while the time spent actively communicating increases drastically for a normal data session. Moreover, we show that misbehaving UEs have to spend an increasing amount of energy to attack the network when the severity of the signaling storms increases since they also suffer from the attacks.

**Keywords** Signaling storms · LTE · Energy consumption · Bandwidth congestion · Radio resource control · Misbehaving mobile applications · M2M systems

---

F. Francois (✉) · O.H. Abdelrahman · E. Gelenbe  
Intelligent Systems and Networks, Department of Electrical and Electronic Engineering,  
Imperial College London, London SW7 2BT, UK  
e-mail: f.francois@imperial.ac.uk

O.H. Abdelrahman  
e-mail: o.abd06@imperial.ac.uk

E. Gelenbe  
e-mail: e.gelenbe@imperial.ac.uk

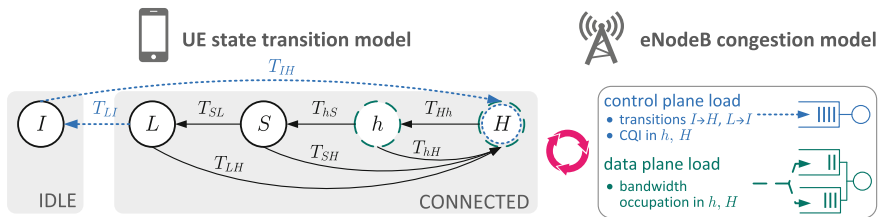
# 1 Introduction

Signaling storms [2, 9] are becoming more frequent and severe due to the large number of smartphones which have access to marketplaces where users can download both malicious [8] and non-malicious [5] applications which adversely interact with the control plane of mobile networks to cause signaling overloads. This paper investigates how a particular type of signaling storms, which leverages the Radio Resource Control (RRC) state machine of a User Equipment (UE), can cause increased energy consumption of both normal and misbehaving UEs, and severe bandwidth congestion and excessive signaling overhead at the eNodeB. This holistic investigation of the different impacts that signaling storms can have is different from previous works which investigate either the impact of different RRC-based attacks on the mobile network in terms of signaling overhead and delay only [1, 10, 12, 13, 16, 17] or the impact of traffic behavior on the energy consumption of the misbehaving UEs only without considering other UEs [3, 14, 18].

## 2 The Mathematical Model

### 2.1 Energy Consumption of the UE's Radio Subsystem for One Data Session

LTE differs from 3G in that the RRC may either be in *IDLE* (*I*) or *CONNECTED* (*C*) mode, and transitions between the 2 modes will cause signaling messages to be exchanged between the LTE UE and the mobile network [1, 7]. In the RRC *C* mode, the radio subsystem of an LTE UE has four decreasing power states: *H*, *h*, *S* and *L* as shown in Fig. 1 where:



**Fig. 1** State transition diagram representing the behavior of a UE in the different LTE RRC states (left) and its effect on the data and control planes of eNodeB (right). Different states will also correspond to different power consumption levels of the radio subsystem.



- In the  $H$  state the UE is actively communicating with the eNodeB.
- In the  $h$  state the UE has finished communicating and if new data is received before a certain timeout  $T_h$ , the UE will become active again (i.e. be in  $H$ ) otherwise the UE will be moved to the  $S$  state.
- The short DRX state  $S$  corresponds to a UE which is in short intermittent reception mode where it sleeps most of the time but wakes up periodically to check whether data is buffered at the eNodeB for its reception or it needs to transmit data to the eNodeB and if so, the UE moves to state  $H$ . If no data is communicated by the UE for a certain timeout  $T_S$ , the UE will be moved to the long DRX state  $L$ .
- The long discontinuous reception mode  $L$  is similar to  $S$  except that the UE wakes up less frequently and if the UE does not communicate within  $T_L$  seconds, it finally enters the idle mode  $I$ .

Through the use of either statistical observations or mathematical models, the following UE behavior can be obtained:

- The *average relative number of times*  $p_{X,Y}$  that a state  $Y$  is visited from state  $X$ , where  $X, Y \in \{I, H, h, S, L\}$ ,  $X \neq Y$ .
- The average time  $\tau_X$  spent in state  $X$  each time it is visited.
- The average time  $T_{XY}$  (shown in Fig. 1) that it takes to accomplish the transition from state  $X$  to state  $Y$ .
- The power consumption of the radio subsystem of the UE during each state occupation and state transition, which are denoted  $\Pi_X$  and  $\Pi_{XY}$ , respectively, for  $X, Y \in \{I, H, h, S, L\}$ .

Given the information above, the total average energy consumption of the radio subsystem of a UE for each complete cycle that starts when the UE exits state  $I$  and ends upon its next departure from  $I$  (referred to here as a *data session*) can be expressed as:

$$E = \sum_{X \in \{I, H, h, S, L\}} n_X \tau_X \Pi_X + \sum_{X, Y \in \{I, H, h, S, L\}} n_X p_{X,Y} T_{X,Y} \Pi_{X,Y} \quad (1)$$

where  $n_X$  denotes the total number of visits to state  $X$  during the cycle, which can be calculated from the relative frequency  $p_{X,Y}$  using the system of linear equations:

$$n_Y = \sum_{X, Y \in \{I, H, h, S, L\}} n_X p_{X,Y} \quad (2)$$

Notice that  $n_I = 1$  since we are considering a single data session only, and  $p_{IH} = p_{Hh} = 1$  because there is only one possible transition out of  $I$  and  $H$ . In order to obtain the remaining  $n_X$  values in terms of the relative frequencies, Eq. (2) and the state transition diagram of Fig. 1 can be used leading to:

$$n_H = n_h = \frac{1}{p_{hS} p_{SL} p_{LI}}, \quad n_S = \frac{1}{p_{SL} p_{LI}}, \quad n_L = \frac{1}{p_{LI}} \quad (3)$$

Since energy is also consumed when moving between power states, it is necessary to calculate the number of times  $n_{XY}$  that a single UE makes the transition from  $X$  to  $Y$  during one data session, which can be calculated from the relationship  $n_{XY} = n_X p_{XY}$  and using (3):

$$\begin{aligned} n_{IH} &= 1, & n_{hH} &= \frac{p_{hH}}{p_{hS} p_{SI} p_{LI}}, & n_{SH} &= \frac{p_{SH}}{p_{SL} p_{LI}}, & n_{LH} &= \frac{p_{LH}}{p_{LI}}, \\ n_{Hh} &= \frac{1}{p_{hS} p_{SL} p_{LI}}, & n_{hS} &= \frac{1}{p_{SL} p_{LI}}, & n_{SL} &= \frac{1}{p_{LI}}, & n_{LI} &= 1 \end{aligned} \quad (4)$$

Now the rate at which the normal UE makes data sessions is simply the inverse of the average session duration:

$$\lambda_n = \left[ \sum_{X \in \{I, H, h, S, L\}} n_X \tau_X + \sum_{X, Y \in \{I, H, h, S, L\}} n_X p_{XY} T_{XY} \right]^{-1} \quad (5)$$

During signaling storms, the time  $\tau_H$  that a UE spends in each visit to state  $H$  will be affected since the number of UEs which are connected to the same eNodeB and in states  $H$  or  $h$  determines the bandwidth allocated to a UE for communication; therefore, signaling storms may lead to congestion in the data plane. On the other hand, control plane congestion affects mainly  $T_{IH}$  and  $T_{LI}$  because significant signaling is required to setup and release RRC connections; the other transition times  $T_{hH}$ ,  $T_{Hh}$ ,  $T_{hS}$ ,  $T_{SL}$ ,  $T_{SH}$  and  $T_{LH}$  involve negligible signaling between the UE and eNodeB and are therefore marginally affected by the attack. This coupling between the state transitions of UEs and signaling and bandwidth congestion at the eNodeB is illustrated in Fig. 1.

## 2.2 Signaling Behavior of Misbehaving UEs

We assume that a misbehaving UE will cause excessive signaling through a continuous cycle of promotions  $I \rightarrow H$  and demotions  $L \rightarrow I$  which is triggered by sending a small amount of data when the UE is in state  $I$  and waiting for demotion back to  $I$  after undergoing the timeouts. The rate  $\lambda_a$  at which a misbehaving UE can attack is then:

$$\lambda_a = \left[ \sum_{X \in \{h, S, L\}} T_X + \sum_{XY \in \{IH, Hh, hS, SL, LI\}} T_{XY} + \tau_I^A + \tau_H^A \right]^{-1} \quad (6)$$

where the sum of the timeout periods  $T_X$  and the transition times  $T_{XY}$  represents the minimum duration that a misbehaving UE must spend before returning to  $I$ . We assume that misbehaving UEs spend negligible amount of time  $\tau_I^A \simeq 0$  in  $I$  during the storm, and  $\tau_H^A$  is the amount of time the misbehaving UE spends in  $H$ , typically to send a small amount of data to trigger a promotion from  $I$  to  $H$ .

**Table 1** General parameters for LTE network

Parameter	Value
Total no. of UEs, $\beta$	175
Total no. of eNode-Bs, $k_{\text{eNode-Bs}}$	1
Rate of CQI messages at eNodeB, $\lambda_c$	$100 \text{ s}^{-1}$
Processing time of CQI messages at eNodeB, $\delta_c$	2 ms
Inter-arrival time distribution of normal traffic bursts	Exponential
Mean inter-arrival time of normal traffic bursts, $E[\Omega]$	60 s
Total bandwidth at eNodeB, $B_e$	150 Mbps
Average size of normal bandwidth requests, $D_N$	1.88 MB
Average size of misbehaving bandwidth requests, $D_A$	10 KB
Timeout for demotion from $h$ state, $T_h$	100 ms
Timeout for demotion from <i>Short DRX</i> state, $T_S$	200 ms
Timeout for demotion from <i>Long DRX</i> state, $T_L$	10 s
Proportion of user uplink traffic, $U$	0.3
No. of signaling messages exchanged during $I \rightarrow H$ promotion, $r_{IH}$	16
Average processing time of each $I \rightarrow H$ promotion message at eNodeB, $\delta_{IH}$	3 ms
No. of signaling messages exchanged during $L \rightarrow I$ demotion, $r_{LI}$	5
Average processing time of each $L \rightarrow I$ demotion messages at eNodeB, $\delta_{LI}$	3 ms
Power consumed when in $I$ state, $\Pi_I$	0.031 W
Power consumed when in $L$ state, $\Pi_L$	1.076 W
Power consumed when in $S$ state, $\Pi_S$	1.091 W
Power consumed when in $h$ state, $\Pi_h$	1.29 W
Power consumed when transmitting at 1Mbps in $H$ , $\pi_H^{tx}$	0.438 W/Mbps
Power consumed when receiving at 1Mbps in $H$ , $\pi_H^{rx}$	0.052 W/Mbps

### 2.3 Control Plane Congestion at eNodeB

Following our approach in [1, 13] for 3G networks, the times  $T_{IH}$  and  $T_{LI}$  that it takes for a single LTE UE to promote from  $I$  to  $H$ , and to demote from  $L$  to  $I$  depend on the signaling queueing delay  $w_e$  at the eNodeB and can be characterised by the following expression:

$$T_{XY} = r_{XY} w_e + \sum_{n=1}^{r_{XY}} (t_{XY}[n] + \delta_{XY}[n]), \quad XY \in \{IH, LI\} \quad (7)$$

where  $r_{XY}$  is the number of messages sent between the UE and eNodeB (which is given in Table 1) during the transition  $X \rightarrow Y$  [7, 17], and  $t_{XY}[n]$  and  $\delta_{XY}[n]$  are the propagation delay and processing time of the  $n$ -th signaling message, respectively.

The signaling queueing delay  $w_e$  at the eNodeB can be calculated as follows: we first obtain the rate of signaling messages  $\gamma_e$  generated by a single normal UE on the eNodeB:

$$\gamma_e^N = \lambda_n[r_{IH} + r_{LI} + n_H \lambda_c(\tau_H^N + \tau_h)] \quad (8)$$

where  $\lambda_n$ , given in (5), is the rate at which a normal UE generates data sessions, and  $\lambda_c$  is the rate at which a UE sends Channel Quality Information (CQI) messages while in  $H$  and  $h$  [6] so that  $n_H \lambda_c(\tau_H^N + \tau_h)$  is the total number of CQI messages sent within a session. Similarly, the total signaling rate of a misbehaving UE during the storm is given by:

$$\gamma_e^A = \lambda_a[r_{IH} + r_{LI} + \lambda_c(\tau_H^A + \tau_h)] \quad (9)$$

From (8) and (9), it is possible to obtain the total rate of signaling messages at the eNodeB due to the whole population of UEs connected to the eNodeB as:

$$\Gamma_e = (1 - Z)\beta\gamma_e^N + Z\beta\gamma_e^A \quad (10)$$

where  $\beta$  is total number of UEs connected to the eNodeB,  $Z$  is the effective proportion of misbehaving UEs in the total population, which are only sending misbehaving traffic and no normal traffic.

From the above, the average signaling queueing delay  $w_e$  at the eNodeB can be obtained by approximating the signaling server at eNodeB by an  $M/M/1$  system [11] as in [13].

## 2.4 Data Plane Congestion at eNodeB

We now proceed with obtaining the time and energy that a UE spends in the  $H$  state, which require modeling the bandwidth congestion between the UE and the serving eNodeB in both the uplink and downlink directions. A bandwidth congestion model can be developed based on the following generic characteristics of mobile networks:

- The total uplink and downlink bandwidth of an eNodeB are different and are shared between the UEs in states  $H$  and  $h$  according to the *allocation and retention priority* (ARP) of the mobile operator.
- The bandwidth occupation in  $H$  is different from that in  $h$ .
- The traffic flow between a UE and its attached eNodeB can be asymmetric, e.g. as in web browsing and video streaming.

Furthermore, we can make the following assumptions to simplify the analysis:

- The data link between a UE and eNodeB is shared using time division duplexing.
- Requests for data communication are assumed to be made up of a fixed proportion of uplink  $U$  and downlink data  $1 - U$ .

- There are two types of bandwidth requests that UEs make during visits to  $H$ : (i) normal bandwidth requests with  $D_N$  amount of data to transfer on average and (ii) misbehaving bandwidth requests with average data transfer size of  $D_A$  with  $D_A \ll D_N$ .
- The total data bandwidth  $B_e$  of the eNodeB (both uplink and downlink) is shared equally among all the UEs which are currently in the  $H$  and  $h$  state.

The bandwidth sharing assumption above is accurate enough to capture the different effects of signaling storms on UEs but it should be noted that in an operational mobile network, users usually obtain different data rates depending on various factors such as the specific ARP used by the network operator, the Quality of Service Class Identifier of data bearers and the agreed data plan with the network operator. It is left for future work to investigate if a more complex bandwidth sharing model is necessary to fully characterize the effect of signaling storms on different types of normal users.

Based on the above discussion, an egalitarian processor sharing model [4, 15] with two classes of bandwidth requests, each requiring a different service demand, can be used to represent bandwidth congestion at the eNodeB. Specifically, the service demand  $\mu_i$  of a bandwidth request of type  $i \in \{N, A\}$  consists of two components: (i) the time it takes to actually transfer  $D_i$  bytes of data, and (ii) the time that the UE waits after active communication while reserving bandwidth, before either going into short DRX after timeout or going back into  $H$  because new data is available before timeout. Hence, the service demand of bandwidth requests is given by:

$$\mu_N = \frac{D_N}{B_e} + \tau_h, \quad \mu_A = \frac{D_A}{B_e} + T_h \quad (11)$$

The average number of concurrent normal and misbehaving bandwidth requests in an egalitarian processor sharing system are [4, 15]:

$$G_N = \frac{\rho_N}{1 - \rho_A - \rho_N}, \quad G_A = \frac{\rho_A}{1 - \rho_A - \rho_N} \quad (12)$$

where  $\rho_i = \lambda_i \mu_i$ , with  $\lambda_i$  denoting the total rate of bandwidth requests of type  $i$ :

$$\lambda_N = (1 - Z)\beta n_H \lambda_n, \quad \lambda_A = Z\beta \lambda_a \quad (13)$$

Note that data plane congestion occurs when  $\rho_N + \rho_A$  is close to 1. Using Little's theorem, the time that a UE of type  $i$  spends in *both*  $H$  and  $h$  per visit is  $\frac{G_i}{\lambda_i}$ , from which the times spent per visit to  $H$  become:

$$\tau_H^N = \frac{G_N}{\lambda_N} - \tau_h, \quad \tau_H^A = \frac{G_A}{\lambda_A} - T_h \quad (14)$$

Finally, the energy spent by the radio subsystem of a UE per visit to  $H$  is given by:

$$E_H^i = \tau_H^i [\Pi_h + U \frac{D_i}{\tau_H^i} \pi_H^{ix} + (1 - U) \frac{D_i}{\tau_H^i} \pi_H^{rx}] \quad (15)$$

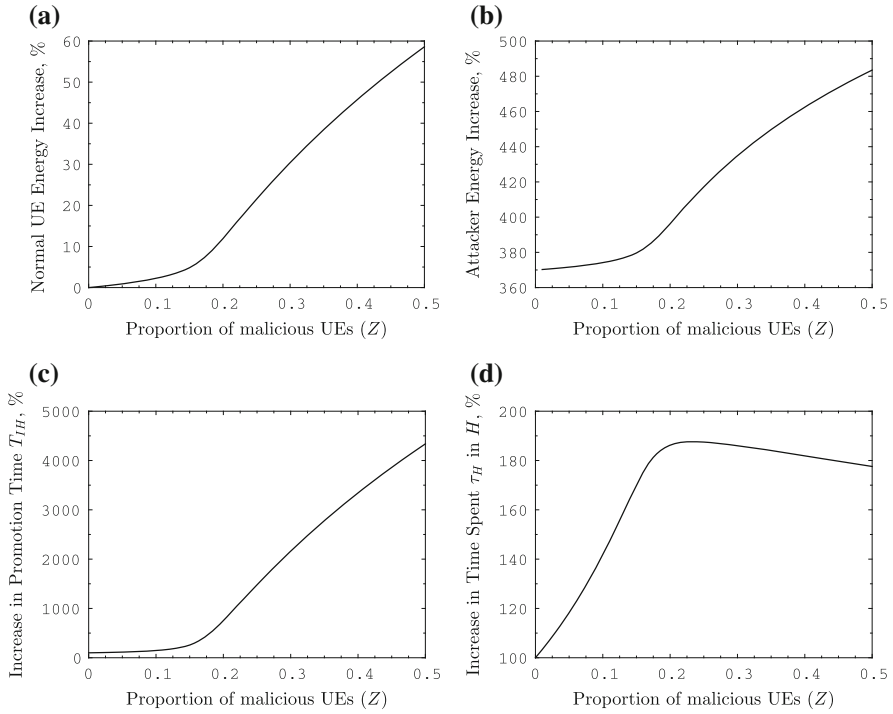
where power consumption is assumed to follow a linear relationship with throughput based on the experimental evaluation in [14]. Here,  $\Pi_h$  is the power consumed when the radio subsystem is awake but not transmitting or receiving data (i.e. the intercept), while  $\pi_H^{ix}$  (resp.  $\pi_H^{rx}$ ) is the increase in power consumption when the bit rate increases by 1Mbps in the uplink (resp. downlink). Notice that since we only consider the average throughput and energy consumption,  $\tau_H^i$  remains only in the intercept of (15).

### 3 Results

Table 1 shows the numerical values which have been chosen for each parameter of the mathematical models during the evaluation. In order to calculate the energy consumption of the radio subsystem of a UE, the estimated values obtained for the power levels  $\Pi_X$  were taken from [14] and are given in Table 1. The power expended by the radio subsystem of the UE when moving from  $I$  to  $H$  was measured to be 1.21W. We assumed that the same power is consumed by the radio subsystem when it is transitioning back from  $L$  to  $I$ . For the other transitions between power states, we use the approximation:  $\Pi_{XY} = 0.5(\Pi_X + \Pi_Y)$ .

In addition, a simple user traffic model which allows us to estimate the transition probabilities  $p_{XY}$  and the average time  $\tau_X$  spent in each visit to state  $X \neq H$  of Fig. 1 was also developed, where we represent data transfer between the UE and eNodeB as traffic bursts of average size  $D_N$  bytes with the idle time between traffic bursts (i.e. time when there is no communication between the UE and eNodeB) being a random variable  $\Omega$  which is independent of the control and data plane congestion. We chose  $\Omega$  to have an exponential distribution of mean 60 s, but any other distribution could be used to obtain the numerical values of  $p_{XY}$  and  $\tau_X$ .

Figure 2a shows that the energy consumed by the radio subsystem of a normal UE to complete one data session increases rapidly with increasing  $Z$  and therefore, signaling storms do not only have a negative impact on the access time of normal users but also on their energy consumption. Note that from Fig. 1, there are four variables which affect energy consumption and change with  $Z$ :  $\tau_H$ ,  $\tau_h$ —the times spent in  $H$ ,  $h$ , and  $T_{IH}$ ,  $T_{LI}$ —the connection setup and release times. Figure 2c, d show that both  $T_{IH}$  and  $\tau_H^N$  increase during signaling storms respectively. However,  $\tau_H^N$  drops after  $Z \simeq 0.2$  because signaling congestion becomes so high that the normal UEs find it difficult to access the network, resulting in lower bandwidth congestion.



**Fig. 2** **a** Shows the percentage increase in energy consumption of radio subsystem of a normal UE for 1 data session when  $Z$  increases, and **b** shows the corresponding percentage increase in energy consumed by a misbehaving UE. The energy consumed by a normal UE when no attack is 15.9J. **c** and **d** Show the percentage increase in the time  $T_{IH}$  spent transitioning and in active communication  $\tau_H^N$  for the normal UE, respectively, with increase in  $Z$ .

Figure 2b shows that the percentage increase in the energy consumed by the radio subsystem of a misbehaving UE during the time period of one normal data session compared to if the misbehaving UE was a normal UE increases with  $Z$ , because similar to the normal data sessions, the misbehaving UEs suffer from increased bandwidth and signaling congestion, and therefore they require more energy in order to setup and release connections. It is also worth mentioning that the increase in energy consumption by the radio subsystem of a misbehaving UE during the time period of one normal data session compared to if the misbehaving UE was idle is quite significant at 2740% when  $Z = 0.01$ . This is due to the fact that the power consumed by the radio subsystem of a UE when it is idle is much lower than the power consumed when the radio subsystem is in other power states, e.g.  $h$  state, as shown in Table 1.

## 4 Conclusions

This paper demonstrates that signaling storms do not only affect the waiting time that normal users experience before accessing the network but also significantly increase the energy consumption of the radio subsystem and the time it takes for a normal UE to communicate the same amount of data after gaining access to the network. These important results were obtained through three coupled mathematical models which capture the adverse effects of signaling storms on the control and data plane of LTE mobile networks including UEs.

**Acknowledgments** The work presented in this paper was supported by the EU FP7 research project NEMESYS (Enhanced Network Security for Seamless Service Provisioning in the Smart Mobile Ecosystem), under grant agreement no. 317888 within the FP7-ICT-2011.1.3 Trustworthy ICT domain.

## References

1. Abdelrahman, O.H., Gelenbe, E.: Signalling storms in 3G mobile networks. In: Proceedings of IEEE International Conference on Communications (ICC), pp. 1017–1022. Sydney, Australia (2014). doi:[10.1109/ICC.2014.6883453](https://doi.org/10.1109/ICC.2014.6883453)
2. Abdelrahman, O.H., Gelenbe, E., Gorbil, G., Oklander, B.: Mobile network anomaly detection and mitigation: the NEMESYS approach. In: E. Gelenbe, R. Lent (eds.) Proceedings of the 28th International Symposium on Computer and Information Sciences (ISCIS), LNEE, vol. 264, pp. 429–438 Springer (2013). doi:[10.1007/978-3-319-01604-7\\_42](https://doi.org/10.1007/978-3-319-01604-7_42)
3. Bontu, C.S., Illidge, E.: DRX mechanism for power saving in LTE. IEEE Commun. Mag. **47**(6), 48–55 (2009). doi:[10.1109/MCOM.2009.5116800](https://doi.org/10.1109/MCOM.2009.5116800)
4. Cohen, J.: The multiple phase service network with generalized processor sharing. Acta Informatica **12**(3), 245–284 (1979). doi:[10.1007/BF00264581](https://doi.org/10.1007/BF00264581)
5. Corner, S.: Angry Birds + Android + ads = network overload. IT Wire (2011). <http://www.itwire.com/business-it-news/networking/47823>
6. ETSI 3GPP: LTE; evolved universal terrestrial radio access (E-UTRA); physical layer procedures (3GPP TS 36.213 version 8.8.0 release 8) (2010)
7. ETSI 3GPP: LTE; evolved universal terrestrial radio access (E-UTRA); radio resource control (RRC); protocol specification (3GPP TS 36.331 version 12.3.0 release 12) (2014)
8. Felt, A.P., Finifter, M., Chin, E., Hanna, S., Wagner, D.: A survey of mobile malware in the wild. In: Proceedings of the 1st ACM W'shop on Security and Privacy in Smartphones and Mobile Devices (SPSM), pp. 3–14. Chicago, Illinois, USA (2011). doi:[10.1145/2046614.2046618](https://doi.org/10.1145/2046614.2046618)
9. Gabriel, C.: DoCoMo demands Google's help with signalling storm. Rethink Wireless (2012). <http://www.rethink-wireless.com/2012/01/30/docomo-demands-googles-signalling-storm.htm>
10. Gelenbe, E., Abdelrahman, O.H.: Time-outs and counters against storms (2014). (Submitted for publication)
11. Gelenbe, E., Pujolle, G.: Introduction to Queueing Networks, 2nd edn. John Wiley and Sons (1998)
12. Gorbil, G., Abdelrahman, O.H., Gelenbe, E.: Storms in mobile networks. In: Proceedings of the 10th ACM Symposium on QoS and Security for Wireless and Mobile Networks (Q2SWinet), pp. 119–126. Montreal, QC, Canada (2014). doi:[10.1145/2642687.2642688](https://doi.org/10.1145/2642687.2642688)



13. Gorbil, G., Abdelrahman, O.H., Pavloski, M., Gelenbe, E.: Modeling and analysis of RRC-based signalling storms in 3G networks. *IEEE Transactions on Emerging Topics in Computing* (2015). doi:[10.1109/TETC.2015.2389662](https://doi.org/10.1109/TETC.2015.2389662)
14. Huang, J., Qian, F., Gerber, A., Mao, Z.M., Sen, S., Spatscheck, O.: A close examination of performance and power characteristics of 4G LTE networks. In: *Proceedings of the 10th ACM International Conference on Mobile Systems, Applications, and Services (MobiSys)*, pp. 225–238. Lake District, UK (2012). doi:[10.1145/2307636.2307658](https://doi.org/10.1145/2307636.2307658)
15. Kelly, F.P.: *Reversibility and Stochastic Networks*. Wiley (1979)
16. Lee, P.P., Bu, T., Woo, T.: On the detection of signaling DoS attacks on 3G/WiMax wireless networks. *Compu. Netw.* **53**(15), 2601–2616 (2009). doi:[10.1016/j.comnet.2009.05.008](https://doi.org/10.1016/j.comnet.2009.05.008)
17. Puttonen, J., Virtej, E., Kesitalo, I., Malkamaki, E.: On LTE performance trade-off between connected and idle states with always-on type applications. In: *Proceedings of the 23rd IEEE International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, pp. 981–985. Sydney, Australia (2012). doi:[10.1109/PIMRC.2012.6362927](https://doi.org/10.1109/PIMRC.2012.6362927)
18. Zhou, L., Xu, H., Tian, H., Gao, Y., Du, L., Chen, L.: Performance analysis of power saving mechanism with adjustable DRX cycles in 3GPP LTE. In: *Proceedings of the 68th IEEE Vehicular Technology Conference (VTC Fall)*, pp. 1–5. Calgary, BC (2008). doi:[10.1109/VETEFC.2008.312](https://doi.org/10.1109/VETEFC.2008.312)

# A State-Dependent Control for Green Computing

Evsey Morozov and Alexander Rumyantsev

**Abstract** In this paper, we consider a state-dependent mechanism which can be used to reduce mean power consumption in modern communication and computer systems. This mechanism allows to change flexibly the service rate so as to increase the energy efficiency of the system. We focus on computer systems with sojourn time dependent service rate, and present a new regenerative proof of the sufficient stability conditions for a wide class of such systems. We develop both performance analysis in stationary regime and asymptotic analysis of the mean power consumption as the switching threshold increases. Some numerical results are presented and discuss in detail.

**Keywords** State-dependent system · Green computing · Power consumption · Stability · Regeneration · QoS requirement

## 1 Introduction

In this paper, we describe a state-dependent queueing system with service rate depending on the sojourn time of an arrived customer. The main motivation of this work is the so-called green computing. Constantly high energy consumption of modern data centers contrasts dramatically with their under-utilization during a considerable fraction of time. This motivates hardware vendors to imply diverse technologies allowing the data center owners to reduce their power budget. These technologies allow switching between the full capacity utilization (or even performance boost

---

E. Morozov (✉) · A. Rumyantsev  
Institute of Applied Mathematical Research of the Karelian Research Centre  
of the Russian Academy of Sciences, Petrozavodsk, Russia  
e-mail: emorozov@karelia.ru

A. Rumyantsev  
e-mail: ar0@krc.karelia.ru

E. Morozov  
Petrozavodsk State University, Petrozavodsk, Russia

states) see e.g. [11], and the low-power states, either on the CPU level [4, 10, 12], or on the level of the machine and its components [1]. However, operating these tools may cause a performance degradation. To the best of our knowledge, there is no commonly recognised scenario that guarantees optimal energy saving provided a QoS constraint. Moreover, as we show below, in some cases such scenario does not exist. In this research, following the papers [7–9], we aim a threshold-based switching policy to minimize the mean power consumption, which is an important metric of the modern server farms working under power constraints. This problem can also be reformulated as the minimization of the long-run average cost under a given policy and is closely connected with the power optimization, discussing for instance in [2, 5, 6]. We also study the dependence between the mean workload and the switching threshold in stable regime.

In Sect. 2 we apply a regenerative approach to prove sufficient stability conditions of the aforementioned state-dependent system, and it is the first main contribution of this paper. The second contribution, given in Sect. 3, is the analytical investigation of the mean power consumption reduction in the system with threshold-based sojourn time dependent service rate. Finally, in Sect. 4 we present and discuss in detail numerical results to illustrate theoretical analysis. These results show that there are non-intuitive scenarios under which the reduction of the mean power consumption is possible, while in other cases such a reduction is impossible.

## 2 Stability Analysis

We consider an initially empty single-server FIFO system with a renewal input with arrival instants  $\{t_n\}$ , inter-arrival times  $\tau_n = t_{n+1} - t_n$  and input rate  $\lambda := 1/\mathbb{E}\tau \in (0, \infty)$ , where  $\tau$  is a generic interarrival time. It is assumed that customer 0 arrives at instant  $t_0 = 0$ . Customers have i.i.d service times  $\{S_n, n \geq 0\}$  with a generic element  $S$  and the mean  $\mathbb{E}S \in (0, \infty)$ . Let  $W_n$  be the workload at instant  $t_n^-$  and  $V_n$  be the sojourn time (the workload at instant  $t_n^+$ ) of customer  $n$ , so  $V_n = W_n + S_n, n \geq 0$ . Assume that there are  $M$  thresholds,

$$0 = x_0 < x_1 < \dots < x_M < x_{M+1} := \infty,$$

such that, if  $V_n \in [x_i, x_{i+1})$ , then the service speed (rate) is  $r_i > 0, i = 0, \dots, M$ . Thus service rates  $r_i = r_i(V_n)$  depend on the accumulated work. It is assumed that service time becomes available at the arrival instants, and service rate is constant during the next inter-arrival time. (Such an opportunity exists in some modern computer systems, where the customer runtime becomes known at the arrival instant [14].) The sequence  $\{V_n\}$  constitutes a Markov chain defined by the following Lindley-type recursion

$$V_{n+1} = (V_n - r(V_n)\tau_n)^+ + S_{n+1}, \quad n \geq 0 \quad ((x)^+ := \max(0, x)).$$

The regenerations of this system are defined as

$$\theta_{n+1} = \min\{t_k > \theta_n : W_k = 0\}, \quad n \geq 0 \quad (\theta_0 := 0),$$

that is, as the arrival times of non-waiting customers. We denote a generic regeneration period by  $\theta$ . The process  $\{W_n\}$  is called *positive recurrent* if  $\mathbf{E}\theta < \infty$ . This condition is crucial for stability analysis [13]. More exactly, define the remaining regeneration time at instant  $t_n$  as

$$\theta(n) := \min_k (\theta_k - t_n : \theta_k - t_n > 0).$$

Then  $\mathbf{E}\theta < \infty$ , if  $\theta(n) \not\rightarrow \infty$  (in probability), as  $n \rightarrow \infty$ . Let  $\rho_i := \lambda \mathbf{E}S / r_i$ ,  $i = 1, \dots, M$ .

**Theorem 1** *Assume that  $\rho_M < 1$  and that*

$$\min_{1 \leq i \leq M} \mathbf{P}(r_i \tau > S) > 0. \quad (1)$$

*Then  $\mathbf{E}\theta < \infty$ .*

*Proof* Denote  $\Delta_n = V_{n+1} - V_n$ . Then for any  $x \geq 0$ ,

$$\mathbf{E}\Delta_n = \mathbf{E}(\Delta_n, V_n \leq x) + \mathbf{E}(\Delta_n, V_n > x). \quad (2)$$

Because

$$-r(V_n)\tau_n \leq S_{n+1} - r(V_n)\tau_n \leq \Delta_n \leq S_{n+1}, \quad (3)$$

then the following (uniform in  $n$ ) bounds hold:

$$-\infty < -\max_{1 \leq i \leq M} r_i \mathbf{E}\tau \leq \inf_{n \geq 0} \mathbf{E}\Delta_n \leq \sup_{n \geq 0} \mathbf{E}\Delta_n \leq \mathbf{E}S < \infty. \quad (4)$$

On the other hand, we obtain for all  $x \geq x_M$  (denoting  $r = r_M$ )

$$\begin{aligned} \mathbf{E}(\Delta_n, V_n > x) &= \int_{y > x} \mathbf{E}((y - r \tau_n)^+ + S_n - y) \mathbf{P}(V_n \in dy) \\ &= \int_{y > x} \mathbf{E}(S_n - y) \mathbf{P}(r \tau_n > y) \mathbf{P}(V_n \in dy) \\ &\quad + \int_{y > x} \mathbf{E}((S_n - r \tau_n) I(r \tau_n \leq y)) \mathbf{P}(V_n \in dy), \end{aligned} \quad (5)$$

where  $I$  is the indicator function. Assume  $V_n \Rightarrow \infty$ . Then, by (3) and (4), for each  $x$ ,

$$\begin{aligned} \mathbf{E}(\Delta_n, V_n \leq x) &= \mathbf{E}(\Delta_n | V_n \leq x) \mathbf{P}(V_n \leq x) \leq \mathbf{E} S \mathbf{P}(V_n \leq x) \rightarrow 0, \\ \mathbf{E}(\Delta_n, V_n \leq x) &\geq - \max_{1 \leq i \leq M} r_i \mathbf{E} \tau \mathbf{P}(V_n \leq x) \rightarrow 0, \quad n \rightarrow \infty. \end{aligned} \quad (6)$$

Moreover, since  $\mathbf{E} \tau < \infty$ , then

$$y \mathbf{P}(y < r \tau) = y \mathbf{P}(\tau > y/r) \rightarrow 0, \quad y \rightarrow \infty,$$

and the middle term in (5)

$$\int_{y>x} \mathbf{E}(S_n - y) \mathbf{P}(y < r \tau_n) \mathbf{P}(V_n \in dy) \rightarrow 0, \quad x \rightarrow \infty. \quad (7)$$

Because  $I(r \tau_n \leq y) \Rightarrow 1$  (in distribution), as  $y \rightarrow \infty$ , then also

$$(S_n - r \tau_n) I(r \tau_n \leq y) \Rightarrow (S - r \tau), \quad y \rightarrow \infty.$$

Since  $-r \tau_n \leq (S_n - r \tau_n) I(r \tau_n \leq y) \leq S_n$ , then the uniform integrability of the family  $\{(S_n - r \tau_n) I(r \tau_n \leq y), y \geq 0\}$  follows. Thus,

$$\mathbf{E}((S_n - r \tau_n) I(r \tau_n \leq y)) \rightarrow \mathbf{E}(S - r \tau), \quad y \rightarrow \infty, \quad (8)$$

and we can take  $y_0 \geq x_M$  so large that  $\mathbf{E}((S - r \tau) I(r \tau \leq y_0))$  is arbitrary close to  $\mathbf{E}(S - r \tau)$ . It remains to take  $n_0$  so large that  $\mathbf{P}(V_n > y_0)$  is arbitrary close to 1 for all  $n \geq n_0$ . Now we can conclude, from (2), (5)–(8), that (recall that  $r = r_M$ )

$$\mathbf{E} \Delta_n \rightarrow \mathbf{E}(S - r \tau) = r \frac{\rho_M - 1}{\lambda} < 0, \quad n \rightarrow \infty, \quad (9)$$

implying  $\limsup_{n \rightarrow \infty} \mathbf{E} \Delta_n < 0$ . It contradicts the assumption  $V_n \Rightarrow \infty$ , implying  $V_n \not\Rightarrow \infty$ . In other words,  $\inf_k \mathbf{P}(V_{n_k} \leq D) \geq \varepsilon$ , for some finite constant  $D$ , some  $\varepsilon > 0$  and a deterministic subsequence of the arrival instants  $\{t_{n_k}\}$ ,  $t_{n_k} \rightarrow \infty$ ,  $k \rightarrow \infty$ . Then, using condition (1), we can prove as in [13] that a non-waiting customer arrives in a finite interval  $[t_{n_k}, t_{n_k} + C]$  with a positive probability, and both the constant  $C$  and the probability do not depend on  $t_{n_k}$  and  $n_k$ . In other words,  $\inf_k \mathbf{P}(\theta(n_k) \leq C) > 0$ . Thus  $\theta(n) \not\Rightarrow \infty$ , and we obtain  $\mathbf{E} \theta < \infty$ .  $\square$

### 3 Performance Analysis of the Mean Power Consumption Model

In some systems the CPU may work in different regimes, alternating between the full capacity performance (e.g., Intel Turbo Boost technology) and partial performance modes, where the frequency and/or voltage decreases (DVFS technology). A widely used system software that operates performance modes is, for instance, a `cpufreq`

linux subsystem [12], which allows switching between two modes with different performance/power consumption. Although switching between these modes takes microseconds, it may give a significant energy saving. Motivated by this scenario, we consider the following particular case of the system, described in Sect. 2, with the Poisson input with rate  $\lambda$  and exponential service time with rate  $\mu = 1/ES$ . Assume that there is a single threshold  $x_M = K$  such that the service rate after arrival of customer  $n$  is  $r(V_n) := r_1 I(V_n \leq K) + r_2 I(V_n > K)$ ,  $n \geq 0$ . We also accept the natural assumption  $r_1 \leq r_2$ . Assume that stability condition  $\rho_2 < 1$  holds, see Theorem 1. Then it follows from [3] that the stationary sojourn time density  $f_{V(K)}$  and waiting time density  $f_{W(K)}$  are

$$f_{V(K)}(x) = \begin{cases} Q_1 e^{-\mu(1-\rho_1)x} + Q_2 e^{\mu\rho_2 x}, & 0 < x \leq K, \\ Q_3 e^{-\mu(1-\rho_2)x}, & x > K, \end{cases} \quad (10)$$

and

$$f_{W(K)}(x) = \begin{cases} Q_1 \rho_1 e^{-\mu(1-\rho_1)x} + Q_2 (1 + \rho_2) e^{\mu\rho_2 x}, & 0 < x \leq K, \\ Q_3 \rho_2 e^{-\mu(1-\rho_2)x}, & x > K, \end{cases} \quad (11)$$

where

$$\begin{aligned} Q_1 &:= Q_3 \frac{1}{\rho_2 - \rho_1 + 1} e^{\mu(\rho_2 - \rho_1)K}, & Q_2 &:= Q_3 \frac{\rho_2 - \rho_1}{\rho_2 - \rho_1 + 1} e^{-\mu K}, \\ Q_3 &:= \mu \left[ \frac{\rho_1 - \rho_2}{\rho_2(\rho_2 - \rho_1 + 1)} e^{-\mu K} - \frac{\rho_1 - \rho_2}{\rho_2(1 - \rho_1)(1 - \rho_2)} e^{-\mu(1-\rho_2)K} \right. \\ &\quad \left. + \frac{1}{(1 - \rho_1)(\rho_2 - \rho_1 + 1)} e^{-\mu(\rho_1 - \rho_2)K} \right]^{-1}, \end{aligned}$$

and moreover,

$$P(W^{(K)} = 0) = \frac{Q_1 + Q_2}{\mu}. \quad (12)$$

Denoting

$$\varphi(c, a) := \int_0^K cx e^{ax} dx = \frac{c}{a^2} (aK - 1) e^{aK} + \frac{c}{a^2},$$

we obtain from (11) the following (convenient for calculation) expression for the mean stationary workload

$$\begin{aligned} EW^{(K)} &= \varphi(Q_1 \rho_1, \mu(\rho_1 - 1)) + \varphi(Q_2 (1 + \rho_2), \mu\rho_2) \\ &\quad - \varphi(Q_3 \rho_2, \mu(\rho_2 - 1)) + \frac{Q_3 \rho_2}{\mu^2 (\rho_2 - 1)^2}. \end{aligned}$$

We remark that this model becomes classic (state-independent)  $M/M/1$  system with traffic intensity  $\rho_2$ , if either  $K = 0$ , or  $\rho_1 = \rho_2 < 1$ , in which case (in an evident notation)

$$\begin{aligned}
f_{V^{(0)}}(x) &= \mu(1 - \rho_2)e^{-\mu(1-\rho_2)x}, \quad x > 0, \\
f_{W^{(0)}}(x) &= \mu\rho_2(1 - \rho_2)e^{-\mu(1-\rho_2)x}, \quad x > 0, \\
\mathbf{E}V^{(0)} &= \frac{1}{\mu(1 - \rho_2)}, \quad \mathbf{E}W^{(0)} = \frac{\rho_2}{\mu(1 - \rho_2)}, \\
\mathbf{P}(W^{(0)} = 0) &= 1 - \rho_2.
\end{aligned} \tag{13}$$

Denote by  $c_0$  the power consumed in an empty state, and let  $c_j$  be the power consumption for the service rate  $r_j$ ,  $j = 1, 2$ . (The power consumption is often linearly connected with the CPU frequency [5], in which case we may take  $c_i = r_i$ .) By the property PASTA,  $\mathbf{P}(W = 0)$  is also the limiting fraction of the idle time. (The same interpretation holds true for other stationary probabilities.) Then the mean power consumption is defined as

$$\begin{aligned}
\mathcal{E}^{(K)} &= c_0\mathbf{P}(W^{(K)} = 0) + c_1\mathbf{P}(W^{(K)} > 0, V^{(K)} \leq K) \\
&\quad + c_2\mathbf{P}(W^{(K)} > 0, V^{(K)} > K).
\end{aligned} \tag{14}$$

We note that

$$\begin{aligned}
\mathbf{P}(W^{(K)} > 0, V^{(K)} \leq K) &= \mathbf{P}(V^{(K)} \leq K) - \mathbf{P}(W^{(K)} = 0, S \leq K) \\
&= \frac{Q_1(e^{\mu(\rho_1-1)K} - 1)}{\mu(\rho_1 - 1)} + \frac{Q_2(e^{\mu\rho_2K} - 1)}{\mu\rho_2} \\
&\quad - \frac{Q_1 + Q_2}{\mu}(1 - e^{-\mu K}),
\end{aligned} \tag{15}$$

while

$$\begin{aligned}
\mathbf{P}(W^{(K)} > 0, V^{(K)} > K) &= \mathbf{P}(V^{(K)} > K) - \mathbf{P}(W^{(K)} = 0, S > K) \\
&= \frac{Q_3e^{-\mu(1-\rho_2)K}}{\mu(1 - \rho_2)} - \frac{Q_1 + Q_2}{\mu}e^{-\mu K}.
\end{aligned} \tag{16}$$

It is easy to see that, for  $K = 0$ , relations (14)–(16) give the following mean power consumption in the system  $M/M/1$  with service rate  $r_2$ ,

$$\mathcal{E}^{(0)} = c_0(1 - \rho_2) + c_2\rho_2, \tag{17}$$

which is consistent with (13).

Now we develop the asymptotic analysis of the mean power consumption  $\mathcal{E}^{(K)}$  as the threshold  $K \rightarrow \infty$ . We consider two main cases: (i)  $\rho_1 > 1$  and (ii)  $\rho_1 < 1$ .

(i) In this case, after some algebra, we obtain from (10) and (12) that

$$\mathbf{P}(V^{(K)} > K) = \frac{Q_3e^{-\mu(1-\rho_2)K}}{\mu(1 - \rho_2)} \rightarrow \frac{\rho_2(\rho_1 - 1)}{\rho_1 - \rho_2}, \tag{18}$$

and  $\mathbf{P}(W = 0) \rightarrow 0$ , as  $K \rightarrow \infty$ . Because

$$\lim_{K \rightarrow \infty} \mathbf{P}(V^{(K)} \leq K) = \frac{\rho_1(1 - \rho_2)}{\rho_1 - \rho_2}, \quad (19)$$

it then follows from (14) that, provided  $\rho_1 > 1$ ,

$$\lim_{K \rightarrow \infty} \mathcal{E}^{(K)} = c_1 \frac{\rho_2(\rho_1 - 1)}{\rho_1 - \rho_2} + c_2 \frac{\rho_1(1 - \rho_2)}{\rho_1 - \rho_2}. \quad (20)$$

(It is important to stress that the limits in (18)–(20) are taken in such a way that first, for each  $K$ , the system achieves stability, and then we let  $K \rightarrow \infty$ .) Relation (20) has clear intuitive interpretation: for each fixed  $K$ , the stationary process  $V^{(K)}$  lives in a *neighborhood of the threshold*  $K$ . At that, the 1st and the 2nd terms in (20) represent the fraction of the consumption when the workload is *below*  $K$  and *above*  $K$ , respectively. An asymmetry of (20) (in comparison with (14)) is caused by the influence of the *lower reflecting barrier*  $W^{(K)} = 0$  in (14), which is non-negligible for  $K$  relatively small. Moreover, in this case, after a routine algebra, we obtain the following asymptotic expression for the mean workload,

$$\mathbf{E}W^{(K)} = K - \frac{1 - \rho_1 - \rho_2}{\mu} \left( \frac{2 - \rho_1}{1 - \rho_2} + \frac{1}{\rho_2(1 - \rho_1)} \right) + o(1), \quad K \rightarrow \infty, \quad (21)$$

that is, the mean workload increases (approximately) linearly in  $K$ .

(ii) In this case we rewrite (14) as (15) and (16)

$$\mathcal{E}^{(K)} = \frac{Q_1 + Q_2}{\mu} (c_0 - c_1 + e^{-\mu K} (c_1 - c_2)) + c_1 + \frac{Q_3 e^{-\mu(1-\rho_2)K}}{\mu(1 - \rho_2)} (c_2 - c_1),$$

implying

$$\lim_{K \rightarrow \infty} \mathcal{E}^{(K)} = c_0(1 - \rho_1) + c_1\rho_1. \quad (22)$$

Moreover, after some algebra, one obtains the following asymptotic expression for the mean workload:

$$\mathbf{E}W^{(K)} = \frac{\rho_1}{1 - \rho_1} + o(1), \quad K \rightarrow \infty. \quad (23)$$

To explain (22) and (23), we note that, by stability condition  $\rho_1 < 1$ , the stationary process  $V^{(K)}$  is tight regardless of  $K$ , and, thus  $\lim_{K \rightarrow \infty} \mathbf{P}(V^{(K)} > K) = 0$ . As a result, the original system approaches the state-independent system  $M/M/1$  with service rate  $r_1$  and traffic intensity  $\rho_1$ .

The developed performance analysis allows in general to select both the rates  $r_i$  and the threshold  $K$  in such a way to reduce the mean power consumption  $\mathcal{E}^{(K)}$  (in comparison with state-independent system with rate  $r_2$ ), keeping a given QoS



requirement. For instance, we could select these parameters to satisfy the upper bound of the expected stationary workload,  $\mathbf{E}W^{(K)} \leq \mathbf{E}W^{(0)}(1 + \delta)$ , for a given  $\delta > 0$ , or, alternatively, the upper bound of the overflow probability

$$\mathbf{P}(V^{(K)} > K_0) \leq \gamma, \quad (24)$$

for a given (small) probability  $\gamma > 0$  and a level  $K_0$ . To illustrate the theoretical analysis, we consider an important particular case

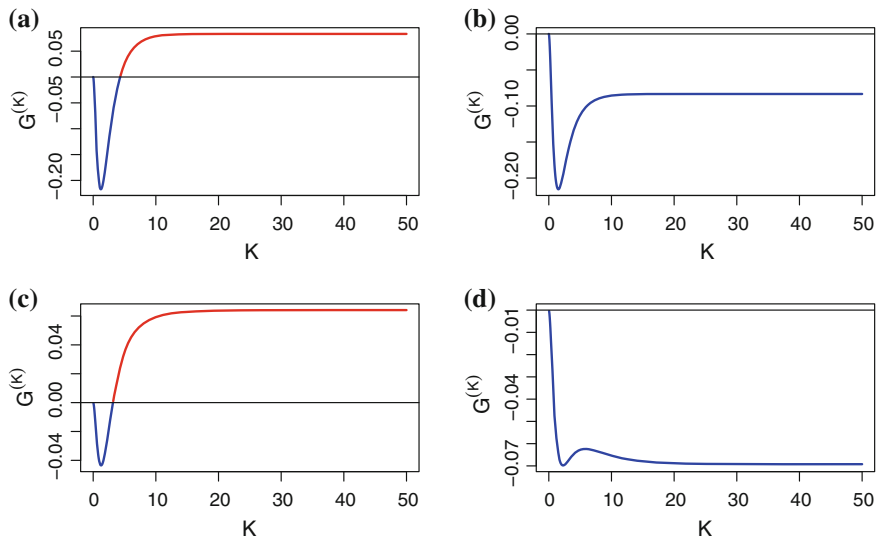
$$\frac{r_1}{r_2} > \frac{c_1 - c_0}{c_2 - c_0}. \quad (25)$$

(This relation has a clear intuitive interpretation and holds true if the power consumed is linearly connected with the CPU frequency.) It is easy to check in this case that, by (17), (20), and (22), the gain  $G^{(K)} := \mathcal{E}^{(0)} - \mathcal{E}^{(K)} \rightarrow C > 0$ , as  $K \rightarrow \infty$  ( $C$  is an explicit constant), and thus there exists a threshold  $K^*$  such that  $G^{(K)} > 0$  for all  $K \geq K^*$ . Note that if  $\rho_1 > 1$ , then we anyway must use a finite threshold  $K$  (and speed  $r_2$ ), to avoid instability. However, as (18) shows, the requirement (24) and positivity of  $G^{(K)}$  may be inconsistent. Conversely, if  $\rho_1 < 1$  then, under condition (25), both requirements (24) and  $G^{(K)} > 0$  are satisfied for all  $K$  large enough.

## 4 Numerical Analysis

We numerically illustrate our performance analysis by Fig. 1, where, for  $\lambda = \mu = c_0 = 1$ ,  $c_1 = 2$ , we depict  $G^{(K)}$  versus  $K$  for the following sets of parameters  $(r_1, r_2, c_2)$ : (a) (0.5, 2, 5.5); (b) (0.5, 2, 4.5); (c) (1.3, 3, 3.5); (d) (1.3, 3, 3.1). These four cases correspond to the following alternatives: (i)  $\rho_1 > 1$  for (a), (b);  $\rho_1 < 1$  for (c), (d); and (ii) condition (25) holds for (a), (c); (25) is violated for (b), (d). As Fig. 1 shows, the gain is positive only under condition (25) (cases (a) and (c)). In general, as we see, a dependence of  $G^{(K)}$  on the parameters is rather complicated (in particular,  $G^{(K)}$  is not monotone, as one might expect).

Now we consider simulation results in more detail. It is worth mentioning that, as it will be shown below, the asymptotic results are achieved in all scenarios with a high accuracy for  $K$  relatively small. In case (a), the workload grows linearly in  $K$ , since the system is unstable below  $K$ . By this reason we are forced to use the higher speed  $r_2$  to stabilize the system, however to have a gain, we must take  $K > 10$ . In case (b),  $G^{(K)}$  is negative for any  $K > 0$ , and thus value  $K = 0$ , that is constant speed  $r_2$ , minimizes the energy consumption. In case (c), the system is stable under speed  $r_1$ , and one can see that the gain becomes positive for  $K$  relatively small. However, if in this case  $\rho_1 \approx 1$ , then the QoS requirement (24) might be essential for the final decision. In case (d), like in case (b), the quantity  $G^{(K)}$  is always negative, and constant speed  $r_2$  becomes less consuming. The main reason of the negativity

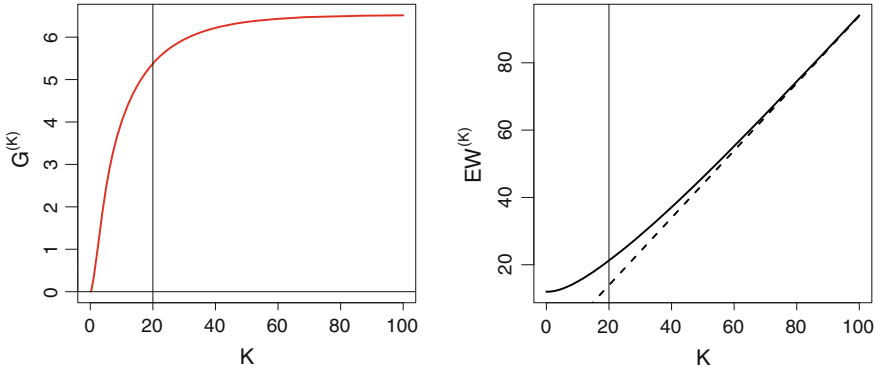


**Fig. 1** The gain  $G^{(K)}$  versus threshold  $K$ . (i)  $\rho_1 > 1$ : (a), (b);  $\rho_1 < 1$ : (c), (d).  
(ii) Condition (25) holds: (a), (c); (25) is violated: (b), (d)

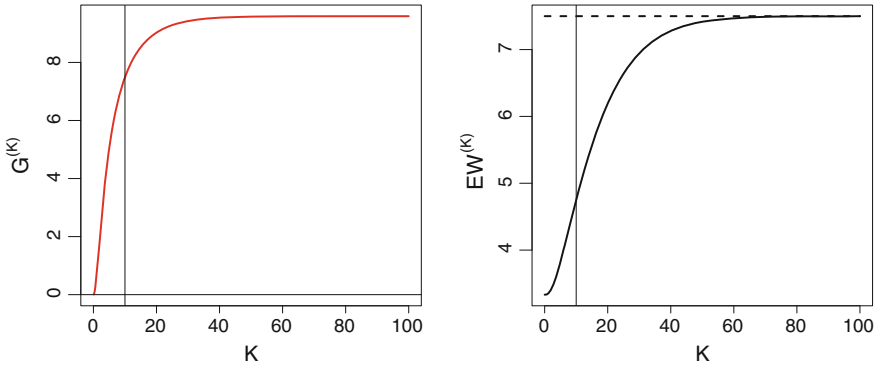
of  $G^{(K)}$  in this case is that the slowly working system becomes idle less often, while the idle state power consumption is significantly less than that in active regimes.

Below we present simulation results for a quad-core Intel Core i7 model of 2014 year. We consider the following settings: the normal working frequency is 3.4 GHz, while the Turbo Boost regime frequency is 3.9 GHz. The idle power consumption (in C1E state) is 26 W, normal power consumption is 84 W, and the Turbo Boost consumption is 105 W. In other words, we take  $r_1 = 3.4$ ,  $r_2 = 3.9$ ,  $c_0 = 26$ ,  $c_1 = 84$ ,  $c_2 = 105$ . Also we take  $\mu = 1$ . Now we consider two scenarios: (i)  $\lambda_1 = 3.6$ , implying  $\rho_1 = \lambda/r_1 > 1$ , implying instability at normal speed (Fig. 2); (ii)  $\lambda = 3$ , implying  $\rho_1 < 1$  (Fig. 3). It is seen that in case (i) there exists a tradeoff between the gain  $G^{(K)}$  and stationary mean workload  $\mathbf{E}W^{(K)}$ . That is, in scenario (i) we have to use the Turbo Boost regime, since the system is unstable, otherwise. In this scenario we compare the system in Turbo Boost regime (with rate  $r_2$ ) with the system, which uses two regimes ( $r_1, r_2$ ) and switching at the threshold  $K$ . It is seen, that relatively small  $K$  (e. g.,  $K = 20$ , see Fig. 2) may give a tradeoff between the mean workload and the gain.

Eventually, in case (ii) the system is stable at normal speed, but switching to a Turbo Boost regime (say, at level  $K = 10$ , see Fig. 3) may be still useful to decrease the mean workload comparing to the system with constant normal speed. On the other hand, this decision will also decrease the gain.



**Fig. 2** The gain  $G^{(K)}$  (left) and mean workload  $EW^{(K)}$ , case (i)  $\rho_1 > 1, \lambda = 3.6$



**Fig. 3** The gain  $G^{(K)}$  (left) and mean workload  $EW^{(K)}$ , case (ii)  $\rho_1 < 1, \lambda = 3$

### 5 Conclusion

We propose and analytically investigate the mean power consumption of the model describing a server with the threshold-based sojourn time dependent service rate. Sufficient stability conditions of such a model are proved by the regenerative method. Moreover, stationary performance analysis is developed for the model with exponential distributions and two different sojourn time dependent service rates. The numerical results are given, which illustrate a complicated dependence between the mean power consumption gain and given system parameters. This work shows an opportunity of an effective state-dependent control for green computing, however a detailed analysis of various scenarios, including energy optimization under QoS constraint, requires a future research.

**Acknowledgments** This research is supported by Russian Foundation for Basic Research, projects 13-07-00008, 14-07-31007, 15-07-02341, 15-07-02354, 15-07-02360, and the Program of Strategic development of Petrozavodsk State University.

## References

1. Advanced Configuration & Power Interface. <http://www.acpi.info/spec.htm>
2. Alonso, M., et al.: Power saving in regular interconnection networks. *Parallel Comput.* **36**, 696–712 (2010) (Elsevier, Amsterdam)
3. Bekker, R., Borst, S.C., Boxma, O.J., Kella, O.: Queues with workload-dependent arrival and service rates. *Queueing Syst.* **46**, 537–556 (2004)
4. Enhanced Intel SpeedStep Technology. <http://www.intel.com/cd/channel/reseller/ASMO-NA/ENG/203838.htm>
5. Gandhi, A., et al.: Optimal power allocation in server farms. *ACM SIGMETRICS Perform. Eval. Rev.* **37**, 157–168 (2009) (ACM, Seattle)
6. Gandhi, A., et al.: Power capping via forced idleness. In: *Proceedings of Workshop on Energy Efficient Design*, pp. 1–6. Austin, Texas (2009)
7. Gandhi, A., Harchol-Balter, M., Adan, I.: Server farms with setup costs. *Perform. Eval.* **67**, 1123–1138 (2010)
8. Hyytiä, E., Penttinen, A., Aalto, S.: Size- and state-aware dispatching problem with queue-specific job sizes. *Eur. J. Oper. Res.* **217**, 357–370 (2012)
9. Hyytiä, E., Richter, R., Aalto, S.: Task assignment in a heterogeneous server farm with switching delays and general energy-aware cost structure. *Perform. Eval.* **75–76**, 17–35 (2014)
10. IBM EnergyScale for POWER7 Processor-Based Systems. <http://www-03.ibm.com/systems/power/hardware/whitepapers/energyscale7.html>
11. Intel TurboBoost Technology. <http://www.intel.ru/content/www/ru/ru/architecture-and-technology/turbo-boost/turbo-boost-technology.html>
12. Linux CPUFreq. <https://www.kernel.org/doc/Documentation/cpu-freq/governors.txt>
13. Morozov, E., Delgado, R.: Stability analysis of regenerative queueing systems. *Autom. Remote Control* **70**(12), 1977–1991 (2009)
14. Tsafirir, D., Etsion, Y., Feitelson, D.G.: Modeling user runtime estimates. Job scheduling strategies for parallel processing. *LNCS* **3834**, 1–35 (2005)

# Environment Friendly Energy Efficient Distributed Data Centers

Ahsan Ali and Oznur Ozkasap

**Abstract** Geographically distributed data centers form a significant technology used by the Internet users to fulfil the demand of storage, processing and large scale computations. Most of the operational cost of such data centers is due to the electricity cost, which affect both service providers and consumers. In this paper, we addressed the problem of energy consumption of data center entities and reviewed state-of-the-art solutions proposed to reduce the electricity cost. We present the full view of the problem by providing the widely used energy consumption and/or operational cost models. We identified key characteristics of efficient techniques proposed for reduction of the electricity cost, carbon emission and financial penalties in case of SLA violations. These techniques include environment friendly cost minimization, energy efficient load migration, job scheduling and resource allocation. We also identified open challenges as guidelines for future research.

**Keywords** Distributed data centers · Environment friendly · Energy efficient

## 1 Introduction

In the present era of technology, cloud data centers are becoming more popular due to the features such as on demand computing resources, user only pay for what they use, multiple users can use the same physical infrastructure, and high computational power. Cloud Service Providers (CSP) are responsible for managing and allocating resources, Scheduling of tasks, power consumption optimization and management of network traffic. The operational cost in data center is affected by the power consumption. A data center's 30–50 % operational cost consists of electricity bills [1].

---

A. Ali (✉) · O. Ozkasap  
Department of Computer Engineering, Koc University, Istanbul, Turkey  
e-mail: aali14@ku.edu.tr

O. Ozkasap  
e-mail: oozkasap@ku.edu.tr

**Table 1** Characterization of cost minimization techniques

Category	Description
Environment friendly cost minimization	Encouraging use of renewable energy sources for power generation
Energy efficient load migration	Migrating load from higher electricity price to cheap electricity price data centers
Energy efficient job scheduling	Utilizing spatio temporal variation of electricity prices
Energy efficient resource allocation	Efficient allocation of data center resources and using energy efficient entities

To reduce the energy cost, different techniques have been proposed, that includes use of energy efficient servers, server consolidation by moving virtual machines (VM), using renewable energy sources such as solar, wind, wave, tidal with traditional electricity grid power, exploiting the temporal and geographical variation of electricity prices. There are two different charging schemes being used in electricity market of US: retail pricing and wholesale pricing. In retail pricing, electricity prices are announced in advance and remain constant for a certain period, while in wholesale market electricity prices vary on 15 min to one hour basis. Different geographical locations and regions exhibit different price variations on daily as well as hourly basis. Therefore, geographically distributed nature of data centers and the spatio-temporal variation of prices is being used by researchers to minimize the energy cost.

We group cost minimization techniques for distributed data centers into four categories (Table 1), identify key characteristics in each category along with the proposed solutions, and discuss open challenges. These aspects distinguish our review in comparison to the previous studies. For instance, green energy aware power management problem for data centers is investigated in [2]. Another prior work [3] focuses mostly on reducing the losses and wastage of electricity by supporting subsystems in data center for minimizing the cost, while [4, 5] focus on energy consumption of data center networks.

## 2 Modelling of Energy Consumption

In the system model, we assume a CSP with  $M$  geographically distributed data centers consisting of  $N$  servers in each. There exist  $K$  regions generating  $\lambda$  requests with deadline restriction  $\mu$  agreed in the SLA. Each request is initially received by a front-end server or scheduler, which assigns the request to a server in a particular data center. The unit price of the electricity in location of data center  $M_j$  is expressed as  $\alpha_j$ . Total electricity consumed in  $M_j$  is the summation of energy consumed by the servers, energy consumed by the cooling systems and the energy consumed by the network elements. Electricity generation from carbon-intensive fossil fuels is

expensive due to high fuel prices and carbon emission taxes. While green energy sources are cheap but highly unpredictable for the unexpected load spikes.

Power usage efficiency (PUE) is a benchmark for calculating the power efficiency of a data center, which is defined as the ratio of overall power consumption to power consumed for performing actual computation by different entities of data centers. Usually the value of PUE is between 1 and 2, the lower the value of PUE the better is the energy efficiency. Qureshi et al. [6] modelled the power consumption of a server  $S_i$  as:

$$P_{O_i} = P_{idle} + (P_{peak} - P_{idle}) \times U_i$$

where  $P_{idle}$  is the average power consumed by the server when there is no operation being performed,  $P_{peak}$  is the power consumption of the server while running at full capacity and  $U_i$  is the utilization of  $S_i$ . The electricity cost of  $S_i$  can be calculated by the product of unit price of electricity and the total power consumed:

$$E_{S_i} = P_{O_i} \times \alpha_i$$

The cooling system of a data center is responsible for maintaining the temperature within a safety zone. Guo et al. [7] modelled the electricity consumption of a cooling system defined as the ratio of power consumed by active servers in data center  $M_j$  to the coefficient of performance (COP).

$$E_{C_j} = \frac{\sum_{i=1}^{i=N} E_{S_i}}{COP}$$

Suppose the cost of transferring a single request from the scheduler to the data center is  $R_c$ . Electricity cost of network element for the data center  $M_j$  can be calculated as the product of the total number of requests being transferred to the data center and the cost of single request:

$$E_{N_j} = \lambda_j \times R_c$$

Total electricity cost of a data center with  $N$  servers can be calculated as the summation of individual electricity cost of all servers, network elements and overall cooling cost. Electricity cost of data center  $M_j$  can be represented as:

$$E_{Tc_j} = \left( \sum_{i=1}^{i=N} E_{S_i} \right) + E_{C_j} + E_{N_j}$$

While providing the services, CSP needs to maintain the quality of service (QoS) mentioned in Service Level Agreement (SLA). QoS includes the fairness among users and the delay constraints. In case of any violation of the SLA, CSP faces a financial penalty which results in an additional cost for them.

### 3 Environment Friendly Cost Minimization

It is suggested that the carbon footprints in the process of electricity generation can be reduced by using the green reusable energy resources such as wind, solar, and hydro. Table 2 represents an overview of the research carried out to minimize the energy cost by using green energy sources with other constraints. *Workload prediction* techniques are being used to find out the future workload based on the previous requests and workload. *Latency* is also taken into consideration for SLA and QoS purposes, which refers to the time required to transfer the user request from user's end to the data center. *Queuing delay* refers to the time a user request waits before it is processed. As data centers host multiple users and applications, it is necessary that each user would get equal time for processing and *fair* resource allocation. *Server provisioning* is the technique to monitor the activities of servers in a data center to avoid overloading or under loading.

The way geographical load balancing (GLB) can enhance the use of green renewable energy and minimize the use of carbon intensive fuels for electricity generation is explored in [8]. Two distributed algorithms for achieving optimal GLB are proposed and the feasibility of powering the data center entirely with on-site renewable energy source (i.e. wind or solar) is explored. It is shown that by efficient use of dynamic electricity pricing GLB provides significant reductions in brown energy use. This work jointly optimizes energy cost and delays keeping in consideration price diversity and network delay diversity. However, study of [8] ignores some aspects such as reliability and availability of on-site renewable energy sources.

A request routing framework named FORTE (Flow Optimization based framework for requesting-Routing and Traffic Engineering) is proposed in [9]. FORTE consists of three algorithms providing three way trade-off among electricity cost, access latency and carbon footprint. Although the work of [8] considered the cost for the carbon emission, [9] argues that the carbon emission taxes are negligible as compared to the overall cost of data center. However, the cost of carbon tax considered by Gao et al. [9] is less than 2 % of overall electricity cost which is far less than the real carbon market [2].

On-line algorithms to minimize number of active servers in each data center to fulfill load requirement are suggested in [10]. Receding horizon control (RHC)

**Table 2** Characteristics of environment friendly cost minimization

Methods	Workload prediction	Latency	Queuing delays	Fairness	Server provisioning	Carbon emission taxes
GLB [8]	×	✓	✓	×	✓	✓
FORTE [9]	×	✓	✓	×	×	✓
RHC/AFHC [10]	✓	✓	✓	×	✓	×
Green fair [11]	✓	✓	✓	✓	×	✓



is proposed to determine the minimum number of active servers by predicting the future workload for homogeneous data centers. It is shown that Performance of RHC is better for homogeneous environment, however it is also proved that in case of heterogeneous setting RHC performs poorly. Average Fixed Horizon Control (AFHC) is suggested for heterogeneous environment which guarantees good performance in such settings. This study also proposed the idea of powering the data centers entirely using green renewable energy sources.

A novel approach of using green renewable energy resources to minimize the queuing delays by scheduling the jobs to the data centers closer to the user is proposed in [11], which incorporates SLAs. The proposed green fair algorithm ensures fairness amongst the consumers providing the service capacity constraints and latency constraints. Optimal electricity, social and latency costs are sought within fairness constraints agreed in the SLAs.

### 4 Energy Efficient Load Migration

Techniques in this category investigate how much workload should be scheduled in the current time slot and how much to be put in a queue and move to another data center in a future time slot. In an event of price change of electricity, the user requests as well as VMs are moved to a data center with a cheaper electricity price. Table 3 characterises the main features of Energy Efficient Load Migration. Delaying the requests might lead to miss the deadline resulting an SLA violation and incurring financial penalty. The SLA penalties increase the overall cost for the CSP. *Bandwidth cost* is the network cost while transferring the user requests or VMs from one data center to another. *Load of data center* refers to the number of requests handled by the data center in the current time slot.

An online algorithm for migrating batch jobs between data centers is proposed in [12]. The proposed solution considers multiple energy sources at each data center with continuous variation of prices and availability. Migration decision is based on current cost as well as the future uncertainty of electricity price and availability.

**Table 3** Characteristics of energy efficient load migration

Methods	SLA penalties	Workload prediction	Latency	Bandwidth cost	Load of data center
Efficient online algorithm [12]	×	×	×	✓	✓
Dispatcher/datacenter level sched. [13]	✓	×	✓	×	✓
CP-LNS [14]	×	✓	×	✓	✓
VR-HM [15]	✓	✓	×	✓	✓

Bandwidth cost, as an important factor in job migration is included in overall cost calculation.

A cost reduction off-line algorithm is proposed in [13] for assignment and migration of requests depending on the electricity price and load of data center. The theoretical information is then used to design an online algorithm for GLB, migration and error prediction. Performance of the suggested algorithm depends on the duration between price fluctuation and the quality of error prediction. This algorithm does not consider server provisioning, heterogeneity, and availability of the renewable sources.

By using the migration of VMs and outside temperature for reducing the cooling cost, a solution for cost minimization problem is proposed in [14]. Inter and intra data center migrations of VMs are used to reduce the cost by following the lowest energy price in the geographically distributed data centers. By using workload prediction, assignment of the VMs for the longer time frame is performed which reduces the migration cost.

A migration method for VMs, namely Virtual machine Resizing Heuristic Migration (VR-HM), to handle server failure and SLA is proposed by Ghoreyshi and Mohammad [15]. For tolerating the high failure rate in cloud environment, a heuristic migration model is presented for online migration of VMs between data centers by considering energy and deadlines constraints when failures occur. There is a local manager for each data center which monitors the utilization, fault occurrence and energy consumption of VMs, and issues orders for VM resizing and migrations.

## 5 Energy Efficient Job Scheduling

The aim of research in this category is to find the data center for processing user requests with minimum possible electricity cost. It is also necessary to meet the QoS constraint while scheduling the requests to data centers. Table 4 summarizes the main features of energy efficient job scheduling with the key common factors to maintain the QoS and minimize the cost. In order to reduce the cost, both *inter data center* and *intra data center* constraints should be considered. *Cooling cost* is also an important factor in the overall electricity cost of the data center.

A latency and cost request dispatching policy is presented in [6]. This price aware optimizer observes the locational time-varying fluctuation of electricity prices and distributes the user requests to data centers with the cheapest price within a maximum radial geographical distance. While routing the requests to data centers, the algorithm uses two parameters: a distance threshold and a price threshold. Apart from reducing the electricity cost by dispatching the request to cheaper data centers, the scheduler also reduces the energy consumption.

Study of [16] optimizes the intra data center workload distribution for the efficiency of cooling system, and divides the data center into three temperature zones based on the airflow: hot, warm and cool. The hot zone requires more power to maintain the temperature as compared to the cool zone. Therefore, requests are assigned

**Table 4** Characteristics of energy efficient job scheduling

Method	SLA penalties	Data center load	Latency	Inter data center	Intra data center	Cooling cost
Distance construction electric price optimizer [6]	✓	×	✓	✓	×	×
Power trade/surge guard [16]	×	✓	×	✓	×	✓
Brenner's algorithm [17]	✓	✓	×	✓	×	×
Energy price driven dispatcher [18]	✓	✓	✓	✓	×	×
GreFar [19]	✓	✓	×	✓	×	×
Cheapest-DC/S [20]	✓	×	×	✓	×	✓
JET [7]	×	✓	×	✓	✓	✓

to servers accordingly to reduce the cooling cost. In order to reduce the overall cost, two schemes are proposed: power trade and surge guard. Power trade dynamically distributes the load between and within zones based on the temperature of each zone. Surge guard improves the response time of servers in case of an unexpected load spike, and uses server provisioning to meet the load requirements.

The electricity cost under multiple electricity market is aimed to be minimized in [17]. The problem is formulated as the minimum cost flow problem and Brenne's algorithm is used to find the solution in polynomial time. The algorithm assigns job requests according to the electricity prices, data center workload and it saves the energy consumption by keeping minimum number of active servers to meet the delay requirement.

An energy-price-driven dispatcher, that forwards the client requests to data centers with cheap electricity price and can meet the latency threshold, is presented in [18]. No request is assigned to the data center if the resource utilization exceeds 80 % which is not being considered in [6]. The results are compared with the policy proposed by Qureshi et al. [6] and random dispatching policy, which show that in terms of cost the policy of [6] is more efficient but it may overload the cheapest data centers as it does not consider the load of data centers while assigning the requests.

GreFar algorithm is proposed by Ren et al. [19] that minimizes the energy cost and ensures fairness between data center consumers with keeping in consideration queuing delay. GreFar schedules requests based on the trade-off between cost-delay and energy-fairness. With the cost-delay parameter, the jobs are scheduled to avoid the queuing delay, to meet the delay constraint or when electricity prices meets the cost constraint. In the energy-fairness scheduling, all the resources are allocated based on a specific fairness weight. The energy cost marginally increases in energy fairness but the average delay decreases. If these two parameters are fine-tuned,

this scheduling scheme will provide better fairness, minimum queuing delay and minimum cost.

A recent study [20] suggests using the outside temperature for the cooling of data centers along with temporal and spatial variation of electricity prices. Two types of algorithms for scheduling the batch jobs are proposed. Immediate scheduling algorithm schedules jobs on the basis of FCFS among the available cheapest data center. Delayed scheduling algorithm schedules the jobs in a future time slot where the electricity price is cheaper. Besides, the SLA penalties are also included in the cost calculation in case of any SLA violations.

As the scheme in [16] only focuses on workload management within data center and also does not consider electricity price variation in geographically distributed data centers, the recent work of [7] proposed a joint inter- and intra-data center workload management scheme called JET, which considers both electricity price variation among the distributed data centers and the efficiency of cooling system. It reduces the cooling cost by minimizing the number of active servers which results in reducing the electricity cost.

## 6 Energy Efficient Resource Allocation

In recent years, several efficient resource allocation techniques have been proposed which consist in multitasking on each server to reduce the energy consumption, efficient dynamic allocation of virtual machines, and adapting the rate of operations of active servers according to the load (DVFS). Resource allocation, when done efficiently, can save the cost for both the user and the provider as well as it can also improve the QoS in terms of response time and related parameters. Table 5 highlights the main features of energy efficient resource allocation. *Virtualization* is the technique of running more than one VM on a server. *Server provisioning* refers to monitor the activities of servers in a data center to avoid overloading or under loading. *DVFS* is the technique to control the CPU frequency according to the current processing and load.

An automated server provisioning system (ACES) proposed in [21] faces three way trade-off between cost, performance, and reliability to perform energy-aware server provisioning. It has two key components, load predictor and optimization framework, to reduce the energy consumption, meet the load requirement and minimize the reliability cost.

A scheduling scheme, taking into consideration electricity price, carbon emission rate, workload, server power efficiency and the deadline of the requests, is presented in [22]. Various meta scheduling policies are provided to maximize the profit, minimize the carbon emission and reduce the power consumption of servers by varying the CPU frequency (DVFS). The meta scheduler finds an optimum CPU frequency for servers in order to process as many requests as possible without missing request deadlines.

**Table 5** Characteristics of energy efficient resource allocation

Methods	SLA penalties	Virtualization	Server provisioning	DVFS
ACES [21]	✓	×	×	✓
Meta scheduling [22]	✓	×	×	✓
ECE-CIS [23]	✓	✓	×	✓
Cost and deadline optimization [24]	✓	✓	✓	×
Fair [25]	✓	×	✓	×
Cloud aware scheduling algorithm [26]	✓	✓	✓	✓
MPC [27]	✓	×	✓	×

A carbon efficient heuristic for VM placement is presented in [23]. The users send their requests with the predefined requirements to the cloud provider. The heuristic places the VM in geographically distributed data centers taking into account the carbon footprint rates, energy sources, QoS and PUE in order to minimize the overall electricity cost.

A resource allocation model in which a CSP and CSC are encouraged to share the information for the efficient resource allocation is presented in [24]. The Asymmetry Algorithm is proposed for allocating the resources to different users according to the capacity and the load. VMs are migrated according to the server utilization and putting servers on standby mode or idle mode. If a requested job deadline is below a threshold, it is immediately scheduled, otherwise the job is not scheduled until it meets the cost constraints.

The cost minimization, fair request rate allocation problem with SLA constraints and spatio temporal variation of electricity price are studied in [25]. Distributed data centers and the scheduler with requests to be allocated are represented as minimum cost multi commodity flow problem. An algorithm named Fair is proposed that is based on the optimization framework, which determines the number of servers to be in active state in each data center to meet the workload and SLA requirements.

The intelligent scheduling is combined with DVFS to utilize cheap electricity as well as to reduce the electricity consumption of the servers in [26], that also considered server consolidation which helps in processing user requests before the defined deadlines.

The cost minimization problem from a resource buffering prospective is tackled in [27]. The novel idea of buffering electricity in batteries while the electricity cost is low and use this buffered power to operate the servers while the prices are higher is proposed. A power management controller per battery is used to make the decision of charging batteries or use them as a power source for servers. They adopt the model

predictive control (MPC) for smart charging under the varying price of electricity. MPC predicts the future electricity cost of server by solving the current electricity price and the battery power levels.

## 7 Conclusions

In this study, we analysed different techniques used to reduce the electricity cost of geographically distributed data centers. First, we presented the model of energy consumption with some basic knowledge of data center infrastructure, energy sources, different electricity markets and carbon emission taxes. We classified the current research to reduce the data center operational cost into four categories and identified the key characteristics in each category. Open issues which need attention are optimizing the use of green energy sources for powering data center entirely and efficient methods for reducing delays along with the cost minimization especially for delay sensitive requests. Data center cooling systems need to be optimized to reduce the cooling cost. These problems have not been fully analysed and can be investigated as future research directions.

**Acknowledgments** This work was partially supported by the COST (European Cooperation in Science and Technology) framework, under Action IC0804, by TUBITAK (The Scientific and Technical Research Council of Turkey) under Grant 109M761, and by HEC (Higher Education Commission of Pakistan).

## References

1. Koomey, J.: In growth in data center electricity use 2005 to 2010. Analytics Press, Oakland. <http://www.analyticspress.com/datacenters.html> (2011)
2. Kong, F., Liu, X.: A survey on green-energy-aware power management for datacenters. *ACM Comput. Surv. (CSUR)* **47**(2) (2014)
3. Mastelic, T., et al.: Cloud computing: survey on energy efficiency. *ACM Comput. Surv. (CSUR)* **47**(2) (2014)
4. Aksanli, B., et al.: Utilizing green energy prediction to schedule mixed batch and service jobs in data centers. *ACM SIGOPS Operating Syst. Rev.* **45**(3), 53–57 (2012)
5. Bianzino, A.P., et al.: A survey of green networking research. *IEEE Commun. Surv. Tutorials* **14**(1), 3–20 (2012)
6. Qureshi, A., et al.: Cutting the electric bill for internet-scale systems. *ACM SIGCOMM Comput. Commun. Rev.* **39**(4) (2009)
7. Guo, Z., et al.: JET: Electricity cost-aware dynamic workload management in geographically distributed datacenters. *Comput. Commun.* **50**, 162–174 (2014)
8. Liu, Z., et al.: Greening geographical load balancing. *ACM SIGMETRICS International Conference on Measurement and Modelling of Computer Systems* (2011)
9. Gao, P.X., et al.: It's not easy being green. *ACM SIGCOMM Comput. Commun. Rev.* **42**(2), 211–222 (2012)
10. Lin, M., et al.: Online algorithms for geographical load balancing. *International Green Computing Conference (IGCC)* (2012)

11. Alawnah, R.Y., et al.: Green and fair workload distribution in geographically distributed data-centers. *J. Green Eng.* **4**, 69–98 (2014)
12. Buchbinder, N., Jain, N., Menache, I.: Online job-migration for reducing the electricity bill in the cloud. *IFIP Networking* 172–185 (2011)
13. Adnan, M.A., Sugihara, R., Gupta, R.K.: Energy efficient geographical load balancing via dynamic deferral of workload. In: *IEEE 5th International Conference on Cloud Computing (CLOUD)*, (2012)
14. Mehta, D., O’Sullivan, B., Simonis, H.: Energy cost management for geographically distributed data centres under time-variable demands and energy prices. In: *IEEE/ACM 6th International Conference on Utility and Cloud Computing (UCC)* (2013)
15. Ghoreyshi, S. M.: Energy-efficient resource management of cloud datacenters under fault tolerance constraints. In: *IEEE Green Computing Conference (IGCC)* (2013)
16. Ahmad, F., Vijaykumar, T.N.: Joint optimization of idle and cooling power in data centers while maintaining response time. *ACM Sigplan Not.* **45**(3) (2010)
17. Rao, L., et al.: Minimizing electricity cost: optimization of distributed internet data centers in a multi-electricity-market environment. *IEEE INFOCOM* (2010)
18. Yamada, H., et al.: Request dispatching for cheap energy prices in cloud data centers. In: *IEEE 2nd International Conference on Cloud Networking (CloudNet)* (2013)
19. Ren, S., He, Y., Xu, F.: Provably-efficient job scheduling for energy and fairness in geographically distributed data centers. In: *IEEE 32nd International Conference on Distributed Computing Systems (ICDCS)* (2012)
20. Guler, Huseyin, B. Barla Cambazoglu, and Ozgur Ozkasap. “Cutting Down the Energy Cost of Geographically Distributed Cloud DataCenters.” *Energy Efficiency in Large Scale Distributed Systems*. Springer Berlin Heidelberg, 279–286 (2013)
21. Guenter, B., Jain, N., Williams, C.: Managing cost, performance, and reliability tradeoffs for energy-aware server provisioning. *IEEE INFOCOM* (2011)
22. Garg, S.K., et al.: Environment-conscious scheduling of HPC applications on distributed cloud-oriented data centers. *J. Parallel Distrib. Comput.* **71**(6), 732–749 (2011)
23. Khosravi, Atefeh, Saurabh Kumar Garg, and Rajkumar Buyya. “Energy and carbon-efficient placement of virtual machines in distributed cloud data centers.” *Euro-Par 2013 Parallel Processing*. Springer Berlin Heidelberg, 317–328 (2013)
24. Joy, J., Krishna Kumar, L.: Cost and deadline optimization along with resource allocation in cloud computing environment. In: *IEEE International Conference on Advanced Computing and Communication Systems (ICACCS)* (2013)
25. Xu, Z., Liang, W.: Minimizing the operational cost of data centers via geographical electricity price diversity. In: *IEEE Sixth International Conference on Cloud Computing (CLOUD)* (2013)
26. Calheiros, R.N., Buyya, R.: Energy-efficient scheduling of urgent bag-of-tasks applications in clouds through DVFS. In: *IEEE 6th International Conference on Cloud Computing Technology and Science (CloudCom)* (2014)
27. Yao, J., Liu, X., Zhang, C.: Predictive electricity cost minimization through energy buffering in data centers. *IEEE Trans. Smart Grid* **5**(1), 320–338 (2014)

# Modeling Power Consumption in Multicore CPUs with Multithreading and Frequency Scaling

D. Cerotti, M. Gribaudo, P. Piazzolla, R. Pincioli and G. Serazzi

**Abstract** The rapid growth of energy requirements in large data-center has motivated several research projects focusing on the reduction of power consumption. Several techniques have been studied to tackle this problem, and most of them require simple power models to estimate the energy consumption starting from known system parameters. It has been proven that the CPU is the component of a server that is most responsible for its total power consumption: for this reason several power models focusing on this resource has been developed. However, only a few accounts for standard CPU features like dynamic frequency scaling and hyperthreading, which can have a significant impact on the estimation accuracy. In this paper, we present the results from a set of experiments focusing on these CPU features, and we propose a simple power model able to provide accurate power estimates by taking them into account.

## 1 Introduction

Power consumption reduction in datacenter is one of the most important research topics that is being addressed in different ways and with different techniques by many scientists both from the industry and the academia. However, most of the

---

D. Cerotti (✉) · M. Gribaudo · P. Piazzolla · R. Pincioli · G. Serazzi  
Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano,  
via Ponzio 34/5, 20133 Milano, Italy  
e-mail: [davide.cerotti@polimi.it](mailto:davide.cerotti@polimi.it)

M. Gribaudo  
e-mail: [marco.gribaudo@polimi.it](mailto:marco.gribaudo@polimi.it)

P. Piazzolla  
e-mail: [pietro.piazzolla@polimi.it](mailto:pietro.piazzolla@polimi.it)

R. Pincioli  
e-mail: [riccardo.pincioli@polimi.it](mailto:riccardo.pincioli@polimi.it)

G. Serazzi  
e-mail: [giuseppe.serazzi@polimi.it](mailto:giuseppe.serazzi@polimi.it)



proposals require a suitable power model that can estimate the energy consumption starting from simple system parameters, like the utilization of its resources, and the time required to complete the considered tasks. Especially when the considered techniques are based on results that exhibit non-linear behavior, the accuracy of the power consumption model is of paramount importance to correctly identify the optimal system configurations that can achieve the target performance with the lowest possible energy budget. In different works, such as [9, 16], the instantaneous power consumption of a server was observed to have a linear relation with the CPU utilization. Such works present also a non-linear estimation, which however requires the computation of a coefficient that must be extrapolated from a large set of measurements collected from the analyzed system. Both models consider the CPU as the main index of the server activity. Since the CPU is involved in managing any server components, there is a strong correlation between the utilization of CPU and the utilization of all other resources. For such reason, the CPU utilization alone is representative of the server activity and thus of the entire power consumption.

The typical use of analytic expressions for the power computation is the inclusion of energy characterization in models of the system defined using suitable formalisms, such as queuing networks, Petri nets, and so on. Quite often the system models are designed to evaluate techniques or policies for energy saving aiming to analyze the benefits obtained on a single server. In most of the cases, the system models can provide estimates not only for the utilization, but also for other configuration settings such as the number of cores used and the frequency at which the system is running. Thus, taking into account widespread commercial power reduction techniques, like dynamic and voltage frequency scaling (DVFS) and hyperthreading, we can provide a more detailed power consumption expressions that can compute more accurate estimates.

In this paper we will focus on multicore CPUs that exploit hyperthreading, i.e., each physical core can execute two or more threads simultaneously. This number of thread is usually referred to as *Simultaneous MultiThreading (SMT) level*. The operating system (OS) can thus concurrently execute as many threads as the product of the number of cores and the SMT level. Moreover, the OS can control the so called *clock governor* by setting the maximum and minimum frequency that the DVFS mechanism can use. Initially, we will present a set of experiments designed to assess the impact on the power consumption of the number of cores used, of the level of SMT exploited and of the maximum clock frequency set. Then, we will show how the models presented in [9] may be affected by large relative errors if the above described features are not considered correctly. We finally propose a new power consumption model that can effectively take into account the considered CPU settings to provide more accurate energy estimations.

The paper is organized as follows. The main related work of the literature is described in Sect. 2. In Sect. 3, we describe the setup of the performed experiments. Section 4 presents and analyses the experimental results. Section 5 proposes various power models of increasing complexity able to capture the impact of hyperthreading and dynamic frequency scaling. The conclusion of the paper is provided in Sect. 6.

## 2 Related Work

As data-centers continue to consume an ever-greater share of the world's electricity, estimated to be 91 billion kWh of electricity in 2013 only [12], a huge amount of researches are dedicated to improve the energy efficiency and develop effective energy-preserving strategies. These researches has produced a wide spectrum of techniques that exploit, albeit in different ways, two basic mechanisms: the dynamic scaling of components performances (Dynamic Speed Scaling) and the dynamic hibernation of components (Dynamic Resource Sleeping), see e.g. [2, 11, 15].

In the first studies that applied power control techniques on specific components of data-centers [9], the power consumption of a whole server was assumed proportional to the CPU utilization. The evidence of such a behavior was investigated and confirmed in further studies, see e.g. [16]. Recently, in order to provide more accurate estimates, authors started to isolate also the contribution of other components to the whole energy consumption [7]. Some of these newer works focus on CPU, see e.g., [17], other focus on memory, e.g. [10], while others on disks, [6].

To the best of our knowledge, only a few works propose power models taking into account hyperthreading and dynamic voltage and frequency scaling (DVFS) and none considers both of them together. For instance, [14] proposes an operating system facility to profile the power requirement of server requests, but without considering hyperthreading nor DVFS. Instead, [19] provides a detailed hyperthread-aware model to profile the power consumption of individual application, but ignore the DVFS impact. Such effect may be significant, as shown in [3] where the authors combine CPU indexes with instantaneous voltage demand information to improve the DVFS setting and achieve an average speedup of 7.3% over Windows Vista default DVFS algorithm.

In the literature, energy consumption behavior in data-centers has been modeled using different evaluation formalisms or metric models. Petri Nets are used in works like [8] where the authors, using a non-Markovian Stochastic Reward Networks modeling approach, investigate power-performance efficiency. Queueing theory has been widely applied for investigating optimal power allocation and load distribution in clouds, like e.g. [5]. In works like e.g.: [1] or [13], queueing theory is exploited for proposing tools to manage energy efficiency in datacenters. However, the effectiveness of such model may be hindered without a proper estimate of the power consumption of an individual server.

## 3 Experimental Setup

We perform the power evaluation experiments on two different servers. The first called *Server1* is a PC with an i3-2120 CPU@3.3GHz. with two CPU cores and 6 GB of RAM, an hard disk drive and an integrated graphic card. *Server2* is a PC with an i7-3770 CPU@3.4GHz with four cores and 16 GB including two hard disk

drives and a solid state disk with a dedicated GeForce GTX 560 graphic card. Both machines have an SMT level of two, meaning, for instance, that the i3-2120 CPU architecture provides two cores to the user, but is capable to concurrently run a total of four threads.

In all experiments the servers run CPU-bound applications, in particular instances of the Sunflow benchmark from the daCapo [4] suite. It renders a set of images using ray tracing, splitting the load into several concurrent threads. The number of threads used by the benchmark is set equal to the maximum number of threads that the server is capable to run concurrently. The server OS is Linux Ubuntu, but with different version: Ubuntu 12.04 for *Server1*, 14.04 for *Server2*. In both versions the maximum CPU frequency can be directly set editing the “scaling\_max\_freq” file. Moreover, it is possible to manually turn on and off each available threads editing the “online” file in “/sys/devices/system/cpu”. It is also allowed to separately set the maximum frequency at which each thread run and switching off the multithreading on a specific CPU core.

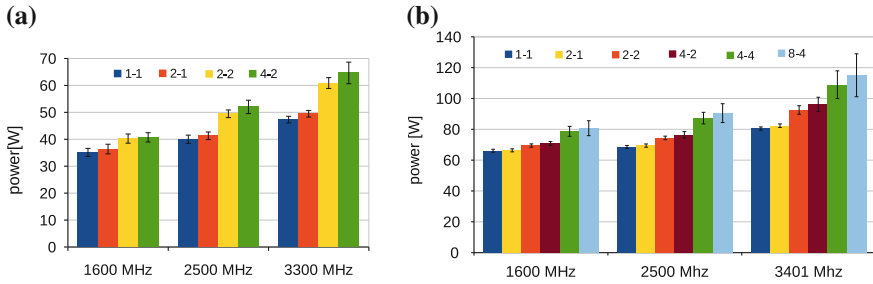
The power consumption during each experiment is periodically measured by the the Yokogawa WT210 digital power meter [18] directly attached to the analysed server. The CPU utilization is monitored by the `io_stat` command which periodically produces CPU statistics calculated as the average among all threads. Finally, the CPU frequency is monitored by inspecting the “/proc/cpu\_info” file.

## 4 DVFS and Hyperthreading

In this Section we analyse the power consumption of a CPU-bound application taking in consideration two main technologies integrated in the CPUs: DVFS and hyperthreading. We expect a significant effect of such technologies on the power consumption. In fact, in a CMOS circuit the dynamic power [2] can be computed as  $P_{dyn}(U) = \alpha \cdot V^2 fr$ , where  $\alpha$  is a positive constant. Thus, changing the clock frequency  $fr$  or the voltage  $V$  will clearly have an energetic impact. Furthermore, in hyperthreading, emulating an additional CPU core by concurrently executing two hardware threads should consume less power than using two physical cores. However, in both cases such effects are not captured by considering the CPU utilization only.

To investigate such phenomena, we measure through the power meter the server power consumption of the Sunflow benchmark fixing in each execution both the maximum CPU frequency and the number of threads used. We perform a full factorial experiment on the machine *Server1* with a maximum CPU frequency ranging on the set {1600, 2500, 3300} MHz. By selecting as maximum 1600 MHz, the frequency scaling is disabled; with 2500 MHz., during the execution of the benchmark, the CPU frequency may span from 1600 to 2500 MHz., and so on. The number of concurrent threads ranges from 1 to 4. For each configuration 20 executions have been performed.

The histograms in Fig. 1a show the resulting power consumption with 95 % confidence intervals. As expected, the power linearly increases with respect to the maximum CPU clock frequency. Such increment grows according to the number



**Fig. 1** CPU-bound application power varying CPU frequency with several CPU cores and hyper-threading: **a** dual-core; **b** quad-core CPU. In the legend, a configuration with a number  $x$  of threads running on  $y$  CPU cores is denoted by the label  $x - y$

of parallel threads used; for instance, using four of them the increment from 1600 to 3300 MHz is slightly more than 50 % (40–63 W). Instead, the power does not increase linearly according to the number of threads used, but it follows a step behavior with a gap of nearly 10 W from the second to third configurations. Indeed, the configurations with 1 or 2 threads involve just a single CPU core, while in the remaining ones both CPU cores work together, thus consuming more power.

An analogous set of experiments was performed on the *Server2* machine with a maximum CPU frequency ranging on the set {1600, 2500, 3401} MHz. and the number of parallel threads ranging from 1 to 8. As shown in Fig. 1b, a similar behavior is obtained. In this case, a further gap is present using two or four CPU cores.

## 5 Power Modeling

The first power model proposed in [9] for the estimate of warehouse-sized data-center consumption was the following:

$$P(U_{CPU}) = P_{Idle} + U_{CPU} \cdot (P_{Max} - P_{Idle}) \quad (1)$$

where  $U_{CPU}$  is the CPU utilization,  $P_{Idle}$  is the power consumed when no user applications are running and  $P_{Max}$  is the maximum power consumption when the server is 100 % utilized. In the same work also a more accurate non-linear estimate was proposed:

$$P(U_{CPU}) = P_{Idle} + (2U_{CPU} - (U_{CPU})^r) \cdot (P_{Max} - P_{Idle}) \quad (2)$$

with  $r$  experimentally computed by data collected from the analyzed system.

We compute the values of the parameters  $P_{Max}$ ,  $P_{Idle}$  and  $r$  by a fitting procedure from the collected experiments. The fitting procedure was performed by the DEPS Evolutionary Algorithm, a variant of the DEPSO algorithm proposed in [20], for

**Table 1** Fitted parameters

Cores	Power model	U Def.	$P_{Idle}$	$P_{Max}$	$\alpha_{th}$	$\alpha_{cr}$	$\eta$	$r$	Error (%)
4	Eq. 1	Eq. 3	32.99	52.884	–	–	–	–	13.756
4	Eq. 2	Eq. 3	32.99	52.931	–	–	–	0.934	13.754
4	Eq. 1	Eq. 4b	28	61.363	–	–	–	–	4.12
4	Eq. 2	Eq. 4b	28	62.855	–	–	–	0.938	4.273
4	Eq. 8	–	30.256	67.143	0.277	0.71	1.533	–	1.14
8	Eq. 1	Eq. 3	65.147	92.761	–	–	–	–	9.943
8	Eq. 2	Eq. 3	64.804	92.675	–	–	–	1.045	9.928
8	Eq. 1	Eq. 4b	60.905	105.293	–	–	–	–	4.712
8	Eq. 2	Eq. 4b	65.225	11.216	–	–	–	0.731	4.456
8	Eq. 8	–	63.708	115.514	0.364	0.723	2.101	–	2.99

non-linear optimization. It integrates together the Differential Evolution and Particle Swarm Optimization techniques. To prevent over-fitting, we use half of the collected measurements to perform the fitting, thus computing the value of the parameters. The remaining half is used to validate the obtained model and compute the resulting mean absolute percentage error (MAPE).<sup>1</sup>

The straightforward application of such models to estimate the power consumption of the experiments in Sect. 4 provides results with a very low MAPE. Indeed, during all the experiments the CPU utilization is nearly 100 %, thus according to Eq. 1 the energetic power consumption should be maximum, even if the server is not fully exploited since only a fraction of the available computational power (i.e. CPU frequency or number of threads and cores) is used.

As a first solution to avoid such problem, we weight the CPU utilization with the fraction of used threads:

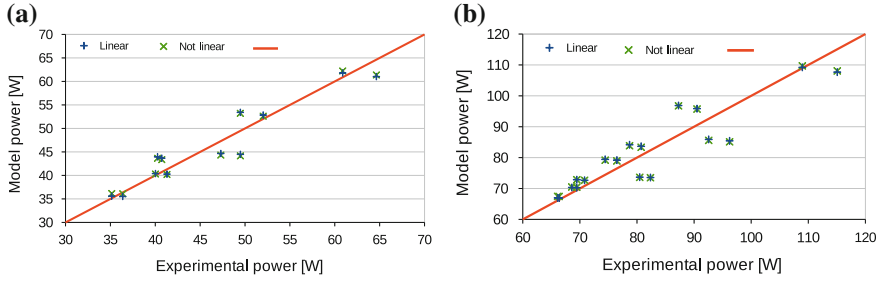
$$U(C_{th}) = U_{CPU} \frac{C_{th}}{T_{th}} \quad (3)$$

where  $C_{th}$  is the number of used threads and  $T_{th}$  is the total number of threads that the CPU is capable to run concurrently. Using Eq. 3 to compute the models of Eqs. 1 and 2 of the dual cores CPU, the values of MAPE obtained are 13.756 and 13.754 %, respectively, as shown in Table 1.

We can improve the results considering also the CPU frequency, thus defining the value of  $U_{CPU}$  as either one of the two following expressions:

$$U(fr, C_{th}) = U_{CPU} \frac{C_{th}}{T_{th}} \frac{fr}{fr_{Max}} \quad (a), \quad U(fr, C_{cr}) = U_{CPU} \frac{C_{cr}}{T_{cr}} \frac{fr}{fr_{Max}} \quad (b) \quad (4)$$

<sup>1</sup>The mean absolute percentage error is defined as:  $MAPE = \frac{1}{N} \sum \left| \frac{A_i - F_i}{A_i} \right|$ , where  $A_i$  is the actual value and  $F_i$  is the estimated one.



**Fig. 2** Comparison of simple linear and non linear models with experimental results: **a** dual-core; **b** quad-core CPU

where  $C_{th}$  ( $C_{cr}$ ) is the number of used thread (cores),  $T_{th}$  ( $T_{cr}$ ) is the total number of threads (cores). Moreover,  $fr$  and  $fr_{Max}$  are the average and maximum CPU frequency, respectively. We will call  $O = \{C_{th}, C_{cr}, fr\}$  the operative parameters which may be set by software applications. Instead  $M = \{T_{th}, T_{cr}, fr_{Max}\}$  is the set of machine parameters which are fixed and depend on the HW characteristics of the server.

We performed the aforementioned fitting procedure between the power values collected in the experiments and the model of Eq. 1 (called Linear) and Eq. 2 (Not linear), for both the expressions in Eq. 4. Figure 2 shows a comparison plot between the experimental and the estimated power consumption fitted using  $U(fr, C_{cr})$ , which achieves the best accuracy between the two expressions in Eq. 4. Both the linear and not linear models achieve nearly the same accuracy, in particular the linear model has a MAPE of 4.12 and 4.712 % for the dual-core and quad-core CPU, respectively.

Our proposal to further improve the results is introducing the scaling factor:

$$\Delta(O, M) = \left( \frac{C_{th}}{T_{th}} \alpha_{log} + \frac{C_{cr}}{T_{cr}} \alpha_{cr} \right) \left( \frac{fr}{fr_{Max}} \right)^\eta \quad (5)$$

where  $\alpha_{th}$ ,  $\alpha_{cr}$  and  $\eta$  are coefficients evaluated by the fitting procedure.

We define the maximum power consumption  $P_{Max}(O, M)$ , dependent on the current operative conditions and the server features, as:

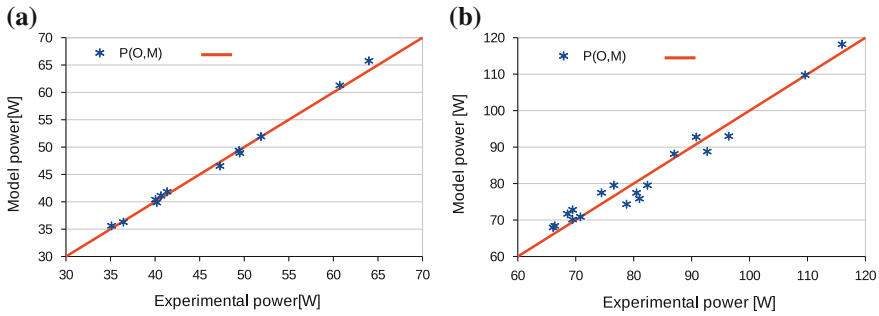
$$P_{Max}(O, M) = P_{Idle} + \Delta(O, M) \cdot (P_{Max} - P_{Idle}) \quad (6)$$

Finally, we define the power value  $P(O, M)$  as:

$$P(O, M) = P_{Idle} + (P_{Max}(O, M) - P_{Idle}) U_{CPU} \quad (7)$$

Substituting Eq. 6 in Eq. 7 we obtain:

$$P(O, M) = P_{Idle} + \Delta(O, M) \cdot (P_{Max} - P_{Idle}) U_{CPU} \quad (8)$$



**Fig. 3** Comparison of proposed model with experimental results: **a** dual-core; **b** quad-core CPU

As before, we perform a fitting procedure with the DEPS Algorithm to compute the coefficients  $\alpha_{th}$ ,  $\alpha_{cr}$ ,  $P_{Idle}$  and  $P_{Max}$ . The values obtained are shown in Table 1. The comparison plot between the resulting estimate and the experiment results is shown in Fig. 3. The accuracy is further improved with a MAPE of 1.14 % for the dual core and 2.99 % for the quad core.

## 6 Conclusion

In this work we have considered the power consumption of CPUs with multiple cores, hyperthreading and dynamic frequency scaling. The next step will be including different types of resources such as disks, memory, and network: even if their contribution to the power budget of a server is much smaller than the CPU it is still important to include them for having a complete picture of the energy consumption of a server. The Graphic Processing Unit (GPU) can require even more energy than the CPU: this device however is needed only for very specific tasks like video-transcoding, and currently it is only marginally used in most of servers in a data-center. Finally, the impact of virtualization should be considered, since the “trap-and-execute” technique employed by most of the virtual machine managers correlates the virtualized resources and the utilization of the CPU on the hosts, leading to different energy foot-prints.

## References

1. Aljohani, A., Holton, D., Awan, I., Alanazi, J.: Performance evaluation of local and cloud deployment of web clusters. In: Network-Based Information Systems (NBIS), 2011 14th International Conference on, pp. 274–278 (2011). doi:[10.1109/NBIS.2011.47](https://doi.org/10.1109/NBIS.2011.47)
2. Beloglazov, A., Buyya, R., Lee, Y.C., Zomaya, A., et al.: A taxonomy and survey of energy-efficient data centers and cloud computing systems. *Adv. Comput.* **82**(2), 47–111 (2011)

3. Bircher, W.L., John, L.: Predictive power management for multi-core processors. In: Proceedings of the 2010 International Conference on Computer Architecture, ISCA'10, pp. 243–255. Springer, New York (2012). doi:[10.1007/978-3-642-24322-6\\_21](https://doi.org/10.1007/978-3-642-24322-6_21). [http://dx.doi.org/10.1007/978-3-642-24322-6\\_21](http://dx.doi.org/10.1007/978-3-642-24322-6_21)
4. Blackburn, S.M., Garner, R., Hoffmann, C., Khang, A.M., McKinley, K.S., Bentzur, R., Diwan, A., Feinberg, D., Frampton, D., Guyer, S.Z., Hirzel, M., Hosking, A., Jump, M., Lee, H., Moss, J.E.B., Phansalkar, A., Stefanović, D., VanDrunen, T., von Dinklage, D., Wiedermann, B.: The dacapo benchmarks: Java benchmarking development and analysis. SIGPLAN Not. **41**(10), 169–190 (2006). doi:[10.1145/1167515.1167488](https://doi.org/10.1145/1167515.1167488)
5. Cao, J., Li, K., Stojmenovic, I.: Optimal power allocation and load distribution for multiple heterogeneous multicore server processors across clouds and data centers. IEEE Trans. Comput. **63**(1), 45–58 (2014). doi:[10.1109/TC.2013.122](https://doi.org/10.1109/TC.2013.122)
6. Chen, D., Goldberg, G., Kahn, R., Kat, R., Meth, K.: Leveraging disk drive acoustic modes for power management. In: Mass Storage Systems and Technologies (MSST), 2010 IEEE 26th Symposium on, pp. 1–9 (2010). doi:[10.1109/MSST.2010.5496993](https://doi.org/10.1109/MSST.2010.5496993)
7. Economou, D., Rivoire, S., Kozyrakis, C.: Full-system power analysis and modeling for server environments. In: Workshop on Modeling Benchmarking and Simulation (MOBS) (2006)
8. Escheikh, M., Jouini, H., Barkaoui, K.: A versatile traffic and power aware performability analysis of server virtualized systems. In: Modelling, Analysis Simulation of Computer and Telecommunication Systems (MASCOTS), 2014 IEEE 22nd International Symposium on, pp. 207–212 (2014). doi:[10.1109/MASCOTS.2014.34](https://doi.org/10.1109/MASCOTS.2014.34)
9. Fan, X., Weber, W.D., Barroso, L.A.: Power provisioning for a warehouse-sized computer. SIGARCH Comput. Archit. News **35**(2), 13–23 (2007). doi:[10.1145/1273440.1250665](https://doi.org/10.1145/1273440.1250665). <http://doi.acm.org/10.1145/1273440.1250665>
10. Kant, K.: A control scheme for batching dram requests to improve power efficiency. In: Proceedings of the ACM SIGMETRICS Joint International Conference on Measurement and Modeling of Computer Systems, SIGMETRICS '11, pp. 139–140. ACM (2011). doi:[10.1145/1993744.1993795](https://doi.org/10.1145/1993744.1993795)
11. Liu, Y., Zhu, H.: A survey of the research on power management techniques for high-performance systems. Softw.: Pract. Exp. **40**(11), 943–964 (2010). doi:[10.1002/spe.952](https://doi.org/10.1002/spe.952). <http://dx.doi.org/10.1002/spe.952>
12. Ltd., D.D.: Dcd industry census 2013: Data center power. Tech. Rep. (2014). <http://www.datacenterdynamics.com/critical-environment>
13. Mitrani, I.: Trading power consumption against performance by reserving blocks of servers. In: M. Tribastone, S. Gilmore (Eds.) Computer Performance Engineering, Lecture Notes in Computer Science, vol. 7587, pp. 1–15. Springer, New York (2013). doi:[10.1007/978-3-642-36781-6\\_1](https://doi.org/10.1007/978-3-642-36781-6_1). [http://dx.doi.org/10.1007/978-3-642-36781-6\\_1](http://dx.doi.org/10.1007/978-3-642-36781-6_1)
14. Shen, K., Shriraman, A., Dwarkadas, S., Zhang, X., Chen, Z.: Power containers: An os facility for fine-grained power and energy management on multicore servers. SIGPLAN Not. **48**(4), 65–76 (2013). doi:[10.1145/2499368.2451124](https://doi.org/10.1145/2499368.2451124). <http://doi.acm.org/10.1145/2499368.2451124>
15. Valentini, G., Lassonde, W., Khan, S., Min-Allah, N., Madani, S., Li, J., Zhang, L., Wang, L., Ghani, N., Kolodziej, J., Li, H., Zomaya, A., Xu, C.Z., Balaji, P., Vishnu, A., Pinel, F., Pecero, J., Kliazovich, D., Bouvry, P.: An overview of energy efficiency techniques in cluster computing systems. Cluster Comput. **16**(1), 3–15 (2013). <http://dx.doi.org/10.1007/s10586-011-0171-x>
16. Vasan, A., Sivasubramaniam, A., Shimpi, V., Sivabalan, T., Subbiah, R.: Worth their watts?—an empirical study of datacenter servers. In: High Performance Computer Architecture (HPCA), 2010 IEEE 16th International Symposium on, pp. 1–10 (2010). doi:[10.1109/HPCA.2010.5463056](https://doi.org/10.1109/HPCA.2010.5463056)
17. von Laszewski, G., Wang, L., Younge, A., He, X.: Power-aware scheduling of virtual machines in dvfs-enabled clusters. In: Cluster Computing and Workshops, 2009. CLUSTER '09. IEEE International Conference on, pp. 1–10 (2009). doi:[10.1109/CLUSTER.2009.5289182](https://doi.org/10.1109/CLUSTER.2009.5289182)
18. Yokogawa Electric Corporation: <http://tmi.yokogawa.com/discontinued-products/digital-power-analyzers/digital-power-analyzers/wt210wt230-digital-power-meters/>



19. Zhai, Y., Zhang, X., Eranian, S., Tang, L., Mars, J.: Happy: Hyperthread-aware power profiling dynamically. In: Proceedings of the 2014 USENIX Conference on USENIX Annual Technical Conference, USENIX ATC'14, pp. 211–218. USENIX Association (2014)
20. Zhang, W.J., Xie, X.F., et al.: Depso: Hybrid particle swarm with differential evolution operator. *IEEE Int. Conf. Syst. Man Cybern.* **4**, 3816–3821 (2003)

**Part III**  
**Network Security**

# A Role and Activity Based Access Control for Secure Healthcare Systems

Naim Alperen Puler, Duygu Karaođlan Altop and Albert Levi

**Abstract** We introduce a novel access control mechanism in order to safeguard privacy of medical data of patients in dynamic environments. Our access control model takes advantages from role-based access control (RBAC) and criticality aware access control (CAAC). In this way, our original approach allows the medical professionals with different roles to be granted access to medical records of patients automatically and without explicit request in case of a medical emergency. In this context, we design secure and privacy aware protocols from initial login to patients' medical data transmission and retrieval by the medical professionals. Moreover, we formally define access control policies for our system. Finally we show the feasibility of our approach by implementation and performance evaluation.

## 1 Introduction

Access control has been an important issue since certain resources are not open for public usage. In a cyber environment, those resources should be reachable by a limited number of subjects and those subjects must be defined within the system. An access control model restricts the unauthorized parties to access resources. It also impedes unauthorized modifications during legitimate accesses to resources by authorized parties [9].

The access control issues become more and more important when the system deals with *private data* of the subjects. The reason behind this importance comes from the fact that any failure during access to information brings about private data leakage which also has some legal consequences [2].

---

N.A. Puler · D.K. Altop · A. Levi (✉)  
Sabancı University, Istanbul, Turkey  
e-mail: levi@sabanciuniv.edu

N.A. Puler  
e-mail: alperenp@sabanciuniv.edu

D.K. Altop  
e-mail: duyguk@sabanciuniv.edu

## 1.1 Motivation

The traditional identity based access control (IBAC) model is not scalable due to requirement of having each user having separate access control rights. Thus, role-based access control (RBAC) models are receiving increasing attention as a generalized approach [8, 10, 11]. In RBAC schemes, system administrators predefine roles and their associated authorized actions according to access policy decisions. Permissions determine the actions which are required to be performed when a particular service is requested. Once role-permission mappings are defined, users are assigned to the abstractions known as “roles” [1]. In case of medical environments, access control can basically be maintained by utilizing RBAC. National Institute of Standard and Technology (NIST) claimed that RBAC addresses the commercial and governmental requirements such as user confidence, personal information privacy, hampering unauthorized distribution of financial assets [4].

RBAC model is more effective as compared to IBAC in the integration of small changes within the system. Roles should be predefined in the system and users are required to be mapped to those roles in order to have access to permitted resources. It is also same in IBAC, except that if one of the role permissions is required to be changed, each user’s permissions should be updated one by one. On the other hand, changing a permission in RBAC just requires a single permission change defined on the role itself. Even though RBAC gives more flexibility over IBAC, it still suffers from certain disadvantages: (i) What if some users in the system should access certain information for a limited amount of time? (ii) What if some of the subjects’ access are required to be changed dynamically due to some critical events? In our proposed model we overcome these problems as detailed in Sect. 2.

## 1.2 Related Work and Our Contributions

Today’s access control models mainly use RBAC principles in order to reduce the number of control operations over a target subject. Zheng et al. [13] defines participation, act and activity in order to obtain a dynamic version of RBAC. Act is defined as an operation of application systems and role is defined as a set of subjects sharing the same access control policies to certain objects. Participation denotes a functional role and co-works with act. It is new abstraction between roles and acts. First, the role of a subject that requests access is found within the system. Then according to that role, subject is granted participation controlled by defined rules in access control policy. If participation of a subject is mapped to requested act in activity cell, then access is granted to user.

A RBAC mechanism is also constructed for cyber-physical systems by Muppavarapu and Chung [6]. They try to reduce the administration overhead, which stems from the role privileges of the individuals by a middleware. They apply a protocol to gain access control credentials and once those credentials are obtained, the protocol communicates with the resource manager in order to perform the requested operation.

The above mentioned two studies [6, 13] do not address the criticality management requirement of our proposed model.

Venkatasubramanian [12] claims that in a medical environment access control should be adaptive, and therefore, dynamic for emergency management. This versatility provides the required privileges to the subjects implicitly for short periods of time. With the use of critical-aware access control, he constructed a model, which behaves like context based access control (CBAC) in normal state. In CBAC, context information of the subject determines the access control. For instance, context can be constrained by time and space. If a subject requests access in different places at the same time, system rejects requests according to the policy of having a subject not to appear in different places at the same time. Other than normal states, when someone experience criticality, it shifts from this model to another, which is more proactive in nature.

Undoubtedly, this work [12] is closely related to our study since it supports criticality management. However, it achieves the regulation of critical situations by applying regular checks over the system in certain periods. Another drawback of [12] is that it tries to automate the responsive actions over patients for a calculated amount of time. This is a medical risk, because treatments cannot be applied to all patients in the same way even if they suffer from same disease. Therefore, we come up with a model which interprets patients' medical information whenever the data are received by system. Under critical circumstances, system dynamically gives extra control to medical professionals in order to recover the patients from their critical diseases.

In a nutshell, our contributions in this paper are as follows:

1. Our model allows the predefined access policies over patients to be changed dynamically according to medical criticalities. More importantly this dynamicity directly affects the access rights of the medical professionals for a prompter treatment.
2. Our system allows real-time interpretation and analysis of medical data for automated management of access rights and policies.
3. System users are able to have more than one access right at the same time.

The remainder of this paper is organized as follows. In Sect. 2, we introduce our access control model. Section 3 evaluates the system performance with respect to security related delay and average time required to interpret medical data. Finally in Sect. 4, we conclude the paper.

## **2 Our Proposed Access Control Model: Role and Activity Based Access Control for Healthcare Systems**

Our access control model is composed of 2 phases: (i) Client login to the system. (ii) Access request from authenticated client. Symbols used in the protocol definition are given in Table 1.

**Table 1** List of identifiers used in our R-ABAC mechanism

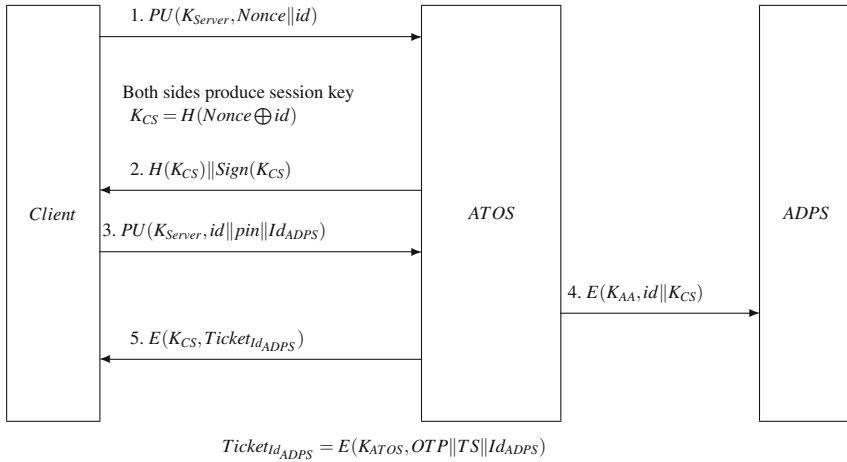
$ATOS$	Authentication and Ticket Obtainment Server	$K_{Server}$	Public Key of ATOS
$ADPS$	Authorization and Data Processing Server	$K_{CS}$	Shared Key between Client and ADPS
$APM$	Access Policy Manager	$M$	Set of Medical Experts
$CCP$	Constraint and Control Policy	$OTP$	One-Time Pad
$D$	Set of Diseases	$P$	Set of Patients
$E(key, plain)$	Symmetric key encryption of plain using given key	$P_j$	Assigned patient to doctor $m_j \in M$
$H$	Set of Health Information	$T$	Set of Time
$Id_{ADPS}$	Identity of ADPS	$TS$	Time Stamp
$K_{AA}$	Pre-shared key between ATOS and ADPS	$PU(key, plain)$	Public key encryption of plain using given key
$K_{ATOS}$	Key which is only known by ATOS	$Ticket_{Id_{ADPS}}$	Ticket assigned to server identifier

## 2.1 Login to System

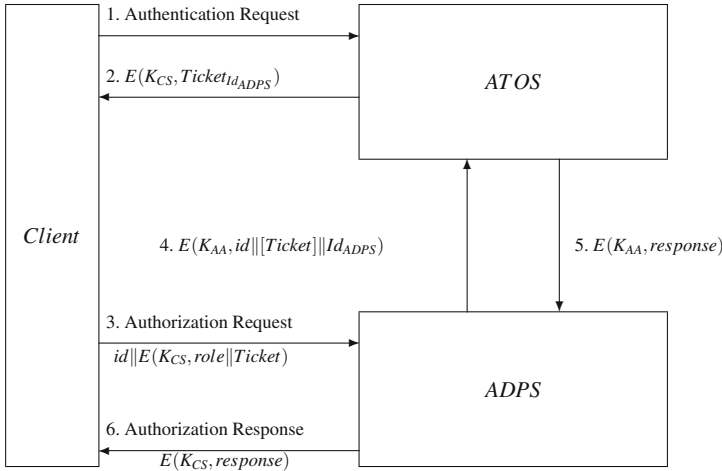
Healthcare systems should be defined carefully since they naturally work with medical data. For this reason, the approach should possess high level of data security. To ensure the security, our protocol requires a trusted third party, which is called authentication and ticket obtainment server (ATOS). Our login protocol is a ticket based one as in Kerberos [7]. Clients who would like to login to the hospital servers, named authorization and data processing server (ADPS), should first identify themselves to ATOS. Our login protocol is composed of two phases: (i) authentication and ticket generation, (ii) ticket validation.

### 2.1.1 Phase 1: Authentication and Ticket Generation

This phase can be visualized in Fig. 1. At the beginning, client sends its identity to ATOS together with a nonce in an encrypted form. Using the nonce and id, both sides compute a session key. After that, ATOS authenticates itself to client by sending a signed message. Then client sends its pin to authenticated ATOS in encrypted form to complete mutual authentication. ATOS also sends the session key to the designated ADPS securely. In the final step, ATOS generates a ticket for the communicating client. This ticket contains information about the identifier of the ADPS, to which the client requested to login. The ticket is going to be checked again by ATOS. Tickets cannot be modified by an unauthorized user since its content is encrypted by a key known only to ATOS.



**Fig. 1** Authentication and ticket generation protocol



**Fig. 2** Ticket validation

**2.1.2 Phase 2: Ticket Validation**

Once the ticket is controlled by ATOS, it will decide whether the client is valid for logging in ADPS or not, as visualized in Fig. 2. If the ticket is validated, ADPS checks the identity of the client and its role within the system. During the protocol, client also generates a session key,  $K_{CS}$ , which is going to be used for the confidentiality and integrity of the private data. If a valid user does not send its role during the login protocol, then it will have limited access over resources, i.e., time and/or resource restriction. In this case, the client can perform certain operations for a restricted amount of time.

## 2.2 Role and Activity Based Access Control

Our proposed R-ABAC model, set definitions and critical state conditions are explained in this subsection.

### 2.2.1 R-ABAC

Access control is based on RBAC, so we designed it as the following. Users are assigned to roles. Each role has certain capabilities over the system. Administrators are capable and responsible for adding and removing users, assigning patients to doctors. Patients are able to transmit their medical data to ADPS and request their medical record. Finally doctors are able to request medical data of assigned patients. However, the most important access policy is “any role other than the health professional or patient, even the administrator, does not have access over health data”. This is important due to the fact that the privacy of the data should be maintained.

Roles defined in R-ABAC are mapped into participations. Here, the roles are more like static groups and the participations provide the dynamicity of the access control mechanism. The reason behind this is that access can be obtained by the users when they participate in an act. The act set is defined according to the control policy. Participations and acts are mapped with each other within a directed graph called the activity cells. The mapping defines which participation instances are able to do predefined acts. While controlling the roles, the system also addresses the problem of time constraints over the roles. As a unique feature of our model, we extend RBAC by assigning users to roles for a given amount of time. Consider the scenario that a doctor is going to access the system for a short period of time. If the administrator adds this new doctor to the system, then (s)he becomes a legitimate user for the system as a doctor from that time on. This is not an expected behaviour of the system, because the system itself handles with private data and should perpetuate the requirements of the access control policy. To overcome such issues, we designed R-ABAC such that the users are valid between certain periods of time.

### 2.2.2 Set Definitions

Proposition 1 and Proposition 2 explain the system requirements over patients and doctors. In a hospital environment, generally patients are assigned to certain doctors, who are responsible for monitoring their health status. Under normal conditions, patient’s medical data are collected and transmitted to the hospital server. Proposition 3 defines a function which gives medical record of a patient at a given time. This health information can be seen only by the patient and the preassigned doctor; meaning that other users cannot access that data.

**Proposition 1**  $\forall p_i \in P$ , the system is responsible for monitoring the medical data of the patient  $p_i$ .



**Proposition 2** *Let  $g$  be a function such that  $g(m_j) = P_j$ , where  $P_j$  is the set of assigned patients to medical expert  $m_j$ . Then  $\forall m_j \in M$ , the system is responsible for revealing the medical information of the patient  $p_i \in P_j \subseteq P$  to the medical expert  $m_j$*

**Proposition 3** *For a given time  $t \in T$ , we define the function  $\theta$  such that  $\theta(p_i, t) = h_{i_t}$ , where  $p_i \in P$  and  $h_{i_t} \in H$  at time  $t$ .*

### 2.2.3 Medical Data Interpretation and Analysis

In our model, we interpret the data with six different vital signs. Those signs are used for the diagnosis of emergency anomalies. A patient's singular medical record consists of body temperature, blood pressure, respiration, oxygen saturation, Electrocardiogram (ECG) and heart rate (pulse). Interpretation starts with the analysis of ECG, which is the most time consuming operation. During this analysis P, Q, R, S and T waves, QRS complex of the ECG data and other wave intervals are inferred. Pulse information is also extracted from ECG. Other five vital signs can be directly interpreted without extra operation. After that, medical data is evaluated with a function by checking the values of interpretation results. This function checks whether the vital signs are normal or there is a problem with the patient's health condition.

For instance, "Hypokalemia" has the following disease characteristics: low amplitude of T wave, prolonged QT interval, flattened or inverted T wave. These information can be directly extracted via ECG interpretation. The disease also shows the vital signs as low respirations and high blood pressure. Once ECG is analysed, other vital signs of the medical data are also taken into account of analysis. The ADPS has predefined set of diseases which has been mapped with certain combination of the interpreted results of vital signs. From the predefined mapping across diseases and vital sign characteristics, potential diseases that a particular patient is suffering from are figured out by the system. If there is not any potential diseases captured, system continues its regular operation. However, if it discovers an anomaly for a patient, the system automatically triggers an emergency action for this patient.

### 2.2.4 Critical State

Patient's medical data may contain vital information, indicating a disease. This disease could threaten the patient's life. As a result, emergency conditions should take over the access control mechanism in order to be able to cure the critical disease. This problem is addressed in Proposition 4.

**Proposition 4** *If the patient  $p_i$  experiences a fatal disease for a given health information  $h_{i_t} \in H$  at time  $t \in T$ . Define function  $f$  such that  $f(h_{i_t}) = d_k \in D$  and  $d_k \neq \emptyset$ .*

Such conditions are the so-called critical state of a patient. Accordingly, the main assumption is that the preassigned medical expert may not be able to take action

against the critical condition at that moment. As a result, the system should give the ability to monitor the patient to some other medical experts. While roles are defined for R-ABAC, doctors' speciality is also defined. This is done due to finding on-line doctors who are assumed to be more knowledgeable to cure certain critical diseases. The diseases are categorized to a specialities of doctors. For instance, "Hypokalemia" is categorized with speciality internal medicine. Once the patient experiences a critical disease, the system should find which medical expert(s),  $m_j \in M$ , are going to be granted with extra access over the patient  $p_i \notin P_j$ . In the "Hypokalemia" case, medical experts who are specialized in internal medicine should be selected for criticality. Predefined number of doctors to be notified is a system parameter. System finds that number of on-line medical experts of internal medicine and sends notification message to selected doctors. Algorithm 1 outlines the process of analysis and system response to criticality. In our proof of concept implementation, doctors are notified via desktop computers; but in real life implementations this notification can be done through mobile devices.

Those doctor(s) should advise the patient  $p_i$  according to health condition so that the patient should recover from critical state to normal state. After the patient  $p_i$  returns to his normal state, extra privileges given to those medical experts are revoked by the system.

---

#### Algorithm 1 Critical State Response

---

**Require:**  $h_i \in H$ ,  $n$ : number of doctors to be notified  
 analysisResult = Analyse( $h_i$ )  
 possibleDiseases = GetPossibleDiseasesFromAnalysis(analysisResult)  
**for** each  $d_k \in$  possibleDiseases **do**  
   doctorsToBeAlarmed = FindOnlineDoctorsWithSpeciality( $d_k$ ,  $n$ )  
   **for** each  $m_j \in$  doctorsToBeAlarmed **do**  
     NotifyDoctor( $m_j$ )  
   **end for**  
**end for**

---

### 3 Performance Evaluation

We have implemented our proposed R-ABAC model using C# programming language. Initially, servers start to communicate with each other in order to generate keys for secure communication. Once their channels are established, they become ready for the incoming clients. We have randomly generated some of the health information such as body temperature, respiration, oxygen saturation and blood pressure. These are also known as primal vital signs of fatal diseases. We obtained real Electrocardiogram (ECG) data for 50 different patients from the publicly available PhysioBank MIMIC II Waveform database [5]. ECG signals are interpreted using MATLAB. We have simulated the hospital environment in a local computer, which has Windows 7

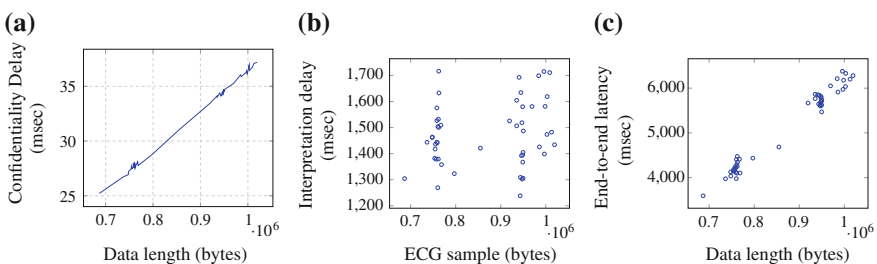
64-bit OS, i7-2600 CPU with 3.40 GHz frequency and 8 GB RAM. Our primary aim is to measure the amount of the consumed system resources. Performance metrics and analyses are given in the following sections.

### 3.1 Performance Metrics

The metrics measured in our system are: (i) confidentiality delay, (ii) end-to-end latency and (iii) ECG signal interpretation delay. Confidentiality delay is the additional time required for symmetric encryption and decryption on private data. End-to-end latency is total time for total time need for entire processing, data transmission and receiving acknowledgement. These time related metrics are important since in a healthcare system, we need fast responses in order to manage critical states. Moreover, the bottleneck of analysis is the interpretation of ECG signal. Thus we also measure it separately. Interpretation of ECG signal is done by first finding QRS complex intervals [3] from the signal. Afterwards interval changes are statically observed and interpreted by system.

### 3.2 Analyses

In the Fig. 3a, confidentiality delay versus data length is shown for 50 patients separately. Since symmetric encryption is applied, overhead caused by security with respect to time is very low. Moreover, as expected, confidentiality delay linearly increases with respect to data length, independent of the medical content. As it is mentioned before, ECG requires a special interpretation and processing. Figure 3b shows time variations of ECG signal processing and interpretation among 50 patients. As can be seen in this figure, ECG signal interpretation delay varies between 1200 and 1700 ms and is independent of the length of the sample. Thus, it changes from patient to patient due to the characteristics within the ECG signal. Figure 3c shows



**Fig. 3** Performance analysis results. **a** Confidentiality delay, **b** ECG interpretation delay, **c** end-to-end latency

the total end-to-end delay of our proposed R-ABAC model. As shown in this figure, the general trend in increase of delay is linear with respect to data length although there are some variations due to patients medical record characteristics.

## 4 Conclusions and Future Work

In this paper, we designed a role and activity based access control mechanism (R-ABAC) for healthcare systems. In our model, we extended classical RBAC by adding dynamicity. Moreover, our system finds out emergency conditions of patient and notifies doctors about this situation in an automated way. We implemented our system and performed simulation based analyses. We conclude that average end-to-end latency is in acceptable level and the additional cost of security is insignificant.

As a future work, our R-ABAC mechanism can be enhanced to integrate several hospitals in a peer-to-peer or hierarchical manner. This feature can help to handle criticalities with the additional assistance of medical experts who are assigned in other hospitals. However, this process would require additional security requirements and improvements to access control model in order not to leak medical information. Another improvement for current server architecture is introducing a hierarchical organization of services. Our ticket based authentication and authorization infrastructure helps to support such a hierarchical organization.

**Acknowledgments** This work was supported by the Scientific and Technological Research Council of Turkey (TÜBİTAK) under grant 114E557.

## References

1. Chakraborty, S., Ray, I.: Trustbac: integrating trust relationships into the rbac model for access control in open systems. In: Proceedings of the Eleventh ACM Symposium on Access Control Models and Technologies, pp. 49–58. ACM (2006)
2. Dwork, C.: A firm foundation for private data analysis. *Commun. ACM* **54**(1), 86–95 (2011)
3. Fraden, J., Neuman, M.: Qrs wave detection. *Med. Biol. Eng. Comput.* **18**(2), 125–132 (1980)
4. Gilbert, M.D.M.: An examination of federal and commercial access control policy needs. In: National Computer Security Conference, 1993 (16th) Proceedings: Information Systems Security: User Choices, p. 107. DIANE Publishing (1995)
5. Kreiseler, D., Bousselet, R.: Automatisierte EKG-Auswertung mit Hilfe der EKG-Signaldatenbank CARDIODAT der PTB. *Biomedizinische Technik/Biomed. Eng.* **40**(1), 319–320 (2009)
6. Muppavarapu, V., Chung, S.M.: Role-based access control for cyber-physical systems using shibboleth. In: Proceedings of DHS Workshop on Future Directions in Cyber-Physical Systems Security, pp. 57–60 (2009)
7. Neuman, B.C., Ts'o, T.: Kerberos: an authentication service for computer networks. *IEEE Commun. Mag.* **32**(9), 33–38 (1994)
8. Osborn, S., Sandhu, R., Munawer, Q.: Configuring role-based access control to enforce mandatory and discretionary access control policies. *ACM Trans. Inform. Syst. Secur. (TISSEC)* **3**(2), 85–106 (2000)

9. Samarati, P., de Vimercati, S.C.: Access control: policies, models, and mechanisms. In: Foundations of Security Analysis and Design, pp. 137–196. Springer, New York (2001)
10. Sandhu, R.: Role hierarchies and constraints for lattice-based access controls. In: Computer Security ESORICS 96, pp. 65–79. Springer, New York (1996)
11. Sandhu, R.S.: Role-based access control. *Adv. Comput.* **46**, 237–286 (1998)
12. Venkatasubramanian, K.K.: Security solutions for cyber-physical systems. Ph.D. thesis, Arizona State University, 2009
13. Zheng, S., Jiang, D., Liu, Q.: A role and activity based access control model for university identity and access management system. In: Fifth International Conference on Information Assurance and Security, 2009, IAS'09, vol. 2, pp. 487–490. IEEE (2009)

# Bandwidth Usage—Based Detection of Signaling Attacks

Mihajlo Pavloski, Gökçe Görbil and Erol Gelenbe

**Abstract** In the last couple of years, both the number of smart devices using mobile networks' services, and the number of security threats for mobile devices have increased rapidly. This growth generates new challenges for mobile network operators. One of the recent challenges is fighting the *signaling attacks* and *storms*, that represent a type of distributed denial of service (DDoS) attacks, which overload the signaling plane of the networks, and threat networks' stability. This paper proposes a detection mechanism for such attacks. A cost function is defined using the low bandwidth usage characteristic and is calculated in a exponential weighted moving average manner to enable real-time detection of attack intervals. The detector is implemented in a simulation environment in a 3G UMTS network and evaluated using metrics of interest such as: detection delay and probability of false positive and false negative detections. Finally, a simple attack mitigation technique is used together with the detector in a network under attack and manages to reduce the signaling load and end-to-end delay to the level of an unattacked network.

**Keywords** Signaling attacks · Detection · Denial of service · Exponential weighted moving average · Mitigation

## 1 Introduction

The use of smart devices and mobile data services in mobile networks record a great increase in the last couple of years. The number of global mobile devices and connections rose for almost half a billion in 2014 out of which smartphones

---

M. Pavloski (✉) · G. Görbil · E. Gelenbe  
Department of Electrical and Electronic Engineering, Intelligent Systems  
and Networks Group, Imperial College, London SW7 2AZ, UK  
e-mail: m.pavloski13@imperial.ac.uk

G. Görbil  
e-mail: g.gorbil@imperial.ac.uk

E. Gelenbe  
e-mail: e.gelenbe@imperial.ac.uk

accounted for 88% of the growth. The mobile data traffic grew 69% from 2013 to 2014, reaching 2.5 exabytes per month [2]. In parallel, the number of security threats for mobile devices is rapidly growing with a tenfold increase of mobile malware attacks per month from August 2013 to March 2014 [1].

*Signaling attacks* that have emerged as novel security threats to mobile networks are instrumented by such mobile malware. Their purpose is to develop a distributed denial of service (DDoS) attack [4, 6] on the control plane of network rather than the data plane [10]. The impact of these attacks can be maximised by using groups of mobile devices—*botnets* [13] and by adapting the attack to the networks' parameters [3]. Similar attacks can also happen due to poor development of smart device applications which use frequent background messages, and are known as *signaling storms* [5, 11]. Both of these attacks cannot be detected by traditional flooding-based attack detection systems.

The exploited vulnerability by these attacks is located in the radio resource control (RRC) part of the system. Whenever a mobile terminal wants to transfer some data, it needs to ask for some communication resources by the network which triggers a signaling procedure called *connection/radio bearer setup* in 3G UMTS networks or *random access* in 4G LTE. This procedure involves exchange of up to 20 signaling messages in the radio access network (RAN) and core network (CN) parts in UMTS, while a smaller number of messages are exchanged in LTE. If this behaviour is repeated by a decent number of mobile terminals in the network it can cause overloading of the signaling servers which leads to service degradation and even system outages [5, 8, 9]. For the mobile user this is manifested in high battery consumption and even unwanted billing.

Our previous work [7, 12] has shown that these attacks can be identified not only by their repetitive pattern but also by their low usage of communication resources in order to evade getting detected by flooding security mechanisms. In this paper we use this characteristic to develop a simple detection mechanism that is capable of identifying attack behaviour in real-time. The paper is organised as follows. Section 2 covers the details of the detector defining the cost function and the decision making. In Sect. 3 we implement and evaluate the detector in the SECSIM simulator [12] in UMTS networks. Furthermore, based on insights from [14] we use a simple mitigation technique on top of the detector and show its influence in reducing the load in the signaling servers in the network and keeping system's stability.

## 2 Detection

The proposed detection mechanism described in this paper is based on the low bandwidth usage characteristic. More precisely, when a malicious mobile terminal gets some communication bandwidth (time and frequency resources), it does not use them or only uses a small portion of them. Since the needed information is available in both the mobile terminal and in the network controller, it can be implemented in both

sides of the system. Our main goal was to develop a simple detector that would not impose any high computational, storage or energy demands, and that would be able to distinguish a malicious behaviour in a real-time manner.

## 2.1 A Cost Function

In the following, we describe how the detection works on a single mobile terminal. There are two input values that the mechanism needs. The first value is the total time that the terminal spends in a ‘High’ state in a given time window  $t_w$ . The ‘high’ state represents a communication state where communication resources are granted to the mobile terminal and is equivalent to states *CELL\_DCH* or *CELL\_FACH* in UMTS or state *Connected* in LTE. These values are denoted with  $t_D$  and  $t_F$  for DCH and FACH states in UMTS. Note that DCH stands for dedicated channel and FACH is a forward access shared channel. The second value is the time at which the mobile terminal spends in ‘high’ state but does not transfer any data (stays Idle), also in a time window  $t_w$ . This is denoted with  $t_{Di}$  and  $t_{Fi}$  respectively for DCH and FACH in UMTS. The ratio  $\frac{t_{Fi}+t_{Di}}{t_F+t_D}$  is calculated whenever resources are (de)allocated, i.e. for every state change. Then we use the exponential weighted moving average (EWMA) to define a cost function  $C$  as:

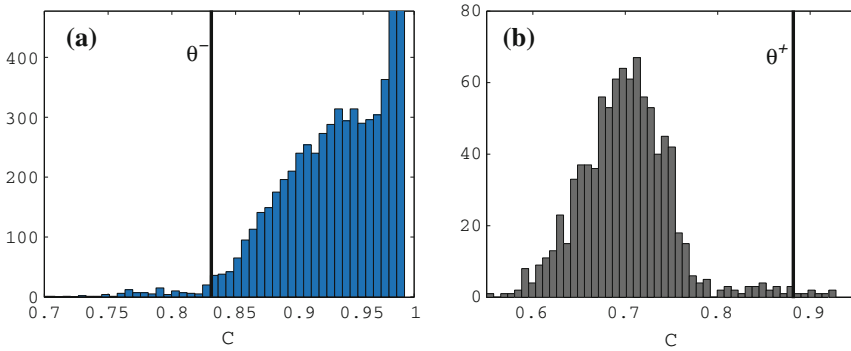
$$C(k) = \alpha \frac{t_{Fi}(k) + t_{Di}(k)}{t_F(k) + t_D(k)} + (1 - \alpha)C(k - 1), \quad (1)$$

where  $k > 0$  is the index of the state change,  $0 \leq \alpha \leq 1$  is a weight parameter and  $C(0) = \frac{t_{Fi}(0)+t_{Di}(0)}{t_F(0)+t_D(0)}$ . This cost function enables detection of attacks on both FACH and DCH channels in UMTS. With suitable adjustments it could be adapted to any network with similar functionality, like LTE. The calculation of the cost function should be run as a background process in each mobile terminal or in a centralised network node that has the needed information for all terminals. That node is the RNC in UMTS and the eNodeB in LTE.

## 2.2 Decision Making

As defined, the cost function has values between 0 and 1 with values closer to 1 indicating an attack. To define when an attack is detected let’s suppose that the mechanism is running for long enough time so that the average value of the cost function  $C_{avg}$  is stable with minor changes. Then we can say that a malicious behaviour is detected if  $C \geq \beta C_{avg}$  where  $\beta$  is a value close to but larger than 1. A suitable choice would be  $\beta = 1.5$  meaning that 50% increase of  $C$  above its average is an indicator of attack. Note that  $C$  is calculated within a time window  $t_w$  while  $C_{avg}$  is calculated from the activation of the mobile terminal.





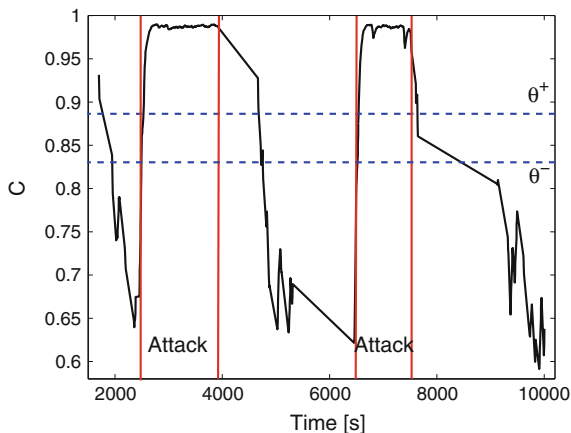
**Fig. 1** Threshold setup based on the histogram of simulated data for **a**  $\text{Prob}\{\text{false negatives}\}=0.01$  and **b**  $\text{Prob}\{\text{false positives}\}=0.01$

There are a couple of problems with this type of decision making. First, if  $C_{avg}$  is higher than a value  $1/\beta$ , an attack cannot be detected. This could happen if an attack is ongoing from the moment of activation of the mobile terminal, thus producing high  $C_{avg}$  value such that  $\beta C_{avg} > 1$ . Another problem appears for heavy traffic users, for example users who use video streaming or voice communication. These terminals use a big portion of the granted resources and therefore have very low  $C_{avg}$ . In this case the product  $\beta C_{avg}$  is still very small and normal usage is often clarified as attack.

To address this issue we propose setting up two thresholds for  $C$ : an *upper threshold*  $\theta^+$  above which we make a decision of an attack, and a *lower threshold*  $\theta^-$  below which we make a decision of normal behaviour. Both decisions are irrespective of  $C_{avg}$ . Setting up these thresholds should be based on offline traffic analysis by the mobile operators. A frequently used thresholding technique is one based on a fixed value for the probability of false positives. Having the probability distribution of  $C$  values for bandwidth requests classified as normal, we set  $\theta^+$  as the threshold above which statistically 1% of normal requests will be declared as attack. Similarly,  $\theta^-$  is set up for a 1% fixed probability of false negatives based only on  $C$  values of bandwidth requests classified as attack. Figure 1 shows the threshold setup on simulated data, while Fig. 2 shows an example of  $C$  in time and setup thresholds. The simulated data for normal traffic is modelled using web metrics used by Google [15] and closely resembles the real life communication in mobile networks. The model of the attack traffic is based on previous research publications.

It is important to highlight the advantages of this kind of decision making compared to a simple single threshold based one. Setting a single threshold on  $C \in [0, 1]$  could be done by fixing the probability of false positives like we did with  $\theta^+$ . Then a decision of attack is done in the case  $C > \theta^+$  and a decision of normal sample is in the case  $C \leq \theta^+$ . The approach we selected allows us to detect attacks with  $C$  samples even below  $\theta^+$ . This is enabled by knowing the average value of  $C$  and gives us the advantage of detecting attack by checking if  $C > \beta C_{avg}$ .

**Fig. 2** An example of the cost function in time and thresholds set to  $\theta^+ = 0.88$  and  $\theta^- = 0.83$ , with two attack intervals



Another important consideration is that decision making should be adjusted to fit different types of mobile terminals. For example, the data traffic of machine to machine communication or sensor communication is small in volume and does not efficiently use the allocated bandwidth, thus resulting in higher  $C$  samples. On the other hand, terminals usually transferring video streaming have small  $C$  samples. Web browsing should be somewhere in between. One approach towards this issue would be to classify terminals by type and assign different values of  $\beta$ ,  $\theta^+$  and  $\theta^-$  to each class.

### 2.3 Mitigation

Previous analytical work in the area [14] has given us some guidelines how to mitigate signaling attacks. The paper suggests that a possible mitigation approach is to delay mobile's requests for bandwidth allocation. However, our simulation analysis have shown that a better approach is to block the whole activity of the malicious mobile terminal for a given period  $t_b$  which is in the range of one to couple of minutes.

## 3 Evaluation and Simulation Results

In this section we briefly describe the developed SECSIM simulator and evaluate the proposed mechanism in terms of detection delay and probability of false detection. Further on, we add a simple attack mitigation technique and show some results of the RNC load and end-to-end delay for unattacked users.

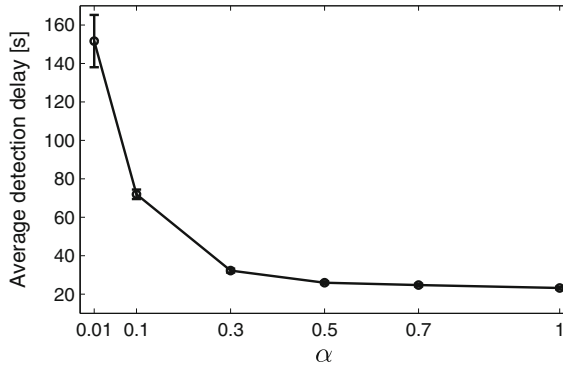
### 3.1 Simulator Description

The SECSIM simulator represents a tool for modeling and simulation of cybersecurity in mobile networks, with a focus on the signaling layer in the radio access part. Based on Omnet++ [16] it is an object-oriented discrete event simulator. It is also modular solution, allowing network components to be easily modified using smaller components—modules. Its current version contains models of functional components of both UMTS and LTE networks, such as: UE, RNC, NodeB, SGSN, GGSN, eNodeB, SGW, Internet hosts etc. For evaluation purposes of this mechanism we will use the UMTS model. In the control plane, the UE model consists of the session management (SM), GPRS mobility management (GMM) and RRC layers. In the data plane, it contains the application layer containing both circuit switched and IP applications, the transport layer with TCP and UDP protocols and a simplified IP layer. MAC and PHY layers are not modelled, while changes in radio conditions are modelled as random variations. The RNC model has the RRC containing a single signaling server, RANAP, NBAP and GTP protocols. The signaling server plays a crucial role in the signaling attacks and their mitigation.

### 3.2 Evaluation

The proposed mechanism allows detection of signaling attack behaviour in a real-time manner. It works with calculating the cost function  $C(k)$  for each data transmission at instance  $k$ . In any case, single data transmissions are not classified as attack/normal but rather a decision is made upon a group of data transmissions in a time window  $t_w$ . This is because single attack transmissions or attacks with low frequency cannot form a signaling attack and cause danger to the network. The mechanism works in real-time because the observation window slides through time and the cost function is calculated using the EWMA averaging.

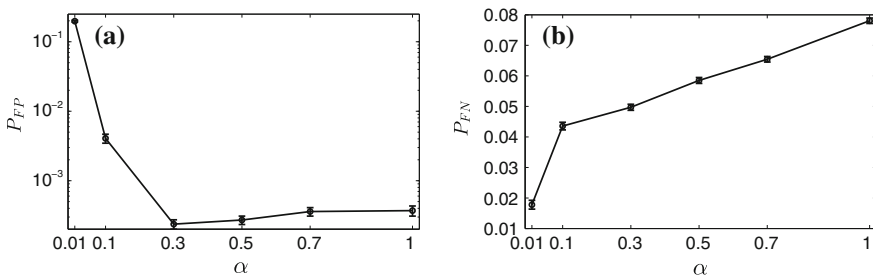
One of the metrics of interest is *detection delay*  $\tau$ —the time delay between attack starts and it is detected. It depends on the moving average parameter  $\alpha$  and the width of the sliding window  $t_w$ . Furthermore, to see the influence of the detector on UEs in unattacked time intervals (or unattacked UEs) we can define the *false positives probability*  $P_{FP}$  as the fraction of time in which an attack is detected but not existing, and the *false negatives probability*  $P_{FN}$  which is the fraction of time in which an attack is ongoing but is not detected. To quantify these metrics, we run the simulation for  $\alpha \in \{0.01, 0.1, 0.3, 0.5, 0.7, 1\}$  in a network of 500 UEs out of which 150 are attacked. The simulation is repeated five times for each  $\alpha$  value with different seeds for the random number generators and results are averaged over all runs. Attacks happen at a random time of the simulation and in intervals of random length. Note that in this case only the attacked UEs are of interest to be analysed, and that the ratio of attacked/total UEs does not influence the performance of the detector.



**Fig. 3** Average detection delay  $\tau$  for  $t_w = 60$  s,  $\beta = 1.5$ ,  $\theta^- = 0.83$  and  $\theta^+ = 0.88$ , 95 % confidence interval used

Figure 3 shows the average detection delay for different values of  $\alpha$ . A smaller value for  $\alpha$  in the calculation of the cost function gives more importance to previous  $C$  values and makes the detection more rigid which is shown by the high values of the detection delay. Increasing  $\alpha$  makes the detector more flexible and improves the detection delay. However, for higher values of  $\alpha$  we expect the cost function to change too rapidly and increase the number of false detections.

As shown on Fig. 4,  $P_{FP}$  is smaller than 0.005 for all values of  $\alpha$  except  $\alpha = 0.001$ . As mentioned earlier, using a very small  $\alpha$  makes the detector un-flexible and has a negative influence on normal users as it wrongly detects normal traffic as attack. For  $\alpha = 0.3$  the detector shows best results for the probability of false positive detections, while as  $\alpha$  increases, the detector is more unstable and  $P_{FP}$  increases again. In the case of the probability of false negative detections, its value increases with the increase of  $\alpha$  for the same mentioned reasons. Note that in our experiments during an attack, the attack traffic is mixed with the traffic of normal applications. This causes the  $P_{FN}$  to have higher values as some parts of the attack interval will still be identified as normal. In case of a deliberate attack with only an attack application installed on mobile devices, the detection will improve its performance.



**Fig. 4** **a** Probability of false positive detection and **b** probability of false negative detection, 95 % confidence interval used

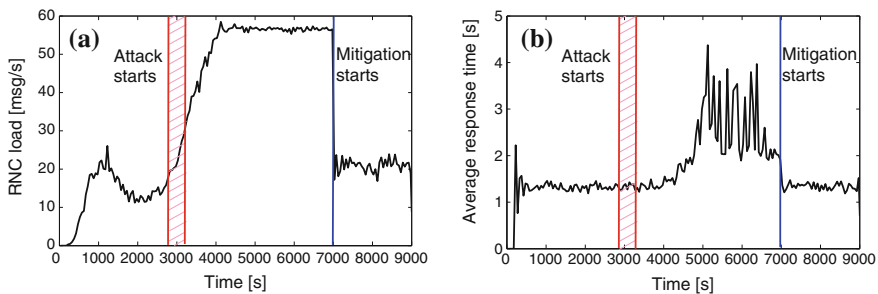
### 3.3 Mitigation

Now we're interested in using the detector with a simple mitigation approach as described in Sect. 2. To mitigate the attack we will use blocking of the attacking UEs for a time duration of  $t_b = 60$  s. The blocking is done immediately when attack is detected.

Since the main purpose of the signaling storms is to overload the signaling servers in the network, we will look at the *load on the signaling server* in the RNC in terms of processed messages per second. High loads on the signaling server are causing delays in the communication of users in the network, both attackers and normal ones. The metric of interest in this case is the *end-to-end delay* measured on application layer experienced by normal users.

The simulation scenario for this purpose is organised as follows. We have 500 UEs among which 150 (25 %) are attackers. All UEs use web browsing communication while attackers use an additional second application to perform the attack. The duration of the experiment is 3 simulated hours. After a period of normal work in the network, attackers start the attack gradually in the period (45, 55 min) which is randomly chosen. The detector runs during the whole experiment and the mitigation starts exactly at 117 min (7000 s). This value is chosen such that there is enough time for the network to get congested (around 62–72 min), and also enough time to show the stabilizing of the network before the simulation ends. Results are shown on Fig. 5.

Results show that in the period without attacks (0–45 min) the network load has a small peak because of the activation of mobile terminals (which includes exchange of signaling messages with the RNC), after which the load stabilizes. The end-to-end delay is also stable. In the period when the attack is ongoing (45–117 min) the load on the RNC starts to increase until a certain point when it reaches a maximum value. At this point the buffers in the signaling server are congested which results in higher delays for the normal users. Using the proposed detector with a simple attack mitigation technique, since the 117 min, manages to decrease the load in the network and stabilize the experienced delay.



**Fig. 5** **a** Load on RNC and **b** End-to-end delay for a single normal UE

## 4 Conclusion

Signaling attacks and storms are a reality in the last couple of years, forcing many mobile operators to look for solutions. These attacks threaten the stability of networks, and on many occasions have managed to reduce the quality of offered services and even cause complete network outages. We have distinguished some basic characteristics of these attacks and used the ‘low bandwidth usage’ to define a detection technique which is capable of detecting attacks in real-time. The technique could be implemented on both mobile terminal and network sides of the system. Furthermore, we have evaluated the proposed detector calculating the average detection delay and the probability of false positive and false negative detection. In all cases, results are satisfactory. Further improvements could be done in dynamically adapting the parameters of the detector ( $\alpha$ ,  $\theta^+$  and  $\theta^-$ ) to the type of the communication of the mobile terminal (human generated, machine to machine communication, sensor data communication etc.).

**Acknowledgments** This work is part of the EU FP7 project NEMESYS (Enhanced Network Security for Seamless Service Provisioning in the Smart Mobile Ecosystem), under grant agreement no.317888 within the FP7-ICT-2011.1.4 Trustworthy ICT domain.

## References

1. Mobile Cyber Threats. Joint Report. <http://securelist.com/analysis/publications/66978/mobile-cyber-threats-a-joint-study-by-kaspersky-lab-and-interpol/> (Oct 2014)
2. Cisco visual networking index: Global mobile data traffic forecast update, 2014–2019. White Paper. [http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white\\_paper\\_c11-520862.pdf](http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/white_paper_c11-520862.pdf) (Feb 2015)
3. Abdelrahman, O.H., Gelenbe, E.: Signalling storms in 3G mobile networks. In: Proceedings of IEEE International Conference on Communications (ICC’14), Communication and Information Systems Security Symposium, pp. 1017–1022. Sydney, Australia. <http://dx.doi.org/10.1109/ICC.2014.6883453> (2014)
4. Gelenbe, E., Loukas, G.: A self-aware approach to denial of service defence. *Comput. Netw.* **51**(5), 1299–1314 (2007). <http://dx.doi.org/10.1016/j.comnet.2006.09.009>
5. Gabriel, C.: DoCoMo demands Google’s help with signalling storm. <http://www.rethink-wireless.com/2012/01/30/docomo-demands-googles-signalling-storm.htm> (Jun 2012)
6. Gelenbe, E., Gellman, M., Loukas, G.: Defending networks against denial of service attacks. In: Carapezza, E. (ed.) Proceedings of the Conference on Optics/Photonics in Security and Defence (SPIE), Unmanned/Unattended Sensors and Sensor Networks, vol. 5611, pp. 233–243. London, UK (October 2004)
7. Gelenbe, E., Abdelrahman, O.H.: Time-outs and counters against storms (Aug 2014), unpublished
8. Gelenbe, E., Görbil, G., Tzovaras, D., Liebergeld, S., Garcia, D., Baltatu, M., Lyberopoulos, G.L.: NEMESYS: enhanced network security for seamless service provisioning in the smart mobile ecosystem. In: Gelenbe, E., Lent, R. (eds.) Information Sciences and Systems 2013—Proceedings of the 28th International Symposium on Computer and Information Sciences, ISCIS 2013, Paris, France, October 28–29, 2013. Lecture Notes in Electrical Engineering, vol. 264, pp. 369–378. Springer, New York. [http://dx.doi.org/10.1007/978-3-319-01604-7\\_36](http://dx.doi.org/10.1007/978-3-319-01604-7_36) (2013)

9. Gelenbe, E., Görbil, G., Tzovaras, D., Liebergeld, S., Garcia, D., Baltatu, M., Lyberopoulos, G.L.: Security for smart mobile networks: The NEMESYS approach. In: 2013 International Conference on Privacy and Security in Mobile Systems, PRISMS 2013, Atlantic City, NJ, USA, June 24–27, 2013. pp. 1–8. IEEE (2013). <http://dx.doi.org/10.1109/PRISMS.2013.6927181>
10. Gelenbe, E., Loukas, G.: A self-aware approach to denial of service defence. *Comput. Netw.* **51**(5), 1299–1314 (2007)
11. Gorbil, G., Abdelrahman, O.H., Gelenbe, E.: Storms in mobile networks. In: Proceedings of the 9th ACM Symposium on QoS and Security for Wireless and Mobile Networks (Q2SWinet' 14). pp. 119–126. <http://dx.doi.org/10.1145/2642687.2642688> (Sep 2014)
12. Gorbil, G., Abdelrahman, O.H., Pavloski, M., Gelenbe, E.: Modeling and analysis of RRC-based signaling storms in 3G networks. *IEEE Trans. Emerg. Top. Comput. Spec. Issue Emerg. Top. Cyber Secur.* **PP**(99), 1–14 (2015). <http://dx.doi.org/10.1109/TETC.2015.2389662>
13. Mulliner, C., Seifert, J.P.: Rise of the iBots: owning a telco network. In: Proceedings of 5th International Conference on Malicious and Unwanted Software (MALWARE' 10). pp. 71–80 (Oct 2010)
14. Pavloski, M., Gelenbe, E.: Mitigating for signalling attacks in UMTS networks. In: Czachrski, T., Gelenbe, E., Lent, R. (eds.) *Information Sciences and Systems 2014*, pp. 159–165. Springer International Publishing, New York. [http://dx.doi.org/10.1007/978-3-319-09465-6\\_17](http://dx.doi.org/10.1007/978-3-319-09465-6_17) (2014)
15. Ramachandran, S.: Web metrics: Size and number of resources. <https://developers.google.com/speed/articles/web-metrics> (May 2010)
16. Varga, A., Hornig, R.: An overview of the OMNeT++ simulation environment. In: Proc. 1st Inter. Conf. on Simulation Tools and Techniques for Communications, Networks and Systems W'shops (Simutools'08). pp. 60:1–60:10 (Mar 2008)

# A BRPCA Based Approach for Anomaly Detection in Mobile Networks

Stavros Papadopoulos, Anastasios Drosou, Nikos Dimitriou,  
Omer H. Abdelrahman, Gokce Gorbil and Dimitrios Tzovaras

**Abstract** Researchers have recently uncovered numerous exploitable vulnerabilities that enable malicious individuals to mount attacks against mobile network users and services. The detection and attribution of these threats are of major importance to the mobile operators. Therefore, this paper presents a novel approach for anomaly detection in 3G/4G mobile networks based on Bayesian Robust Principal Component Analysis (BRPCA), which enables cognition in mobile networks through the ability to perceive threats and to act in order to mitigate their effects. BRPCA is used to model aggregate network data and subsequently identify abnormal network states. A major difference with previous work is that this method takes into account the spatio-temporal nature of the mobile network traffic, to reveal encoded periodic characteristics, which has the potential to reduce false positive rate. Furthermore, the BRPCA method is unsupervised and does not raise privacy issues due to the nature of the raw data. The effectiveness of the approach was evaluated against three other methods on two synthetic datasets for a large mobile network, and the results show that BRPCA provides both higher detection rate and lower computational overhead.

---

S. Papadopoulos (✉) · O.H. Abdelrahman · G. Gorbil  
Imperial College London, London, UK  
e-mail: s.papadopoulos@imperial.ac.uk

O.H. Abdelrahman  
e-mail: o.abd06@imperial.ac.uk

G. Gorbil  
e-mail: g.gorbil@imperial.ac.uk

A. Drosou · N. Dimitriou · D. Tzovaras  
CERTH, ITI, Thessaloniki, Greece  
e-mail: drosou@iti.gr

N. Dimitriou  
e-mail: nikdim@iti.gr

D. Tzovaras  
e-mail: tzovaras@iti.gr



# 1 Introduction

Mobile networks have evolved rapidly in recent years both in terms of speed and market penetration, and users are becoming increasingly reliant on mobile devices for both personal and business use. At the same time, they have become susceptible to attacks due to the rise in cyber-crime and the fact that attackers are more organized and technically savvy than ever. This paper focuses on anomaly detection in mobile networks using aggregated control plane data (i.e. signaling) or data plane information from Call Detail Records (CDR).

## 1.1 Related Work

This section reviews the literature on detection of attacks against mobile networks and subscribers. There are a number of anomaly detection techniques that have been proposed for signaling denial of service attacks (SDoS) in 3G/4G networks. One of the first approaches was introduced in [13] using the cumulative sum (CUSUM) method, which is able to detect SDoS attacks that cannot be identified with traditional flooding-based detection systems, and is designed to make it hard for the attackers to evade detection. Another approach was presented in [14] which applies statistical techniques on time-series of unidimensional feature distributions. Specifically, empirical distributions of individual features over a predefined time interval are produced based on traffic counter values for the network subscribers. Then, the algorithm detects change points in the feature timeseries by comparing the examined distributions with previously observed ones within a time window. While this approach can cope with the non-stationarity of network traffic, it does not incorporate the observation time to increase classification accuracy, and it processes data on a per user basis which may lead to technical and privacy problems.

For the detection of attacks against mobile subscribers, [20] proposed SMS-Watchdog which uses multiple information theoretic metrics to identify abnormal activities of SMS users based on their normal past behaviors. In [12], the use of statistical metrics for the identification of mobile devices that participate in SMS-flooding attacks was suggested, showing to outperform SMS-Watchdog on a simulation dataset. One limitation of the aforementioned approaches is that they require the actual CDRs instead of aggregate traffic, and in some cases also the context of the communications [6], thus demanding high computational power and also raising privacy concerns. The use of aggregated CDR data for the detection of anomalies, as performed in the present work, has not been considered in the literature.

In addition to the above threats, security vulnerabilities of the existing femtocell architecture can be exploited for various illicit activities [2, 8]. Indeed, a compromised femtocell can be utilized to either passively monitor traffic, or send traffic on the behalf of the mobiles attached to it, effectively enabling attackers to steal user information, intercept communications, send spam or premium SMS, and even

launch DoS attacks against the core network [8]. Thus, anomaly detection in femtocell architecture has recently received some attention in the literature. A novel concept called cellpot was proposed in [15] whereby multiple femtocells act as honeypots that monitor traffic and detect anomalies related to SMS spam, mobile theft and mobile malware. A similar system was also suggested in [5] which monitors traffic from femtocells and utilizes the *Snort* open-source intrusion detection software.

## 1.2 Motivation

Traffic flows in computer and mobile networks have been studied extensively (e.g. [4]), and they are known to be non-stationary, exhibiting steep variations between certain hours of the day, with the majority of traffic variables having periodicity of 24 h following the diurnal cycle. Based on these observations and going beyond existing works which treat only the spatial characteristics (i.e. shape) of the traffic variables (e.g. curve tilt, variance), this paper proposes a BRPCA-based approach to exploit the spatio-temporal nature of the data. Unlike existing techniques which are tailored for specific attacks, the proposed approach can be deployed in many mobile network components and is able to combine information from various heterogeneous traffic variables to detect numerous types of attacks. Furthermore, the BRPCA method does not require training as it exploits the global structure of the measurements. In summary, the contributions of the present work are twofold: (a) evaluation of three existing anomaly detection methods for mobile networks, and (b) development and evaluation of a novel approach which is shown to outperform the state of the art on two synthetic mobile network datasets.

## 2 System Overview

The proposed anomaly detection system processes raw data collected by traffic monitoring sensors deployed within the mobile network. The raw data is filtered through a data preprocessing unit which utilizes Shannon's entropy in order to omit traffic variables that have entropy lower than a user-defined threshold. The anomaly detector then analyses data originating directly from the monitoring system and computes anomaly scores in realtime. Four anomaly detection techniques are examined in this paper: the first three use CUSUM, Local Outlier Factor (LOF) and Hidden Markov Model (HMM), while the fourth is our proposed BRPCA-based approach. Finally, the score produced by the anomaly detector is fed into a binary classifier which determines whether there is an ongoing attack. Note that semi-supervised anomaly detection algorithms such as HMM utilise some of the input data as ground truth of normal behavior.

## 2.1 Problem Definition

Anomaly detection refers to the process of identifying instances that do not conform to normal behavior. We formulate this problem using two matrices: a matrix  $\mathbf{G} = \{g_{i,j} | i \in [1, K], j \in [1, Y]\}$  which serves as ground truth for semi-supervised algorithms and contains only normal traffic instances; and a matrix  $\mathbf{E} = \{e_{i,j} | i \in [1, K], j \in [1, Z]\}$  describing the current traffic instances to be classified. Here  $g_{ij}$  and  $e_{i,j}$  are the elements of the matrices  $\mathbf{G}$  and  $\mathbf{E}$ ,  $K$  is the number of traffic variables,  $Y$  is the number of observations of normal traffic and  $Z$  is the number of observations that need to be evaluated. In both  $\mathbf{G}$  and  $\mathbf{E}$ , each row corresponds to a *traffic variable* and each column to an observed *traffic instance*. It should be noted that the time intervals between consecutive instances are fixed. We denote the observation sequence of a traffic variable from  $\mathbf{E}$  as  $\mathbf{E}_{row}(i) = \{e_{ij}, \forall j \in [1, Z]\}$ , where  $i \in [1, K]$  and a traffic instance as  $\mathbf{E}_{col}(j) = \{e_{ij}, \forall i \in [1 \dots K]\}$ , where  $j \in [1, Z]$ . Based on these definitions, the anomaly detection problem can be posed as the identification of the traffic instances  $\mathbf{E}_{col}(n)$  which deviate from normal behavior.

## 2.2 Evaluated State-of-the-Art Techniques

This section presents a brief overview of the application of CUSUM, LOF and HMM methods in the detection of network traffic anomalies. The CUSUM test was used in the detection of SDoS attacks [13], and it is based on the premise that anomalous events are likely to trigger several consecutive irregular measurements. CUSUM aggregates these discrepancies in consecutive measurements and signals an anomaly if a certain threshold is exceeded. However, the approach presented in [13] does not have a mechanism to reset the anomaly score which continues to affect future values even if network conditions are restored. To remedy this effect, if the score is low for a number of consecutive measurements we reset it to a default value (set to 5 in our experiments). The opposite case is also problematic, i.e. if the network conditions are not restored for a large period of time, the anomaly score will start to drop since no changes are detected. In addition, the presence of a small number of very large deviations can significantly affect the performance of the method. For these reasons, the CUSUM method is utilized in a semi-supervised manner in this paper, where annotated normal traffic instances are used to calculate the CUSUM model's parameters, and new instances are then checked for large deviations with respect to the normal traffic.

As mentioned previously, each traffic instance is modeled as a point in a  $K$ -dimensional space. The LOF method [3] operates by comparing the spatial density around a given point with the density around its  $k$  nearest points, and computes a score which indicates if the examined point resides in a low-density area or not. In the case of normal traffic instances, the LOF score is typically  $\sim 1$ , whereas abnormal instances exhibit significant deviations from this base [3]. Although the LOF method

can be unsupervised, we apply it in a semi-supervised manner by using normal traffic instances in order to define the outlier score of every new traffic instance, with respect to this (normal) traffic.

HMM is a popular technique that has been widely applied for anomaly detection [11, 19]. The considered approach is based on a Continuous Density HMM (CDHMM) [16] and consists of two steps: (a) a training step, where matrix  $\mathbf{D}$  is used to train the HMM, and (b) an evaluation step, where the HMM computes the anomaly score for every traffic instance of matrix  $\mathbf{E}$ . Each state of the HMM model corresponds to a specific time of the day, and therefore the initial state distribution was set to uniform in all training examples, since it is equally possible for an observation sequence to begin with any state. The Viterbi and Baum-Welch algorithms [1, 16] and their combination were considered in the training of the model, and the Baum-Welch algorithm was finally selected since it performed better. The trained HMMs were then used to compute the probability of each instance in the observation matrix, i.e.  $P(\mathbf{E}_{col}(j)|\lambda)$ , where  $\lambda$  denotes the parameter set of the model learned in the training phase. This probability is obtained using the forward part of the forward-backward algorithm outlined in [16].

Finally we note that the threshold, above which an anomaly score is considered to represent malicious behaviour, is manually defined for each of the algorithms presented in this paper.

### 2.3 Bayesian Robust Principal Component Analysis

This section describes how we applied BRPCA [7] in order to perform anomaly detection in mobile networks. BRPCA exploits the linear dependency of measurements across different traffic variables in order to reduce dimensionality of the data. The rationale behind the use of BRPCA is that although numerically  $\mathbf{E}$  will have full rank due to small perturbations across different traffic variables, in reality its columns are linearly dependent and its actual rank is much lower. Indeed, under normal circumstances each traffic variable in the mobile network is expected to vary in a similar fashion compared to the others during the course of a day. Based on this assumption, the observation matrix can be written as  $\mathbf{E} = \mathbf{L} + \mathbf{Q}$ , where  $\mathbf{L} \in \mathbb{R}^{K \times Z}$  corresponds to the low rank component, and  $\mathbf{Q} \in \mathbb{R}^{K \times Z}$  models small magnitude perturbations. A common method to recover the low rank matrix  $\mathbf{L}$  is PCA [1, 10] computed using Singular Value Decomposition (SVD) [17]. Initially, by applying SVD,  $\mathbf{E}$  is decomposed as:

$$\mathbf{E} = \mathbf{D}\mathbf{A}\mathbf{W}^T = \sum_{i=1}^r \lambda_i \mathbf{d}_i \mathbf{w}_i^T \quad (1)$$

where  $r$  is the rank of  $\mathbf{E}$ ,  $\mathbf{A} \in \mathbb{R}^{r \times r}$  is a diagonal matrix containing the singular values  $\lambda_i$  of  $\mathbf{E}$  sorted in descending order, and  $\mathbf{D} = [\mathbf{d}_1, \dots, \mathbf{d}_r] \in \mathbb{R}^{K \times r}$  and

$\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_r] \in \mathbb{R}^{r \times Z}$  contain the left and right singular vectors, respectively. To extract the low rank component  $\mathbf{L}$ , singular values close to zero and the respective singular vectors are ignored yielding:  $\mathbf{L} = \sum_{i=1}^{r'} \lambda_i \mathbf{d}_i \mathbf{w}_i^T$ , where  $r'$  is the number of singular values exceeding a predefined threshold  $\epsilon$  close to zero, and is also the rank of  $\mathbf{L}$ . If the linear assumption on the observation vectors  $\mathbf{E}_{row(1)}, \dots, \mathbf{E}_{row(K)}$  holds, then it is anticipated that  $r' \ll r \leq \min\{K, Z\}$ .

For mobile network traffic, the traditional PCA will readily recover the low rank component  $\mathbf{L}$  from the observation matrix  $\mathbf{E}$  in the absence of anomalous events. However, if there is an attack, e.g. malicious premium SMS messages in SMS traffic, the linearity assumption is violated and PCA will fail since arbitrary magnitude observations can change dramatically the singular values and vectors of  $\mathbf{E}$  even if they constitute a small percentage of the elements in the observation matrix. However, in order to detect attacks on mobile networks it is crucial to recognize and isolate anomalous observations. To accomplish this goal, we model the observation matrix as proposed in the BRPCA method:

$$\mathbf{E} = \mathbf{L} + \mathbf{S} + \mathbf{Q} \quad (2)$$

where as before  $\mathbf{L}$  and  $\mathbf{Q}$  correspond to the low rank (i.e. normal traffic) and noise components, whereas  $\mathbf{S} \in \mathbb{R}^{K \times Z}$  is a sparse matrix representing anomalous events that cause an arbitrary change in the observations. To recover the terms of (2), the Bayesian model used by BRPCA is:

$$\mathbf{E} = \underbrace{\mathbf{D}(\mathbf{Z}\mathbf{A})\mathbf{W}}_{\mathbf{L}} + \underbrace{\mathbf{B} \circ \mathbf{X}}_{\mathbf{S}} + \mathbf{Q} \quad (3)$$

where  $\mathbf{Z}$  is an  $N \times N$  diagonal matrix with binary entries,  $\mathbf{A} \in \mathbb{R}^{N \times N}$  is a diagonal matrix,  $\mathbf{D} = [\mathbf{d}_1, \dots, \mathbf{d}_N] \in \mathbb{R}^{K \times N}$ ,  $\mathbf{W} = [\mathbf{w}_1, \dots, \mathbf{w}_Z] \in \mathbb{R}^{N \times Z}$  and the parameter  $N$  corresponds to the largest rank that can be inferred for  $\mathbf{L}$ . The decomposition of  $\mathbf{L}$  is similar to SVD, however the matrix corresponding to the singular values is  $\mathbf{Z}\mathbf{A}$  instead of a single matrix. This permits BRPCA to decouple rank learning from singular value learning. The rank of  $\mathbf{L}$  is inferred from  $\mathbf{Z}$  and is set equal to  $\|\mathbf{Z}\|_0$ , while the magnitude of the singular values is deduced from  $\mathbf{A}$ . The sparse component  $\mathbf{S}$  is factorized into  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_Z] \in \mathbb{R}^{K \times N}$  and  $\mathbf{B} = [\mathbf{b}_1, \dots, \mathbf{b}_Z] \in \{0, 1\}^{K \times Z}$ , where  $\circ$  denotes Hadamard (pointwise) product. Notice again that this model separates the learning of sparseness from the learning of values, such that the zero component in  $\mathbf{S}$  is exactly zero. To recover the matrices of (3), BRPCA applies a Markov Chain Monte Carlo scheme in order to perform posterior inference, in which the joint posterior distribution of the model's parameters are approximated by a set of samples, iteratively drawn for each random variable, using the conditional posterior distribution given the most recent values of all the other parameters. Further details on the algorithm can be found in [7].

For the detection of time instances with anomalies, the calculated sparse matrix  $\mathbf{S} = \{s_{i,j} | i \in [1, K], j \in [1, Z]\}$  is evaluated. Each component of the sparse matrix

$s_{i,j}$  represents the anomaly score of the  $i$ -th traffic variable on the  $j$ -th instance. A traffic instance  $E_{col}(j)$  is considered anomalous if there exists at least one  $s_{i,j}$  for  $i \in [1, K]$  such that  $s_{i,j} > T_h$ , where  $T_h$  is the anomaly score threshold.

The proposed BRPCA approach is able to efficiently identify anomalies in the mobile network traffic. The reason for this is that the different traffic variables exhibit linear dependence, which is the main assumption of BRPCA, and therefore instances that deviate from this linearity, either positively or negatively, are easily identified as anomalies. However, if the linear dependence between the traffic variables does not hold, such as when the majority of the traffic instances in the dataset are anomalous, then the BRPCA method cannot work efficiently.

### 3 Performance Evaluation

The effectiveness of the proposed approach was evaluated and compared with other methods by performing experiments on two synthetic datasets. The first dataset represents a SDoS attack scenario in which 100 days of signalling traces were generated using a normal distribution model fitted from 5-day real traces collected from a 3G/4G mobile network of a European telecommunications provider. The traffic variables were collected from core network components and are in the form of aggregate counters, where each traffic variable corresponds to the number of signaling requests of each type that were sent to the examined core network component within a pre-defined time interval. A small percentage of anomalous instances ( $\sim 1\%$ ) were then manually inserted into the dataset so as to reflect the attacks and possible network component failures outlined in [18].

For the second dataset, the mobile network simulator described in [9] was extended to generate traffic traces from a network with two compromised femtocells which send 1–4 premium SMS on behalf of each user that attaches to them. The network has an area of  $5 \times 5 \text{ km}^2$ , fully covered by 7 macrocells, and with 13 *open access* femtocells distributed within the area, each with a randomly assigned range of 20 or 50 m. There are 10,000 UEs that move within the boundaries of the simulated area according to a *femtocell-aware* random waypoint mobility model as follows. Each femtocell has a rectangular *attraction* area and an attraction probability, such that when a UE decides to move, it has a probability  $p_f$  of choosing femtocell  $f$ . If the UE does not select any of the femtocell areas, which occurs with probability  $1 - \sum_f p_f$ , then it moves to a random destination that may still be covered by a femtocell. The amount of time the UE spends in its chosen destination before repeating this process is drawn from a probability distribution specific to the cell type. The UEs have a diurnal cycle which is also followed by the compromised femtocells to avoid raising suspicions. There are three classes of SMS users in the network: light, moderate and heavy. Each UE generates messages to chosen destinations, and also responds to received messages with a given probability. The destination address (number) of a generated message determines its type, which can be either in-network mobile, out-network mobile, premium, or other. In-network mobiles are naturally represented by

**Table 1** Comparison of the different anomaly detection methods

	Speed	Accuracy	Unsupervised
CUSUM	✓	–	–
LOF	✓	✓	–
HMM	–	✓	–
BRPCA	✓	✓	✓

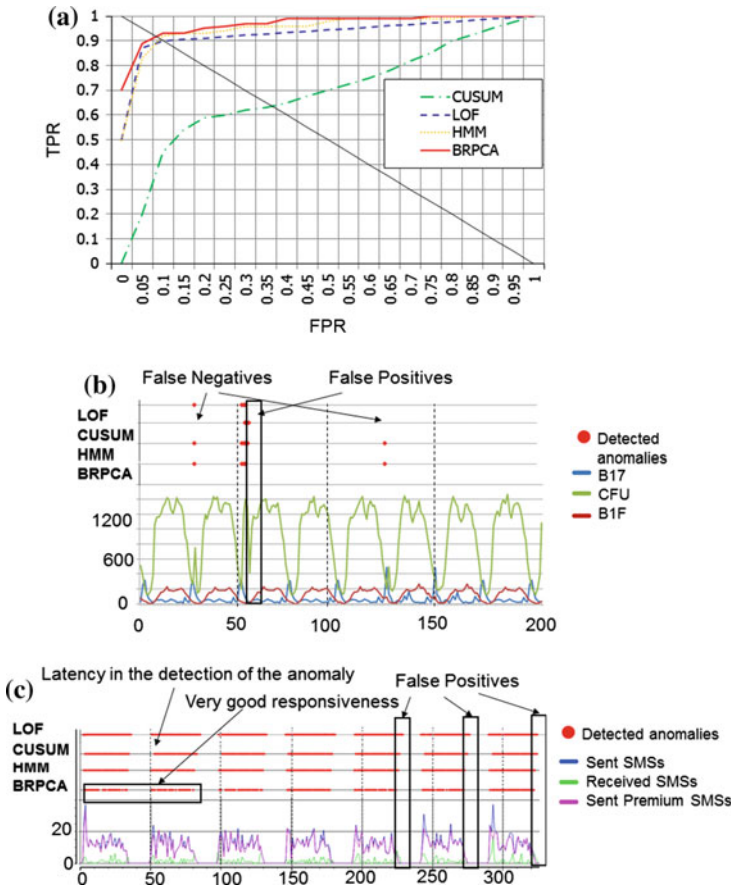
the UEs explicitly simulated, while the rest of the destinations are represented by servers outside the mobile network. The aggregated data in this case represent the number of sent, received, and sent premium SMSs.

Table 1 presents a high level comparison of the different anomaly detection methods considered in this paper. We found the CUSUM method to be very fast, but it performed poorly in the scenario under investigation. Although LOF and HMM performed well, they involve a training phase that can be computationally intensive. Furthermore, the training procedure also has the risk of over-fitting thus and making inaccurate generalizations. On the other hand, the BRPCA on top of being fast and very accurate, it is also unsupervised since it exploits the correlations between the different traffic variables.

Figure 1a illustrates the performance of the four methods on the SDoS dataset, using ROC curves. The results show clearly that the BRPCA approach outperforms the other methods, since its area under curve is larger. On the other hand, CUSUM performs the worst while LOF and HMM appear to be reliable but are inferior to BRPCA. Moreover, when comparing the methods' curves at the point of equal error rate, the BRPCA, HMM, and LOF perform much better than CUSUM, while the BRPCA maintains its superior performance over HMM and LOF.

Figure 1b shows detection results (upper half) at different time instants during the first 200h of the SDoS dataset. The lower half of the figure depicts the temporal behavior of the following signaling messages: B17 (General data transfer with circuit duplex asynchronous), CFU (Call Forwarding) and B1F (General data transfer with circuit duplex synchronous). One can see that BRPCA and HMM clearly outperform the other methods, accurately identifying periods in which the distribution of the signals changes, with fewer number of false positive and false negative alerts. Specifically, CUSUM has a larger number of false positives and higher latency in detecting when the network is under attack. Also, both LOF and CUSUM generate false negatives in this scenario, since they are unable to identify some of the increases in the network signaling, particularly in the CFU signal. HMM performs well and is able to detect the time periods of anomalies but with a small number of false positives. The BRPCA method achieves the best performance in this case, since it is able to detect all the anomalous instances with zero false positives.

Detection results for the compromised femtocell dataset are presented on the upper half of Fig. 1c. The lower half of the figure shows the SMS activity over time in terms of sent, received, and sent premium messages. Again, the CUSUM method suffers from high latency in the detection of anomalies, due to the fact that it requires more time to accumulate high anomaly scores. Furthermore, compared to BRPCA,



**Fig. 1** Evaluation of the LOF, CUSUM, HMM, and BRPCA methods. **a** ROC curves for the SDoS dataset. **b** Detection results for the first 200h of the SDoS dataset. **c** Detection results at different time instants on the compromised femtocell dataset

all the other methods generate a larger number of false positives (highlighted in the figure). BRPCA shows very good responsiveness in identifying periods in which the number of premium SMS is large or small, unlike the other methods which continue to report anomalies when the number of premium SMS is small in one time period and large in the time periods right before and after it. This also contributes to the reduction of false positives reported by BRPCA.

## 4 Conclusions

Anomaly detection approaches for mobile networks have traditionally considered only the spatial characteristics of the network traffic. However, various studies have shown that network traffic variables are non-stationary and exhibit 24h periodicity.



Thus the temporal characteristics of mobile network traffic convey valuable information for the detection of anomalies. Based on this premise, this paper proposed a novel BRPCA-based anomaly detection approach which represents network traffic as a sequence of traffic variable vectors, thus allowing to exploit the spatio-temporal characteristics of the network traffic. The efficiency of the proposed approach was evaluated against three state-of-the-art methods on two synthetic datasets corresponding to a signalling DoS attack against the mobile network core and a femtocell-based attack targeting mobile users. The experimental results indicate that BRPCA outperforms the other methods on both datasets.

**Acknowledgments** This work was partially supported by the European Commission through project FP7-ICT-317888-NEMESYS. The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the European Commission.

## References

1. Bishop, C.M.: Pattern recognition and machine learning, vol. 1. Springer, Berlin (2006)
2. Borgaonkar, R., Redon, K., Seifert, J.P.: Security analysis of a femtocell device. In: Proceedings of 4th International Conference on Security of Information and Networks (SIN), pp. 95–102. ACM, Sydney (2011)
3. Breunig, M.M., Kriegel, H.P., Ng, R.T., Sander, J.: LOF: identifying density-based local outliers. SIGMOD Rec. **29**(2), 93–104 (2000)
4. D’Alconzo, A., Coluccia, A., Romirer-Maierhofer, P.: Distribution-based anomaly detection in 3G mobile networks: from theory to practice. Int. J. Netw. Manag. **20**(5), 245–269 (2010)
5. David, S., Harrison, D., Price, R., Fretheim, S.: Do-it-yourself cellular intrusion detection system. LMG Security, [http://lmgsecurity.com/whitepapers/DIY-Cellular-IDS\\_2013-08-01.pdf](http://lmgsecurity.com/whitepapers/DIY-Cellular-IDS_2013-08-01.pdf) (2013)
6. Delany, S.J., Buckley, M., Greene, D.: SMS spam filtering: methods and data. Exp. Syst. Appl. **39**(10), 9899–9908 (2012)
7. Ding, X., He, L., Carin, L.: Bayesian robust principal component analysis. IEEE Trans. Image Process. **20**(12), 3419–3430 (2011)
8. Golde, N., Redon, K., Borgaonkar, R.: Weaponizing femtocells: the effect of rogue devices on mobile telecommunications. In: Proceedings of NDSS, pp. 1–16 (2012)
9. Gorbil, G., Abdelrahman, O.H., Pavloski, M., Gelenbe, E.: Modeling and analysis of RRC-based signalling storms in 3G networks. IEEE Trans. Emerg. Topics Comput **PP**(99), 1 (2015)
10. Jolliffe, I.T.: Principal Component Analysis, 2nd edn. Springer, Berlin (2002)
11. Joshi, S.S., Phoha, V.V.: Investigating hidden Markov models capabilities in anomaly detection. In: Proceedings of 43rd Annual Southeast Regional Conference, vol. 1, pp. 98–103. ACM, Kennesaw (2005)
12. Kim, E.K., McDaniel, P., La Porta, T.: A detection mechanism for SMS flooding attacks in cellular networks. In: Security and Privacy in Communication Networks, vol. 106, pp. 76–93. Springer, Berlin (2013)
13. Lee, P.P.C., Bu, T., Woo, T.: On the detection of signaling DoS attacks on 3G wireless networks. In: Proceedings of INFOCOM, pp. 1289–1297. IEEE, Alaska (2007)
14. Lee, P.P.C., Bu, T., Woo, T.: On the detection of signaling DoS attacks on 3G/WiMax wireless networks. Comput. Netw. **53**(15), 2601–2616 (2009)
15. Liebergeld, S., Lange, M., Borgaonkar, R.: Cellpot: a concept for next generation cellular network honeypots. In: Proceedings of NDSS, pp. 1–6 (2014)

16. Rabiner, L.: A tutorial on hidden Markov models and selected applications in speech recognition. *Proc. IEEE* **77**(2), 257–286 (1989)
17. Strang, G.: *Linear Algebra and its Applications*, 4th edn. Brooks/Cole (2006)
18. Traynor, P., et al.: On cellular botnets: measuring the impact of malicious devices on a cellular network core. In: *Proceedings of 16th Conference Computer and Communications Security (CCS)*, pp. 223–234. ACM, Chicago (2009)
19. Wang, F., et al.: A HMM-based method for anomaly detection. In: *Proceedings of 4th International Conference on Broadband Network and Multimedia Technology*, pp. 276–280. IEEE, Shenzhen (2011)
20. Yan, G., Eidenbenz, S., Galli, E.: SMS-Watchdog: Profiling social behaviors of sms users for anomaly detection. In: *Proceedings of 12th International Symposium Recent Advances in Intrusion Detection (RAID)*, pp. 202–223. Springer, Saint-Malo (2009)

# LBP-DCT Based Copy Move Forgery Detection Algorithm

Beste Ustubioglu, Guzin Ulutas, Mustafa Ulutas, Vasif Nabiyevev and Arda Ustubioglu

**Abstract** Increase on the availability of the image editing software makes the forgery of the digital image easy. Researchers proposed methods to cope with image authentication in recent years. We proposed a passive image authentication technique to determine the copy move forgery. First, the method divides the image into overlapping blocks. It uses LBP (Local Binary Pattern) to label each block. Labeled blocks are transformed into frequency domain using DCT (Discrete Cosine Transform). Sign values of the first fifteen coefficients of the zigzag scanned block plus average Y, Cb, Cr values constitutes the feature vector for the block. Finally, the feature vectors are lexicographically sorted and element-by-element similarity measurement is used to determine the forged blocks. Experimental results show that the method has higher accuracy ratios and lower false negative values under some post processing operation compared to other DCT based methods. Our method can also detect multiple copy move forgery.

**Keywords** Copy move forgery · LBP · DCT

---

B. Ustubioglu (✉) · G. Ulutas · M. Ulutas · V. Nabiyevev · A. Ustubioglu  
Karadeniz Technihal University, Trabzon, Tukey  
e-mail: bgencturk@ktu.edu.tr

G. Ulutas  
e-mail: gulutas@ktu.edu.tr

M. Ulutas  
e-mail: ulutas@ktu.edu.tr

V. Nabiyevev  
e-mail: vasif@ktu.edu.tr

A. Ustubioglu  
e-mail: ardaustubioglu@ktu.edu.tr

## 1 Introduction

The demand for high quality images resulted in mass production of gadgets capable of capturing and sharing media through on-line services. Digital images are widely used in many fields such as medicine for diagnosis, evidence in judiciary and in press to report news. Users want to edit images to improve quality by image editing applications. Easy to use linear and non-linear image editing tools can be used to tamper images. Thus, approving the fidelity of digital images is a challenging problem. Researchers suggest techniques to examine the originality of digital images. Techniques reported in the literature can be grouped into active and passive methods.

Digital watermarking or digital signatures fall into the first category. The methods in this category require additional information to be transmitted and they also need key management procedures. Passive forgery detection techniques use only some features of the image to judge the originality of the image. They don't necessitate any extra information and the image can be authenticated even if it has not been taken with a specially equipped camera. The advantages of the passive techniques make them attractive to researchers in recent years.

One of the most popular forgery techniques is the copy move forgery since even a novice can select, copy and paste on the same image by freely available image editing tools. This type of forgery copies a portion of the image and pastes it into another region on the same image to hide an object or to replicate an object. Fridrich et al.'s work is the first attempt in the literature to detect the copy move forgery operation [1]. Their method is not robust to against some post processing operations such as scaling, translation and rotation. After this work, Hanry and Farid proposed a method with principal component analysis [2]. Their method aims to decrease the dimension of the feature vector. Experimental results show that their method is robust to against the JPEG compression attack and additive noise. Luo et al. extract seven features from the overlapping blocks and use these vectors to determine the forgery [3]. Their method calculates the average intensity values from the RGB (red, green, blue) channels and extracts diagonal ratios from the blocks. Results show that the method can detect forged regions even if the post processing operations has been applied after the forgery operations. In recent years, many research based on DCT are proposed to make their methods robust against JPEG compression [4–8]. In 2011, Huang et al. quantize DCT coefficients after zigzag scanning [4]. The method also uses first  $k$  elements of the feature vector to decrease the dimension of the vector. Results show that the method can detect forgery operation even if the forged image has been compressed using JPEG transform. After this work, Cao et al. suggested to divide the DCT coefficients extracted from a sub-block into four regions [5]. Gaussian blurring or noise addition operation has no effect on the performance of the method.

We proposed a method based on LBP and DCT. Our method is the first one to combine LBP with DCT in copy move forgery. Results show the effectiveness of the method to detect the copy move forgery compared to other DCT based methods. Our

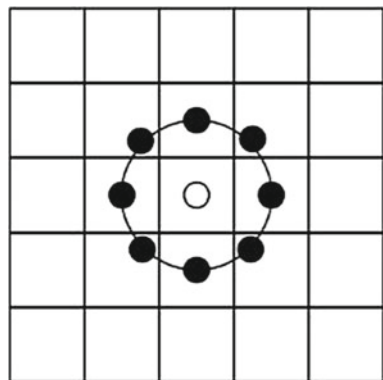
method divides the image into overlapping blocks of size  $8 \times 8$ . LBP value is obtained from each block and DCT is applied on the LBP values. After zigzag scanning of the DCT coefficients, we extract the sign information of the first 15 elements of the vector. Feature vectors of size  $1 \times 15$  will be compared to determine the forgery. Experimental results show that the method gives higher accuracy ratios and lower false negative values compared to other DCT based works in the literature [4, 5]. The rest of the paper is organized as follows. Section 2 gives a brief introduction to LBP approach. The method is explained with details in Sect. 3. The experimental results and conclusions are given in Sects. 4 and 5 respectively.

## 2 General View of the LBP

Ojala et al. assume that a texture has locally two complementary aspects, a pattern and its strength [9]. Their method works in a  $3 \times 3$  pixel block of the images. The central pixel of the block is used to binarize the other pixels in the current block. The eight neighborhoods of the center pixel becomes 1 and 0s. They are multiplied by powers of two and then summed to obtain a label for the center pixel. A total of 256 different labels can be generated by the method for a center pixel. Ojala et al. proposed a generalized version of LBP after years [6]. Generic LBP has no restriction on the number of sampling points and size of the neighborhood. However basic LBP uses eight pixels of a  $3 \times 3$  block. Assume that gray level image denoted by  $I$  and a pixel at  $x$ th row and  $y$ th column denoted by  $i_c = I(x, y)$ . Circular neighborhood of point  $(x, y)$  with evenly spaced  $P$  sampling points and radius  $R$  will be used. Let any point  $i_p$  denotes the gray value of sampling point from the  $P$  points. Sampling point  $(x, y)$  at will be calculated as in (1). Local neighborhood for  $R = 1$  and  $P = 8$  is shown in Fig. 1.

$$i_p = I(x_p, y_p), p = 0, \dots, P - 1 \quad x_p = x + R \cos(2\pi p/P) \quad y_p = y - R \sin(2\pi p/P) \tag{1}$$

**Fig. 1** Neighborhood of center pixel for  $R = 1$  and  $P = 8$



Assume that  $s(t)$  is the step function. If  $t \geq 0$ , the function returns one otherwise the function returns zero. Using the step function, LBP value of a point with radius  $R$  and  $P$  sample points can be calculated as in (2).

$$LBP(x_c, y_c) = \sum_{p=0}^{p-1} s(i_p - i_c)2^p \quad (2)$$

### 3 Proposed Method

The proposed method consists of two stages: (i) feature extraction and (ii) matching and marking.

**Feature Extraction** The method converts the test image into YCbCr space before the feature extraction to make the method more robust against JPEG compression attacks. Assume that the test image of size  $N \times M$  in YCbCr space denoted by  $I$ . The test image is then divided into overlapping sub blocks of size  $8 \times 8$ . Each block of the image denoted by  $B_i, i = 1 \dots (N - 7)(M - 7)$  is used to extract the features. For each block the algorithm will apply the following steps.

*Step 1* Calculates the average intensity values of the three channels (Y, Cb, Cr) for the current block. Let average values for each channel denoted by  $Y_{avg}, Cb_{avg}, Cr_{avg}$ .

*Step 2* Obtain the LBP values for each pixel on the current block with Radius = 2 and Points = 8 using generic LBP operator given in (2). Let the new block with LBP labels generated for the current block  $B_i$  be  $L_i$ . In this step we use the Y channel of the current block.

*Step 3* The block that contains LBP labels  $L_i$  is transformed into frequency domain using DCT. The new block after the transformation is denoted by  $D_i = DCT(L_i)$ .

*Step 4* The block  $D_i$  in the frequency domain is zigzag scanned and first fifteen elements of the ZigZag scanning operation are extracted to create the vector  $V_i$ . The algorithm takes the first a few coefficients to make the method more robust against compression attacks.

*Step 5* Vector  $S_i = \{s_{ij} | j = 1 \dots 15\}$  which contains sign information of the DCT coefficients, is determined as in (3). Corresponding vector  $V_i = \{v_{ij} | j = 1 \dots 15\}$  is used to determine the sign vector.

$$\forall v_{ij} < 0 \Rightarrow s_{ij} = -1 \quad \forall v_{ij} > 0 \Rightarrow s_{ij} = 1 \quad \forall v_{ij} = 0 \Rightarrow s_{ij} = 0 \quad (3)$$

*Step 6* Sign vector  $S_i$  and average intensity values of the channels  $y_{avg}, Cb_{avg}, Cr_{avg}$  constitute the corresponding feature vector  $F_i = \{f_{ij} | j = 1 \dots 18\}$  of size  $1 \times 18$  as in (4).

$$\forall j \in [1 - 15] \Rightarrow f_{ij} = s_{ij} \quad f_{i16} = Y_{avg}, f_{i17} = Cb_{avg} \quad f_{i18} = Cr_{avg} \quad (4)$$

Feature extraction algorithm generates a corresponding block for each block, which contains LBP labels and then transform the LBP block to frequency domain

using DCT. DCT transform ensures the method robustness against JPEG compression attacks. We use only sign information from the DCT coefficients that can take values from  $(\{-1, 0, 1\})$ . The matching algorithm to eliminate comparison of the blocks with different color characteristic uses last three elements of the vector. Feature vectors of each block constitutes a matrix denoted by  $A$  of size  $[(N-7)(M-7), 18]$ . The matrix is lexicographically sorted to make the similar vectors closer. Lexicographically sort operation before the matching procedure decreases the processing time because a vector will be compared to only predefined amount of neighboring feature vectors to judge the similarity.

**Matching and Marking** The proposed method will determine the similarity between the vectors in this section and generate shift vectors from the similar vectors. Matrix  $A$  that contains feature vectors is consulted to judge the similarity. Each feature vector is compared to  $t_n$  feature vectors to determine the similarity. Processing time of the matching algorithm will be decreased because the algorithm uses fixed number of neighboring feature vectors for comparison. If any two vectors are similar, the algorithm will calculate the corresponding shift vector and records the vector into a list denoted by  $sh$ . Matching algorithm for a feature vector denoted by  $F_i$  can be given in the form of steps as the following. Vector  $F_i$  will be compared to  $t_n$  feature vectors. Following algorithm will be applied to all feature vectors in the matrix  $A$ .

*Step 1* Last three elements of the vector  $F_i$  is compared to next vector  $F_j$  determine the color characteristic of the vectors. If the following condition is satisfied, the algorithm continues with the next step, otherwise step 1 is repeated to evaluate the next vector. The algorithm to determine if two vectors correspond to blocks with similar color distribution uses threshold  $t_c$ .

$$\forall (|f_{j16} - f_{i16}| < t_c) \wedge (|f_{j17} - f_{i17}| < t_c) \wedge (|f_{j18} - f_{i18}| < t_c) \Rightarrow \text{goto step 2.} \\ \text{otherwise} \Rightarrow k = k + 1; \text{goto step 1}$$

*Step 2* If two vectors are similar, the next condition is about the how two blocks are far away. This step ensures that similar vectors do not represent a smooth region. Assume that upper left coordinate of the vectors  $F_i$  and  $F_j$  be  $(x_i, y_i)$  and  $(x_j, y_j)$  respectively. The following criterion is used to test the distance between blocks is appropriate or not. The method necessitates that the distance between the two block must be at least threshold  $t_d$ .

$$\forall \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} \geq t_d \Rightarrow \text{goto Step 3, otherwise} \Rightarrow \text{goto step 1}$$

*Step 3* The following condition will be tested for  $k = 1 \dots 15$  to determine the similarity between the elements of the vectors.

$$\forall (f_{ik} - f_{jk}) = 0 \Rightarrow cnt = cnt + 1$$

If two vectors are the same, value of  $cnt$  will be 15. The method compares  $cnt/15$  with predefined threshold value  $t_s$  to test the similarity. If  $cnt/15 > t_s$ , the algorithm will continue with the next step. Otherwise, next step to be executed will be 1.

*Step 4* Shift vector between two vectors is calculated and inserted into list  $sh$ . The same list also accommodates the upper left coordinates of the corresponding blocks. The vector  $[[x_i - x_j || y_i - y_j | x_i y_i x_j y_j]$  will be inserted into list  $sh$  to represent the current comparison. After insertion, the algorithm will continue with the step 1 to evaluate the next vector. The algorithm will sign the corresponding blocks of the image if the amount of the same shift vectors exceeds predefined threshold value  $t_{shift}$ . The method must choose a smaller value for  $t_{shift}$  to detect the small copied and pasted regions and vice versa.

The list  $sh$  will be scanned to determine the same shift vectors. Assume that the amount of vectors whose first two elements are the same be  $f$ . If the value of  $f$  is bigger than  $t_{shift}$ , the marking algorithm signs the blocks corresponding to these vectors.

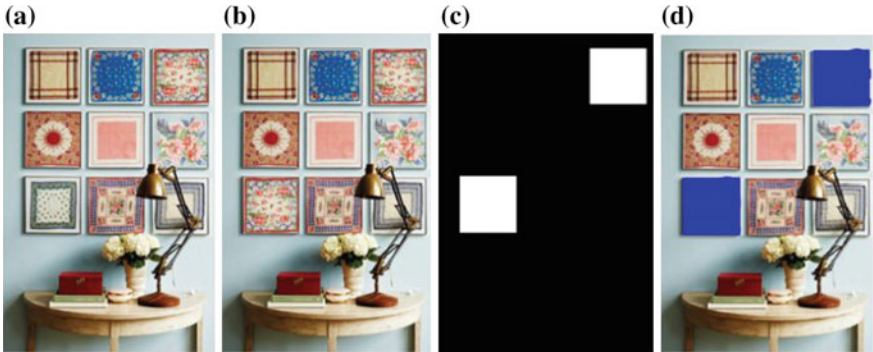
## 4 Experimental Results

This section gives the detailed analysis to show the effectiveness of the method. All the experiments will be realized on a 2.4 GHz dual core (4 threads with hyper-threading) Intel Core i5 processor. The tampered images were created by an open source image editing software, GIMP, using images of size  $200 \times 200$  and  $300 \times 350$  pixels from Google image search . In our experiments we set all the parameters as  $t_n = 50$ ,  $t_c = 2.8$ ,  $t_d = 40$ ,  $t_s = \frac{13}{15} = 0.86$ ,  $t_{shift} = 200$ . The success of the detection method is measured with a metric called accuracy ratio  $p$  in  $[0-1]$  range. An accuracy ratio of 1 corresponds to detection of all copied and pasted regions. Accuracy ratio decreases if the method marks regions not forged as forged. False negative ratio,  $f$  is also used during experiments to have a measure of the regions detected as forged whereas they are not. The success of the detection algorithm improves as the value of  $f$  approaches to zero. Let copied and pasted regions in a fake image be  $D_1$  and  $D_2$  respectively, whereas copied and pasted regions detected by the algorithm be  $R_1$  and  $R_2$  respectively. The accuracy ratio of the algorithm is calculated by using (5).

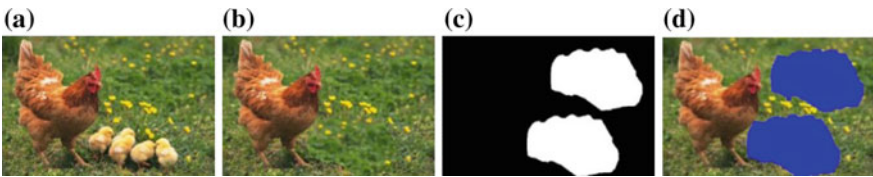
$$p = \frac{|D_1 \cap R_1| + |D_2 \cap R_2|}{|D_1| + |D_2|}, f = \frac{|D_1 \cup R_1| + |D_2 \cup R_2|}{|D_1| + |D_2|} - p \quad (5)$$

The first experiment is realized to show the visual results of the detection algorithm when the copied and pasted region has regular shape. The original image given in Fig. 2a is changed using the mask given in Fig. 2c to create the forged image given in Fig. 2b. The visual result of the detection algorithm (Fig. 2d) shows that the method can detect the copied and pasted region. Accuracy ratio and false negative value are also reported to be approximately 1.00 and 0.01 respectively.

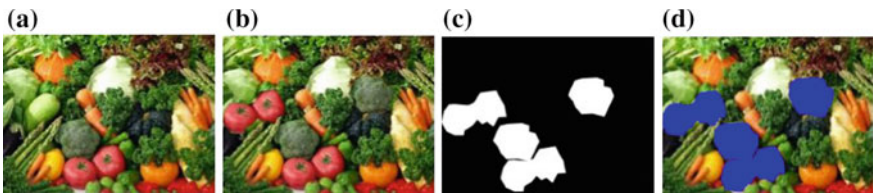




**Fig. 2** a Original image b Forged image c Mask of the forgery operation d Visual result of the detection algorithm



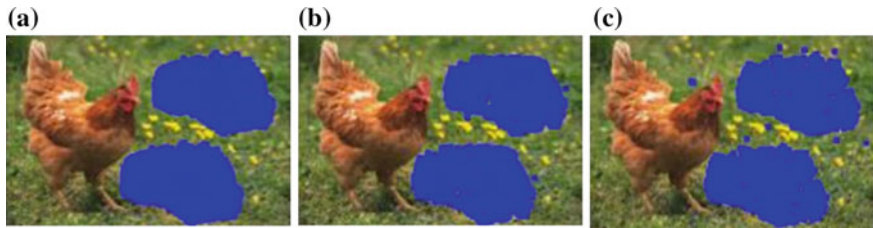
**Fig. 3** a Original image b Forged image c Mask of the forgery operation d Visual result of the detection algorithm



**Fig. 4** a Original image b Forged image c Mask of the forgery operation d Visual result of the detection algorithm

The second experiment gives an idea about the capability of the method when the copied and pasted regions have non-regular shape. Original image, forged image, mask image are given in Fig. 3a, b and c respectively. Visual result of the method shown in Fig. 3d designates that the method can detect the forged regions even if the forged regions have non-regular shape. Accuracy ratio and false negative values are approximately 0.99 and 0.04 respectively.

Multiple copy move forgery is also realized as the third experiment to show the effectiveness of the method when the forged image has more than one region. Multiple region of the original image given in Fig. 4a are used to create the forgery. Forged image given in Fig. 4b is created using the mask image given in Fig. 4c. The result of the detection algorithm shown in Fig. 4d demonstrates copied and forged regions.



**Fig. 5** Visual result of the detection algorithm for different quality factors

**Table 1** Comparison of the methods under JPEG compression attack

QF	Proposed method		[4]	
	p	f	p	f
Q = 90	0.999	0.008765	0.996	0.009
Q = 80	0.987	0.02135	0.98	0.027
Q = 50	0.969	0.05955	0.729	0.277
Q = 40	0.9513	0.1121	0.729	0.277

The method gives approximately 0.97 accuracy ratio even if the original image has multiple forged regions. We also test the proposed method when the forged region has been compressed after the forgery operation. JPEG compression algorithm will be used with various quality factors (QF) to show the effectiveness of the method under post processing operation.

Figure 5 shows the visual detection results of the method. Results show that the method can detect forgery approximately 0.92 accuracy ratio even if it has been compressed with 30 quality factor. False negative values for three results also smaller than 0.1.

The proposed method is compared to similar DCT based methods in the literature to show the effectiveness of the method under post processing operations [4, 5]. Three experiments are realized for comparison purposes. First experiment given in Table 1 shows the capability of the method when the forged image has been compressed with different quality factors. The method shows higher accuracy ratios and lower false negative values compared to [4] as can be seen in Table 1. The figure does not contain the result of [5] since it does not exhibit robustness against JPEG compression. Another experiment compares the method with similar works in the literature when the Gaussian blurring is used as post processing operation. The tampered images are blurred by a Gaussian filter with parameters (*windows size*,  $w = 5$ ,  $\sigma = 0.5$ ,  $\sigma = 1$ ) and ( $w = 3$ ,  $\sigma = 0.5$ ,  $\sigma = 1$ ). Table 2 shows the  $p$  and  $f$  values of the similar works and the proposed method [4, 5]. The proposed method has higher accuracy compared to other works even though with larger blurring radius. In the last experiment, we test the proposed method on the AWGN distorted images. Forged images are distorted

**Table 2** Comparison of the methods under Gaussian Blurring attack

	Proposed method		[4]		[5]	
	p	f	p	f	p	f
w = 3 $\sigma = 0.5$	0.993	0.081	0.922	0.091	0.9584	0.0516
w = 3 $\sigma = 1$	0.998	0.053	0.915	0.089	0.9464	0.0546
w = 5 $\sigma = 0.5$	0.998	0.062	0.888	0.162	0.93	0.055
w = 5 $\sigma = 1$	0.984	0.064	0.86	0.162	0.934	0.059

**Table 3** Comparison of the methods under AWGN attack

	Proposed method		[4]		[5]	
	p	f	p	f	p	f
25 db	0.9924	0.0368	0.8967	0.1068	0.95	0.05
40 db	1	0.0399	0.9987	0.0282	1	0.0348

with different additive white Gaussian noise (SNR = 25 and 40 dB). The method leads to higher accuracy ratios even if the forged image has been distorted with 25 dB noise as can be seen in Table 3.

## 5 Conclusion

We have proposed a new method that is the first one in the literature to combine LBP with DCT in copy move forgery. The method works without any digital watermarks or signature information. Compared with the DCT based methods in the literature, the method detects the copied and pasted regions with higher accuracy ratios and lower false negative values even if the image has undergone some post processing operations. The results also show that sign information of the DCT coefficients is enough to determine the forgery. We believe that the method also contribute to the DCT based copy move forgery detection works in the literature [4, 5].

## References

1. Fridrich, J.: Detection of Copy-Move Forgery in Digital Images. Digital Forensic Research Workshop, Cleveland, OH (2003)
2. Popescu, A.C., Farid, H.: Exposing digital forgeries by detecting duplicated image regions, technical report TR2004-515, Department of Computer Science, Dartmouth College (2004)

3. Luo, W., Huang, J., Qiu, G.: Robust detection of region duplication forgery in digital images. *Int. Conf. Pattern Recogn.* **4**, 746–749 (2006)
4. Huang, Y.: Improved DCT-based detection of copy-move forgery in images, *Forensic Sci. Int.* **206**(1–3), 178–184 (2011)
5. Cao, Y., Gao, T., Fan, L., Yang, Q.: A robust detection algorithm for copy-move forgery in digital images. *Forensic Sci. Int.* **214**, 33–43 (2012)
6. Ojala, T., Pietikäinen, M., Mäenpää, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(7), 971–987 (2002)
7. Ghorbani, M., Firouzmand, M., Faraahi, A.: DWT-DCT (QCD) based copy-move image forgery detection. In: *International Conference on Systems, Signals and Image Processing (IWSSIP)*, pp. 1–4 (2011)
8. Kumar, S., Desai, J., Mukherjee, S.: A fast DCT based method for copy move forgery detection. In: *2nd IEEE International Conference on Image Information Processing (ICIIP)*, pp. 649–654 (2013)
9. Ojala, T., Pietikäinen, M., Harwood, D.: A comparative study of texture measures with classification based on feature distributions. *Pattern Recogn.* **19**(3), 51–59 (1996)

# Smart Mobile Ecosystem Security: Existing Solutions, MNO Requirements and Business Model

George Lyberopoulos, Helen Theodoropoulou, Konstantinos Filis and Ioanna Mesogiti

**Abstract** Mobile Network Operators (MNOs) are starting to investigate ways of securing their networks to defend against the significant increase of malicious attacks targeting smart mobile devices and access/core mobile networks. This increase signifies a new tendency: that of attacking the smart mobile ecosystem. In order for MNOs to be prepared to protect their networks and subscribers, they need to investigate the market and locate efficient security solutions that can be applied to their networks. However, since the market in this field is quite immature, and since new solutions are on their way, it may be proved quite beneficial for the MNOs to start setting their requirements, so that security solution vendors can take these specific requirements into consideration. The implementation of these security solutions will require significant investments from MNOs forcing them to develop new business models for offering new security services.

## 1 Introduction

It is a fact that malicious software targeting mobile devices, otherwise known as mobile malware, is on the rise. According to the latest ALU's Motive Security Labs malware report for the second half of 2014 [1], infections in the smart mobile ecosystem accelerated with an increase of 25 % in 2014, compared with 20 % for 2013. The percentage of infected mobile devices in the 2nd half of 2014 was 0.68 %, which means that, worldwide, about 16 million mobile devices are infected by malware.

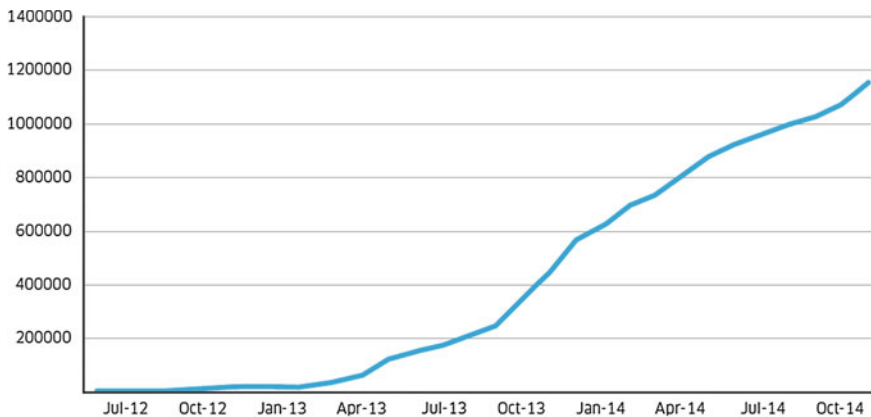
---

G. Lyberopoulos (✉) · H. Theodoropoulou · K. Filis · I. Mesogiti  
COSMOTE - Mobile Telecommunications SA, Athens, Greece  
e-mail: glimperop@cosmote.gr

H. Theodoropoulou  
e-mail: etheodorop@cosmote.gr

K. Filis  
e-mail: cfilis@cosmote.gr

I. Mesogiti  
e-mail: imesogiti@cosmote.gr



**Fig. 1** Android malware samples growth since 2012 [1]

Even though these numbers may appear low, what we need to take into consideration is the fact that the rate of increase of Android malware samples soared to 161 % in 2014, as shown in Fig. 1, substantiating the prediction that the amount of malware attacks against mobile devices will grow significantly over time [3]. Mobile malware is also continuously increasing in sophistication with more robust command and control (C&C) protocols. Regarding the percentage of infections with respect to the OS, the majority is split 50/50 between Android devices and Windows/PCs (i.e. PCs connected to the mobile network via dongles and mobile Wi-Fi devices or simply tethered through smartphones).

It is a fact that no OS can be considered as immune nowadays. According to a research released in April 2015, approximately 1,500 iOS applications were exposed to a vulnerability that could let a hacker bypass HTTPS security and steal passwords and other sensitive data from iOS devices [2]. In the real world, malware can reach the mobile device via Bluetooth, SMS and MMS messaging, e-mail or Internet access and is installed either by taking advantage of the vulnerabilities found on mobile devices, or with the user's full consent.

The main driving force behind MNO's interest in network security is the fact that, within the mobile ecosystem, security threats cannot be contained merely within mobile devices. They can easily affect mobile networks, too. Indicatively, in Japan, NTT DoCoMo experienced a signaling flood that disrupted network access in January 2012, caused by a VoIP OTT application running on Android phones [5]. In fact, the transition of mobile networks to IP brings additional security threats, such as DoS and DDoS attacks, ping floods, DNS hijacking, IP port scanning [4], which may result in the interception of subscriber data, limit subscriber access (causing congestion), and/or compromise the overall network security of the network, since some of the core elements' functionality may be lost. In addition, the cost of a potential cyber crime against the MNO's business must include the damage on the company brand name and reputation, the affected productivity, the loss of intellectual property and business sensitive information, recovering from the attack, the list not

being exhaustive. Therefore, traditional security methods are not safe anymore and new ones are needed.

This paper serves the purpose of presenting the current mobile ecosystem security solutions, the security requirements that an MNO may have and the business model that can result in a profitable investment for such a solution. The material included in this paper is organized as follows: Sect. 2 gives an overview of the existing and planned mobile ecosystem security solutions by major telecommunications vendors. In Sect. 3 we present the MNO requirements towards an efficient security solution. In Sect. 4 we elaborate on a business model that can be considered by the MNO in order to protect its investment, while in Sect. 5 we draw some concluding remarks.

## 2 MNO Requirements

The actors (and users) of a smart mobile ecosystem security solution are mainly the mobile end-users and the MNO security analysts. Therefore, the “security tools” can be logically divided into two main parts: the user equipment-related and the mobile network-related ones.

### 2.1 *User Related Requirements*

The operational requirements for a security framework, which has to address the security issues/attacks against the user, are the following:

- Availability of a lightweight real-time anomaly detection engine running on various end-user devices and various OSs.
- No or extremely low impact on the end-user device performance (e.g. battery life, processing power) caused by the anomaly detection/protection engine.
- Real-time notification of an attack to the security analyst and immediate notification to the mobile user with specific instructions.
- Notification regarding the availability of an anomaly-detection/protection engine SW updated version.
- The anomaly detection engine should be offered to the user free of charge.
- The anomaly detection engine has to be intrinsically secure (it should not be supposed to be vulnerable to attacks).
- The whole process must be transparent to the mobile user.

### 2.2 *System Related Requirements*

The operational requirements for a security framework, which has to address the security issues and attacks against the mobile network, are summarized below:

- Availability of a real-time detection framework (set of collaborative tools) for the presentation and analysis of abnormal behaviour that may be identified either on user devices or the mobile network. Indicatively, the graphical environment (GUI) should:
  - Be secure, intuitive and user-friendly
  - Perform attack attribution in real time
  - Be parameterized by the security analyst and offer the capability to run existing analytical queries and/or create new ones
  - Provide information regarding the severity, geographical extent, etc. of the attack and steps to follow for the resolution of the attack
  - Generate reports.
- Zero impact on access and core network availability and performance end-to-end.
- Always rely upon up to date data in order to identify new attack types.
- Be equipment (vendor) independent (so as to be applicable in a multi-vendor mobile network).
- Be able to collect data from various sources. Indicatively from terminals, (e.g. log file(s) generated upon detection of an anomalous event), the network (e.g. traffic statistics, Gn interface, signalling data) and external sources, e.g. external databases.
- Capability of “learning” from past experience (e.g. past attack patterns, problems resolution).
- Provide communication interfaces and protocols as well as security and access control.
- Provide high availability (99.99999 %)
- Provide support on a 24 × 7 basis
- Provide high reliability (low false positive anomaly detection rate)
- Be scalable/expandable and interoperable (so as to serve large networks and be able to incorporate new network architectures/technologies)
- Be intrinsically secure (it should not be supposed to be vulnerable to attacks which can invalidate its results and usage altogether).
- Be cost efficient.

### 3 Existing Solutions

In this section we present the security solutions of some of the major key players in the market. We need to take into consideration that this is still a very immature market and most of the solutions are still under development and/or testing and hence not still consolidated. According to Gartner research and advisory firm, mobile network based security solutions will not mature before 2020 [6, 7]. Therefore, at the present time these solutions do not cover all of the MNO requirements. On the other hand, we are not aware of existing or under development third-party security services. We believe that in order for such services to become attractive to MNOs,



their providers must adopt a business model that will be able to satisfy all of the MNOs' requirements offering a solution that will be fully interoperable with all network entities and smoothly incorporated to any mobile network.

The need for extra security for mobile networks, since they are now becoming very similar to common IP-based networks, has also attracted a lot of research effort, with the NEMESYS project being part of this effort. NEMESYS develops an infrastructure to collect and analyze data about cyber-attacks targeting mobile devices and networks, track and detect abnormal behaviours and anomalies and provide early warning of attacks on mobile devices and networks by deploying mobile honeypots so as to be probed, attacked or compromised. Moreover, NEMESYS system includes advanced visual analytics tools to help the security analyst identify complex attack phenomena through hypothesis formulation and testing, attack attribution and correlation analysis.

### ***3.1 Alcatel-Lucent (ALU)***

Motive Security Guardian provides both a network-based infection-detection platform and a security analytics solution. It allows service providers to pinpoint and analyze infections in their subscribers' home networks and mobile devices, then take action to protect both the network and subscribers. Motive Security Guardian incorporates the following functionalities [8]:

*Network-Based Intrusion Detection System*, which detects malicious traffic originating from the subscriber home network or device using a specialized traffic-sensing and intrusion-detection software.

*Alert Reporting Cluster*, which is typically deployed in the service provider's data center and is responsible for processing and storing events from the sensors, notifying the subscribers about security alerts.

*Virtualization*, which allows the solution to be deployed on standard, off-the-shelf hardware in a virtual environment such as CloudStack or OpenStack.

*Mobile App and Web Portal* that provides more information about the threat and step-by-step instructions for removing it. A mobile application is also included to remediate infections on mobile devices (Fig. 2).

### ***3.2 Nokia Solutions and Networks (NSN)***

NSN's "Mobile Guard" security solution uses data related to voice, SMS and mobile broadband, and analyzes network traffic patterns to detect malware. Mobile Guard also notifies the user, blocks the affected services on the network and helps subscribers cleanse their smart devices [9–11]. It correlates suspicious network traffic patterns to known threats, building on trusted malware intelligence from its partner F-Secure.

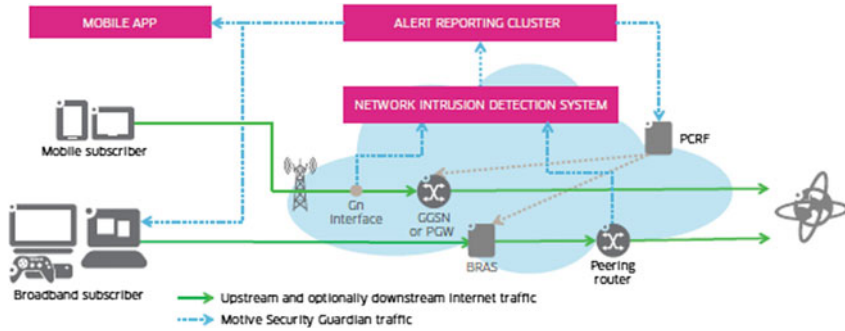


Fig. 2 ALU’s mobile ecosystem security solution [8]

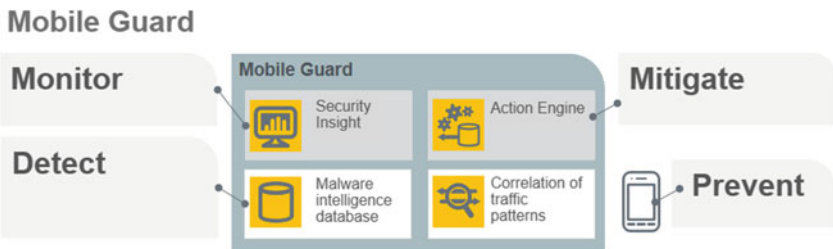


Fig. 3 NSN’s Mobile Guard principle [11]

NSN’s method is self-learning and detects new malware earlier than conventional signature recognition, enabling immediate action to improve protection for smart devices (Fig. 3).

Mobile Guard automatically prevents premium rate SMS messages from being sent, or blocks unauthorized mobile payments, etc. The solution also alerts the subscriber about the contamination, and can optionally offer software to disinfect the device and provide further protection to complement the network-based defense.

### 3.3 Ericsson

Up to now, Ericsson has been relatively quiet about mobile network security, unlike its rivals NSN and ALU. Information retrieved from [12] indicates that Ericsson aims at a holistic perspective for planning the security services and functionality to be implemented in nodes, sub-networks, and at the network level (Fig. 4).

Ericsson believes that a defense-in-depth strategy using perimeter protection is needed, putting different security measures/features at different places in the network, e.g. in each and every node. These features can be complemented with security functionality at site and network levels.

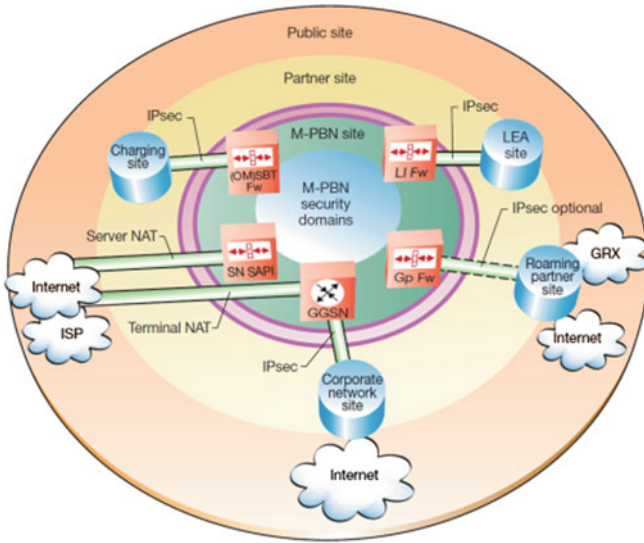


Fig. 4 Ericsson’s security related solution [12]

## 4 Business Model

A successful attack-affecting part or the whole mobile network- will have a negative impact on the MNO business and will cause undesirable degradation on customers’ experience. To make things worse, it is envisaged that the customers will blame the operator whom they are accustomed to trust for security issues; especially if they have purchased their mobile devices from the particular MNO.

The impact of such events may lead to increased churn rates, lower profit, resulting in a negative impact on the MNO’s brand name. Loss of profit may also occur in cases where the affected users refuse to pay high bills (e.g., in cases where a malware has initiated voice calls or sent SMSs to premium numbers). In the long term, if fraud via mobile malware gets out of control, it may lead to a trust loss which may slow down the growth of the overall mobile business.

To cope with the emerging security threats by investing in an advanced security solution is a matter of utmost importance for the MNOs who need a holistic, flexible, effective, proactive, defensive and affordable security strategy for the massive and undoubtedly complex mobile ecosystem. In case of an attack, MNOs should be prepared to respond immediately so as to secure both their networks and their subscribers, thus defending their reputation and ensuring the viability of mobile business.

The advantages for an MNO in incorporating an advanced security solution for both its network infrastructure and mobile devices will be multifold, such as:

- Rapidly identify traffic that is involved in malicious attacks and reduce time between infection and mitigation.
- Warn its customers early enough and prevent intrusions against their mobile devices, achieving wider market share and reducing churn.
- Differentiate from competition by protecting their customers from security risks, misuse and fraud while offering high service experience.
- Save huge effort and expenditures, in the coming years, that would have been required for security “repairs” and downtimes.

A security solution could also generate new and recurring revenues for an MNO by selling “security packages” or by bundling security to up-sell other services (m-banking, m-commerce or any other mobile transaction requiring advanced security). Although further investment would be required from the MNO side for dedicated customer care and support, however, pioneering MNOs in adopting such kind of solutions can expect to obtain a significant competitive advantage that will “reward” them with an additional list of benefits, such as:

- Churn reduction and market share increase, since their subscribers will be spared from bill shocks and offered peace of mind that their security is taken care of
- Generation of new revenues by addressing new markets, such as selling security packages, and security sensitive applications (m-banking/m-commerce/m-payments)
- Promotion of the company brand/reputation.

An indicative MNO business model for the incorporation/deployment of a security solution is depicted in Fig. 5, by making use of the Business Model Canvas [13], a strategic management and entrepreneurial tool for developing new or documenting existing business models.

The above business model assumes that the MNO will offer a new security service to its subscriber base and in particular to the users of smartphones who require enhanced security. Typically, this would reflect 50 % of the smartphone owners. The key partners of the MNO in this business model will be the mobile ecosystem security solution vendors, including infrastructure vendors (NSN, Ericsson, ALU, Huawei, ZTE, etc.) or third-party security service providers. The value proposition of the MNO is to provide its subscribers through various commercial channels (internet, communication media, MNO’s sales network) with Security as a Service (SecaaS) that enhances the security of their mobile devices (early detect malicious applications, send notifications/instructions to the mobile user, etc.); being available and supported on a 24 × 7 basis. At the same time, the MNO can even more strengthen the excellent relationship it has established with its subscribers.

The viability of the business model for the provision of the SecaaS solution will depend on:

- The initial investment (CAPEX/OPEX) required by the MNO for a complete end-to-end solution and the investment required for future upgrades. These costs could be reduced if the solution were adopted at a Group of Companies’ level and










 <p><b>KEY PARTNERS</b></p> <p>Mobile ecosystem security solution vendors including infrastructure vendors (NSN, Ericsson, ALU, Huawei, ZTE, etc.)</p>	 <p><b>KEY ACTIVITIES</b></p> <p>Early detection of malicious applications and notification to subscribers. Early detection of anomalies in the network, cause identification and mitigation.</p>	 <p><b>VALUE PROPOSITIONS</b></p> <p>Deliver to subscribers a new service that enhances the security of their mobile devices (Security as a Service). The MNO is satisfying the need of the subscribers for additional security that would result in minimizing mobile device infections, protecting them from "bill shocks" and ensuring a "peace of mind"</p>	 <p><b>CUSTOMER RELATIONSHIPS</b></p> <p>In general an MNO has established an excellent relationship with its subscribers. This relationship can be strengthened even more by providing 24 h per day 7 days per week support for new security related service.</p>	 <p><b>CUSTOMER SEGMENTS</b></p> <p>The subscriber base of the MNO and in particular the users of smartphones who require enhanced security. Typically this would reflect 50% of the smartphone owners.</p>
	 <p><b>KEY RESOURCES</b></p> <p>HW and SW required for the provisioning of the new service.</p>		 <p><b>CHANNELS</b></p> <p>Internet Television Radio MNO network of shops</p>	
 <p><b>COST STRUCTURE</b></p> <ul style="list-style-type: none"> <li>•HW and SW cost for the provisioning of the new service</li> <li>•Cost for supporting the service</li> </ul>		 <p><b>REVENUE STREAMS</b></p> <p>Monthly fee for the service</p>		

Fig. 5 MNO Business Model for a security solution

- The expected revenues coming from: monthly fees for the provisioning of the service, up-selling other services by bundling security, market share increase (brand name enhancement because of network/devices protection from future malware infections, avoidance of churn reduction), etc.

## 5 Conclusions

MNOs need to consider installing advanced security solutions to protect their subscribers' mobile devices as well as their networks from emerging threats, caused by malicious activity explicitly directed against the smart mobile ecosystem. Towards this end, MNOs need to investigate the mobile security market for major key players and their approaches. Such a search reveals that the available solutions are quite few, since this is still a very immature market with most of the solutions still under testing and hence not consolidated. It is expected that smart mobile ecosystem security solutions will not mature before 2020, hence MNOs need to set their requirements as early as possible, in order to be taken into consideration by the developers of the security solutions. These solutions will require significant investments from MNOs but they could also generate new and recurring revenues for them by selling "security packages" or by bundling security to up-sell other services. Therefore, MNOs should also start making their own business plans in order to protect these investments.

**Acknowledgments** The current study is part of the Project NEMESYS (Enhanced Network Security for Seamless Service Provisioning in the Smart Mobile Ecosystem) which has received funding from the European Union Seventh Framework Programme (FP7) under grant agreement 317888.

**Disclaimer** The views expressed in this article are those of the authors and do not necessarily represent the views of the company.

## References

1. Motive Security Labs (ALU) malware report—H2 (2014)
2. <http://sourcedna.com/blog/20150420/afnetworking-vulnerability.html>
3. NEMESYS Project, <http://www.nemesys-project.eu/nemesys/> available technical Deliverables D1.1, D2.1
4. Paolini, M.: Wireless Security in LTE Networks
5. <http://www.reuters.com/article/2012/01/27/us-docomo-idUSTRE80Q1YU20120127>
6. NEMESYS Project, <http://www.nemesys-project.eu/nemesys/> available technical Deliverable D8.3.2
7. <http://www.gartner.com/newsroom/id/2828722>
8. <http://www.alcatel-lucent.com/solutions/security-guardian>
9. <http://nsn.com/news-events/press-room/press-releases/new-network-based-security-solution-from-nsn-helps-protect-smart-device-users-against-malw>
10. [http://nsn.com/sites/default/files/document/nsn\\_mobile\\_guard\\_executive\\_summary.pdf](http://nsn.com/sites/default/files/document/nsn_mobile_guard_executive_summary.pdf)
11. <http://ru.networks.nokia.com/file/29471/nsn-mobile-guard>
12. [http://www.ericsson.com/ma/res/thecompany/docs/publications/ericsson\\_review/2004/2004125.pdf](http://www.ericsson.com/ma/res/thecompany/docs/publications/ericsson_review/2004/2004125.pdf)
13. <https://canvanizer.com/new/business-model-canvas>

# Undermining Isolation Through Covert Channels in the Fiasco.OC Microkernel

M. Peter, M. Petschick, J. Vetter, J. Nordholz, J. Danisevskis  
and J.-P. Seifert

**Abstract** System designers have come to recognize the merits of building critical systems on top of small kernels for their ability to provide strong isolation at system level. This is due to the fact that enforceable isolation is the prerequisite for any reasonable security policy. Towards this goal we examine some internals of Fiasco.OC, a microkernel of the prominent L4 family. Despite its recent success in certain high-security projects for governmental use, we prove that Fiasco.OC is not suited to ensure strict isolation between components meant to be separated. Unfortunately, in addition to the construction of system-wide denial of service attacks, our identified weaknesses of Fiasco.OC also allow covert channels across security perimeters with high bandwidth. We verified our results in a strong affirmative way through many practical experiments. Indeed, for all potential use cases of Fiasco.OC we implemented a full-fledged system on its respective archetypical hardware: Desktop server/workstation on AMD64 x86 CPU, Tablet on Intel Atom CPU, Smartphone on ARM Cortex A9 CPU. The measured peak channel capacities ranging from  $\sim 13,500$  bits/s (Cortex-A9 device) to  $\sim 30,500$  bits/s (desktop system) clearly falsify Fiasco.OC's isolation guarantee.

---

M. Peter · M. Petschick · J. Vetter · J. Nordholz · J. Danisevskis · J.-P. Seifert (✉)  
TU Berlin, 10587 Berlin, Germany  
e-mail: jean-pierre.seifert@telekom.de

M. Peter  
e-mail: peter@sec.t-labs.tu-berlin.de

M. Petschick  
e-mail: matthias@sec.t-labs.tu-berlin.de

J. Vetter  
e-mail: julian@sec.t-labs.tu-berlin.de

J. Nordholz  
e-mail: jnordholz@sec.t-labs.tu-berlin.de

J. Danisevskis  
e-mail: janis@sec.t-labs.tu-berlin.de

## 1 Introduction

Over the past years, electronic devices have found their way into virtually every aspect of our daily lives, changing the way we communicate, work, and spend our leisure time. Even the latest wave—the triumph of smartphones and tablets—could qualify as a technological revolution. On the downside, it is obvious that radical changes of that magnitude involve many security risks.

As if the threat from criminal circles was not bad enough, the situation has been further worsened by targeted attacks—also known as cyberwar. Long on financial, technical, and even social resources and expertise, the entities behind these attacks were long thought to not have difficulties overcoming standard security measures, cf. [4]. Recent reports on highly advanced attacks corroborate this assumption [1, 2].

To counter such attacks, efforts have been launched to improve the security of existing operating systems, with SELinux being a prominent example. Not only has that proven a lengthy process, but it has also raised questions regarding further increases in system complexity and limited backward compatibility with existing software. Consequently, completely rebuilding the system architecture from ground up offers a worthwhile alternative path. Building systems on small secure kernels has been long acknowledged both in industry, security agencies, cf. [7], and of course in academia [10, 11], yet adopting them was limited by insufficient resources and lacking performance of considered systems.

Owing to their architectural elegance and small size, microkernels are often automatically embraced as “secure”, yet this blind trust is often too optimistic. In this paper, we will show that Fiasco.OC, a popular microkernel of the L4-family, does not deliver on that very promise. This is particularly interesting as Fiasco.OC is used in multiple projects for its assumed ability to enforce strict isolation, among them secure laptops [6] and secure mobile devices [9, 16].

Following the classical research directions on covert channels, cf. [3], we identify so far unknown weaknesses in Fiasco.OC’s kernel memory subsystem. These substantially undermine all efforts to isolate subsystems in security critical systems built on top of Fiasco.OC. In the rest of the paper, we develop real-world covert channels based on those weaknesses found in Fiasco.OC’s kernel memory management.

## 2 Fiasco.OC Background

Fiasco.OC [5], a member of the venerable L4 family developed at the TU Dresden (Germany), has many attractive features. Besides its relatively small size—between 20 kSLOC and 35 kSLOC depending on the configuration—it supports the construction of secure systems by offering a security model based on capabilities.



## 2.1 Memory Management

Like the first L4 kernel, cf. [11], Fiasco.OC seeks to restrict its functionality to mechanisms, upon which policies can be implemented by user-level servers depending on the specific needs of the applications at hand. L4 pioneered this principle by moving the management of user-level memory completely out of the kernel. To that end, all L4-related kernel provide three mechanisms: first, page faults are exported to user-level, usually by having the kernel synthesize a message on behalf of the faulting thread. Second, a mechanism whereby the right to access page can be delegated (L4 terminology: *map*) between tasks and, finally, a mechanisms to revert that sharing (*unmap*). Since pages can be recursively mapped, the kernel needs to track this operation, otherwise the unmap might not completely revoke all derived mappings. The kernel data structure used for this purpose is usually called *mapdb*, an abbreviation for *mapping database*.

The situation is quite different for kernel memory, for which Fiasco.OC provides no direct management mechanism. When, in the course of a syscall, kernel memory is requested or released, the kernel turns to its internal allocators, each of which follows a fixed allocation strategy. Contrary to the microkernel philosophy, the user cannot implement its own. To prevent individual tasks from monopolizing the limited kernel memory, Fiasco.OC uses a quota mechanism. Each task is associated with a quota object, which represents the amount of kernel memory that is available for all activities in that task.<sup>1</sup>

Intuitively, properly set quotas should ensure that each subsystem can use the share of kernel memory allocated for it regardless of the activity of other subsystems. However, due to fragmentation, it may happen that the allocators cannot find a contiguous memory range large enough to accommodate the requested object. We will show that the combination of Fiasco.OC's design and implementation gives an attacker the opportunity to bring about fragmentation on purpose, which, in turn, allows him to tie down kernel memory far beyond what his quota should allow, effectively rendering the quota mechanism useless.

## 2.2 Fiasco.OC Implementation

At the lowest level, kernel memory is managed by a buddy allocator. Since the size of kernel objects varies markedly, ranging from 32 bytes to 16 kB, and is not in all cases a power of two, allocating them directly from the buddy allocator would cause fragmentation over time. Thus, many objects are not directly allocated from the buddy allocator but are managed through slab allocators, which are supplied with memory by the buddy allocator. Slab allocators, eighteen in total, accommodate only objects of the same type.

---

<sup>1</sup>Quota objects can be shared among multiple tasks.

Of the ten slab allocators for mapping trees (described below), eight are vulnerable to user-induced object placement. Of the remaining slab allocators, only three are of interest: *ipc\_gate*, *irq\_factory*. The other allocators hold objects that have associated objects, for which an attacker cannot easily influence the placement.

Fiasco.OC's implementation of the `mapdb`<sup>2</sup> seeks to minimize its size. Each physical frame that is used as user memory is tracked through compact representation of a tree—the mapping tree—in a depth-first pre-order encoding. Each mapping can be represented by two machine words holding a pointer to its task, the virtual address of the mapping and the depth in the mapping tree.<sup>3</sup> While this representation saves space compared to other implementations using pointer-linked data structures, it brings about an object that potentially grows and shrinks considerably, depending on the number of mappings of that page.

If the number of mappings exceeds the size identifier of the mapping tree, the kernel allocates a bigger tree, into which it moves the existing mappings along with the new one. Conversely, when a thread revokes mappings, the kernel invalidates tree entries. Shrinking the tree, which involves moving it into a smaller data structure, takes place when the number of active entries is smaller than a quarter of the tree's capacity.

The combination of object whose size can be easily manipulated from user-level prompting the allocation of ever larger containers (mapping trees) and the known allocation strategy within slabs allows it to mount an attack on Fiasco.OC's kernel memory subsystem.

### 2.3 Unintended Channels

While the quota accounts for objects created on behalf of an agent, it does not capture the unused space in the slabs. A malicious agent is capable of bringing about a situation where this empty space accounts for more than 50% by deliberately choosing the order in which objects are created and destroyed.

The amount of memory that can be tied down depends on four factors: the number of vulnerable slab caches, the number of objects held in their individual slabs, the order in which slabs are attacked, and, for objects with variable size such as mapping trees, on the minimal size required for the object to stay in a certain slab.

In addition to the allocation channels there is another implementation artifact that can be misused as a communication channel. Fiasco tracks user-level pages in tree-like data structures, so called mapping trees. Although the mapping tree structure is dynamically resizable, Fiasco imposes an upper bound, thus limiting the number of mappings that can exist of a physical page in the system. L4Re, the user-level OS framework running on top of Fiasco.OC, provides a number of services, among them

---

<sup>2</sup>mapping database.

<sup>3</sup>Since a page address is always aligned to the page size, the lower bits of the mapping address can be used for the depth in the mapping tree.

a runtime loader, which is not unlike the loader for binaries linked against dynamic libraries on Linux (ld.so). Experimentally, we found that 18 pages are shared this way among all subsystems that are started through the regular L4Re startup procedure.

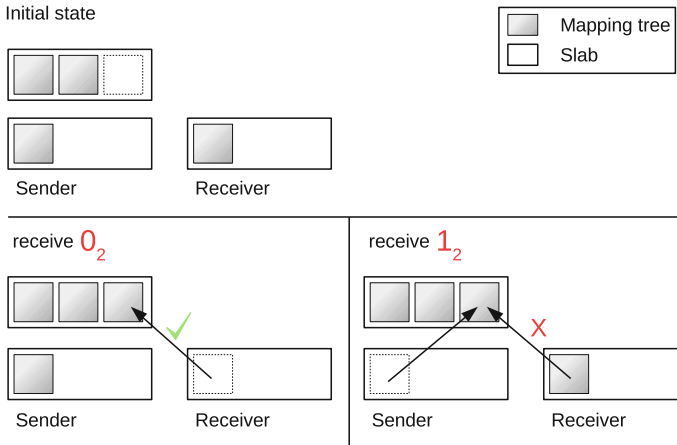
### 3 Channel Construction

In this section, we explain how to construct a high bandwidth covert channel. We also devise more sophisticated channels to work around problems impeding stability and transmission rates, which allow us to establish a reliable and fast covert channel even in difficult conditions. Two of the proposed channels rely on the fact that a malicious thread is able to use up more kernel memory than its quota allows. The remaining channel exploits a limit in a shared data structure.

The Page Table Channel (PTC) can be modulated in two ways: for sending a  $1_2$ , the sender allocates page tables from the kernel allocator, for transferring a  $0_2$ , it waits for one interval. To facilitate page table allocation, a helper task is created by the sender and pages are mapped into its address space. These pages are placed in such a way that each page requires the allocation of a new page table. The receiver can detect the amount of free memory in the kernel allocator by performing the same steps as the sender. The number of page tables available to the receiver is inversely proportional to the number of page tables held by the sender. This knowledge can be used to distinguish between the transmission of a  $1_2$  and a  $0_2$ . At the end of every interval, the sender has to release the page tables it holds. Unmapping the pages from the helper task is not sufficient because Fiasco.OC does not release page tables during the lifetime of a task. Instead, the helper task has to be destroyed and recreated.

The slab channel (SC) uses contention for object slots in slabs to transfer data. One slab is selected to hold mapping trees for the transmission. This slab is filled with mapping trees until only one empty slot remains, which is used for the actual transmission. Figure 1 shows the principle. To transfer a  $1_2$ , the sender needs to fill this last slot. Assuming a slab of size 4KB for the transmission, it causes a mapping tree residing in a slab of 2KB to grow, until the it exceeds its maximum size and the kernel moves it into a bigger slab—the one intended for transmission. The receiver can determine which bit was sent by performing the same operation as the sender. If this fails, the receiver interprets this as a  $1_2$  and as  $0_2$  otherwise.

The Mapping Tree Channel (MTC) exploits the fact that applications started by L4Re share 18 read-only pages. A size constraint imposed by the maximal size of the mapping tree—the data structure whereby an individual page is tracked—limits the number of times a page can be mapped to 2047. When this maximum is reached, any attempt to create new mappings of the page will fail. In the most basic form, the communication partners agree on one shared page, which they then use as a conduit. In the preparatory phase, the mapping tree of the chosen page will be filled such that only room for a single entry remains. Creating these mappings is possible because the shared pages are regular and are not subject to mapping restrictions. To transfer a  $1_2$ , the sender creates a mapping of the chosen shared page, maxing out



**Fig. 1** Transmission in the slab channel

its mapping count. The receiver, concurrently trying to create a mapping, fails as result. Conversely, a  $0_2$  is indicated by no mapping created by the sender so that the receiver’s operation succeeds. Compared to the other two channels, the MPC incurs low overhead. Unlike the SC, an MPC transmission requires at most three operations. In contrast to the PTC, no costly task destruction is necessary.

## 4 Evaluation

To evaluate the feasibility of our approach, we ran a number of a experiments on three different hardware platforms:

- PandaBoard [19], ARM Cortex, A9 1.0Ghz, 2 cores.
- MinnowBoard [18], Intel Atom E640, 1.0Ghz, 1 core, 2 hyperthreads.
- Desktop-grade PC, AMD Athlon X2, 2.4Ghz, 2 cores.

All three platforms are equipped with 1 GB of RAM. In our measurement setups, we used Fiasco.OC/L4Re version r54. Both sender and receiver agents were implemented as native L4Re applications, with the system setup such that they did not have direct communication channels between them. All transmission rates reported in this section reflect the number of bits transmitted per time interval, including checksum bits. The “channel states” row in our tables lists the number of states  $n$  a channel needs to support in order to transmit  $\log_2 n$  bits per interval.

Our first setup implements the page table channel, as introduced in Sect. 3 as our most basic transmission method. The surprising result that uni-processor setups yield higher bandwidths than multi-processor ones (Table 1) is down to Fiasco.OC using *read-copy-update* (RCU) as synchronization for object destruction, which lengthens

**Table 1** PTC results

Platform		UP		MP	
AMD	Channel states (#)	2	32	2	16
	Period (clock ticks)	2	3	5	5
	Capacity (bits/s)	500	1666	200	800
Minnowboard	Channel states (#)	2	4	2	16
	Period (clock ticks)	2	2	6	6
	Capacity (bits/s)	500	1000	167	667

**Table 2** SC results

Platform		UP	TSC + MP
AMD	Channel states (#)	2	2
	Period (clock ticks)	2	1
	Capacity (bits/s)	500	1000
Minnowboard	Channel states (#)	2	2
	Period (clock ticks)	2	1
	Capacity (bits/s)	500	1000
Pandaboard	Channel states (#)	2	–
	Period (clock ticks)	2	–
	Capacity (bits/s)	500	–

the time needed to destruct kernel objects. The PTC with its helper task, which needs to be destroyed for each transmission cycle, is affected. The slab channel operates in principle similar to the PTC but does not require a potentially costly object destruction. The transmission rate is solely limited by the amount of data that can be written to and read from the channel within a transmission interval, which depends on the time required to create and delete mappings (Table 2). The last column shows our final optimization of this transmission method, which increases the frequency by replacing the KIP clock with the high resolution TSC for synchronization purposes.

We will now take a look at how well the self-synchronizing transmission works. All tests used the MTC for both synchronization and data transfer.

The results (Table 3) show that for all three systems, the channel capacity grows with the number of channels. Transmitting a bit requires two or three operations, depending on the value to be transmitted. The synchronization overhead for two transmission steps is five operations. So, going from one to two data channels increases the number of operations between 44 % (6.5/4.5) and 54 % (8.5/5.5). The measured increased data capacities (54 % (Pandaboard, 5408/3511), 56 % (Minnowboard, 3840/2448), 48 % (AMD, 14738/9957)) are commensurate.

In contrast, the channel capacities scale much poorer with the number of (logical) cores used. We attribute this to resource conflicts in the memory subsystem (cache, memory bandwidth) as each operation has to scan through 16kB mapping tree. In the

**Table 3** Throughput depending on the number of transmission channels and number of cores used

Nr. of channels (#)	Throughput (bits/s)		
	Panda-board	Minnow-board	AMD
1	3511	2448	9957
1 + 1	5605	2414	11,940
2	5408	3840	14,738
2 + 2	8782	3815	18,592
4	7449	5368	19,974
4 + 4	12,166	5420	25,492
8	9207	6697	24,902
8 + 4	13,569	–	–
8 + 8	–	6881	30,582
16	10,457	7637	28,416

case of the Minnowboard, we do not see any scaling at all, which was to be expected as the used Atom E640 processor only features a single core with hyperthreading.

## 5 Related Work

Ristenpart et al. [15] showed that constructing a cross-VM side-channel is possible in Amazon’s EC2, prompting further research on virtual machine infrastructure. Wu et al. [21] use the EC2 as a host for covert communication. Exploiting peculiarities of the underlying x86 platform, they were able to achieve transmission rates of up to 100 bits/s. Xu et al. [23] also leverage cache crosstalk to construct a covert channel with channel capacities of up to 262 bits/s. Suzaki et al. [17] examined various virtual machines and discovered a side-channel in KVM based on the *memory deduplication* feature, which merges same-content memory pages, allowing for increased memory utilization. Based on these findings, Xiao et al. [22] constructed a covert channel with bandwidths of up to 1000 bits/s yet with a large memory footprint of  $\sim 400$  MB. Lipinski et al. [13] drew on work done by Ristenpart et al. [15] and improved the method of hard disk contention achieving a 1000 times higher steganographic bandwidth compared to the 0.1 bits/s by Ristenpart et al. Okamura et al. [14] examined the load-based covert channel on the Xen hypervisor. Changes in execution time due to sharing of a physical CPU led to an information leakage. The constructed covert channel had bandwidth of  $\sim 0.5$  bits/s. Lin et al. [12] propose two covert communication channels based on the last PID and a temporary file in the Linux OS. They construct three channel mechanisms to counter mitigation techniques. The authors achieve bandwidths of up to 40 bits/s. In comparison, we achieved significantly higher bandwidths of up to 5000 bits/s, on a system that emphasizes isolation.

## 6 Conclusions

It was shown that Fiasco.OC, a widely known and used third-generation microkernel, cannot deliver on isolation expectation in the face of a determined attacker. We have identified three shared facilities whose control mechanisms can either be rendered ineffective (kernel allocator, object slabs) or, worse, who lack them at all (mapping trees). In our experiments, we showed the feasibility of using them as a means of communication and achieved maximal channel capacities of up to  $\sim 30,000$  bits/s, outstripping previously found channels in VM environments by a fair margin.

The identified channels raise concerns about Fiasco.OC's role in devices with security requirements. For example, the covert channel rates achieved on the Pandaboard (up to  $\sim 13,500$  bits/s) lend plausibility to some disconcerting attack scenarios against phones that offer encrypted VoIP calls. Specifically, a Trojan horse may eavesdrop on the raw audio data of an encrypted VoIP connection as they are processed unencrypted in a protected compartment.<sup>4</sup> The bandwidth of the covert channels needed to exfiltrate the call data into an unprotected compartment is sufficient for voice calls if compressed with modern algorithms, which need  $\sim 4000$  bits/s.

Although an exhaustive list of design changes suitable to counter our attacks is beyond the scope of the paper, we would like to touch on the topic. As a communication step requires both conspirators to execute, it seems advisable to introduce a mechanism along the lines of work done by Wu et al. [20], whereby the switch rate between isolated domains can be controlled. We acknowledge that such a scheme may have a negative impact on throughput in scenarios where the activity pattern of subsystem is bursty.

The emergence of seL4 [8], the first general purpose kernel for which a formal correctness proof could be produced, brings within reach the prospect that systems can be constructed on error-free kernels.

Small kernels have the tendency to inspire users confidence. Yet, as demonstrated this trust might be misguided. Instead, we encourage designers of security-conscious systems to scrutinize the merits of the OS kernels they use on reasonable grounds, not promises or perceptions.

**Acknowledgments** This research was supported by the Helmholtz Research School on Security Technologies.

## References

1. Apt1: Exposing one of China's cyber espionage units (2013)
2. Apt28: A window into russia's cyber espionage operations? (2014)
3. Bishop, M.A.: Computer Security. Addison-Wesley Professional, Art and Science (2002)

---

<sup>4</sup>Unlike regular voice calls, VoIP calls are not handled by the baseband processor but by the application processor.

4. Clarke, R., Knake, R.: *Cyber War: The Next Threat to National Security and What to Do About It*. Ecco (2012)
5. Fiasco.oc website (2014). <http://os.inf.tu-dresden.de/fiasco/>
6. Genua cyber-top. <https://www.genua.de/loesungen/security-laptop-cyber-top.html>
7. Karger, P., Zurko, M., Bonin, D., Mason, A., Kahn, C.: A vmm security kernel for the vax architecture. In: *Proceedings of the IEEE Computer Society Symposium on Research in Security and Privacy*, 2–19 (1990)
8. Klein, G., Elphinstone, K., Heiser, G., Andronick, J., Cock, D., Derrin, P., Elkaduwe, D., Engelhardt, K., Kolanski, R., Norrish, M., Sewell, T., Tuch, H., Winwood, S.: *Sel4: formal verification of an os kernel*. In: *Proceedings of the ACM SIGOPS 22nd Symposium on Operating Systems Principles, SOSP '09*, 207–220, New York, USA (2009). (ACM)
9. L4re-based simko 3 security smartphone wins bsi approval for classified data (2013). <http://www.kernkonzept.com>
10. Liedtke, J.: Improving ipc by kernel design. *SIGOPS Oper. Syst. Rev.* **27**(5), 175–188 (1993)
11. Liedtke, J.: On micro-kernel construction. *SIGOPS Oper. Syst. Rev.* **29**(5), 237–250 (1995)
12. Lin, Y., Ding, L., Wu, J., Xie, Y., Wang, Y.: Robust and efficient covert channel communications in operating systems: design, implementation and evaluation. In: *Proceedings of the IEEE 7th International Conference on Software Security and Reliability-Companion (SERE-C)*, 45–52 (2013)
13. Lipinski, B., Mazurczyk, W., Szczypiorski, K.: Improving hard disk contention-based covert channel in cloud computing environment. *arXiv preprint* (2014). [arXiv:1402.0239](https://arxiv.org/abs/1402.0239)
14. Okamura, K., Oyama, Y.: Load-based covert channels between xen virtual machines. In: *Proceedings of the 2010 ACM Symposium on Applied Computing, SAC '10*, 173–180, New York, USA (2010). (ACM)
15. Ristenpart, T., Tromer, E., Shacham, H., Savage, S.: Hey, you, get off of my cloud: exploring information leakage in third-party compute clouds. In: *Proceedings of the 16th ACM Conference on Computer and Communications Security, CCS '09*, 199–212, New York, USA (2009). (ACM)
16. Simko–sichere mobile kommunikation (2014). <http://security.t-systems.de/loesungen/mobile-simko>
17. Suzuki, K., Iijima, K., Yagi, T., Artho, C.: Memory deduplication as a threat to the guest os. In: *Proceedings of the Fourth European Workshop on System Security, EUROSEC '11*, 1:1–1:6, New York, USA (2011). (ACM)
18. Technical features - Minnowboard (2014). <http://www.minnowboard.org/technical-features/>
19. Technical features–Pandaboard (2014). <http://pandaboard.org/content/platform>
20. Wu, J., Ding, L., Lin, Y., Min-Allah, N., Wang, Y.: Xenpump: a new method to mitigate timing channel in cloud computing. In: *Proceedings of the IEEE 5th International Conference on Cloud Computing (CLOUD)*, 678–685 (2012)
21. Wu, Z., Xu, Z., Wang, H.: Whispers in the hyper-space: high-speed covert channel attacks in the cloud. In: *Proceedings of the 21st USENIX Conference on Security Symposium, Security'12*, 9–9, Berkeley, USA (2012). (USENIX Association)
22. Xiao, J., Xu, Z., Huang, H., Wang, H.: Security implications of memory deduplication in a virtualized environment. In: *Proceedings of the 43rd Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, pp. 1–12 (2013)
23. Xu, Y., Bailey, M., Jahanian, F., Joshi, K., Hiltunen, M., Schlichting, R.: An exploration of 12 cache covert channels in virtualized environments. In: *Proceedings of the 3rd ACM Workshop on Cloud Computing Security Workshop, CCSW '11*, 29–40, New York, USA (2011). (ACM)



**Part IV**  
**Smart Algorithms**

# A Hybrid Movie Recommender Using Dynamic Fuzzy Clustering

Fatih Gurcan and Aysenur Akyuz Birturk

**Abstract** We propose an online hybrid recommender strategy named content-boosted collaborative filtering with dynamic fuzzy clustering (CBCF<sub>dfc</sub>) based on content boosted collaborative filtering algorithm which aims to improve the prediction accuracy and efficiency. CBCF<sub>dfc</sub> combines content-based and collaborative characteristics to solve problems like sparsity, new item and over-specialization. CBCF<sub>dfc</sub> uses fuzzy clustering to keep a certain level of prediction accuracy while decreasing online prediction time. We compare CBCF<sub>dfc</sub> with pure content-based filtering (PCB), pure collaborative filtering (PCF) and content-boosted collaborative filtering (CBCF) according to prediction accuracy metrics, and with online CBCF without clustering (CBCF<sub>onl</sub>) according to online recommendation time. Test results showed that CBCF<sub>dfc</sub> performs better than other approaches in most cases. We also evaluate the effect of user-specified parameters to the prediction accuracy and efficiency. According to test results, we determine optimal values for these parameters. In addition to experiments made on simulated data, we also perform a user study and evaluate opinions of users about recommended movies. The user evaluation results are satisfactory. As a result, the proposed system can be regarded as an accurate and efficient hybrid online movie recommender.

**Keywords** Recommender systems · Hybrid online movie recommendation

## 1 Introduction

Finding desirable information becomes more difficult as the size and diversity of information in the Internet increases. Recommender systems have emerged to overcome this difficulty. Today, they are used intensively in most of the domains

---

F. Gurcan · A.A. Birturk (✉)  
METU Department of Computer Engineering, Ankara, Turkey  
e-mail: birturk@metu.edu.tr

F. Gurcan  
e-mail: fgurcan@gmail.com

including book, movie, restaurant, news and music recommendation. Several well-known websites using recommender systems are Amazon, imdb, MovieLens and Yahoo! MUSIC. Burke defined a recommender system as “any system that produces individualized recommendations as output, or has the effect of guiding the user in a personalized way to interesting or useful objects in a large space of possible options” [6]. Recommender systems are information retrieval tools helping users in their information seeking tasks and guiding them in a large space of possible options. Many hybrid recommender systems are proposed so far to overcome shortcomings born of pure content-based (PCB) and pure collaborative filtering (PCF) systems. Most studies on recommender systems aim to improve the accuracy and efficiency of predictions. With collaborative filtering (CF) techniques, recommendation of an object is based on previously given ratings to that object by other users. Practically, CF techniques provide a basis for recommendation of objects, regardless of whether their content can be represented in a way that is useful for recommendation. In contrary to content based filtering (CBF), CF does not run into some problems such as limited content analysis, synonymy, over-specialization and concept drift. On the other hand, they have some challenges including new item, sparsity, scalability and gray sheep. These challenges are briefly explained below:

- **New Item:** New item, also known as “first rater”, problem is a weak side of CF techniques. With CF techniques, until a new object is rated by enough number of users, the object cannot be recommended [4]. This problem can be solved by using hybrid techniques [10].
- **New User:** If the recommender does not know the preferences of a user, it cannot recommend any objects to the user. Several techniques are used to overcome this challenge. Some systems use demographics of users while some other systems use more advanced techniques such as item popularity and item entropy.
- **Sparsity:** In most domains, number of previously given ratings per object is generally very low [9]. This fact causes the problem of sparsity. Sparsity is not related only with the overall proportion of ratings, the distribution of the given ratings also causes sparsity. Several techniques are proposed to overcome sparsity problem. These include using demographic data of users [11], Singular Value Decomposition (SVD) [11] and EM learning algorithms [12].
- **Scalability:** In big user and object spaces, making dynamic recommendations requires large computation times. To get rid of efficiency-concerned problems, various model-based techniques are proposed. Model-based techniques predict ratings of previously unrated objects according to a created model. Frequently used model-based techniques are Bayesian networks, Clustering Algorithms and Bayesian classifiers [2].
- **Over-specialization:** In some circumstances, users may rarely receive a recommendation differing from his previously experiences. This problem can be solved by adding some serendipity to the prediction algorithm. Genetic algorithms in the context of information filtering are proposed by some researchers to overcome over-specialization [13].

- **Concept Drift:** Concept drift is used for the long term changes in the attitudes and behaviors of users while they were interacting with the recommender system. Since the profile of a user is created cumulatively, the recommender system needs to adopt itself quickly to the changes in the user's behaviors. Most popular technique that partially overcomes this problem is giving more weight to the newer observations [8].
- **Gray Sheep:** CF systems are compelled by users having extraordinary behaviors that do not fit to any category. Claypool et al. call this type of users as gray sheep users [7]. CBF techniques are used to overcome this problem.
- **Synonymy:** In some certain domains, objects should not be recommended if they are too similar to previously recommended objects. For instance, in news recommendation domain, different news describing same event should not be recommended several times. To overcome this problem some systems, such as DailyLearner, filter out objects those are too similar to previously recommended objects [5].

Neighborhood-based CF algorithm is the most frequently used CF algorithm. Melville et al. [3] explains working mechanism of neighborhood-based CF in three steps:

1. By using a similarity measuring method weight all users with respect to similarity with the active user.
2. Select  $n$  users that have the highest similarity with the active user. These selected users form the neighborhood of the active user.
3. Compute a prediction from a weighted combination of the selected neighbors' ratings.

## 2 Motivation

Our recommendation algorithm is based on Content-Boosted Collaborative Filtering (CBCF) algorithm which is proposed by Melville et al. [3]. In our work, we extend and improve their proposed system by adding some new techniques and components. Melville et al.'s work has some limitations:

- Their Bayesian classifier was only using maximum likelihood decision rule which yields inaccurate predictions compared to average likelihood decision rule.
- Their proposed algorithm was working only offline and did not cover online recommendation.
- In their work, the effect of parametric values like number of selected neighbors in collaborative filtering or weight of content-based ratings to the accuracy and efficiency are not considered.
- They made their experiments only with Movielens datasets and they lack a user interface for the evaluation of user preferences.

The major contribution of this paper can be considered as the enhancement of content-boosted collaborative filtering (CBCF) algorithm proposed by Melville et al. We named our recommender as  $CBCF_{dfc}$  which stands for content-boosted collaborative filtering with dynamic fuzzy clustering. Motivations for the techniques used in our approach ( $CBCF_{dfc}$ ) are listed below:

- Why a hybrid approach? Both of CF and CBF have pluses and minuses. In our approach we created a hybrid recommender combining strong sides of CF and CBF. By employing CBF, we overcome new item, sparsity and gray sheep problems. Similarly, by employing CF we overcome limited content analysis, concept drift, over-specialization and synonymy problems. Another domain-specific motivation for using CBF is the text-based and utilizable on the dataset we work on.
- Why a model-based approach? Memory-based methods make accurate predictions, however they suffer from high online computation times in big databases [2]. To be able to make recommendations in reasonable time, we developed a model-based approach.
- Why fuzzy clustering? Clustering is a widely used technique in model-based method recommender systems. In our approach we use fuzzy clustering because, contrary to hard clustering, fuzzy clustering generates more proper clusters. Consequently, proper clusters of users supposed to yield more accurate predictions.

### 3 Our Method

Our system, basically, aims to predict ratings of previously unrated movies for a sparse user-movie matrix. We implemented pure content-based filtering (PCB), pure collaborative filtering (PCF) and content-boosted collaborative filtering (CBCF). All of these approaches work offline and incapable of online recommendation.

PCB's prediction process is based on movie content data and active user's profile. PCB uses naïve Bayesian classifier (NBC) to gather information from movie content data and active user's previously given ratings. We use two decision rules in NBC named as maximum-likelihood and average-likelihood. PCF bases its prediction process on previously given ratings to a movie by other users. Our PCF uses similarity-based neighborhood algorithm and it uses Pearson correlation coefficient as similarity measure. CBCF bases its prediction process on the output ratings of PCB (i.e. pseudo-ratings). CBCF performs same neighborhood-based collaborative filtering algorithm with PCF. After applying CF, differently from PCF, it combines predicted ratings of its CF part with pseudo-ratings. Clustering is applied to the final predictions of CBCF for scalability. As the clustering algorithm we use fuzzy c-means clustering. Each user in the user space has memberships to a number of clusters instead of a single cluster. Fuzzy c-means clustering algorithm is an iterative algorithm trying to converge an optimal grouping of users. In order to minimize offline cluster initialization time and uniformly distribute users, we propose a new initialization method.

In order to evaluate user opinions about recommended movies we implemented a graphical user interface. Our proposed interface asks users to rate at least 15 movies than according to these ratings it displays a recommendation list to users. Eventually, for each recommended movie, it asks the opinion of the user for an evaluation. The problem space we deal with can be formulated as a matrix of users versus movies. In this matrix, each cell represents a user's rating on a specific movie. Under this formulation, our research problem reduces to predicting values for empty cells of the original matrix and listing the ones having highest predicted ratings. Original matrix is supposed to be very sparse, because each user is supposed to rate a small percentage of movies. We use a single numeric rating for each movie with high values representing interest and low values representing disinterest. We employ explicit user ratings on a scale of 1–5. Each approach (i.e. PCB, PCF, CBCF, CBCF<sub>onl</sub> and CBCF<sub>dfc</sub>) tries to fill the matrix by its own prediction method. By actual matrix we mean the original matrix composed of actual ratings. By pseudo-matrix we mean the user-movie matrix which is the output of PCB. PCB analyzes movie descriptions to identify movies that are of particular interest to the user. Although using demographic data of users is possible, PCB only uses movie content data. Our CBCF approach is an effective recommendation framework that combines content based and collaborative techniques. CBCF uses a content-based predictor to enhance existing user data, and then provides personalized suggestions through collaborative filtering.

First, we perform CF on full user-movie matrix composed of pseudo-ratings. Then we combine these CF predictions with pseudo-ratings. We use two parameters to determine neighborhood of each user: Neighborhood size (NS) and similarity threshold (ST). NS is a user-specified positive integer and set to 30 by default. ST is a user-specified decimal number on a scale of 0–1 inclusive and set to 0.6 by default. According to our neighborhood selection algorithm, NS users with having similarities more than ST to the active user are chosen. In simple terms, NS and ST limit the size of neighborhoods. Although some systems develop different approaches, in our design each user has the same neighborhood for all movies. CBCF performs collaborative filtering on pseudo-ratings instead of actual ratings. The collaborative filtering algorithm we use is the same with PCF's. NS and ST are user-specified parameters for also CBCF.

Then, we combine predicted ratings of CF with pseudo-ratings. We use a parameter named content weight (CW) which determines weight of pseudo-ratings in final predictions. At this point, we introduce an additional parameter: self weighting threshold (SW). For each user, SW indicates minimum number of rated movies for CW to be able to exactly affect the combination. There is a need for a parameter like SW because the accuracy of content-based predictions are directly proportional to the number of movies rated by the user.

We apply fuzzy clustering to the final predictions of CBCF for scalability concerns. Our new model based strategy using dynamic fuzzy clustering is named CBCF<sub>dfc</sub>. Fuzzy c-means (FCM) is a method of clustering which allows one piece of data to belong to two or more clusters. General structure of fuzzy c-means algorithm (FCM) is given below:

- $N$  users are determined as cluster centers.
- Fuzzy partitioning is achieved through an iterative optimization of the objective function by calculating cluster centers and membership of each user to the clusters.
- This iteration stops when the difference between two iterations become smaller than a user-specified value.

Number of clusters ( $NC$ ), degree of fuzziness ( $m$ ), sensitivity threshold ( $\epsilon$ ) are user-specified parameters. FCM minimizes intra-cluster variance as well, but has a major problem, the results depend on the initial choice of cluster centers. To improve the performance of clustering we proposed a new method based on maximizing the distances between cluster centers. For the initialization of fuzzy clustering algorithm we developed two initialization methods: random initialization and max-distance initialization. In random initialization method, we randomly select  $NC$  users as cluster centers. In max-distance initialization method, we iteratively select cluster centers according to their distance to initial cluster centers. As distance measure, we use Euclidean distance. Number of initial clusters ( $NIC$ ) is also a user-specified integer equal to or less than  $NC$ . We also developed two methods for the selection of initial  $NIC$  clusters with max-distance method: selection of random users and selection of distinguishable users. Random selection is straightforward. To select distinguishable users, we select the users having maximum standard deviation values.

For an active user  $u_a$ , we first find fuzzy nearest cluster, which is the cluster  $P$  to which the membership  $t_{pa}$  is maximum. Then we take the users having maximum memberships to  $P$  and calculate similarity of these users to  $u_a$ . We developed a new method which uses multiple clusters in the calculation instead of using single cluster.  $NAC$  is a user-specified parameter denoting number of active clusters used in the calculation. Top- $NAC$  nearest clusters are selected according to the membership of the active user to  $NC$  clusters.

## 4 Evaluation and Results

Our hybrid approach is a unified model, which employs some user-specified parameters in order to get more accurate and efficient recommendations. We use well-known MovieLens rating data that was collected by GroupLens Research Project and the IMDB movie information for experiments (100,000 ratings for 1682 movies by 943 users). MovieLens dataset have following characteristics:

- Each user has rated at least 20 movies.
- Information about movies is compatible with Internet Movie Database (imdb).
- Simple demographics about users (e.g. gender, profession) are available.
- Simple content data about movies (e.g. release date, genre) is available.
- All ratings are integer values on a scale of 1–5.

Experiments are conducted with a subset of 7 training sets and corresponding test sets. Metrics used are Mean Absolute Error (MAE) and Receiver Operating Characteristic (ROC). The MAE metric is a kind of prediction accuracy metric that measures

the average absolute deviation between a predicted rating and an actual rating. The ROC area metric is a classification accuracy metric that can be represented equivalently by plotting the fraction of true positives (TPR = true positive rate) versus the fraction of false positives (FPR = false positive rate). The range of ROC sensitivity is between 0 and 1, where 0.5 is random and 1 is perfect. We use ROC-4 in our proposed system. The overviews of offline and online evaluations are given in the Figs. 1 and 2 below.

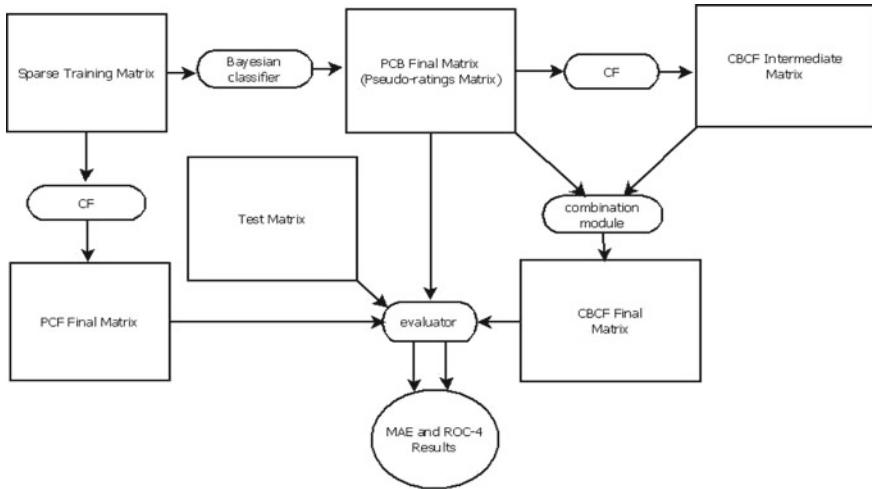


Fig. 1 Overview of offline evaluation

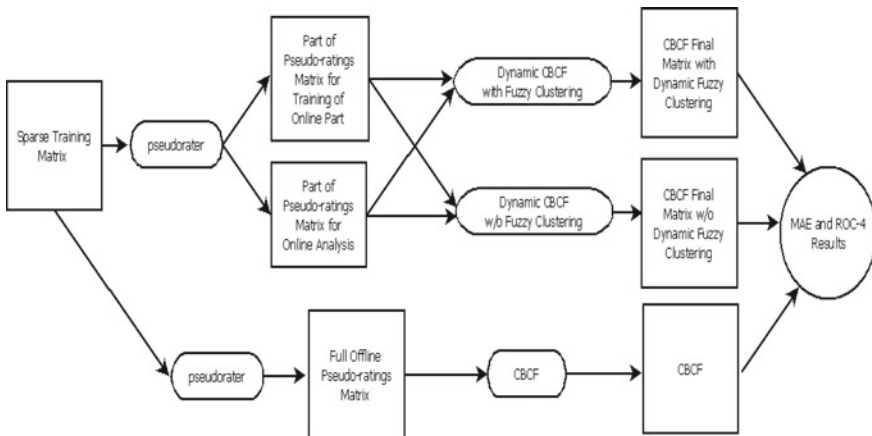


Fig. 2 Overview of online evaluation



Approach	MAE					ROC-4				
	u1	u2	u3	u4	u5	u1	u2	u3	u4	u5
<b>PCB</b>	0,95	0,927	0,919	0,918	0,93	0,6093	0,6073	0,602	0,6093	0,605
<b>PCF</b>	0,897	0,871	0,866	0,864	0,878	0,6083	0,6021	0,5996	0,5979	0,5949
<b>CBCF</b>	0,864	0,848	0,834	0,845	0,848	0,6138	0,6121	0,6044	0,6094	0,6055

Fig. 3 Comparison of CBCF with PCB and PCF

First, we enhance the naive Bayesian classifier (NBC) proposed in [1] in order to get a better classification of movies. In addition to maximum-likelihood that Melville et al. used, we use a different decision rule named average-likelihood. By means of test results, we see that NBC using average-likelihood outperforms NBC using maximum-likelihood on both accuracy metrics. We also use additive smoothing with a user-specified parameter (AV) to overcome anomalies caused by existence of zeros in the multiplication of the likelihood estimates. Test results indicate the optimal value of AV as a decimal number close to 1. We compare accuracy of CBCF using average-likelihood with pure content-based filtering (PCB) and pure collaborative filtering (PCF). According to test results, CBCF outperforms PCB and PCF significantly on both metrics (Fig. 3).

Secondly, in order to be able to make online recommendations efficiently, we use a fuzzy clustering model (i.e. fuzzy c-means clustering) that combines model and memory-based techniques. Although, fuzziness increases offline initialization time of clusters, it does not pose a problem because we are interested only in online recommendation time. We compare  $CBCF_{dfc}$  (CBCF with dynamic fuzzy clustering) with  $CBCF_{onl}$  (online CBCF without clustering) and CBCF (online CBCF) in terms of accuracy and efficiency metrics. Results are surprising; in addition to dramatic decline from  $O(n)$  to  $O(1)$  in online recommendation time,  $CBCF_{dfc}$  performs even better than  $CBCF_{onl}$  on accuracy metrics. To sum up, accuracy of  $CBCF_{dfc}$  is comparable with the accuracy of the  $CBCF_{onl}$  and CBCF, on the other side scalability of  $CBCF_{dfc}$  is considerably better than the scalability of  $CBCF_{onl}$  (Fig. 4).

Approach	MAE				ROC-4				online prediction time(s)			
	u4	u5	ua	ub	u4	u5	ua	ub	u4	u5	ua	ub
<b>CBCF</b>	0,846	0,849	0,852	0,858	0,6238	0,6183	0,6160	0,6190	-	-	-	-
<b><math>CBCF_{onl}</math></b>	0,822	0,823	0,829	0,847	0,6137	0,6130	0,6181	0,6228	8,10	8,19	8,10	8,17
<b><math>CBCF_{dfc}</math></b>	0,820	0,827	0,831	0,846	0,6183	0,6153	0,6197	0,6258	2,32	2,33	2,32	2,35

	MAE	ROC-4	time(s)
<b><math>CBCF_{dfc}</math> vs. CBCF</b>	2,40%	0,08%	-
<b><math>CBCF_{dfc}</math> vs. <math>CBCF_{onl}</math></b>	0,14%	0,47%	249%

Fig. 4 Comparison of  $CBCF_{dfc}$  with  $CBCF_{onl}$  and CBCF

Thirdly, we propose a new cluster initialization method in order to increase accuracy and efficiency. In addition to random selection of clusters ( $CBCF_{dfc;rand}$ ) we iteratively select cluster centers according to their distance to initial cluster centers ( $CBCF_{dfc;dist}$ ).  $CBCF_{dfc;dist}$ 's offline cluster construction time is more than  $CBCF_{dfc;dist}$ 's cluster construction time. However, since we are not interested with offline time, it does not pose a problem. In all experiments,  $CBCF_{dfc;dist}$  performed better than  $CBCF_{dfc;rand}$  in accuracy and online efficiency. Also, in addition to random selection of initial clusters ( $CBCF_{dfc;dist;rand}$ ), we select more distinguishable initial users by picking the users having maximum standard deviation values ( $CBCF_{dfc;dist;stdev}$ ). Test results were meaningful,  $CBCF_{dfc;dist;stdev}$  is marginally better than  $CBCF_{dfc;dist;stdev}$ .

Fourthly, we evaluate the effect of user-specified parameters those significantly effect the accuracy and efficiency of the recommendations. These parameters are adjustment value (AV) in NBC, neighborhood size (NS) and similarity threshold (ST) in collaborative filtering engine, weight of content-based ratings (CW) in combination part of CBCF and number of clusters (NAC) in clustering part. AV is used for additive smoothing in NBC and according to test results choosing AV a relatively sizeable value increases accuracy as it is explained above. NS and ST are used to limit the number of neighbors of a user. Test results show that, effects of these parameters to the accuracy are similar. As the number of neighbors increases we see that accuracy increases until some point then remains almost constant. On the other side, increase of NS negatively affects efficiency of  $CBCF_{dfc}$ . The reason is clear; online recommendation time of  $CBCF_{dfc}$  is linearly dependent with calculation time of neighborhood selection in CF. So for NS and ST there is a trade-off between accuracy and efficiency. As the fourth parameter, CW, determines weight of content-based and collaborative characteristics of CBCF. If CW is a relatively big number than CBCF gets closer to PCB, otherwise CBCF gets closer to PCF. By means of various experiments we see that there is an optimal interval for CW on both accuracy metrics, so it is reasonable to keep CW in this optimal interval. As the last parameter, we employ NAC which corresponds number of active clusters used for the computation of final predictions in  $CBCF_{dfc}$ . Increase of NAC marginally improves accuracy, however negatively effects efficiency. As a result keeping NAC small is reasonable.

Fifthly, we evaluate opinions of users about recommended movies (Fig. 5). For user evaluation we use feedbacks obtained by the user interface. Feedbacks are grouped under three categories; positive, negative and neutral. 68 % of users agreed with recommended movies to them and 12 % of users disagreed. Received feedbacks

	<b>mean</b>	<b>st.dev.</b>
<b>positive attitude</b>	68%	9%
<b>negative attitude</b>	12%	5%
<b>neutral attitude</b>	21%	7%

Fig. 5 Evaluation results of 35 users

are quite satisfactory in comparison with other studies in recommender systems literature. In conclusion, our system  $CBCF_{dfc}$  can be evaluated as an accurate and efficient online recommender for movie domain.

## 5 Conclusion and Future Work

We enhanced the naive Bayesian classifier (NBC) by using average likelihood decision rule and adjustment value in order to get a better classification of movies. Our NBC outperformed Melville et al.'s significantly on both accuracy metrics. We applied fuzzy clustering to  $CBCF$  for online predictions which sharply reduced online recommendation time. We employed a new technique in the calculation of final predictions with multiple clusters. A new cluster initialization method was proposed for fuzzy c-means clustering which increased classification accuracy and efficiency of online predictions, marginally. The user-specified parameters (AV, NS, ST, CW, NAC) significantly affect the accuracy and efficiency of the recommender. We implemented a user interface in order to evaluate opinions of users.

In  $CBCF_{dfc}$ , in addition to user-based similarity, item-based (i.e. movie-based) similarity can be used to improve the accuracy of the system. Item-based clustering applied with user-based clustering can decrease online recommendation time for each user.

## References

1. Fuzzy c-means clustering. [http://home.dei.polimi.it/matteucc/Clustering/tutorial\\_html/cmeans.html](http://home.dei.polimi.it/matteucc/Clustering/tutorial_html/cmeans.html) (2006)
2. Adomavicius, G., Tuzhilin, A.: Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *IEEE Trans. on Knowledge and Data. Engineering* **17**(6), 734–749 (2005)
3. Melville, P., Mooney, R.J., Nagarajan, R.: Content-boosted collaborative filtering for improved recommendations. In *Proceedings of the 18th National Conference on AI* (2002)
4. Avery, C., Zeckhauser, R.: Recommender systems for evaluating computer messages. *Commun. ACM* **40**(3), 88–89 (1997)
5. Billsus, D., Pazzani, M.: User modeling for adaptive news access. *User Model. User-Adap. Inter.* **10**(2–3), 147–180 (2000)
6. Burke, R.: Hybrid recommender systems: Survey and experiments. *User Model. User-Adap. Inter.* **12**(4), 331–370 (2002)
7. Claypool, M., Gokhale, A., Miranda, T., Murnikov, P., Netes, D., Sartin, M.: Combining content-based and collaborative Filters in an online newspaper. *Algorithms and Evaluation, In Workshop on Recommender Systems* (1999)
8. Koren, Y.: Collaborative Filtering with temporal dynamics. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining* (2009)
9. Linden, G., Smith, B., York, J.: Amazon.com recommendations: Item-to-item collaborative Filtering. In *IEEE Internet Computing* (2003)
10. Papagelis, M.: Crawling the algorithmic foundations of recommender technologies. Master's thesis, University of Crete (2005)

11. Pazzani, M.: A framework for collaborative, content-based and demographic Filtering. *AI Rev.* **13**(5–6), 393–408 (1999)
12. Popescul, R., Ungar, L.H., Pennock, D.M., Lawrence, S.: Probabilistic models for unified collaborative and content-based recommendation in sparse-data environments. In: *Proceedings of the 17th Conference on Uncertainty in Artificial Intelligence (2001)*
13. Sheth, B., Maes, P.: Evolving agents for personalized information Filtering. In: *Proceedings of the 9th IEEE Conference on Artificial Intelligence for Applications (1993)*

# Generating Minimum Height ADSs for Partially Specified Finite State Machines

Robert M. Hierons and Uraz Cengiz Türker

**Abstract** In earlier work, the problem of generating a preset distinguishing sequence from a finite state machine (FSM) was converted into a Boolean formulae to be fed into a SAT solver, with experiments suggesting that such approaches are required as the size of input alphabet grows. In this paper we extend the approach to the minimum height adaptive distinguishing sequence construction problem for partially specified FSMs (PSFMSs), which is known to be an NP- HARD problem. The results of experiments with randomly generated PSFMSs and case studies from the literature show that SAT solvers can perform better than a previously proposed brute-force algorithm.

**Keywords** Model based testing · Finite state machines · Adaptive distinguishing sequence

## 1 Introduction

The potential practical benefits of automatically deriving test sequences from a finite state machine (FSM) has led to the development of a number of approaches and their use in various fields such as sequential circuits [1], lexical analysis [2], software design [3], communication protocols [3–6], object-oriented systems [7], and web services [8]. The purpose of generating such test sequences is to decide whether an *implementation under test (IUT)* conforms to its specification  $M$ , where the IUT conforms to  $M$  if they have the same behaviour.

In testing one might use a *checking sequence (CS)*: an input sequence that is guaranteed to determine whether the IUT conforms to specification  $M$  as long as the IUT satisfies certain well-defined conditions. The typical assumption made is that

---

R.M. Hierons · U.C. Türker (✉)  
Department of Computer Science, Brunel University, Uxbridge, Middlesex, UK  
e-mail: uraz.turker@brunel.ac.uk

R.M. Hierons  
e-mail: rob.hierons@brunel.ac.uk

the IUT can be described by an FSM  $N$  with at most  $m$  states (some predefined  $m$ ). A CS  $\bar{x}$  is applied to the IUT and the output sequence produced is compared to that produced by  $M$  when executed with  $\bar{x}$ . If the output sequences are identical then the IUT is deemed to be correct; otherwise it is faulty.

Many CS construction techniques use distinguishing sequences (DSs): a class of input sequences that can be used to identify the state of the FSM  $M$ . Essentially, a CS leads to a different output sequence from any two states of  $M$ . As one can construct a relatively short CS when using a DS [9],<sup>1</sup> most CS generation approaches rely on the existence of such sequences. While other approaches such as *Unique Input Output (UIO) sequences* or *Characterising Sets (W-Set)* can be used to identify the current state of the IUT, these lead to longer CSs [10].

FSM specifications are often Partially Specified FSMs (PSFSMs): some state-input combinations do not have corresponding transitions [11, 12]. Much of the FSM based testing literature applies the *Completeness Assumption* [13, 14] that a PSFSM can be completed to form an FSM. This might be achieved, for example, by adding transitions with null output. However, there are situations in which one cannot complete a PSFSM and generate a CS from the resultant FSM [15]. For example, there being no transition from state  $s$  with input  $x$  might correspond to the case in which  $x$  should not be received in state  $s$  and testing should respect this restriction. This might be the case if the tests are to be applied by a context that cannot supply unspecified inputs [15]. It has been observed that it is possible to test the IUT via another PSFSM (tester PSFSM) that may never execute the missing transitions, which partially bypasses the need for the completeness assumption [16]. Nevertheless, in the FSM based testing literature we know of only one paper [17] in which the CS generation problem is addressed for PSFSMs. Although the previously proposed method [17] provided a polynomial time algorithm, the algorithm assumes that DSs are known in advance but does not report how one can derive DSs for a PSFSM.

There are two types of DSs: Preset Distinguishing Sequences (PDSs) and Adaptive Distinguishing Sequences (ADS). A PDS is an input sequence for which different states of  $M$  produce different output sequences. On the other hand, an ADS can be thought as a decision tree (ADSs are defined in Sect. 2). There are some benefits to using an ADSs rather than a PDS. Türker and Hierons show that checking the existence and computing a PDS from a PSFSM is a PSPACE-complete problem. They also showed that, for a given PSFSM  $M$  with  $n$  states and  $p$  inputs, the existence of an ADS can be decided in time of  $O(pn \log n)$  [18].

As the length of the checking sequence determines the duration/cost of testing, Türker and Yenigün [19] investigated corresponding optimisation problems. For completely specified FSMs they provided three notions of the “cost” of an ADS: (i) the height of the ADS (MINHEIGHTADS problem), (ii) the sum of the depths of all leaves in the ADS (*external path length*) (MINADS problem), and (iii) the sum, over the leaves, of the product of the depth and weight of the leaf (MINWEIGHTADS prob-

---

<sup>1</sup>While the upper bound on PDS length is exponential, test generation takes polynomial time if there is a known PDS.

lem). They showed that constructing a minimum ADS with respect to one of these metrics is NP- COMPLETE and NP- HARD to approximate.

As far as we are aware, no previous work has investigated the problem of generating a minimum height ADS from a PSFSM. On the other hand, Gunice et al. produced an algorithm that receives a completely specified FSM and an integer  $\ell$  and converts the PDS generation problem into a Boolean formula to be fed into a SAT solver [20]. This paper investigates the use of SAT solvers to generate a minimum height ADS from a PSFSMs. We encode the minimum height ADS generation problem as a Boolean formula and use a SAT solver to check the satisfiability of this formula. We also report on the results of initial experiments. The experiment subjects included randomly generated PSFSMs and PSFSMs drawn from a benchmark. The results suggest that in terms of scalability and the amount of time required to construct an ADS, SAT solvers are better than the brute-force approach.

The paper is structured as follows: Sect. 2 introduces terminology and notation used throughout the paper. Section 3 presents the SAT formulation and Sect. 4 presents the results of experiments. Finally, in Sect. 5 we conclude.

## 2 Preliminaries

A PSFSM  $M$  is defined by tuple  $(S, X, Y, \delta, \lambda, D)$  where  $S = \{s_1, s_2 \dots s_n\}$  is the finite set of states,  $X = \{a, b, \dots, p\}$  and  $Y = \{1, 2, \dots, q\}$  are finite sets of inputs and outputs,  $D \subseteq S \times X$  is the domain,  $\delta : D \rightarrow S$  is the transition function, and  $\lambda : D \rightarrow Y$  is the output function. If  $(s, x) \in D$  then  $x$  is *defined* at  $s$ . Given input sequence  $\bar{x} = x_1x_2 \dots x_k$  and  $s \in S$ ,  $\bar{x}$  is *defined at*  $s$  if there exist  $s_1, s_2, \dots, s_{k+1} \in S$  such that  $s = s_1$  and for all  $1 \leq i \leq k$ ,  $x_i$  is defined at  $s_i$  and  $\delta(s_i, x_i) = s_{i+1}$ .  $M$  is *completely specified* if  $D = S \times X$  and otherwise is *partially specified*. If  $(s, x) \in D$  and  $x$  is applied when  $M$  is in state  $s$ ,  $M$  moves to state  $s' = \delta(s, x)$  and produces output  $y = \lambda(s, x)$ . This defines *transition*  $\tau = (s, x/y, s')$  with *label*  $x/y$ , *start state*  $s$ , and *end state*  $s'$ .

We use juxtaposition to denote concatenation. The transition and output functions can be extended to input sequences as follows in which  $\varepsilon$  is the empty sequence,  $x \in X$ ,  $\bar{x} \in X^*$ , and  $x\bar{x}$  is defined at  $s$ :  $\delta(s, \varepsilon) = s$  and  $\delta(x\bar{x}) = \delta(\delta(s, x), \bar{x})$ ;  $\lambda(s, \varepsilon) = \varepsilon$  and  $\lambda(s, x\bar{x}) = \lambda(s, x)\lambda(\delta(s, x), \bar{x})$ . If there exists  $\bar{x} \in X^*$  defined in  $s$  and  $s'$  such that  $\lambda(s, \bar{x}) \neq \lambda(s', \bar{x})$ , then  $\bar{x}$  *distinguishes*  $s$  and  $s'$ . We now define *Preset DSs* and *Adaptive DSs*.

**Definition 1** Given PSFSM  $M$ ,  $\bar{x} \in X^*$  is a *Preset Distinguishing Sequence* for  $M$  if all distinct states of  $M$  are distinguished by  $\bar{x}$ .

**Definition 2** Let  $M$  be an FSM with  $n$  states. An *adaptive distinguishing sequence* is a rooted tree  $T$  with  $n$  leaves; the nodes are labeled with input symbols, the edges are labeled with output symbols and the leaves are labeled with distinct states such that: (1) output labels of edges emanating from a common node are different. (2) for

every leaf of  $T$ , if  $\alpha, \beta$  are the input-output sequences respectively formed by the node–edge labels on the path from the root node to the leaf and if the leaf is labeled by a single state  $s$ , then  $\lambda(s, \alpha) = \beta$ .

An ADS defines an experiment ending in a leaf. Applying ADS  $\mathcal{A}$  in  $s \in S$  leads to the input/output sequence that labels both a path from the root of  $\mathcal{A}$  to a leaf and a path of  $M$  with start state  $s$ . From the definition, the input/output sequences for distinct states differ and so  $\mathcal{A}$  distinguishes the states of  $M$ .

### 3 SAT Formulation for the Minimum Height ADS Problem

In this section we formulate the constraints of an ADS in the form of Boolean formulae to be fed into a SAT solver. The set of formulae generated depends on the PSFSM, and also on the height of the ADS that we would like to find. The question asked to the SAT solver is thus: is there an ADS of height  $\ell$ ? We will now explain how we convert this question into a Boolean formula. For each class of formula that we introduce, we provide the number of clauses generated.

We use  $\neg, \wedge, \vee$ , and  $\Rightarrow$  to denote negation, conjunction, disjunction, and implication, respectively. Additionally we used an operator called *exactly-one-OR* ( $\nabla(O_1, O_2, \dots, O_m)$ ) introduced in [20]. This operation evaluates to true if exactly one of its operands is true. Otherwise it evaluates to false. The formulae and the number of clauses to construct an ADS are given in Table 1.

The algorithm begins by forming the input clauses ( $\varphi_1$ ). These clauses use Boolean variables ( $X_{i,l,x}$ ) to let the SAT solver guess an input sequence for each state. The guessed input sequences are checked to see if they form an ADS. If the length we are

**Table 1** The proposed formulae and the number of clauses introduced

Formula	Number of clauses
$\varphi_1 = \bigwedge_{l \leq \ell, i \in S} (\nabla \{X_{i,l,x} \mid x \in X\})$	$\ell n p (p - 1) / 2 + 1$
$\varphi_2 = \bigwedge_{i \in S} \{S_{i,0,i}\}$	$n$
$\varphi_3 = \bigwedge_{l \leq \ell, i \in S} (\nabla \{S_{i,l,j} \mid j \in S\})$	$\ell n (n (n - 1) / 2 + 1)$
$\varphi_4 = \bigwedge_{x \in X} (\bigwedge_{i, j \in S} (X_{i,0,x} \Rightarrow X_{j,0,x}))$	$p n^2$
$\varphi_5 = \bigwedge_{i, j \in S, l < \ell, x \in X} ((S_{i,l,j} \wedge X_{i,l,x}) \Rightarrow S_{i,l+1,k})$	$\ell n^2 p$
$\varphi_6 = \bigwedge_{l \leq \ell, i \in S} (\nabla \{Y_{i,l,k} \mid k \in Y\})$	$\ell n (q (q - 1) / 2 + 1)$
$\varphi_7 = \bigwedge_{i, j \in S, l \leq \ell, x \in X} ((S_{i,l,j} \wedge X_{i,l,x}) \Rightarrow Y_{i,l,y})$	$\ell n^2 p q$
$\varphi_8 = \bigwedge_{i, j \in S, l < j, l \leq \ell, y \in Y} ((Y_{i,l,y} \wedge Y_{j,l,y}) \Rightarrow E_{i,j,l})$	$\ell n ((n - 1) / 2) q$
$\varphi_9 = \bigwedge_{i, j \in S, l < j, l \leq \ell, y, y' \in Y, y \neq y'} ((Y_{i,l,y} \wedge Y_{j,l,y'}) \Rightarrow \neg E_{i,j,l})$	$\ell n ((n - 1) / 2) q^2$
$\varphi_{10} = \bigwedge_{l < \ell} (\bigwedge_{i, j \in S, i < j} ((\bigwedge_{z \leq l} E_{i,j,z}) \Rightarrow \nabla (\{\chi_{i,j,l+1,x} \mid x \in X\})))$	$\ell^2 n^2 (p (p - 1) / 2 + 1)$
$\varphi_{11} = (\bigwedge_{i, j \in S, l < \ell, x \in X} (\chi_{i,j,l,x} \Rightarrow (X_{i,l,x} \wedge X_{j,l,x})))$	$\ell n^2 p$
$\varphi_{12} = \neg (\bigvee_{i, j \in S, i < j} (\bigwedge_{l \leq \ell} E_{i,j,l}))$	$\ell n (n - 1) / 2$



trying is  $\ell$ , for each  $l < \ell$ , for each state  $s_i \in S$  and for each input  $x \in X$ , we generate a Boolean variable  $X_{i,l,x}$ . Variable  $X_{i,l,x}$  should be true only if *at step*  $l$  (after we have applied  $l - 1$  inputs) the input  $x$  is used if we started from state  $s_i$ . Since we are considering deterministic PSFSMs, the query ensures that at each step only one input is applied. Note that a given exactly-one-OR operator with  $m$  operands introduces  $m(m - 1)/2 + 1$  clauses. Therefore  $\varphi_1$  introduces  $\ell n(p(p - 1)/2 + 1)$  new clauses.

The algorithm uses variable  $S_{i,l,j}$  being true to represent the first  $l$  inputs taking  $M$  from state  $s_i$  to state  $s_j$ . The formula  $\varphi_2$  ensures that when  $l = 0$  (the application of the ADS has not started),  $S_{i,l,j}$  is true if and only if  $i = j$ . As the PSFSM has  $n$  states,  $n$  clauses are introduced by  $\varphi_2$  (i.e. for  $l = 0$ ).

As the PSFSM is deterministic, we have ensured that when we start applying an ADS from a state  $s$ , there is a unique ‘current state’ at level  $l$ . This restriction is enforced by formula  $\varphi_3$ . As we use the exactly-one-OR operator  $\ell$  times,  $\varphi_3$  introduces  $\ell(n^2(n - 1)/2 + 1)$  clauses. Besides, note that from each initial state, the same input should be applied; formula  $\varphi_4$  enforces this. Since there are  $p$  inputs and  $n$  states,  $\varphi_4$  introduces  $pn^2$  clauses.

In order to represent the ending states of transitions of the PSFSM, the algorithm uses another set of clauses ( $\varphi_5$ ). Let  $t = (s_j, x/y, s_k)$  be a transition of  $M$ . No matter which state we started from, at a step  $l \leq \ell$ , if the current state is  $s_j$  and the input guessed for step  $l$  is  $x$ , then for step  $l + 1$  the current state should be  $s_k$ . The formula  $\varphi_5$  enforces this. Note that for each pair of states, for each input symbols and for each step we introduce a clause. Therefore the number of clauses introduced in  $\varphi_5$  is  $\ell n^2 p$ .

In the next step we trace outputs, using variables  $Y_{i,l,y}$ , constrained to be true if we observe output  $y$  at step  $l$  when we start at  $s_i$ . We have three concerns similar to those we had for states. First, we have to ensure that, for each starting state and each step, the guessed input produces exactly one output. Second, we have to ensure that this is the correct output. The first concern is handled using formula  $\varphi_6$ . Note that we need to construct  $\ell n$  exactly-one-OR operator for output symbols therefore  $\varphi_6$  introduces  $\ell n q(q - 1)/2 + 1$  clauses. For the second and third concerns, we use the transition information of  $M$ . Again let  $t = (s_j, x/y, s_k)$  be a transition of  $M$ . No matter which state we started from, if the current state is  $s_j$  at step  $l \leq \ell$  and the input guessed at step  $l$  is  $x$ , then the output produced in step  $l$  must be  $y$ . This constraint is enforced by  $\varphi_7$ . As we need to introduce a clause for each pair of states and for each input and output and for each step, we need to introduce  $\ell n^2 pq$  clauses in the formula  $\varphi_7$ .

In order to check that states are distinguished, the algorithm uses variables  $\{E_{i,j,l} | i, j \in S, i < j, l \leq \ell\}$ .  $E_{i,j,l}$  is true if, starting from states  $s_i, s_j$ , we observe the same output at level  $l$ . In order to achieve this we use output variables. If variables  $Y_{i,l,y}$  and  $Y_{j,l,y}$  are true then  $E_{i,j,l}$  should be true. This is achieved by the formula  $\varphi_8$ . Note that in  $\varphi_8$  we consider states with indexes  $i < j$  therefore the number of clauses introduced is  $\ell n(n - 1)/2$ . Thus the number of clauses introduced by  $\varphi_8$  is  $\ell n(n - 1)/2q$ . Conversely, we have to consider the case when output symbols are not identical: in such cases  $E_{i,j,l}$  should be false. Formula  $\varphi_9$  achieves this.

Since for every pair of outputs and steps, we need to compare outputs,  $\varphi_9$  introduces  $\ell n(n-1)/2q^2$  clauses.

Note that for a given ADS, the sequences retrieved from the ADS for two different states  $s_i$  and  $s_j$  share a common prefix  $\bar{x}$  if  $\lambda(s_i, \bar{x}) = \lambda(s_j, \bar{x})$ . We introduce a Boolean variable  $\chi_{i,j,l,x}$  to preserve this condition. Variable  $\chi_{i,j,l,x}$  is set to true if at level  $l$ , states  $s_i$  and  $s_j$  have not been distinguished and we then apply input  $x$ . This condition is enforced using formula  $\varphi_{10}$ . If outputs observed from a pair of states until step  $l$  are the same, then the same input should be applied at step  $l+1$ . Note that there are  $\ell^2 n^2$  exactly-one-OR introduced. Thus the number of new clauses introduced by formula  $\varphi_{10}$  is  $\ell^2 n^2 (p(p-1)/2 + 1)$ .

After the input to be applied for a pair of states is selected, we need to set the corresponding variables to true. We use formula  $\varphi_{11}$  to set these variables to true. In  $\varphi_{11}$  we use  $\chi_{i,j,l,x}$  as follows:  $\chi_{i,j,l,x}$  implies  $X_{i,l,x}$  and  $X_{j,l,x}$ . Since there are  $p$  inputs  $\ell$  steps and  $n$  states  $\varphi_{11}$  introduces  $n^2 \ell p$  number of clauses.

The algorithm ends by checking if all states produce unique outputs in response to the guessed input sequences. This is achieved by formula  $\varphi_{12}$  which introduces  $n(n-1)/2$  clauses. The overall formula for checking the existence of an ADS of length  $\ell$  is

$$\phi_1 \wedge \phi_2 \wedge \phi_3 \wedge \phi_4 \wedge \phi_5 \wedge \phi_6 \wedge \phi_7 \wedge \phi_8 \wedge \phi_9 \wedge \phi_{10} \wedge \phi_{11} \wedge \phi_{12}$$

Note that the number of clauses introduced by each formula can be given by a polynomial function (see Table 1). Therefore the proposed algorithm can construct the formulae using polynomial time and space.

## 4 Experimental Evaluation

We conducted a set of experiments to compare the performance of the brute-force algorithm (called BF below) as given in [21] and the SAT based approach (called SAT below) outlined in this paper.

All PSFSMs were constructed as described in [18]. We constructed four sets of FSMs with 14 inputs and 2 outputs and with 10, 20, 30 and 40 states. Moreover, in order to explore how the performance varied with respect to varying input sizes, we also construct three additional sets of PSFSMs with 30 states, 2 outputs, and input alphabets of size 64, 130 and 260. We used 100 PSFSMs in each of these seven sets and so a total of 700 PSFSMs. All machines were strongly connected, minimal, and had an ADS. In the experiments we set a time limit of 300s: if an approach did not derive an ADS in 300s, it terminated.

The experiments were carried out using MiniSat 2.2.0 on a machine with a 3.30GHz Intel Core I5-4590 and 8GB RAM running Windows 7 Enterprise. For each PSFSM, we measured the time taken by BF and by SAT. Table 2 summarises the results of our experimental study.

**Table 2** Computation times used for randomly generated FSMs

Algorithm	$n = 10, p = 14$	$n = 20, p = 14$	$n = 30, p = 14$	$n = 40, p = 14$
BF (s)	3.390	12.409	154.442	--
SAT (s)	0.801	1.309	1.920	1.341
Algorithm	$n = 30, p = 14$	$n = 30, p = 66$	$n = 30, p = 130$	$n = 30, p = 260$
BF (s)	154.442	296.298	--	--
SAT (s)	1.920	8.100	19.101	56.214

**Table 3** Computation times required for constructing ADSs for Benchmark FSMs

Name	$ S $	$ X $	SAT (s)	BF (s)
ex1	20	$2^9$	79.240	--
ex4	14	$2^6$	7.904	--
ex6	8	$2^5$	5.427	261.372
opus	10	$2^5$	4.137	287.634

The results are as expected. The BF algorithm gets slower and slower as the number of states and inputs increases. Although we can observe the same trend in the SAT-based approach, it requires much less time to construct an ADS.

Since randomly generated PSFMSs need not be representative of real PSFMSs, we used PSFMSs drawn from an ACM/SIGDA benchmarks [22] and repeated the experiment on these. Table 3 presents the size of the PSFMSs and the time required to compute ADSs.

The results from the benchmarks are similar to the results with randomly generated PSFMSs. These results have two implications. First, the SAT method outperformed the brute-force method especially as the size of the input alphabet grows. Note that in Tables 2 and 3, there are cases (indicated --) where we could not present computation time values. In two of the cases this is because none of the experiments returned an ADS within the 300s. However, for  $n = 30$  and  $p = 130$ , the BF approach did construct an ADS for one of the 100 FSMs (in 287.45 s). Although the SAT based approach scales well when compared to the brute-force approach, clearly, we should investigate heuristics for deriving minimum height ADSs from PSFMSs.

## 5 Conclusion

In this paper we addressed the problem of deriving minimum height adaptive distinguishing sequences (ADSs) from partially specified deterministic finite state machines (PSFMSs).

We proposed an algorithm that converts the bounded ADS generation problem for PSFMSs into a Boolean formula to be fed into a SAT solver. We carried out a set of

experiments. The results of experiments suggest that the time required to construct short ADSs is lower when the proposed SAT based ADS construction method is used.

As the class of completely specified FSMs is a subset of the class of PSFSMs, the proposed approach can also be used for constructing minimum height ADSs for completely specified FSMs. For completely specified FSMs we might compare the proposed approach with a previously defined exponential algorithm [23]. Moreover, it would be interesting to investigate SAT based approaches and heuristics of other problems related to ADSs such as the MinWeight and MinADS problems introduced in Sect. 1.

**Acknowledgments** This research is sported by The Scientific and Technological Research Council of Turkey under the grant reference no B.14.2.TBT.0.06.01-219-115543.

## References

1. Friedman, A.D., Menon, P.R.: Fault Detection in Digital Circuits. Computer Applications in Electrical Engineering Series. Prentice-Hall (1971)
2. Aho, A.V., Sethi, R., Ullman, J.D.: Compilers: Principles, Techniques, and Tools. Addison-Wesley (1986)
3. Chow, T.S.: Testing software design modeled by finite-state machines. *IEEE Trans. Softw. Eng.* **SE-4**(3), 178–187 (1978)
4. Holzmann, G.J.: Design and Validation of Computer Protocols. Prentice Hall (1991)
5. Lee, D., Yannakakis, M.: Principles and methods of testing finite-state machines—a survey. *Proc. IEEE* **84**(8), 1089–1123 (1996)
6. Sabnani, K., Dabhura, A.: A protocol test generation procedure. *Comput. Netw.* **15**(4), 285–297 (1988)
7. Binder, R.V.: Testing Object-Oriented Systems: Models, Patterns, and Tools. Addison-Wesley (1999)
8. Haydar, M., Petrenko, A., Sahraoui, H.: Formal verification of web applications modeled by communicating automata. In: Formal Techniques for Networked and Distributed Systems—FORTE 2004, vol. 3235, pp. 115–132. Springer (2004)
9. Lee, D., Yannakakis, M.: Testing finite-state machines: State identification and verification. *IEEE Trans. Comput.* **43**(3), 306–320 (1994)
10. Ural, Hasan: Formal methods for test sequence generation. *Comput. Commun.* **15**(5), 311–325 (1992)
11. Tsai, P.-C., Wang, S.-J., Chang, F.-M.: FSM-based programmable memory BIST with macro command. In: Memory Technology, Design, and Testing, 2005. MTDT 2005. IEEE International Workshop on, pp. 72–77 (2005)
12. Zarrineh, K., Upadhyaya, S.J.: Programmable memory BIST and a new synthesis framework. In: Fault-Tolerant Computing, 1999. Digest of Papers. Twenty-Ninth Annual International Symposium on, pp. 352–355 (1999)
13. Yannakakis, M., Lee, D.: Testing finite state machines: Fault detection. *J. Comput. Syst. Sci.* **50**(2), 209–227 (1995)
14. Yevtushenko, N., Petrenko, A.: Synthesis of test experiments in some classes of automata. *Autom. Control Comput. Sci.* **4** (1990)
15. Petrenko, Alexandre, Yevtushenko, Nina: Testing from partial deterministic FSM specifications. *IEEE Trans. Comput.* **54**(9), 1154–1165 (2005)

16. Rho, J.-K., Hachtel, G., Somenzi, F.: Don't care sequences and the optimization of interacting finite state machines. In: Computer-Aided Design, 1991. ICCAD-91. Digest of Technical Papers. IEEE International Conference on, pp. 418–421 (1991)
17. da Silva Simão, A., Petrenko, A.: Generating checking sequences for partial reduced finite state machines. *Test. Softw. Commun. Syst.* 153–168 (2008)
18. Hierons, R.M., Türker, U.C.: Distinguishing sequences for partially specified FSMs. In: NASA Formal Methods—6th International Symposium, NFM 2014, Houston, TX, USA, April 29 - May 1, 2014. Proceedings, pp. 62–76 (2014)
19. Türker, U.C., Yenigün, H.: Hardness and inapproximability of minimizing adaptive distinguishing sequences. *Formal Methods Syst. Des.* **44**(3), 264–294 (2014)
20. Güniçen, C., Türker, U.C., Ural, H., Yenigün, H.: Generating preset distinguishing sequences using SAT. In: Computer and Information Sciences II—26th International Symposium on Computer and Information Sciences, London, UK, 26–28 September 2011, pp. 487–493 (2011)
21. Gill, A.: Introduction to The Theory of Finite State Machines. McGraw-Hill, New York (1962)
22. Brglez, F.: ACM/SIGMOD benchmark dataset. <http://cbl.ncsu.edu:16080/benchmarks/Benchmarks-upto-1996.html>
23. Türker, U.C., Ünliyurt, T., Yenigün, H.: Lookahead-Based Approaches for Minimizing Adaptive Distinguishing Sequences. In: Testing Software and Systems—26th IFIP WG 6.1 International Conference, ICTSS 2014, Madrid, Spain, September 23–25, 2014. Proceedings, pp. 32–47 (2014)

# Line-Search Aided Non-negative Least-Square Learning for Random Neural Network

Yonghua Yin

**Abstract** Recently, Timotheou has formulated the learning problem of the random neural network (RNN) into a convex non-negative least-square problem that can be solved to optimality. By incorporating this work of problem formulation and the line-search technique, this paper designs a line-search aided non-negative least-square (LNNLS) learning algorithm for the RNN, which is able to find a nearly optimal solution efficiently. (The source code is available at [www.yonghuayin.icoc.cc](http://www.yonghuayin.icoc.cc).) Numerical experiments based on datasets with different dimensions have been conducted to demonstrate the efficacy of the LNNLS learning algorithm.

## 1 Introduction

In [1], Gelenbe developed a gradient-descent learning algorithm for the random neural network (RNN), and the RNN with the algorithm achieves acceptable performance in various applications [2–5]. However, one drawback is that the searching process in the algorithm may become trapped into local minimals, especially when handling high-dimensional problems. In addition, the efficacy of the algorithm is strongly related to initial values of parameters. To make an improvement, Timotheou has recently formulated the learning problem of the RNN into a convex non-negative least-square problem [2, 3] that can be solved to optimality. It also means that there is no local minimal in the searching process regardless of the number of dimensions. On the other hand, the line-search technique has been demonstrated to be efficient for solving unconstrained minimization problems [6–8]. For example, in [6, 7], the line-search technique has been exploited to select the step size in the iterative procedure such that a fast-convergence performance has been achieved.

In this paper, we incorporate the idea of the non-negative least-square problem formulation in [2, 3] and the line-search technique in [7] and then design an efficient learning algorithm for the RNN, which is termed the line-search aided non-negative

---

Y. Yin (✉)

Intelligent Systems and Networks Group, Department of Electrical and Electronic Engineering, Imperial College, London SW7 2BT, UK  
y.yin14@imperial.ac.uk

© Springer International Publishing Switzerland 2016  
O.H. Abdelrahman et al. (eds.), *Information Sciences and Systems 2015*,  
Lecture Notes in Electrical Engineering 363,  
DOI 10.1007/978-3-319-22635-4\_16

181

least-square (LNNLS) learning algorithm. This LNNLS learning algorithm is not affected by local minimal or sensitive to initial values of parameters. In addition, it selects the appropriate learning step size dynamically during the searching process (or say, iterative procedure). These features allow it to find a nearly optimal solution efficiently for the learning problem of the RNN.

The organization of this paper is as follows. Section 2 presents briefly the mathematical model of the RNN and describes the non-negative least-square formulation of the learning problem. In Sect. 3, the detailed procedure of the LNNLS learning algorithm is provided. Numerical results are presented in Sect. 4 to substantiate the efficacy of the LNNLS learning algorithm. Section 5 concludes this paper and points out the possible future directions.

## 2 RNN Model and Problem Formulation

This section presents briefly the mathematical model of the RNN and then describes the non-negative least-square formulation of the learning problem.

### 2.1 Mathematical Model

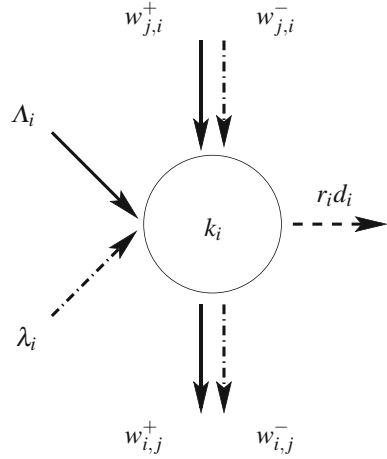
Let us consider a RNN with  $N$  neurons, where  $k_i(t) \geq 0$  denotes the potential (or say, state) of the  $i$ th neuron at time  $t$  with  $i = 1, \dots, N$  and correspondingly vector  $k(t) = [k_1(t) \ k_2(t) \ \dots \ k_N(t)]$  denotes the state of the whole RNN. All neurons are fully connected and exchange positive (or say, excitatory) and negative (or say, inhibitory) signals in the form of unit amplitude spikes (by firing). Each positive signal increases the potential of the receiving neuron by 1, while each negative signal reduces the potential that is larger than zero by 1. The  $i$ th neuron is excited when  $k_i(t) > 0$  or idle when  $k_i(t) = 0$ . When excited, the  $i$ th neuron fires according to an independent exponential distribution with rate  $r_i$ . In addition, its potential is reduced by 1 each time it fires. The fired spike heads for the  $j$ th neuron as a positive signal with probability  $p_{i,j}^+$  or as a negative signal with probability  $p_{i,j}^-$ , where  $j = 1, \dots, N$ , or it departs from the network with probability  $d_i$ . Correspondingly,

$$\sum_{j=1}^N (p_{i,j}^+ + p_{i,j}^-) + d_i = 1.$$

All neurons also receive signals from the outside world. Positive and negative signals arrive at the  $i$ th neuron from the outside world according to Poisson processes of rates  $\Lambda_i$  and  $\lambda_i$ , respectively.

Let  $q_i = \lim_{t \rightarrow \infty} \text{Prob}(k_i(t) > 0)$  denote the stationary probability of the  $i$ th neuron being excited. Based on [1], the signal flows in the RNN can be described by the following system of nonlinear equations:

**Fig. 1** Steady-state representation of a neuron in the RNN



$$q_i = \begin{cases} \frac{\lambda_i^+}{r_i + \lambda_i^-}, & \text{if } \lambda_i^+ < r_i + \lambda_i^-, \\ 1, & \text{if } \lambda_i^+ \geq r_i + \lambda_i^-, \end{cases} \quad (1)$$

where  $\lambda_i^+ = \Delta_i + \sum_{j=1}^N q_j w_{j,i}^+$ ,  $\lambda_i^- = \lambda_i + \sum_{j=1}^N q_j w_{j,i}^-$ ,  $w_{j,i}^+ = r_j p_{j,i}^+$ ,  $w_{j,i}^- = r_j p_{j,i}^-$  and  $i = 1, \dots, N$ . For better understanding, Fig. 1 shows the steady-state representation of a neuron in the RNN, where  $k_i = \lim_{t \rightarrow \infty} k_i(t)$ .

### 2.2 Non-negative Least-Square Formulation

Let us consider a dataset  $\{(\underline{x}_c, \underline{y}_c) | c = 1, \dots, C\}$  with its name defined as *Example*, where  $\underline{x}_c = [\underline{\Delta}_c, \underline{\lambda}_c]$  and  $\underline{y}_c = [y_{1,c} \dots y_{N,c}]$ . In addition,  $\underline{\Delta}_c = [\Delta_{1,c} \dots \Delta_{N,c}]$  and  $\underline{\lambda}_c = [\lambda_{1,c} \dots \lambda_{N,c}]$ . The objective for the RNN to learn the *Example* dataset is to adjust the weights  $\{w_{j,i}^+, w_{j,i}^- | i, j = 1, \dots, N\}$  such that, when input  $\underline{x}_c$  is given, each element in the RNN output vector  $\underline{q}_c = [q_{1,c} \dots q_{N,c}]$  becomes as close as possible to the corresponding element in the desired output vector  $\underline{y}_c$ , where  $c = 1, \dots, C$ .

Based on [2, 3], by assuming  $\lambda_i^+ < r_i + \lambda_i^-$  and  $d_i < 1$ , the above learning objective can be formulated into the following non-negative least-square problem:

$$\begin{aligned} \text{Minimize } f(\underline{w}) &= \text{Minimize } \frac{1}{2} \|A\underline{w} - b\|_2^2, \\ \text{Subject to } \underline{w} &\geq 0, \end{aligned} \quad (2)$$

where  $A \in R^{NC \times 2N^2}$ ,  $b \in R^{NC \times 1}$  and  $\underline{w} \in R^{2N^2 \times 1}$ . Vector  $\underline{w}$  consists of the weights  $w_{i,j}^+$  and  $w_{i,j}^-$  and is given by



$$\begin{aligned}\underline{w}(h_{i,j}^+) &= w_{i,j}^+, \\ \underline{w}(h_{i,j}^-) &= w_{i,j}^-, \end{aligned}$$

where  $h_{i,j}^+ = N^2 + (i-1)N + j$  and  $h_{i,j}^- = (i-1)N + j$ . In addition, matrix  $A$  and vector  $b$  are given by

$$\begin{aligned} A(h_{i,c}, h_{i,j}^+) &= \frac{y_{i,c}}{1-d_i}, \quad \forall j \neq i, \\ A(h_{i,c}, h_{i,j}^-) &= \frac{y_{i,c}}{1-d_i}, \quad \forall j \neq i, \\ A(h_{i,c}, h_{i,j}^+) &= \frac{y_{i,c}}{1-d_i} - y_{j,c}, \quad \forall j = i, \\ A(h_{i,c}, h_{i,j}^-) &= \frac{y_{i,c}}{1-d_i} + y_{i,c}y_{j,c}, \quad \forall j = i, \\ A(h_{i,c}, h_{j,i}^+) &= -y_{j,c}, \quad \forall j \neq i, \\ A(h_{i,c}, h_{j,i}^-) &= y_{i,c}y_{j,c}, \quad \forall j \neq i, \\ A(h_{i,c}, \text{otherwise}) &= 0, \\ b(h_{i,c}) &= \Lambda_{i,c} - y_{i,c}\lambda_{i,c}, \end{aligned}$$

where  $h_{i,c} = (c-1)N + i$  and  $c = 1, \dots, C$ .

### 3 LNNLS Learning Algorithm

The solution of a standard least-square problem can be obtained directly by the pseudoinverse method [6]. However, due to the non-negative constraint of  $\underline{w}$  in problem (2), the solution needs to be searched via an iterative procedure. Based on [2], in the  $m$ th iteration, where  $m = 1, \dots, M$  with  $M$  being the maximum number of iterations, the following weight-update formula can be exploited:

$$\underline{w}^{(m+1)} = P[\underline{w}^{(m)} - \eta \nabla f(\underline{w}^{(m)})], \quad (3)$$

where

$$\nabla f(\underline{w}^{(m)}) = A^T(A\underline{w}^{(m)}) - A^T b, \quad (4)$$

$\eta > 0$  is the step size and

$$P[\underline{w}(h)] = \begin{cases} \underline{w}(h), & \text{if } \underline{w}(h) > 0, \\ 0, & \text{otherwise,} \end{cases}$$

with  $h = 1, \dots, 2N^2$ .

The iterative procedure by using (3) may show a high dependency on the value of step size  $\eta$ , where a too large step size may cause oscillation and a too small one may cause slow convergence. In this paper, we extend the line-search technique that is originally designed for solving unconstrained minimization problems [6–8] and apply it to selecting  $\eta$  dynamically and appropriately. Based on Eq. (3) and the line-search technique, the LNNLS learning algorithm for the RNN can finally be designed and its detailed procedure is provided in Algorithm 1, where

$$f(\underline{w}) = \frac{1}{2} \|A\underline{w} - b\|_2^2 \quad (5)$$

is used to measure the learning performance. Note that the influence of the way of initializing  $\underline{w}^{(1)}$  on the efficacy of the LNNLS learning algorithm can be negligible in most cases if  $M$  is sufficiently large. However, initializing  $\underline{w}^{(1)}$  with the zero value may be an acceptable choice, which is further illustrated in Sect. 4.

---

**Algorithm 1** The LNNLS learning algorithm
 

---

```

Construct  $A$  and  $b$ ;
initialize  $\underline{w}^{(1)}$  and calculate  $f(\underline{w}^{(1)})$  via (5);
 $\eta_{\text{best}} \leftarrow \eta$ ,  $\underline{w}_{\text{best}} \leftarrow \underline{w}^{(1)}$ ,  $f_{\text{best}} \leftarrow f(\underline{w}^{(1)})$ ;
 $\eta \leftarrow 1.5$ ;  $\vartheta \leftarrow 2$ ;  $m \leftarrow 1$ ;
while  $m \leq M$  do
  calculate  $\nabla f(\underline{w}^{(m)})$  via (4);
   $\underline{w}^{(m+1)} \leftarrow P[\underline{w}^{(m)} - \eta \nabla f(\underline{w}^{(m)})]$ ;
  calculate  $f(\underline{w}^{(m+1)})$  via (5);
  while  $f(\underline{w}^{(m+1)}) < f_{\text{best}}$  do
     $\eta_{\text{best}} \leftarrow \eta$ ;
     $\underline{w}_{\text{best}} \leftarrow \underline{w}^{(m+1)}$ ;
     $f_{\text{best}} \leftarrow f(\underline{w}^{(m+1)})$ ;
     $\eta \leftarrow \eta \vartheta$ ;
    calculate  $\nabla f(\underline{w}^{(m)})$  via (4);
     $\underline{w}^{(m+1)} \leftarrow P[\underline{w}^{(m)} - \eta \nabla f(\underline{w}^{(m)})]$ ;
    calculate  $f(\underline{w}^{(m+1)})$  via (5);
  end while
  while  $f(\underline{w}^{(m+1)}) > f_{\text{best}} \& \eta > 10^{-20}$  do
     $\eta \leftarrow \eta / \vartheta$ ;
    calculate  $\nabla f(\underline{w}^{(m)})$  via (4);
     $\underline{w}^{(m+1)} \leftarrow P[\underline{w}^{(m)} - \eta \nabla f(\underline{w}^{(m)})]$ ;
    calculate  $f(\underline{w}^{(m+1)})$  via (5);
    if  $f(\underline{w}^{(m+1)}) < f_{\text{best}}$  then
       $\eta_{\text{best}} \leftarrow \eta$ ;
       $\underline{w}_{\text{best}} \leftarrow \underline{w}^{(m+1)}$ ;
       $f_{\text{best}} \leftarrow f(\underline{w}^{(m+1)})$ ;
    end if
  end while
   $\eta \leftarrow \eta_{\text{best}}$ ;
   $\underline{w}^{(m+1)} \leftarrow \underline{w}_{\text{best}}$ ;
   $f(\underline{w}^{(m+1)}) \leftarrow f_{\text{best}}$ ;
   $m \leftarrow m + 1$ ;
end while

```

---

*Remark 1* The LNNLS learning algorithm (i.e., Algorithm 1) is designed for the RNN with only output neurons. But, by using similar procedure of the weight-initialization algorithm in [2], the LNNLS learning algorithm can be easily extended to the case where the RNN is composed of both output and non-output neurons. Suppose the RNN has  $N_{\text{out}}$  output and  $N_{\text{out}}^-$  non-output neurons. The weights and desired outputs of non-output neurons  $y_{i_{\text{out}}^-,c}$ ,  $\forall i_{\text{out}}^- \in N_{\text{out}}^-, c \in C$  in the LNNLS learning algorithm with weight initialization are initialized for  $L$  times, where the detailed procedure is given in Algorithm 2.

---

**Algorithm 2** The LNNLS learning algorithm with weight initialization

---

Initialize  $\underline{w}^{(1)}$  and  $y_{i_{\text{out}}^-,c}$ ,  $\forall i_{\text{out}}^- \in N_{\text{out}}^-, c \in C$ ;  
 $l \leftarrow 1$ ;  
**while**  $l \leq L$  **do**  
    obtain  $\underline{w}_{\text{best}}$  via Algorithm 1 with  $\underline{w}^{(1)}$  and  $y_{i_{\text{out}}^-,c}$ ,  $\forall i_{\text{out}}^- \in N_{\text{out}}^-, c \in C$ ;  
    solve (1) using  $\underline{w}_{\text{best}}$  for  $q_{i_{\text{out}}^-,c}$ ,  $\forall i_{\text{out}}^- \in N_{\text{out}}^-, c \in C$ ;  
     $\underline{w}^{(1)} \leftarrow \underline{w}_{\text{best}}$ ;  
     $y_{i_{\text{out}}^-,c} \leftarrow q_{i_{\text{out}}^-,c}$ ,  $\forall i_{\text{out}}^- \in N_{\text{out}}^-, c \in C$ ;  
**end while**

---

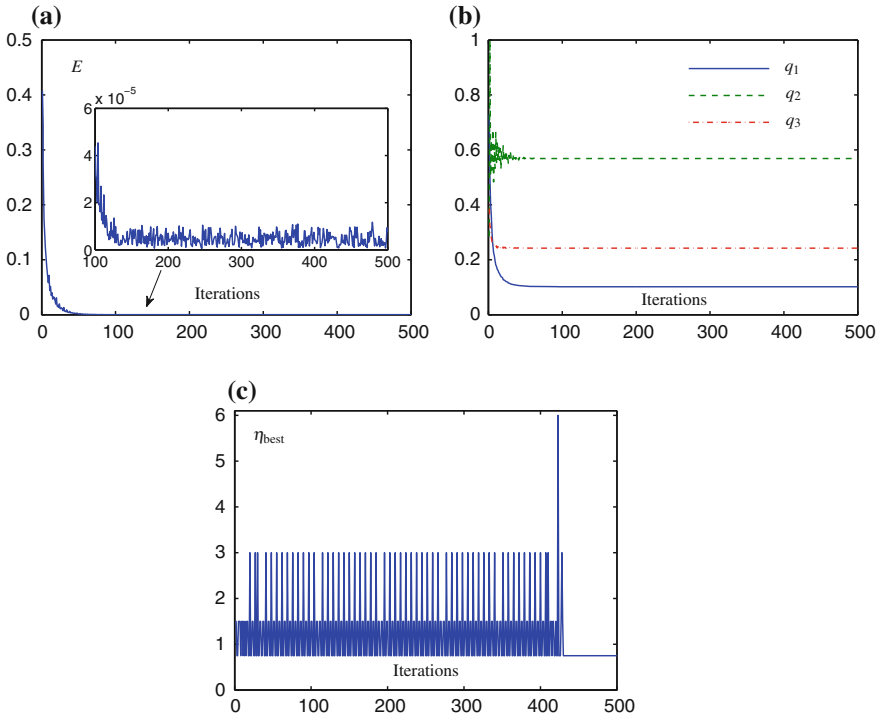
## 4 Numerical Verifications

In this section, numerical experiments are conducted to demonstrate the efficacy of the LNNLS learning algorithm for the RNN. (The source code of Algorithm 1 in MATLAB language is available at [www.yonghuayin.icoc.cc](http://www.yonghuayin.icoc.cc).) All of the numerical experiments are conducted in a MATLAB R2014a environment, which is operated on a personal computer (CPU: Intel i7-4770 3.40 GHz; memory: 8.00 GB). Note that, for simplicity,  $d_i$  with  $i = 1, \dots, N$  are set as zero. In addition, the root mean square error (RMSE) defined as

$$E = \sqrt{\frac{1}{NC} \sum_{i=1}^N \sum_{c=1}^C (y_{i,c} - q_{i,c})^2} \quad (6)$$

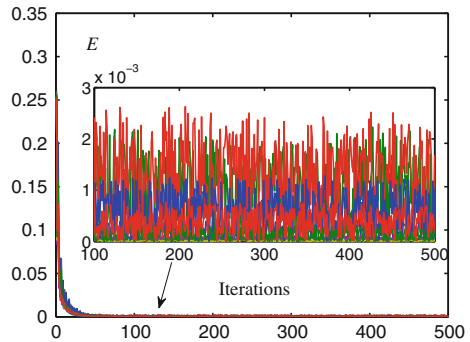
is used to measure the performances of the RNN and learning algorithms.

First, let us consider a simple dataset named as *Simple* with  $N = 3$  and  $C = 1$ . In addition,  $\underline{\Delta}_{1,1} = [0.5934 \ 0.5501 \ 0.9935]$ ,  $\underline{\lambda}_{1,1} = [0.1814 \ 0.1896 \ 0.8415]$  and  $\underline{y}_1 = [0.1018 \ 0.5684 \ 0.2422]$ , the values of which are randomly generated among range  $[0, 1]$ . Then, the LNNLS learning algorithm is exploited for the RNN to learn the *Simple* dataset. The numerical results are shown in Figs. 2 and 3. In the numerical experiment of Fig. 2, initial values of weights (i.e.,  $\underline{w}^{(1)}$ ) are set as zero. From Fig. 2a, we can see that the RMSE decreases rapidly to tiny values during the iterative process.

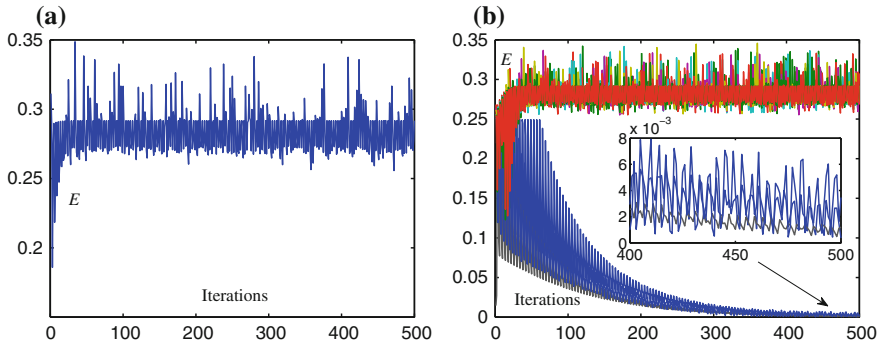


**Fig. 2** Learning performance of the RNN equipped with the LNNLS algorithm for the *Simple* dataset under “zero” initial weights. **a** RMSE versus number of iterations, **b** RNN outputs versus number of iterations, **c** step size versus number of iterations

**Fig. 3** Learning performance of the RNN equipped with the LNNLS algorithm for the *Simple* dataset under random initial weights

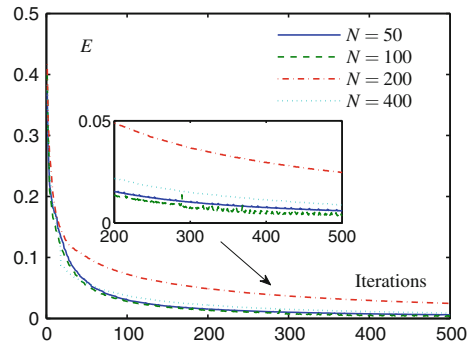


To be more specific, the RMSE decreases to  $4.3980 \times 10^{-5}$  at the 100th iteration and reaches the smallest value of  $5.5626 \times 10^{-7}$  at 246th iteration. Figure 2b illustrates that each RNN output converges rapidly to the corresponding desired output. In the numerical experiment of Fig. 3, initial values of weights are randomly selected and 10 trials are conducted. It can be seen from the results in Fig. 3 that, under different



**Fig. 4** For comparison: learning performance of the RNN equipped with the projected gradient algorithm in [3, 9] for the *Simple* dataset under “zero” and random initial weights. **a** “Zero” initial weights, **b** random initial weights

**Fig. 5** Learning performance of the RNN equipped with the LNNLS algorithm for datasets with  $N = 50, 100, 200, 400$



initial weights, all the RMSEs decrease rapidly to tiny values and become smaller than  $3.0 \times 10^{-3}$  after 100 iterations, demonstrating a stable learning performance of the RNN equipped with the LNNLS learning algorithm.

For comparison, the projected gradient algorithm in [3, 9] is also exploited for the RNN to learn the *Simple* dataset. The results are given in Fig. 4, where we can see that this algorithm is highly affected by the values of initial weights. Comparing results in Figs. 2, 3 and 4, we can conclude that the LNNLS learning algorithm proposed in this paper outperforms the one in [3, 9] for the RNN learning.

Then, let us consider datasets with higher dimensions. The numerical experiments in Fig. 5 are conducted based on datasets with  $N = 50, 100, 200, 400$ , where the values of inputs and desired outputs are randomly generated among range  $[0, 1]$ . In addition,  $C = 1$ . All results (including Figs. 2 through 5) demonstrate that, by exploiting the LNNLS learning algorithm, the RMSEs converge towards zero in general as the number of iterations increases.

## 5 Conclusion

By incorporating the idea of non-negative least-square formulation and the line-search technique, this paper has designed an effective leaning algorithm, i.e., the LNNLS learning algorithm, for the RNN. Numerical results on learning high-dimensional datasets have demonstrated the efficacy of the RNN equipped with the LNNLS learning algorithm.

In the further work, we will further optimize the LNNLS learning algorithm, such as reducing its computational complexity and lowering its requirement for computer memory. In addition, we will investigate the feasibility of this learning algorithm for solving the energy-distribution problem in the energy packet network [10, 11] that is a network model based on the theory of the RNN or more general the G-networks [12].

**Acknowledgments** This work is funded by an Imperial College Ph.D. Scholarship.

## References

1. Gelenbe, E.: Learning in the recurrent random neural network. *Neural Comput.* **5**(1), 154–164 (1993)
2. Timotheou, S.: A novel weight initialization method for the random neural network. *Neurocomputing* **73**(1), 160–168 (2009)
3. Timotheou, S.: Nonnegative least squares learning for the random neural network. In: *Artificial Neural Networks-ICANN*, pp. 195–204 (2005)
4. Gelenbe, E., Hussain, K.F.: Learning in the multiple class random neural network. *IEEE Trans. Neural Netw.* **13**(6), 1257–1267 (2002)
5. Gelenbe, E., Sungur, M.: Random network learning and image compression. In: *IEEE International Conference on Neural Networks*, pp. 3996–3999 (1994)
6. Zhang, Y., Yin, Y., Guo, D., Yu, X., Xiao, L.: Cross-validation based weights and structure determination of Chebyshev-polynomial neural networks for pattern classification. *Pattern Recogn.* **47**(10), 3414–3428 (2014)
7. Zhang, Y., Yang, Y., Cai, B., Guo, D.: Zhang neural network and its application to Newton iteration for matrix square root estimation. *Neural Comput. Appl.* **21**(3), 453–460 (2012)
8. Diniz-Ehrhardt, M.A., Martinez, J.M., Raydn, M.: A derivative-free nonmonotone line-search technique for unconstrained optimization. *J. Comput. Appl. math.* **219**(2), 383–397 (2008)
9. Lin, C.J.: Projected gradient methods for nonnegative matrix factorization. *Neural comput.* **19**(10), 2756–2779 (2007)
10. Gelenbe, E.: Energy packet networks: smart electricity storage to meet surges in demand. In: *5th International ICST Conference on Simulation Tools and Techniques*, pp. 1–7 (2012)
11. Gelenbe, E.: Energy packet networks: adaptive energy management for the cloud. In: *2nd International Workshop on Cloud Computing Platforms*, pp. 1 (2012)
12. Gelenbe, E., Labeled, A.: G-networks with multiple classes of signals and positive customers. *Eur. J. Oper. Res.* **108**(2), 293–305 (1998)

# A Novel Concise Specification and Efficient F-Logic Based Matching of Semantic Web Services in Flora-2

Shahin Mehdipour Atae and Zeki Bayram

**Abstract** We propose a novel concise specification of semantic web services conforming to the WSMO standard using the Flora-2 language, as well as a precise logical definition of what it means for a goal to match a web service. Our innovative usage of Flora-2 allows very short but expressive descriptions of both goals and web service capabilities, which are then used by a matching engine to discover which web services can satisfy a given goal. The matching engine, using the meta-level F-logic inferencing capabilities of the underlying Flora-2 reasoner, is very efficient and has a very concise definition itself.

## 1 Introduction

In service-oriented architecture (SOA), web services are defined, registered, invoked, and interconnected via some pre-agreed specifications [1]. Web service discovery is the process of finding one or more appropriate web service(s) among a possibly large pool of diverse web services. The non-intelligent way is to manually refer to web service repositories and use traditional text retrieval techniques to find some candidates for the specified application. On the other hand it can be done (semi) automatically by applying certain AI techniques such as logical inference. The latter approach requires the availability of rich semantic description of web service capabilities, user requirements and other related aspects of web services (such as non-functional properties), commonly in the form of logical statements in some appropriate form of logic, and in this case discovery amounts to proving certain logical inferences.

The Web Service Modeling Ontology (WSMO) [2] is a meta-ontology for describing relevant aspects of semantic web services that facilitates the logical description of web services, user requests in the form of goals, common vocabulary in the form of ontologies, and bridging the heterogeneities among web services and goals through mediation. Several languages of varying expressive power, such as WSML-Rule,

---

S.M. Atae (✉) · Z. Bayram  
Eastern Mediterranean University, TRNC via Mersin 10, Famagusta, Turkey  
e-mail: shahin.mpa@gmail.com

WSML-Flight, and WSML-Full [3] have been proposed to specify web services according to WSMO, all based on Frame-logic (F-logic) [4].

Flora-2 by Michael Kifer et al. [5] is a powerful language for knowledge representation and reasoning. It is based on F-logic, HiLog [6], and Transactional logic [7]. Frames are fundamental structures in Flora-2. They provide a means for defining classes and objects logically. In the case of objects, a frame can be shown generally in form of `objectId[attribute1-> value1, attribute2-> value2, ..., attributen-> valuen]`.

In this work, we use Flora-2 as a specification language for semantic description of web service components according to WSMO and implement a matching engine based on inference in F-logic in order to discover web services that can satisfy user requests specified in the form of goals. Our semantic specification is very concise since it makes use of the underlying Flora-2 syntax to the highest degree possible. The implementation of the matcher is also very compact since it makes effective use of the meta-level capabilities of the Flora-2 system.

The rest of this paper is structured as follows. In Sect. 2, we describe the semantic specification of goals and web services in WSMO, as well as the logic we have used in the implementation of our service matching engine. In Sect. 3, we demonstrate the power and practicality of our scheme through a simple but realistic scenario (first described in [8]) and show how our solution fulfills all the requirements needed to specify and discover web services/goals. We present a short survey of related work and compare it to ours in Sect. 4. Finally, in Sect. 5 we present the conclusion and future work in the area of semantic web service discovery based on our semantic web service specification approach.

## 2 Semantic Specification of Goals and Web Services and the Matching Process

The functionality of a WSMO web service is defined under the *capability* tag (element) which contains four axioms: pre-condition, assumption, post-condition and effect [9]. Pre and post conditions represent the internal state of the web service, whereas assumption and effect represent the state of the outside world (environment). A WSMO web service guarantees its post-condition and effect if its pre-condition and assumption are true. This feature of WSMO web services is used for discovery and selection purposes. For the sake of simplicity, we only consider web service pre-conditions (shown by *web.pre*) and web service post-conditions (shown by *web.post*), since assumptions and effects can be handled in a similar way.

Logically, the functionality of a web service can be shown by the following formula (the arrow represents the *implication* operator).

$$\forall x : web.pre(x) \Rightarrow web.post(x) \quad (1)$$



This formula means that for all instantiations of the free variables in the formula (represented by  $x$ ), if the pre-condition is true, the web service guaranties that the post-condition will also be true after the web service has finished its execution.

The definition of a logical match between a goal and a web service can be described precisely with the formula below:

$$\begin{aligned} \forall x_i \forall y_i : ((goal.pre(x_i) \Rightarrow web.pre(y_i)) \\ \wedge \\ (goal.pre(x_i) \wedge (web.pre(y_i) \Rightarrow web.post(y_i)) \Rightarrow goal.post(x_i)) \end{aligned} \quad (2)$$

This formula should be shown to be a valid statement in F-logic before we can say that the web service completely satisfies the functional requirements of the goal. In this formula,  $x_i$  represents the free variables in the goal and  $y_j$  represents the free variables in the web service.

Informally, the formula above checks that the goal pre-condition logically implies the pre-condition of the web service (hence guarantying that the web service has all it needs before it gets executed) and that the goal pre-condition, together with the implicit statement of the web service functionality (that the web service pre-condition implies the web service post-condition) logically implies the goal post-condition, thus guarantying that the goal will get the desired result with the execution of the web service.

In our implementation of web service specifications and the matcher, we diverge slightly from the logical definition given above in order to take advantage of the meta-logical capabilities of Flora-2 i.e. insertion of new facts which is a meta-logic operation. Specifically, the post-condition of web services can contain the *insert* predicate of Flora-2, so that the post-condition, instead of just being stated as being true, is made to be true by insertion of facts into the knowledge base. Then the post-condition of the goal can be tested against the new knowledge base.

Listing 1 depicts how our matching logic is implemented in Flora-2. Predicate `%match` in line 1 takes two variables, `?goal` representing a goal object and `?WS` representing a web service object as its parameters. In line 2, a new variable, `?module` is defined and assigned to the web service tag. In line 3, a new Flora-2 module with the same name as the web service tag is created and the description of the goal object is loaded into it. In line 4, the pre-condition of the goal (*goal.pre*) is inserted into the created module. Then in line 5, by calling `%applyWebService(?WS)` the Flora-2 reasoner attempts to prove the web service functionality specified in the form of an *if-then-else* statement in the knowledge base module. If `%applyWebService(?WS)` is proven, this means that the pre-condition of the web service is logically implied by the pre-condition of the goal, and moreover the actions specified in the post-condition of the web service have been carried out. In line 6, the variable `?gPost` is assigned to the goal post-condition, and in line 7 its validity is checked against the current knowledge base. If the check succeeds, this means that the goal post-condition is logically implied by the web service post-condition. It should be clear that the `%match` predicate indeed verifies

the validity of the formula (2) that we defined as the meaning of a successful match between a goal and a web service.

**Listing 1** The `%match` and `%matcher` predicates

```

1: %match(?goal, ?WS) :-
2:     ?module=?WS.tag, // Specifies the name of module
3:     %loadGoal(?goal, ?module),
4:     %insertGoalPre(?goal, ?module), // Inserts
                                     goal.pre into the KB.
5:     %applyWebService(?WS), // If ws.pre gets true
                               then ws.post will be
                               inserted into the KB.
6:     ?goal[post -> ?gPost]@?module,
7:     ?gPost. // Checks whether goal.post is implied
               by the ws.post.
8: %applyWebService(?ws) :-
    ?X = ?ws.def, ?X. // Tries to prove the web
                      service definition (ws.def).
9: %matcher(?goal, ?WS) :-
    \if %match(?goal, ?WS)
    \then writeln(['Goal', ?goal,
                  'matches', ?WS, '. '])@\prolog
    \else writeln(['Goal', ?goal,
                  'does not match', ?WS, '. '])@\prolog.

```

### 3 Use Case: Medical Appointment Finder

In this section we show how our approach can be used to describe the scenario in [8]. We will see that same results are achievable. However, our matching engine is much simpler.

The use case scenario is as follow: A patient named *Philip* wants to make an appointment with a specialist doctor (*ophthalmologist*) in *Montpellier* hospital located in a city of *France*. His preferred dates for this appointment are the days either before 19th or after 23rd (excluded) of the month. The patient should provide some basic information about himself, as well as a description of what he desires. Listing 2 shows a sample goal for this scenario rewritten in our specification format.

The patient provides the specialty *ophthalmology*, his name *Philip*, the hospital name he wishes to get the appointment from (i.e. *Montpellier*), and his age (which

is requested by the web service) in the form of an appointment request. What he wants is an appointment date either before the 19th or after the 23rd and an available specialist doctor's name.

For the web service side, web service pre and post-conditions have been given in listing 3. The web service uses a local database of Doctor instances containing information about doctors (i.e. *doctor1* and *doctor2*) and some general facts (i.e. *Montpellier* hospital is in *Paris*), and these are also depicted in listing 3.

**Listing 2** Goal specification for the appointment use case

```

1: o_G03:c_Goal. //o_G03 is an object of the concept
      c_Goal
2: o_G03[
3: pre -> ${RequestAppointment[ specialty->Ophthalmology,
4:           patientName -> Philip,
5:           appointmentDate -> ?_,
6:           hospitalName -> MontpellierHospital,
7:           age -> 22],
8:           livesIn(Philip,Paris)},
9:post -> ${Appointment[appointmentDate -> ?Date,
10:          doctorName -> ?DN,
11:          patientName -> Philip,
12:          hospitalName -> MontpellierHospital
13:          ],
14:          ((?Date < 19); (?Date > 23))}].

```

The web service provides some placeholders for its inputs while checking them over some predefined criteria (like, the patient must be at least 19 years old). Moreover, it checks whether the patient lives in the same city as the location of candidate hospital. After successful unification of inputs, the web service inserts all the possible appointments into the specified module (in this case @WS01) which is the common knowledge base between the web service and the goal. In this example, just the doctors with the specialty of *ophthalmology* who are working in *Montpellier* hospital are inserted into this module.

By referring to listing 1 again, we can see that all these actions take place through the call to `%applyWebService(?WS)` at line 5 in the `%match` predicate. At line 7, the Flora-2 reasoning engine attempts to prove the goal post-condition. At this point, the appointment date is checked to verify that it conforms to the constraints specified by the patient in the post-condition of the goal (i.e. either before the 19th or after the 23rd). This checking filters out those doctors who are not available during the requested dates.

**Listing 3** Web service specification for the appointment use case

```

1: doctor1[
2:     doctorName → Robert,
3:     specialty → Neurology,
4:     hospitalName → MontpellierHospital,
5:     availableDate → 22
6: ]:Doctor.

7: doctor2[
8:     doctorName → Green,
9:     specialty → Ophthalmology,
10:    hospitalName → MontpellierHospital,
11:    availableDate → 10
12: ]:Doctor.
13: hospital(MontpellierHospital, Paris).
14: o_WSOL:c_WebService[
15:   tag → WSOL,
16:   def →
17:     ${\if (RequestAppointment[specialty → ?DS]@WSOL,
18:     RequestAppointment[patientName → ?PN]@WSOL,
19:     RequestAppointment[appointmentDate→?Date]@WSOL,
20:     RequestAppointment[hospitalName → ?HN]@WSOL,
21:     RequestAppointment[age → ?X]@WSOL,
22:     (?X > 18),
23:     livesIn(?PN,?city)@WSOL,
24:     hospital(?HN,?city),
25:     ?doctor:Doctor[doctorName → ?DN,
26:                     specialty → ?DS,
27:                     hospitalName → ?HN,
28:                     availableDate → ?Date])
29:   \then (
30:     ?post = ${Appointment[appointmentDate→?Date,
31:                           doctorName → ?DN,
32:                           patientName → ?PN,
33:                           hospitalName → ?HN]@WSOL},
34:     %insert{?post}
35:   \else \false }].

```

If we changed the last line in listing 3 to `hospital(MontpellierHospital, Berlin)`, the match would fail since the web service pre-condition would not be satisfied. Similarly, if we changed line 11 in listing 3 to `availableDate → 21`, again the match would fail, but this time due to the fact that the goal's post-condition would not be satisfied.

## 4 Related Work

A good explanation of WSMO and WSML can be found in [9] and [3]. The main available resource for Flora-2 is its user's manual containing useful examples of Flora-2 code [5].

In [10], the authors propose a framework for semantic web service discovery using FIPA multi-agents. They have a broker architecture and deal with OWL-S [16] rather than WSMO, as we do. In [11] the work presented is similar to ours, however the authors use WSML to specify goals and web services, which is very verbose. Furthermore, they do not state the proof commitments that are needed for a successful match in a logical way.

The closest and most comparable work to ours is [12], where the authors use Flora-2 to present a logical framework for automated web service discovery. Moreover, they use WSMO specification as the conceptual description of web services as we do. However, their specification of web services and goals are very involved, and they resort to Transaction logic for proof commitments. Our specifications, as apparent in the realistic example given in Sect. 3, are quite intuitive and simple. Furthermore, we make use of only Frame-logic for stating our proof commitments, which itself is equivalent to first-order predicate logic [4], and is much more accessible to the reader. The simplicity of our approach can be an important facilitator in its adaptation by industry (after necessary enhancements to deal with the web environment).

There are several surveys and reviews about semantical as well as non-semantical web service discovery proposals, which give a general overview of this field of study [13–15]. Many of the surveyed proposals are based on OWL-S. As M. Kifer et al. stated in [12], such approaches rely on subsumption reasoning [17] and due to the lack of rules in OWL, they are not able to exactly guarantee goal post-conditions.

## 5 Conclusion and Future Work

We have demonstrated how Flora-2 can be used as a convenient and expressive way to model semantic web services conforming to WSMO. We have also shown that reasoning for semantic web service discovery can be done very effectively by relying on the underlying Flora-2 engine and its meta-level capabilities. Our work can readily be the basis of implementing strategies for semantic web service composition and choreography. We intend to concentrate our efforts for realizing this goal in the near future.

## References

1. OASIS: Reference Architecture Foundation for Service Oriented Architecture Version 1.0. Committee Specification 01 (2012)
2. Bruijn, J., Bussler, C., Domingue, J., Fensel, D.: Web Service Modeling Language. <http://www.w3.org/Submission/WSMO/> (2005)
3. Fensel, D., Kifer, M., Bruijn, J.: D16.1v1.0 WSML Language Reference (WSML Final Draft). <http://www.wsmo.org/TR/d16/d16.1/v1.0/> (2008)
4. Kifer, M., Lausen, G., Wu, J.: Logical foundations of object oriented and frame-based languages. *J. ACM (JACM)* **42**(4), 741–843 (1995)
5. Kifer, M., Yang, G., Wan, H., Zhao, C.: Flora-2: User's Manual, Version 1.0 (Cherimoya). Department of Computer Science, Stony Brook University, Stony Brook, NY 11794–4400, U.S.A. (2014)
6. Chen, W., Kifer, M., Warren, D.: HiLog: a foundation for higher-order logic programming. *J. Logic Program.* **15**(3), 187–230 (1993). doi:[10.1016/0743-1066\(93\)90039-J](https://doi.org/10.1016/0743-1066(93)90039-J)
7. Kifer, M., Bonner, A.: Logic Programming for Database Transactions in Logics for Databases and Information Systems. Kluwer Academic Publication, Dordrecht (1998)
8. Sharifi, O., Bayram, Z.: Matching Goal and Semantic Web Services in Flora-2: A Logical Inference Based Discovery Agent. submitted for publication (2015)
9. Fensel, D., Lausen, H., Polleres, A., de Bruijn, J., Stollberg, M., Roman, D., Domingue, J.: Enabling Semantic Web Services, The Concepts of WSMO, pp. 63–81. Springer, Heidelberg. [http://dx.doi.org/10.1007/978-3-540-34520-6\\_6](http://dx.doi.org/10.1007/978-3-540-34520-6_6). Accessed 1 Jan 2007
10. Neiat, A., Mohsenzadeh, M., Forsati, R., Rahmani, A.: An agent-based semantic web service discovery framework. In: International Conference on Computer Modeling and Simulation, 2009 (ICCMS '09), pp. 194–198. IEEE (2009). doi:[10.1109/ICCMS.2009.75](https://doi.org/10.1109/ICCMS.2009.75). ISBN:978-0-7695-3562-3
11. Sirbu, A., Toma, I., Roman, D.: A Logic-based Approach for Service Discovery with Composition Support, pp. 101–116. Whitestein Series in Software Agent Technologies and Autonomic Computing, Emerging Web Services Technology (2007)
12. Kifer, M., Lara, R., Polleres, A., Zhao, C., Keller, U., Lausen, H., Fensel, D.: A logical framework for web service discovery. In: Proceedings of the ISWC 2004 workshop on Semantic Web Services: Preparing to Meet the World of Business Applications (2004)
13. Duy Ngan, L., Kirchberg, M., Kanagasabai, R.: Review of semantic web service discovery methods. In: 6th World Congress on Services (SERVICES-1), pp. 176–177 (2010). doi:[10.1109/SERVICES.2010.85](https://doi.org/10.1109/SERVICES.2010.85). ISBN:978-1-4244-8199-6
14. Kster, U., Lausen, H., Knig-Ries, B.: Evaluation of Semantic Service Discovery? A Survey and Directions for Future Research. Emerging Web Services Technology, Volume II, Whitestein Series in Software Agent Technologies and Autonomic Computing 2008, pp. 41–58
15. Malaimalavathani, M., Gowri, R.: A survey on semantic web service discovery. In: International Conference on Information Communication and Embedded Systems (ICICES), pp. 222–225. IEEE (2013). doi:[10.1109/ICICES.2013.6508208](https://doi.org/10.1109/ICICES.2013.6508208). ISBN:978-1-4673-5786-9
16. Martin, M., Burstein, M., Hobbs, J.: OWL-S: Semantic Markup for Web Services. <http://www.w3.org/Submission/OWL-S/> (2004)
17. Patel-Schneider, P., Hayes, P., Horrocks, I.: OWL Web Ontology Language Semantics and Abstract Syntax. <http://www.w3.org/TR/owl-semantics/>, W3C Recommendation, W3C (2014)

# Fast Frequent Episode Mining Based on Finite-State Machines

Stavros Papadopoulos, Anastasios Drosou and Dimitrios Tzovaras

**Abstract** Frequent Episode Mining (FEM) techniques play an important role in data mining, and have multiple applications, spanning from identifying user marketing habits to performing anomaly detection in computer networks. Most of the FEM approaches exhaustively search for frequent patterns, while using a threshold to efficiently reduce the search space. While this approach provides efficient results in small datasets, it fails in large datasets due to heavy processing, which leads to low performance. This paper, proposes a fast frequent episode mining method which utilizes Finite-State Machines (FSM). Initially, a FSM is created based on a subset of the data, in order to approximate the type and frequency of the most dominant episodes. Instead of applying traditional exhaustive search procedures, the parsing of the dataset is herein guided by the proposed FSM approach. Experimental results show that the proposed approach has better time performance than the traditional FEM algorithms, while still maintaining high accuracy.

## 1 Introduction

Frequent Episode Mining (FEM) [1] is a well known set of methods for discovering interesting patterns in data sequences, and also a popular field in data mining. The main limitation of the previously proposed apriori and pattern growth-based FEM methods, is that they exhaustively search the space of all the possible episodes for frequent appearances, based only on a user-defined threshold to trim it. This fact renders the previously proposed FEM methods not suitable for application on large

---

S. Papadopoulos (✉)  
Imperial College London, London, UK  
e-mail: s.papadopoulos@imperial.ac.uk

A. Drosou · D. Tzovaras  
CERTH, ITI, Thessaloniki, Greece  
e-mail: drosou@iti.gr

D. Tzovaras  
e-mail: tzovaras@iti.gr

datasets. To overcome this issue, this paper proposes the use of Finite-State Machines (FSM), which drive the search for frequent episodes, without the need to consider all the possible episodes. An FSM, models the probabilities of event appearances and transitions, while a transition on the FSM checks the most probable event sequences, for the generation of the set of frequent episodes. When compared to traditional FEM methods, the proposed FSM-FEM method runs much faster, while it is still able to identify the most frequent episodes.

Recently, the research community has started to focus on the identification of other types of patterns, which are not necessary frequent, but significant in some other sense [3, 4]. This paper is not focused on the identification of significant patterns, but rather on the identification of frequent patterns, which have multiple applications, e.g. [1, 5]. This paper does not consider simultaneous events.

## 2 Related Work

One of the earliest works on FEM was presented by Mannila et al. [11]. The authors proposed WINEPI, a breadth-first search based method to mine frequent episodes using a window-based frequency definition, which limits the size of the discovered patterns. A modification of windows frequency was proposed by Iwanuma et al. [9], called Head frequency. Based on this frequency definition, Ding et al. [8] proposed the MINEPI+ algorithm, which is a depth-first approach for discovering all frequent episodes. In the same paper, the authors also proposed a second algorithm called EMMA, which utilizes memory anchors to accelerate the mining task. Pei et al. proposed prefixspan [12] a sequential pattern mining method which utilizes prefix-projection to reduce the number of candidate for the subsequence generation.

While the aforementioned methods find the most interesting patterns, they are not scalable to a large dataset. Towards this end, this paper introduces a fast FEM algorithm based on a FSM model, which is created based on a small dataset, and is afterwards used to guide the search for the identification of the most significant patterns. Experimental demonstration reveals that the proposed method is much faster than traditional FEM algorithms.

## 3 Notation

Let us denote as  $E$  the set which includes all the possible events, and  $T$  the set of all possible time instances. A single sequence of events  $S$  is defined as an ordered sequence of events occurring at specific time instances:  $S = \langle \langle e_1, t_1 \rangle, \langle e_2, t_2 \rangle, \dots, \langle e_n, t_n \rangle \rangle$ , where  $e_i \in E$ ,  $t_i \in T$ , and  $t_{i-1} \leq t_i$ . The same event can be repeated at different time instances. A specific event in the sequence is denoted as  $S[i] = e_i$ . A sequence database is a set of sequences, denoted as  $S_{all} = \{S_1, S_2, \dots, S_N\}$ . A pattern  $P$  is defined as an ordered sequence of events,



without including the time parameter:  $P = \langle e_1, e_2, \dots, e_k \rangle$ . The  $i$ th event in the pattern sequence is denoted as  $P[i]$ .

**Definition 1** A sequence  $S_i = \langle \langle e_1, t_1 \rangle, \langle e_2, t_2 \rangle, \dots, \langle e_n, t_n \rangle \rangle$  is a subsequence of sequence  $S_j = \langle \langle e'_1, t'_1 \rangle, \langle e'_2, t'_2 \rangle, \dots, \langle e'_k, t'_k \rangle \rangle$ , denoted as  $S_i \leq S_j$ , if and only if there exist a sequence of integers  $1 \leq l_1 < l_2 < \dots < l_n \leq k$  such that  $S_i[l_q] = S_j[l_q]$ ,  $\forall q \in \{1, 2, \dots, n\}$ .

**Definition 2** A pattern  $P = \langle e_1, e_2, \dots, e_n \rangle$  occurs in sequence  $S_j = \langle \langle e'_1, t'_1 \rangle, \langle e'_2, t'_2 \rangle, \dots, \langle e'_k, t'_k \rangle \rangle$  (denoted as  $P \leq S_j$ ) if and only if there exist a sequence of integers  $1 \leq l_1 < l_2 < \dots < l_n \leq k$  such that  $P[l_q] = S_j[l_q]$ ,  $\forall q \in \{1, 2, \dots, n\}$ . The specific sequence of integers  $\langle l_1, l_2, \dots, l_n \rangle$ , defines an occurrence of  $P$  in sequence  $S_j$ .

It should be noted that pattern  $P$  can have more than one occurrences in a sequence  $S_j$ , as well as in the set of all sequences  $S_{all}$ . A specific occurrence in sequence  $S_j$ , is denoted as  $\langle j, \langle l_1, l_2, \dots, l_n \rangle \rangle$ . The total number of occurrences a pattern in the  $S_{all}$  is called the *Support* of pattern  $P$ . Due to the fact that direct counting of the *Support* of a pattern in a sequence is not feasible for large sequences and patterns [1], different methods have been proposed in the literature in order to efficiently calculate different *Support* metrics under different constraints. Some of the most known support definitions include: window-based [11], minimal occurrences [11], head and total frequency [9], and repetitive *Support* [5].

This work also proposes the use of the *appearances Support*, which takes into account multiple sequences:

**Definition 3** Given a set of sequences  $S_{all}$ , the *appearances Support* of pattern  $P$  is defined as the number of sequences in  $S_{all}$  in which the pattern  $P$  appears at least one time.

The *appearances Support* definition complies with the apriori property, and thus, enables the creation of efficient algorithms for the calculation of frequent patterns.

## 4 Creation of the Finite-State Machine

In this work, each state in the FSM represents a unique event in the sequence. The FSM can be represented as a transition matrix  $T$ , in which the element in the  $i$ th row and  $j$ th column is one only and only if there is a direct connection from the  $i$ th state to the  $j$ th, or else it is zero.

Algorithm 1 presents the procedure followed for the creation of the FSM, based on the *Support* of all the pairs of possible episode sequences. The value of the *thresholdFSM*, trims the FSM in order to reduce the search space of the FEM, while keeping the most significant event pairs in the model. The value of *thresholdFSM*=0 creates an FSM with all the possible transitions. The FSM is created for a subset of sequences  $S_{sub} \subset S_{all}$ , in order to increase its creation speed. The *thresholdFSM*

<p><b>Data:</b> The list of sequences <math>S_{sub} \subset S_{all}</math>, the list of possible events <math>E = \{e_1, e_2, \dots, e_n\}</math>, a threshold value <math>thresholdFSM</math></p> <p><b>Result:</b> The FSM in the form of a transition matrix <math>T</math></p> <p><b>Procedure</b> <math>trainFSM(S_{sub}, E, thresholdFSM)</math></p> <pre> <math>T \leftarrow 0</math> <b>for all</b> <math>e_i, e_j \in E</math> <b>do</b>   <b>if</b> <math>supp(e_i, e_j) \geq thresholdFSM</math> <b>then</b>     <math>T(e_i, e_j) \leftarrow 1</math>   <b>end</b> <b>end</b> <b>return</b> <math>T</math> </pre>
---

**Algorithm 1:** Algorithm for the training of the FSM, based on the dataset  $S_{sub} \subset S_{all}$ .

is used to reduce the total size of the FSM, in order to search only for the patterns that might be frequent, and remove the patterns that are not. Generally, the value of  $thresholdFSM$  is set equal to the *Support trimming threshold* (see Algorithm 3) The reason for this is that the algorithm does not need to search for patterns which are already below the threshold, and thus, these patterns can be safely removed from consideration. This removal takes place through the  $thresholdFSM$  parameter.

For the selection of the appropriate subset of sequences, the set  $S_{sub}$  is iteratively increased, until a stable configuration of FSM trained from this dataset is identified. Thus, a sequence of FSMs  $FSM_{all}$  is created, one for each iteration of the algorithm:  $FSM_{all} = \{FSM_1, FSM_2, \dots, FSM_{|FSM_{all}|}\}$ , where  $FSM_i$  is the FSM created on the  $i$ th iteration.

For the detection of stable FSM configurations, the difference between consecutive FSMs in  $FSM_{all}$  is calculated, in order to create a sequence of differences. Specifically, since each FSM is a directed weighted graph, graph matching methods are applied in order to compute their distances. In particular, the graph edit distance [6] metric is utilized. Thus, the sequence of differences is defined as follows:  $D_{all} = \{D_1, D_2, \dots, D_{|D_{all}|}\}$ , where  $D_i = GED(FSM_i, FSM_{i+1})$ , and  $GED(FSM_i, FSM_{i-1})$  denotes the graph edit distance between the two FSMs  $FSM_i, FSM_{i-1} \in FSM_{all}$ .

Algorithm 2 presents the procedure followed in order to create the most appropriate set  $S_{sub}$ , and the subsequent calculation of the final FSM. The method proposed by Kim et al. [10] is used for the online detection of the point from which the sequence has reached the steady state.

## 5 Traversal of the FSM

A traversal method is applied on the identified FSM, in order to generate the sequence of candidate frequent patterns. Two functions are used, which represent the state transitions, the  $move\_forward(T, P)$  and the  $move\_backward(P)$ , where  $P$  the current candidate pattern. The  $move\_forward(T, P)$  generates a new pattern based on the

**Data:** The list of all the sequences  $S_{all}$ , the list of possible events  $E = \{e_1, e_2, \dots, e_n\}$ , a threshold value  $thresholdFSM$

**Result:** The FSM in the form of a transition matrix  $T$

**Procedure**  $createFSM(S_{all}, E, thresholdFSM)$

```

 $S_{sub} \leftarrow \emptyset$ 
 $T \leftarrow \emptyset$ 
do
  choose randomly an  $S_i \in S_{all}$ 
   $S_{sub} \leftarrow S_{sub} \cup S_i$ 
   $T \leftarrow trainFSM(S_{sub}, E, thresholdFSM)$ 
  update the sequence of differences  $D_{all}$  with the previously found  $T$ 
while  $D_{all}$  has not reached steady state;
return  $T$ 

```

**Algorithm 2:** Algorithm for the creation of the most representative FSM.

last pattern, which can be either larger or smaller than the old pattern, based on the list of available moves. A specific characteristic of the  $move\_forward(T, P)$  is that it is not allowed to generate the same pattern  $P$  twice, and thus, each pattern created by this method must be new. In case that the pattern that is returned is empty, the FSM method is unable to find new patterns, and thus, the algorithm terminates. The  $move\_backward(P)$  method reduces the size of the pattern by 1 by removing the last event in the pattern  $P$ , and moves to the previous FSM state. It should be noted that the  $move\_forward(T, P)$  function can call the  $move\_backward(P)$  if no next moves are found from the current pattern.

Finally, we utilize the *maximal frequent patterns* [7], which reduces the set of frequent patterns by removing redundant information. This set of comprised of all the frequent patterns  $P$  that do not have a super-pattern  $P'$  ( $P \preceq P'$ ) in *FrequentPatterns*.

The computational complexity of the proposed approach has to do with the traversal of the FSM, and thus its size. Its size depends on the number of frequent patterns of size one, or in other words the number of states in the FSM, denoted as  $V$ , and the number of edges of the FSM, or in other words the number of frequent patterns of size two, denoted as  $E$ . The algorithm checks for each starting state the possible transitions on the FSM, leaving out regions of the FSM that do not represent frequent patterns. In the worst case, the algorithm will check all the edges for each starting state and not leave out any region of the FSM. In the worst case, the number of iterations of the algorithm is  $V * E$ . On average, the algorithm will check only half of the edges for each state, i.e. the average number of iterations of the algorithm is  $1/2 * V * E$ . In the best case, were the dataset is comprised only of patterns of size 1, the algorithm will only check each vertex once. Thus, in the best case, the average number of iterations of the algorithm is  $V$ . This means that the complexity of the algorithm is  $O(V * E)$ , with best case complexity  $O(V)$ .

```

Data: The list of sequences  $S_{all} = \{S_1, S_2, \dots, S_N\}$ , the set of events  $E$ , the threshold, and the
thresholdFSM
Result: The set of frequent sequences FrequentPatterns
Main  $FSM_{fem}(S_{all}, E, threshold)$ 
   $T \leftarrow createFSM(S_{all}, E, thresholdFSM)$ 
   $FrequentPatterns \leftarrow \emptyset$ 
   $P \leftarrow \emptyset$ 
   $P \leftarrow move\_forward(T, P)$ 
  while  $P \neq \emptyset$  do
    if  $supp(S_{all}, P) \geq threshold$  then
       $FrequentPatterns \leftarrow FrequentPatterns \cup P$ 
    else
       $P \leftarrow move\_backward(P)$ 
    end
     $P \leftarrow move\_forward(T, P)$ 
  end
  return FrequentPatterns

```

**Algorithm 3:** FSM-FEM algorithm for the calculation of the frequent event patterns.

Algorithm 3 presents the proposed FSM-FEM algorithm for the calculation of the frequent event patterns. The candidate patterns are generated from the FSM, and are checked based on their *Support* value. The algorithm results in the set of *Frequent-Patterns*.

## 6 Identification of Frequent Patterns in DNA Sequences

This section applies the proposed FSM-FEM method on multiple DNA sequences, in order to identify frequent patterns of amino-acid sequences.

The utilized dataset represents real DNA sequences from patients with CLL [2], and it is a proprietary dataset. The data represent amino-acid sequences from Complementarity Determining Regions 3 (CDR3) of B-cell lymphocytes. These data were collected from 1182 patients, and thus, there are in total 1182 sequences. The sequences vary in length, from 11 to 30 amino-acids. There are in total 20 different amino-acids. In this dataset, it is important to identify frequent patterns in specific positions, and thus each amino-acid is representing not only by its identifier (i.e. a specific letter of the alphabet), but also by its position in the sequence (e.g. first, second etc.). Each sequence in the dataset corresponds to one patient.

The results of the application of the FSM-FEM method on the DNA sequence dataset are presented in Table 1. This table shows a subset of the most frequent patterns, while a total of around 7000 frequent patterns were discovered. The *Support*

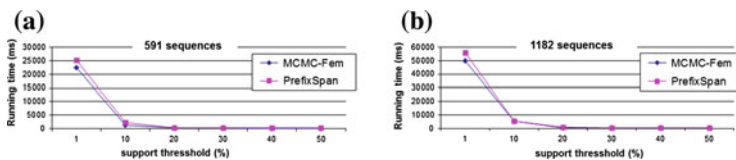
**Table 1** The most frequent patterns identified by applying the proposed FSM-FEM approach on the DNA sequence dataset using *Support* threshold equal to 10% of the total sequences

#	Frequent pattern	Support	Support threshold	Percentage of appearances (%)
1	$\langle C - 0, A - 1, R - 2, Y - 12 \rangle$	119	118	10.06
2	$\langle C - 0, A - 1, R - 2, Y - 14 \rangle$	169	118	14.29
3	$\langle R - 2, G - 3 \rangle$	209	118	17.68
4	$\langle R - 2, G - 5 \rangle$	128	118	10.82

The first letter of each event is the amino acid, and the second is the corresponding position in the sequences

utilized for this calculation is the *appearances Support*, while the FSM-FEM was applied with *Support* threshold equal to 10% of the total sequences. The identified patterns are indeed characteristic of patients with CLL. Patterns 1 through 3 in Table 1 represent a specific subset of patients with high similarity both in their sequences but also to their clinical image (i.e. subset 4 according to [2]), while patterns 4 through 6 another subset of patients with many similarities (i.e. subset 6 according to [2]). It should also be noted that the patterns 1 through 3 are “subset 4-specific” and “CLL-specific” [2]. Thus, the proposed FSM-FEM method was able to efficiently identify the patterns that exist in the different CLL subgroups. Around 15% of the total dataset was utilized for the creation of the FSM.

Figure 1 presents the comparison of the prefixspan [12] and the FSM-FEM with respect to their running times, based different the *Support* thresholds. The DNA dataset is partitioned in two datasets, one comprised of 591 sequences, and a second comprised of all the 1182 sequences. As shown in both cases the FSM-FEM is faster, but for large *Support* thresholds the difference is not so large. The reason for this is that the number of patterns with high frequency is relatively small, and thus, for small sequences the prefixspan and the FSM-FEM have comparable running times. The difference, however, increases for frequent patterns with *appearances Support* larger than 1%.



**Fig. 1** The running time of the prefixspan [12] and the FSM-FEM based on different *Support* thresholds. **a** Utilization of 1182 sequences. **b** Utilization of 591 sequences

**Table 2** The most frequent patterns identified by applying the proposed FSM-FEM approach on the NBC website visits sequence dataset using *Support* threshold equal to 1 %

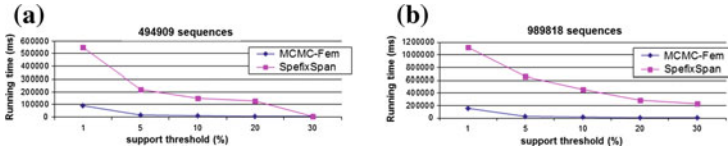
#	Frequent pattern	Support	Support threshold	Percentage of appearances (%)
1	<i>(frontpage, frontpage, frontpage)</i>	108,876	9898	10.99
2	<i>(bbs, bbs)</i>	72,063	9898	7.28
3	<i>(misc, misc)</i>	58,746	9898	5.93
4	<i>(health, health, health)</i>	22,980	9898	2.32

## 7 Identification of Frequent Patterns of Visits on the NBC Website

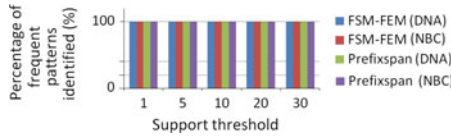
The data comes from Internet Information Server (IIS) logs for msnbc.com and news-related portions of msn.com for the entire day of September, 28, 1999 (Pacific Standard Time).<sup>1</sup> This dataset is freely available from the UCI machine learning repository [13]. Each sequence in the dataset corresponds to page views of a user during that 24 h period. Each event in the sequence corresponds to a user’s request for a page. Requests are not recorded at the finest level of detail, i.e. at the level of URL, but rather, they are recorded at the level of page category (as determined by a site administrator). Any page requests served via a caching mechanism were not recorded in the server logs and, hence, not present in the data. This dataset is comprised of 989,818 users, with an average number of visits per user equal to 5.7. The results of the application of the FSM-FEM method on the NBC website visits sequence dataset are presented in Table 2. This table shows a subset of the most frequent patterns, while a total of around 1600 frequent patterns were discovered. For the calculation of these results, a *Support threshold* equal to 1 % was used, while for the calculation of the support, the *appearances Support* was utilized. As shown in this table, the vast majority of the users visit the “frontpage” first, and navigate in it (pattern 1 in Table 2), until they decide what to see next. The most visited webpages are by far the “bbs” and the “on-air” (patterns 2 and 3), while next in popularity comes “misc”, “local”, and “health” (patterns 4, 5, and 6). Around 15 % of the total dataset was utilized for the creation of the FSM.

Figure 2 presents the comparison of the prefixspan [12] and the FSM-FEM with respect to their running times, based on different *Support* thresholds. Two datasets are extracted from the initial NBC dataset, one comprised of all the 989,818 sequences, and one comprised of 494,909 sequences. For a large threshold, the number of patterns identified is very small (around 10 patterns) and thus, the running times of the two algorithms are similar. As the *threshold* is getting smaller, and the number of patterns increases, the difference in the running times between the prefixspan and the FSM-FEM is getting larger.

<sup>1</sup>This data is available thanks to msnbc.com.



**Fig. 2** The running time of the prefixspan [12] and the FSM-FEM based on different *Support* thresholds. **a** Utilization of 989,818 sequences. **b** Utilization of 494,909 sequences



**Fig. 3** Comparison of the percentage of total frequent patterns identified by FSM-FEM and prefixspan, for different support thresholds and the different datasets. Both methods are able to identify 100 % of the patterns in each case

**Table 3** Frequent patterns and example sequences that have these patterns, for both the NBC and DNA datasets

#	Frequent pattern	Example sequence	Dataset
1	$\langle \textit{frontpage}, \textit{frontpage}, \textit{frontpage} \rangle$	$\langle \textit{frontpage}, \textit{bbs}, \textit{frontpage}, \textit{bbs}, \textit{frontpage}, \textit{frontpage}, \textit{frontpage}, \textit{frontpage}, \textit{frontpage} \rangle$	NBC visits
2	$\langle \textit{health}, \textit{health}, \textit{health} \rangle$	$\langle \textit{health}, \textit{health}, \textit{tech}, \textit{health}, \textit{frontpage}, \textit{health}, \textit{local}, \textit{local}, \textit{health} \rangle$	NBC visits
3	$\langle C - 0, A - 1, R - 2, Y - 14 \rangle$	$\langle \textit{CAEAVVPAAMVPNYYYYYMDVW} \rangle$	DNA
4	$\langle R - 2, G - 3 \rangle$	$\langle \textit{CARGLLTRREGVYMDIW} \rangle$	DNA

Figure 3 shows a comparison with respect to the total frequent patterns identified by FSM-FEM and prefixspan, for different support thresholds and the different datasets. It is apparent from this figure that both methods are able to efficiently identify 100 % of the patterns in each case. It should be noted that since the prefixspan is a method that identifies the complete set of frequent patterns, this figure also illustrates a comparison between the patterns identified by FSM-FEM and the ground truth. Table 3 shows a small set of frequent patterns and example sequences that have these patterns, for both the NBC and DNA datasets. The patterns identified in the sequences are highlighted in bold.

## 8 Conclusions

This paper presented a fast FEM method for the identification of frequent patterns in multiple sequences. The proposed approach is based on FSM methods, which guide the search of the frequent patterns. This fact reduces significantly the search space of the algorithm, and increases its running speed, while maintaining its accuracy.

Experimental demonstration of the proposed approach on real datasets revealed its capabilities in detecting frequent in a efficient and scalable manner. The proposed FSM-FEM method was compared with the traditional FEM methods, and was found to be superior with respect to its running time, while still being able to identify the most frequent patterns. Specifically, the FSM-FEM method runs more than 2–5 times faster than other methods depending on the dataset, a fact that renders it more suitable for mining frequent patterns in large datasets.

**Acknowledgments** This work has been partially supported by the European Commission through project FP7-ICT-317888-NEMESYS funded by the 7th framework program. The opinions expressed in this paper are those of the authors and do not necessarily reflect the views of the European Commission.

## References

1. Achar, A., Sastry, P.S., et al.: Pattern-growth based frequent serial episode discovery. *Data Knowl. Eng.* **87**, 91–108 (2013)
2. Agathangelidis, A., Darzentas, N., Hadzidimitriou, A., Brochet, X., Murray, F., Yan, X.-J., Davis, Z., van Gastel-Mol, E.J., Tresoldi, C., Chu, C.C., et al.: Stereotyped B-cell receptors in one-third of chronic lymphocytic leukemia: A molecular classification with implications for targeted therapies. *Blood* **119**(19), 4467–4475 (2012)
3. Aggarwal, C.C., Han, J.: *Frequent pattern mining*. Springer 1st edition (2014)
4. Castro, N.C., Azevedo, P.J.: Significant motifs in time series. *Stat. Anal. Data Min.: ASA Data Sci. J.* **5**(1), 35–53 (2012)
5. Ding, B., Lo, D., Han, J., Khoo, S.-C.: Efficient mining of closed repetitive gapped subsequences from a sequence database. In: *Data Engineering, 2009. ICDE'09. IEEE 25th International Conference on*, pp. 1024–1035. IEEE (2009)
6. Gao, X., Xiao, B., Tao, D., Li, X.: A survey of graph edit distance. *Pattern Anal. Appl.* **13**(1), 113–129 (2010)
7. Gouda, K., Zaki, M.: Efficiently mining maximal frequent itemsets. In: *Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference on*, pp. 163–170. IEEE (2001)
8. Huang, K.-Y., Chang, C.-H.: Efficient mining of frequent episodes from complex sequences. *Inf. Syst.* **33**(1), 96–114 (2008)
9. Iwanuma, K., Takano, Y., Nabeshima, H.: On anti-monotone frequency measures for extracting sequential patterns from a single very-long data sequence. In: *Cybernetics and Intelligent Systems, 2004 IEEE Conference on*, vol. 1, pp. 213–217. IEEE (2004)
10. Kim, M., Yoon, S.H., Domanski, P.A., Payne, W.V.: Design of a steady-state detector for fault detection and diagnosis of a residential air conditioner. *Int. J. Refrig.* **31**(5), 790–799 (2008)
11. Mannila, H., Toivonen, H., Verkamo, A.I.: Discovery of frequent episodes in event sequences. *Data Min. Knowl. Discov.* **1**(3), 259–289 (1997)
12. Pei, J., Han, J., Mortazavi-Asl, B., Pinto, H., Chen, Q., Dayal, U., Hsu, M.-C.: Prefixspan: Mining sequential patterns efficiently by prefix-projected pattern growth. In: *2013 IEEE 29th International Conference on Data Engineering (ICDE)*, pp. 215. IEEE Computer Society (2001)
13. UCI. Machine learning repository. <http://archive.ics.uci.edu/ml> (2010)



# Hybrid Heuristic Algorithms for the Multiobjective Load Balancing of 2D Bin Packing Problems

Muhammed Beyaz, Tansel Dokeroglu and Ahmet Cosar

**Abstract** 2D Bin packing problem (2DBPP) is an NP-hard combinatorial optimization problem. Multiobjective versions of this well-known industrial engineering problem can occur frequently in real world application. Recently, Hybrid Evolutionary Algorithms have appear as a new area of research with their ability to combine alternative heuristics and local search mechanisms together for higher quality solutions. In this study, we propose a set of novel multiobjective hybrid genetic and memetic algorithms that make use of the state-of-the-art metaheuristics and local search techniques for minimizing the number of bins while also maintaining the load balance. We analyze the optimization time and the resulting solution quality of the proposed algorithms on an offline 2DBPP benchmark problem set with 500 instances. Using these results of exhaustive experiments, we conclude that the proposed hybrid algorithms are robust with their ability to obtain a high percentage of the optimal solutions.

## 1 Introduction

The two-dimensional Bin Packing Problem (2DBPP) consists of planning a set of rectangular items into a fixed width and height 2D bins orthogonally, without overlapping, while minimizing the number of bins [1–5]. The 2DBPP is an intractable optimization problem and widely faced during the industrial manufacturing processes. Proposed algorithms find a pareto-optimal solution for both the minimal number of bins with the most efficiently load-balanced placing of the rectangular items. rotating

---

M. Beyaz · T. Dokeroglu (✉) · A. Cosar  
Middle East University Computer Engineering Department, Universities Street,  
6800 Ankara, Turkey  
e-mail: tansel@ceng.metu.edu.tr

M. Beyaz  
e-mail: muhammed.beyaz@ceng.metu.edu.tr

A. Cosar  
e-mail: cosar@ceng.metu.edu.tr

objects (Orientation: whether the objects can be rotated or not is a key aspect of the 2DBPP) in packing creates better results but objects may not be rotatable in every problem definition. The textile industry can change the orientation of single color shirts by rotating the shirts while the process is in cutting phase, because there is no difference between rotation or not. However, the orientation of fragile items is important in shipping. If an item can be rotated then it is called as non-oriented or orientation-free. If an object of problem cannot be rotated it is called as oriented or orientation-fix. Online and offline are two categories of 2DBPP according to the availability of information about all items. Online bin packing (OnBP) means that objects arrive one by one and there is no way to know the complete input sequence, so it must be inserted into a bin immediately without waiting other items. Load balancing of 2DBPP, is the stabilization of total moments of rectangle items on the left of Centre of Gravity (CG) of bins with total moments of rectangles on the right of CG of bins. Euclidean Center of bin is considered as CG of a bin.

In the proposed algorithms, we use solution methods inspired from Evolutionary Algorithms (EA). Reproduction, mutation, recombination and selection are key mechanisms of EA that are used for solving NP-Hard optimization problems. BPP, Travelling Salesman and Quadratic Assignment Problem [2, 6] are well-known challenging NP-Hard problems modelled and solved successfully with EAs. Genetic Algorithm (GA) and Memetic Algorithm (MA) are the most efficient approaches of EAs. GA mimics the natural evolution process and has the ability to find (near-) optimal solutions in a large search space. Survival of the fittest individual is a rule allowing the best solution in each iteration to converge to a (near-) optimal solution in practical times. In GA, parents mate and produce offsprings and the best individuals are selected to survive to the next generation. MA is another growing area of EA. The fittest of individuals is selected as the solution of optimization problem. It mimics natural evolution process but it may differ from GA by performing individual learning which is also known as meme(s).

We propose two different Multiheuristic Multiobjective GA (MH-MOGA) that optimize both the minimal number and the load-balancing of the bins. The first proposed multiobjective algorithm, MHO-MOGA, uses heuristics: FNF, FFF, BFDH, UTS and LGFof to solve oriented multi objective 2D offline BPPs. The second proposed multiobjective algorithm, MHNO-MOGA, uses heuristics: FNF, FFF, BFDH, UTS, LGFof and LGFi and tries to find a pareto-optimal solution (minimal number of bins and most effective load-balancing of items) for non-oriented multiobjective 2D offline BPPs.

## 2 Formulation of the Load-Balancing Problem

Multiobjective 2DBPP with load balancing [7–9] tries to minimize Eq. 1

$$(C/2 + LB/2) \tag{1}$$

where

$$C = \sum_{j=1}^n c_j \quad (2)$$

$$LB = \sum_{j=1}^C \left| \sum_{i=1}^B p_{ij} d_i m_{ij} \sqrt{(x_{ij} + (w_{ij}/2) - x_{CG})^2 + (y_{ij} + (h_{ij}/2) - y_{CG})^2} \right| \quad (3)$$

which is subject to

$$x_i + (w_i w_i^x) + (h_i h_i^x) \leq x_k + (1 - le_{ik}), \quad \forall i, k, i < k \quad (4)$$

$$x_k + (w_k w_k^x) + (h_k h_k^x) \leq x_i + (1 - ri_{ik}), \quad \forall i, k, i < k \quad (5)$$

$$y_i + (w_i w_i^y) + (h_i h_i^y) \leq y_k + (1 - un_{ik}), \quad \forall i, k, i < k \quad (6)$$

$$y_k + (w_k w_k^y) + (h_k h_k^y) \leq y_i + (1 - ab_{ik}), \quad \forall i, k, i < k \quad (7)$$

$$le_{ik} + ri_{ik} + un_{ik} + ab_{ik} \leq p_{ij} + p_{kj} - 1, \quad \forall i, k, i < k \quad (8)$$

$$\sum_{j=1}^C p_{ij} = 1, \quad \forall i \quad (9)$$

$$\sum_{i=1}^B p_{ij} \leq M c_j, \quad \forall j \quad (10)$$

$$x_i + (w_i w_i^x) + (h_i h_i^x) \leq W_j + (1 - p_{ij})M, \quad \forall i, j \quad (11)$$

$$y_i + (w_i w_i^y) + (h_i h_i^y) \leq H_j + (1 - p_{ij})M, \quad \forall i, j \quad (12)$$

$$w_i^x, w_i^y, h_i^x, h_i^y, le_{ik}, ri_{ik}, ab_{ik}, un_{ik}, p_{ij}, c_j \in 0, 1, \quad \forall i, k, i < k \quad (13)$$

$$x_i, y_i \geq 0, \quad \forall i \quad (14)$$

$$m_{ij} \in -1, 1, \quad \forall i \quad (15)$$

$$x_{CG} \geq (W/2) \quad (16)$$

$$y_{CG} \geq (H/2) \quad (17)$$

$B$  total number of rectangles

$C$  total number of bins

- LB* the total sum of load balancing  
 $w_i, h_i$  width and height of rectangle  $i$   
 $d_i$  weight of rectangle  $i$   
 $W_j, H_j$  width and height of bin  $j$   
 $x_i, y_i$  left-bottom corner of rectangle  $i$  as coordinate  
 $x_{CG}$   $x$  coordinate of center of gravity of bin which is equal to  $(W/2)$   
 $y_{CG}$   $y$  coordinate of center of gravity of bin which is equal to  $(H/2)$   
 $w_i^x, w_i^y$  width of rectangle  $i$  is parallel to  $X$  and  $Y$  axis  
 $h_i^x, h_i^y$  height of rectangle  $i$  is parallel to  $X$  and  $Y$  axis  
 $le_{ik}$  rectangle  $i$  is placed on the left side of rectangle  $k$   
 $ri_{ik}$  rectangle  $i$  is placed on the right side of rectangle  $k$   
 $ab_{ik}$  rectangle  $i$  is placed above rectangle  $k$   
 $un_{ik}$  rectangle  $i$  is placed under rectangle  $k$   
 $p_{ij}$   $p_{ij} = 1$  if rectangle  $i$  is placed in bin  $j$  otherwise  $p_{ij} = 0$   
 $m_{ij}$   $m_{ij} = 1$  if  $(x_{ij} + (w_{ij}/2) - x_{CG}) \geq 0$  otherwise  $m_{ij} = -1$   
 $c_j$   $c_j = 1$  if bin  $j$  is used otherwise  $c_j = 0$   
 $M$  an arbitrarily large number used in Bin- $M$  constraints

### 3 Proposed Algorithms

The chromosome is an array of values representing a possible solution to a 2DBPP. There are two parts in the chromosome. Rectangle items and a heuristics part that keeps the heuristics. Permutation encoding which is a form of keeping width-height of rectangular items and processing sequence between rectangles is used to keep the identification of rectangles. Gene (rectangle item) packing is done in two different ways. If Heu1 is equal to Heu2, then all genes are packed as a whole by using Heu1 and the result of Heu1 shows the number of required bins for solution. If Heu1 is different than Heu2, then Heu1 packs the first half of the genes and Heu2 packs the second half of the genes. The sum of required bins of Heu1 and Heu2 shows the number of required bins for solution. Elitist selection that gives higher chance to better chromosomes in the population is used in the proposed algorithms.

Single point crossover is used in our algorithms. Three different mutation operators are used in the algorithms in accordance with the orientation possibility of the items. The proposed mutation operators work on the rectangular items part of a chromosome. Swap mutation, rotation mutation (for non-oriented items), and swap-rotation mutation are the mutation operators.

*The proposed MA* mimics the natural evolution process. Our algorithms consider load balancing with center of gravity to each bin. In order to calculate center of gravity, each bin's center point is selected as Center of Gravity (CG). When a rectangle box is inserted, the Euclidean distance of its center to bins CG is calculated and multiplied by the weight of rectangle. This calculated value is CG of a rectangle. If sign of  $x$  coordinate of rectangle is minus then CG of rectangle is subtracted from total CG

of Bin. If sign of  $x$  coordinate of rectangle is plus then CG of rectangle is added to total CG of Bin. When all rectangles are inserted to bins, absolute values of total CG of bins are added. This calculated value is called as CG of a chromosome. The load balance of a bin is explained in Eq. 18 and load balance of a chromosome is explained in Eq. 19.

$$LB_{Bin} = \sum_{j=1}^{\#Rect} d_j m_{ij} \sqrt{(x_{ij} + (w_{ij}/2) - x_{CG})^2 + (y_{ij} + (h_{ij}/2) - y_{CG})^2} \quad (18)$$

$$LB_{Chromosome} = \sum_{i=1}^{\#Bin} \left| \sum_{j=1}^{\#Rect} d_j m_{ij} \sqrt{(x_{ij} + (w_{ij}/2) - x_{CG})^2 + (y_{ij} + (h_{ij}/2) - y_{CG})^2} \right| \quad (19)$$

Multiheuristic Oriented Multiobjective GA (MHO-MOGA) is proposed to solve oriented multiobjective offline 2DBPP. FNF, FFF, BFDH, LGFof and UTS are applied as heuristics on the base of a GA. Swap mutation operation is used to keep orientation of items. Multiheuristic Non-Oriented Multiobjective GA (MHNO-MOGA) is developed for optimization of non-oriented items. Each individual uses one or two of the heuristics: FNF, FFF, BFDH, UTS, LGFof and LFGi to pack rectangles into bins. At the beginning of GA, each individual picks one of the heuristics. At the next phase, an individual can have two different heuristics. Each heuristic is applied to the corresponding part of rectangle list. Rotation mutation and swap-rotation mutation are used to change orientation of rectangles. Best result of GA becomes the solution of the problem.

## 4 Experimental Setup and Results

Two well known offline 2DBPP instance sets are used for the experiments (Berkey-Wang and Martello-Vigo) [10, 11]. Experimental setup consists of UTS, LFGi and GA whose heuristics are FNF, FFF, BFDH and LGFof. First, we apply GA to the problem and later the best result's rectangle list is given as input to UTS and LFGi. The parameter setting for the population size and the number of generations are decided to be 60 and 40 respectively. These parameter settings are used in all of the proposed algorithms throughout the experiments. The general results of MHNO-MOGA experiments are listed in Tables 1 and 2. We compared our results with LFGi algorithm.

In order to analyze runtime and efficiency of the algorithms, we randomly picked five different item size (20, 40, 60, 80, 100) 2DBPP. Each test is run for five times. Best values of FNF, FFF, BFDH, LGFof and LFGi are used as results. For the proposed algorithms and UTS, we used the average values of the experiments. Comparisons of algorithms according to 500 instance test setup are also explained in detail. The

**Table 1** Result of MHNO-MOGA Ro for Berkey-Wang instances

Class	Num of rect	LGF <sub>i</sub> b	LGF <sub>i</sub> cg	Pr. algo b	Pr. algo cg
1	20	67	230.9	68	67.5
	40	131	392.6	131	218.7
	60	199	620.5	197	359.6
	80	271	877.5	270	584.6
	100	317	1123.7	320	676.6
	Av.	197	649	197.2	381.4
2	20	10	256.8	10	1.7
	40	20	530.6	20	6.4
	60	25	993.6	25	160.7
	80	32	1107.9	31	134.1
	100	39	1078.7	39	195.7
	Av.	25.2	793.5	25	99.7
3	20	50	918.7	49	241.3
	40	97	1706.4	95	643.1
	60	139	2785.7	139	1073
	80	189	3805.2	192	1709
	100	227	4422.1	229	2211.8
	Av.	140.4	2727.6	140.8	1175.6
4	20	10	1794.8	10	0.8
	40	19	3940.2	19	8
	60	26	4109.7	25	174.8
	80	33	5558.4	33	162.9
	100	39	7418.4	40	360.7
	Av.	25.4	4564.3	25.4	141.4
5	20	62	2097.4	60	555.3
	40	117	5039.1	117	2167.1
	60	180	7031	177	3292
	80	246	9068.5	242	4477
	100	290	11640	286	5573.6
	Av.	179	6975.2	176.4	3213
6	20	10	4788.4	10	8
	40	19	13578.4	19	14.6
	60	23	15069.6	22	202.6
	80	30	20176	30	999.9
	100	35	25989.9	34	1995.2
	Av.	23.4	15920.5	23	644.1

runtime analysis of FNF, FFF, BFDH, UTS, LGF<sub>of</sub> and MHO-MOGA for oriented multiobjective random picked tests are shown in Table 3. MHO-MOGA is reported to be the most time consuming algorithm.

**Table 2** Result of MHNO-MOGA Ro for Martello-Vigo instances

Class	Num of rect	LGF <sub>i</sub> b	LGF <sub>i</sub> cg	Pr. algo b	Pr. algo cg
7	20	55	1648.6	53	1145.2
	40	109	3717.4	107	2968.7
	60	161	6041	155	3869.6
	80	223	6989.4	221	6229.4
	100	271	9330.1	264	7259.2
	Av.	163.8	5545.3	160	4294.4
8	20	56	1858.4	55	873.5
	40	111	3877.2	107	2052.2
	60	163	5440.7	155	3055.2
	80	221	7640.6	214	4818.7
	100	269	8981.9	264	6412.7
	Av.	164	5559.8	159	3442.5
9	20	143	3012.3	143	2060.1
	40	275	6081.2	275	4432.6
	60	435	10005.9	435	7661.3
	80	573	13963.7	573	9919.7
	100	693	16765.3	693	12324
	Av.	423.8	9965.7	423.8	7279.5
10	20	41	3800.9	43	396.2
	40	75	6710.7	75	732.4
	60	104	9447.3	105	1885.1
	80	133	12136.1	134	2937.9
	100	163	15835.2	165	4615.5
	Av.	103.2	9586	104.4	2113.4

**Table 3** Runtime of algorithms for oriented multiobjective problems in msec

Rect	FNF	FFF	BFDH	UTS	LGF <sub>of</sub>	MHO-MOGA
20	4.1	4.6	4.7	81.7	64.0	17515
40	4.7	4.8	5.0	317.8	29.9	88024
60	5.3	6.4	7.8	27761.4	1529.8	3034228
80	7.6	8.4	9.2	3050.4	334.6	4664450
100	10.6	12.4	17.4	3746.9	291.1	677750
Av.	6.5	7.3	8.8	6991.6	449.9	1696393.4

The general results of second experiment are listed in Tables 4 and 5. We compared our results with LGF<sub>i</sub> algorithm.

Result (bin/cg) analysis of FNF, FFF, BFDH, UTS, LGF<sub>of</sub> and MHO-MOGA for oriented multiobjective random picked tests are shown in Table 6.

**Table 4** Result of MHNO-MOGA SwRo for Berkey-Wang instances

Class	Num of rect	LGF <sub>i</sub> b	LGF <sub>i</sub> cg	Pr. algo b	Pr. algo cg
1	20	67	230.9	66	76.9
	40	131	392.6	125	202.8
	60	199	620.5	200	359.2
	80	271	877.5	264	630.7
	100	317	1123.7	320	711.7
	Av.	197	649	195	396.3
2	20	10	256.8	10	1
	40	20	530.6	20	8.7
	60	25	993.6	25	133
	80	32	1107.9	31	131.8
	100	39	1078.7	39	245.5
	Av.	25.2	793.5	25	104
3	20	50	918.7	49	260.3
	40	97	1706.4	97	581.1
	60	139	2785.7	141	1165.3
	80	189	3805.2	190	1775.7
	100	227	4422.1	228	2335.5
	Av.	140.4	2727.6	141	1223.6
4	20	10	1794.8	10	0.5
	40	19	3940.2	19	22.6
	60	26	4109.7	25	97.1
	80	33	5558.4	33	346.4
	100	39	7418.4	39	633.6
	Av.	25.4	4564.3	25.2	220
5	20	62	2097.4	60	719.4
	40	117	5039.1	116	2027
	60	180	7031	177	3444.2
	80	246	9068.5	245	5043.8
	100	290	11640	286	5154.2
	Av.	179	6975.2	176.8	3277.7
6	20	10	4788.4	10	4.9
	40	19	13578.4	19	51.3
	60	23	15069.6	22	112.3
	80	30	20176	30	502
	100	35	25989.9	34	1356.8
	Av.	23.4	15920.5	23	405.5



**Table 5** Result of MHNO-MOGA SwRo for Martello-Vigo instances

Class	Num of rect	LGF <sub>i</sub> b	LGF <sub>i</sub> cg	Pr. algo b	Pr. algo cg
7	20	55	1648.6	53	1015.6
	40	109	3717.4	108	2612.5
	60	161	6041	156	3880.6
	80	223	6989.4	220	6126.4
	100	271	9330.1	267	6899.4
	Av.	163.8	5545.3	160.8	4106.9
8	20	56	1858.4	53	785.3
	40	111	3877.2	107	2170.4
	60	163	5440.7	155	3044
	80	221	7640.6	214	4521.9
	100	269	8981.9	265	6237.6
	Av.	164	5559.8	158.8	3351.8
9	20	143	3012.3	143	1885.9
	40	275	6081.2	275	4241.6
	60	435	10005.9	435	6762.7
	80	573	13963.7	573	9355.1
	100	693	16765.3	693	11598.4
	Av.	423.8	9965.7	423.8	6768.7
10	20	41	3800.9	42	263.2
	40	75	6710.7	75	971
	60	104	9447.3	102	2616.8
	80	133	12136.1	133	3837.2
	100	163	15835.2	166	4662.4
	Av.	103.2	9586	103.6	2470.1

Results (bin/cg) of FNF, FFF, BFDH, UTS, LGF<sub>of</sub> and MHO-MOGA for oriented single objective 500 problem set are shown in Table 7.

Superiority of MHO-MOGA to FNF, FFF, BFDH, UTS and LGF<sub>of</sub> for oriented multiobjective 500 problem set are shown in Table 8.

Runtime analysis of FNF, FFF, BFDH, UTS, LGF<sub>of</sub>, LGF<sub>i</sub> and MHNO-MOGA for non-oriented multiobjective random picked tests are shown in Table 9. MHNO-MOGA is reported to be the most time consuming algorithm.

Result (bin/cg) analysis of FNF, FFF, BFDH, UTS, LGF<sub>of</sub>, LGF<sub>i</sub> and MHNO-MOGA for non-oriented multiobjective random picked tests are shown in Table 10.

Results (bin/cg) of FNF, FFF, BFDH, UTS, LGF<sub>of</sub>, LGF<sub>i</sub> and MHNO-MOGA (r) for non-oriented multiobjective 500 problem set (according to continuous lower bound) are shown in Table 11.

Superiority of MHO-MOGA (r) versus FNF, FFF, BFDH, UTS, LGF<sub>of</sub> and LGF<sub>i</sub> for non-oriented multiobjective 500 problem set are shown in Table 12.

**Table 6** Results (bin/cg) of algorithms for oriented multiobjective problems

Rect	FNF	FFF	BFDH	UTS	LGFof	MHO-MOGA
20	11	9	9	9	9	9
	310.4	167.3	151.3	491.1	301.4	116.6
40	17	13	13	13	12	12
	62.4	59.8	50	64.5	33.2	19.6
60	53	47	47	47	47	46
	1207.5	1024.7	1008.5	1103.9	1021.1	813.7
80	30	24	23	24	24	24
	1704.5	1850.1	1440.7	900.9	1728.7	421
100	44	31	31	30	30	29
	1623.9	1391.2	1098.2	1088.9	1331.2	769.9
Av.	31	24.8	24.6	24.6	24.6	24
	981.5	898.6	749.7	729.9	883.1	428.2

**Table 7** Results (bin/cg) of heuristics and MHO-MOGA for oriented multiobjective 500 problem set

Total	FNF	FFF	BFDH	UTS	LGFof	Pr. alg.
Bin	9489	7591	7514	7521	7430	7389
CG	340,954	333,247	319,187	297,483	347,076	126,581

**Table 8** MHO-MOGA versus heuristics for oriented multiobjective 500 problem set

	FNF	FFF	BFDH	UTS	LGFof
MHO-MOGA (%)	100	100	100	100	97.2

**Table 9** Runtime of algorithms for non-oriented multiobjective problems in msec

Rect	FNF	FFF	BFDH	UTS	LGFof	LGF <sub>i</sub>	MHNO-MOGA
20	4.1	4.6	4.7	81.7	64.0	65.0	87,615
40	4.7	4.8	5.0	317.8	29.9	30.1	39,027
60	5.3	6.4	7.8	27761.4	1529.8	2416.4	1545,971
80	7.6	8.4	9.2	3050.4	334.6	257.4	645,442
100	10.6	12.4	17.4	3746.9	291.1	324.3	257,878
Av.	6.5	7.3	8.8	6991.6	449.9	618.6	515186.6

**Table 10** Results (bin/cg) of algorithms for non-oriented multiobjective problems

Rect	FNF	FFF	BFDH	UTS	LGFof	LGF <sub>i</sub>	MHNO-MOGA
20	11	9	9	9	9	8	8
	310.4	167.3	151.3	491.1	301.4	150.9	60
40	17	13	13	13	12	12	12
	62.4	59.8	50	64.5	33.2	51.9	15
60	53	47	47	47	47	46	46
	1207.5	1024.7	1008.5	1103.9	1021.1	901.4	700.2
80	30	24	23	24	24	22	21
	1704.5	1850.1	1440.7	900.9	1728.7	887.4	821.3
100	44	31	31	30	30	29	28
	1623.9	1391.2	1098.2	1088.9	1331.2	1131.2	575
Av.	31	24.8	24.6	24.6	24.6	23.4	23
	981.5	898.6	749.7	729.9	883.1	624.6	434.3

**Table 11** Results (bin/cg) of heuristics and MHNO-MOGA (r) for non-oriented multiobjective 500 problem set

Total	FNF	FFF	BFDH	UTS	LGFof	LGF <sub>i</sub>	Pr. Alg.
Bin	9489	7591	7514	7521	7430	7226	7175
CG	340,954	333,247	319,187	297,483	347,076	311,434	113,925

**Table 12** MHO-MOGA(r) versus heuristics for non-oriented multiobjective 500 problem set

	FNF	FFF	BFDH	UTS	LGFof	LGF <sub>i</sub>
MHNO-MOGA(r) (%)	100	100	100	100	100	95

**Table 13** Results (bin/cg) of heuristics and MHNO-MOGA(sr) for non-oriented multiobjective 500 problem set

Total	FNF	FFF	BFDH	UTS	LGFof	LGF <sub>i</sub>	Pr. Alg.
Bin	9489	7591	7514	7521	7430	7226	7165
CG	340,954	333,247	319,187	297,483	347,076	311,434	111,623

Results (bin/cg) of FNF, FFF, BFDH, UTS, LGFof, LGF<sub>i</sub> and MHNO-MOGA (sr) for non-oriented multiobjective 500 problem set (according to continuous lower bound) are shown in Tables 11 and 13.

Superiority of MHNO-MOGA (sr) versus FNF, FFF, BFDH, UTS, LGFof and LGF<sub>i</sub> for non-oriented multiobjective 500 problem set are shown in Table 14.

**Table 14** MHO-MOGA (sr) versus heuristics for non-oriented multiobjective 500 problem set

	FNF	FFF	BFDH	UTS	LGFof	LGF <sub>i</sub>
MHNO-MOGA(sr) (%)	100	100	100	100	100	96

## 5 Conclusions and Future Work

In this study, two novel robust algorithms for the multiobjective optimization of offline 2DBPP are proposed. Our experimental results show that while the well known heuristics sometimes produce better results for minimizing the number of bins in 2DBPP they are not efficient for solving the multiobjective load balancing problem of 2D bins while also minimizing bins. MHO-MOGA and MHNO-MOGA give better results not only for minimum number of bins but also for the load balancing of 2D bins. MHNO-MOGA makes use of rotation and swap-rotation mutation operators. MHNO-MOGA with rotation mutation outperforms LGF<sub>i</sub> heuristic for 95.0 % of the problems. MHNO-MOGA with swap-rotation mutation outperforms LGF<sub>i</sub> heuristic for 96.0 % of the problems and 97.2 % of the benchmark problems for the LGF<sub>of</sub> heuristic.

## References

1. Johnson, D.S.: Near-optimal bin-packing algorithms. Ph.D. thesis (1973)
2. Dokeroglu, T., Cosar, A.: Optimization of one-dimensional bin packing problem with island parallel grouping genetic algorithms. *Comput. Ind. Eng.* **75**, 176–186 (2014)
3. Coffman, E.G., Shor, P.W.: Average-case analysis of cutting and packing in two dimensions. *Eur. J. Oper. Res.* **44**, 134–144 (1990)
4. Dokeroglu, T.: Hybrid teaching-learning-based optimization algorithms for the Quadratic Assignment Problem. *Comput. Ind. Eng.* **85**, 86–101 (2015)
5. Lodi, A., Martello, S., Monaci, M.: Two-dimensional packing problems: A survey. *Eur. J. Oper. Res.* **141**(2), 241–252 (2002)
6. Tosun, U., Dokeroglu, T., Cosar, A.: A robust island parallel genetic algorithm for the quadratic assignment problem. *Int. J. Prod. Res.* **51**(14), 4117–4133 (2013)
7. Blum, C., Schmid, V.: Solving the 2d bin packing problem by means of a hybrid evolutionary algorithm. *Procedia Comput. Sci.* **18**(0), 899–908 (2013). International Conference on Computational Science (2013)
8. Fernandez, A., Gil, C., Banos, R., Montoya, M.G.: A parallel multiobjective algorithm for two-dimensional bin packing with rotations and load balancing. *Expert Syst. Appl.* **40**(13), 5169–5180 (2013)
9. Thapatsawan, P., et al.: Development of a stochastic optimisation tool for solving the multiple container packing problems. *Int. J. Prod. Econ.* **140**(2), 737–748 (2012)
10. Smith, J.E.: Coevolving memetic algorithms: A review and progress report. *Syst., Man, Cybern., Part B: Cybern.* **37**(1), 6–17 (2007)
11. Berkey, J.O., Wang, P.Y.: Two-dimensional finite bin-packing algorithms. *J. Oper. Res. Soc.* **38**, 423–429 (1987)

**Part V**  
**Stochastic Modelling and Computer**  
**Networks**

# Smoothing the Input Process in a Batch Queue

F. Ait Salaht, H. Castel Taleb, J.M. Fourneau, T. Mautor and N. Pekergin

**Abstract** We state some stochastic comparison results for the loss probabilities and the time to live in a tandem network with two queues. We consider such a network under batch arrivals and constant service time. We show that making the service capacity of the first queue finite will improve the performance. More precisely, we derive stochastic bounds for the accumulated number of losses without any assumptions on the input traffic.

## 1 Introduction

We are interested in tandem networks with finite buffers and fixed capacity of service receiving batches of customer. For such systems, it is generally assumed that reducing the variability of the batch of arrivals while conserving the average will decrease the losses in the network. Here we prove such a result in a strong way using stochastic comparison and sample-paths for the queues. We refer to [7] for theoretical issues of the stochastic comparison method. The system we address is depicted in Fig. 1.

More precisely, we add an input buffer at the entrance (i.e. before the queue) and we study how this extra buffer smoothens the input process and helps to reduce the losses in the network (see Fig. 2). Such an extra queue at the entrance has strong similarities with a leaky bucket. The model is in discrete time. The service capacity is  $S_2$  packets per slot.  $H(t)$  will denote the arrivals during time  $t$ . This extra queue adds at least a delay of one time slot. For the sake of readability we prefer studying the two following systems: the first one (i.e. model 1 in the following), which contains an input queue with a infinite buffer and a finite service capacity  $S_1$  and the second one

---

F.A. Salaht · H.C. Taleb  
SAMOVAR, CNRS Télécom Sud Paris, Évry Cedex, France

J.M. Fourneau (✉) · T. Mautor  
PRiSM, UVSQ, CNRS, Versailles, France  
e-mail: Jmf@prism.uvsq.fr

N. Pekergin  
LACL, Université de Paris Est Créteil, Créteil, France

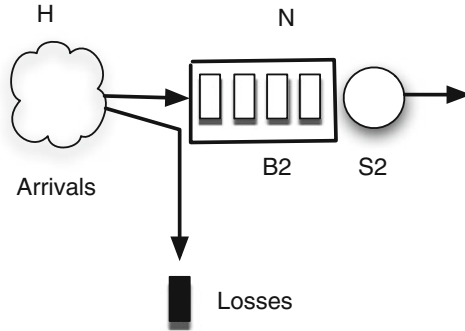


Fig. 1 A batch queue

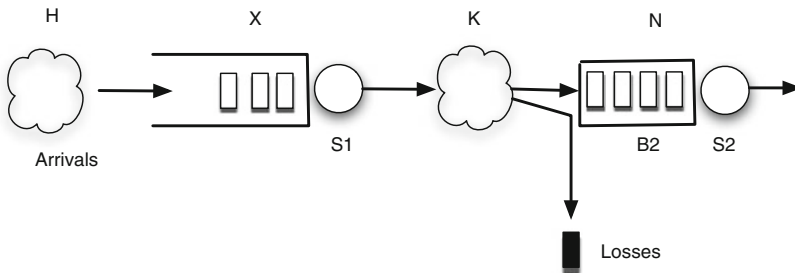


Fig. 2 Description of the network. Note that the cloud icon represents the batch moving from queue 1 to queue 2, it is not an extra buffer

(i.e. model 2) with the same topology but where the service capacity  $S1$  is infinite. Note that this only requires that  $S1 \geq \max(H(t))$  for all  $t$ . Note that in model 2 the batches of arrivals are never modified by the queue. The arrivals are just delayed by one time slot. Thus model 2 is equivalent to the isolated queue we have represented in Fig. 1.

We assume that  $S1 > E(H(t))$  for all  $t$ . Such an extra queue does not change the expectation of the size of batches entering queue 2 but, under some simple conditions, it makes this process less variable. We also assume that  $S1 > S2$  because this is the only configuration where we can lose customers at the entrance of queue 2. Note that as the proof is based on the sample-paths comparisons and structural properties of the queues (Work Conserving, Tail Drop), we do not assume any property for the sequence of arrivals. We only require that the sequence does not depend on the state of the system.

Note that even if the result seems rather intuitive, the proof is rather complex and, to the best of our knowledge, is original for batch queues. Such a result is already known for point arrival processes [2]. Smoothing the input process is also the aim for leaky bucket mechanism, which has been extensively studied for ATM networks (see for instance [3]). Note that our result is established for a very general assumptions about the arrivals. This result is the first step to provide stochastic bounds

for tandem queueing networks with batches of customers and generalize the approach introduced in [1] for a single queue with batch arrivals and constant service. The aim is to improve the accuracy of the approximate analysis given in [6]. The paper is organized as follows. Section 2 is devoted to a presentation of the two models. In Sect. 3, we present the proof to compare the sample-paths.

## 2 Models

First, we have to be more precise about the synchronizations between the queues and the buffers as the models are in discrete time. We assume that the customers leaving the first queue at time  $t$  enter the intermediate buffer at time  $t + 1$ . In a queue, the arrivals take place before the departures. Let us now introduce some notation. Let  $t$  be an arbitrary discrete time instant and let  $H(0) \dots H(t - 1)$  be an arbitrary sequence of arrivals between 0 and  $t - 1$ . For all time  $t$  and sequence  $H$  we define for the model with finite  $S1$ :

- $\mathbf{X}^H(t)$  is the number of customers in the extra queue (i.e. queue 1),
- $\mathbf{K}_F^H(t)$  is the size of the batch between queue 1 and queue 2,
- $\mathbf{N}_F^H(t)$  is the size of queue 2,
- $\mathbf{SP}_F^H(t)$  is the number of losses in queue 2 from 0 to  $t$ ,
- $\mathbf{SS}_F^H(t)$  is the number of customers which exits station 2 from 0 to  $t$ .

Similarly,  $\mathbf{K}_I^H(t)$ ,  $\mathbf{N}_I^H(t)$ ,  $\mathbf{SP}_I^H(t)$  and  $\mathbf{SS}_I^H(t)$  will represent the same quantities when the first queue has an infinite service capacity  $S1$ . Both dynamics can be described by simple equations. The assumptions about the arrivals before the departure can be observed in the equation for  $\mathbf{N}_F^H(t + 1)$ . As usual  $x^+ = \max(x, 0)$ .

Finite service capacity input queue:

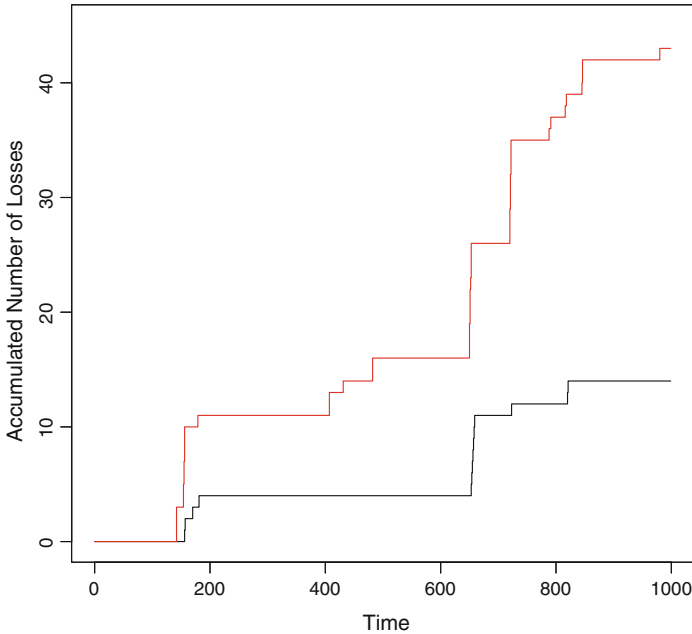
$$\begin{aligned} \mathbf{X}^H(t + 1) &= (\mathbf{X}^H(t) + H(t) - S1)^+, \\ \mathbf{K}_F^H(t + 1) &= \min(S1, \mathbf{X}^H(t) + H(t)), \\ \mathbf{N}_F^H(t + 1) &= \min(B2, (\mathbf{N}_F^H(t) + \mathbf{K}_F^H(t) - S2)^+, \\ \mathbf{SP}_F^H(t + 1) &= \mathbf{SP}_F^H(t) + (\mathbf{N}_F^H(t) + \mathbf{K}_F^H(t) - S2 - B2)^+, \\ \mathbf{SS}_F^H(t + 1) &= \mathbf{SS}_F^H(t) + \min(S2, \mathbf{N}_F^H(t) + \mathbf{K}_F^H(t)). \end{aligned}$$

If  $S1$  is infinite the first queue is always empty at the end of the time slot but it requires one time slot to cross queue 1. Thus, we have:  $\mathbf{K}_I^H(t + 1) = H(t)$ .

Infinite service capacity input queue:

$$\begin{aligned} \mathbf{K}_I^H(t + 1) &= H(t), \\ \mathbf{N}_I^H(t + 1) &= \min(B2, (\mathbf{N}_I^H(t) + \mathbf{K}_I^H(t) - S2)^+, \\ \mathbf{SP}_I^H(t + 1) &= \mathbf{SP}_I^H(t) + (\mathbf{N}_I^H(t) + \mathbf{K}_I^H(t) - S2 - B2)^+, \\ \mathbf{SS}_I^H(t + 1) &= \mathbf{SS}_I^H(t) + \min(S2, \mathbf{N}_I^H(t) + \mathbf{K}_I^H(t)). \end{aligned}$$





**Fig. 3** Two sample-paths for the cumulated number of losses. The system with finite value for  $S1$  is in *black* while the system with infinite capacity input queue is drawn in *red*

We do not make any assumption about traffic  $H$ . We only need that both queues receive the same traffic for each time period. Thus, the traffic is not dependent of the state of the queue.

We also assume that all the queues and the buffers are empty at  $t = 0$ . Before proceeding with the proof, we present in Fig. 3 a sample path for both systems obtained by the simulation engine in the Xborne project [4].

Clearly, as one may expect, the accumulated number of losses since time 0 for the input queue with a finite capacity server is always smaller than this number for an input queue with an infinite capacity server. Thus, the input buffer with a finite capacity server seems to make the input process easier to manage for the second queue. We want to prove this sample-path property.

However the sample paths show that during a short amount of time, the numbers of losses can be larger for the system with infinite capacity as it can be seen in Fig. 4 between time instants 655 and 660. The losses in the system with finite capacity server are delayed because of the queue. Thus they may happen. The sample-paths never cross but the difference between the numbers of losses in both systems is not monotone. This is the intuition behind the use of sample-path arguments using the accumulated number of losses since time 0.

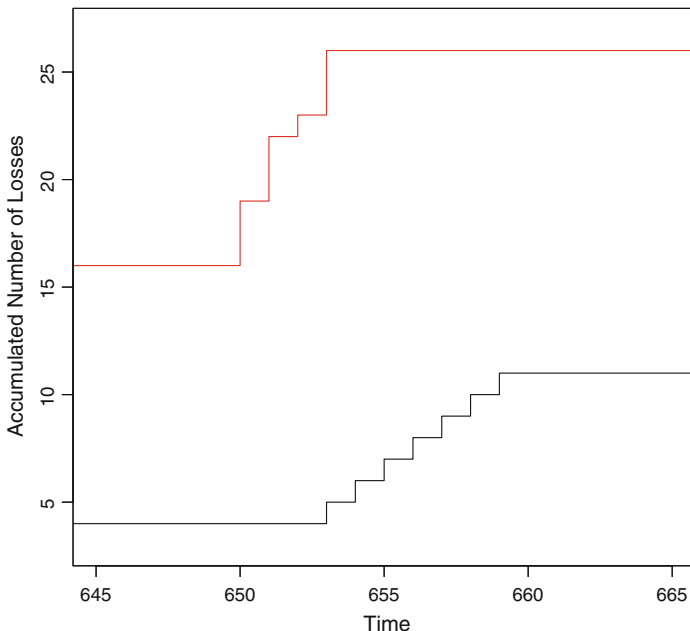


Fig. 4 Details of the previous sample-paths between time instants 645 and 665

### 3 Sample-Path Comparison

We begin with three technical lemmas.

**Lemma 1** (Conservation Law) *for all  $t$ , we have:*

$$\begin{aligned}
 & \mathbf{X}^H(t) + \mathbf{K}_F^H(t) + \mathbf{N}_F^H(t) + \mathbf{SP}_F^H(t) + \mathbf{SS}_F^H(t) \\
 &= \sum_{i=0}^{t-1} H(t) \\
 &= \mathbf{K}_I^H(t) + \mathbf{N}_I^H(t) + \mathbf{SP}_I^H(t) + \mathbf{SS}_I^H(t).
 \end{aligned}$$

*Proof* In both models, for all time instant  $t$  the total numbers of customers which enter the network are equal to  $\sum_{i=0}^{t-1} H(t)$ . Furthermore the first quantity represents the decomposition of this number of customers: some are still in the first queue, or the buffer, or the second queue or they have been lost, or they have left the network after being served at queue 2. Similarly, the last expression describes the same decomposition for the second model. □

**Lemma 2** (Delay at queue 1) *for all  $t$ ,  $\mathbf{X}^H(t) + \mathbf{K}_F^H(t) \geq \mathbf{K}_I^H(t)$ .*

*Proof* Clearly, as we have:  $\mathbf{X}^H(t) + \mathbf{K}_F^H(t) = H(t-1) + \mathbf{X}^H(t-1)$  and  $\mathbf{K}_I^H(t) = H(t-1)$ . □

**Lemma 3** *Let  $H$  be an arbitrary sequence of arrivals, let  $t_0$  be an arbitrary instant and let  $G$  be another sequence built as follows:*

$$\forall t \neq t_0, \quad G(t) = H(t) \text{ and } G(t_0) = H(t_0) + 1,$$

then  $\mathbf{SP}_F^H(t) \leq \mathbf{SP}_F^G(t) \leq \mathbf{SP}_F^H(t) + 1$  for all  $t$ .

*Proof* By inspection of the values of  $t$  compared to  $t_0$

- $t < t_0$ . Before  $t_0$  the arrival sequences are equal and the initial states are also equal. Therefore in both models, the same customers are lost due to buffer overflow. Thus,  $\mathbf{SP}_F^H(t) = \mathbf{SP}_F^G(t)$  for all  $t < t_0$ .
- $t = t_0$ . The extra customer enters the queue. Without loss of generality, we assume that it enters at the end of a batch and it joins the last occupied position in the queue. It leaves the queue at time  $t_1 > t_0$ . From  $t_0$  to  $t_1 - 1$ , the sample paths generated by sequence  $H$  and  $G$  are the same. Furthermore, the inputs of queue 2 for sequences  $H$  and  $G$  are also equal for this time period.

$$\mathbf{K}_F^G(t) = \mathbf{K}_F^H(t), \quad \forall t \in t_0, \dots, t_1 - 1.$$

Similarly,

$$\mathbf{N}_F^G(t) = \mathbf{N}_F^H(t) \text{ and } \mathbf{SP}_F^H(t) = \mathbf{SP}_F^G(t), \quad \forall t \in t_0, \dots, t_1.$$

- After  $t_1$ , the output sequence out of queue 1 associated with input sequence  $H$  is composed of  $l$  batches of size  $S1$  followed by a batch of size  $f < S1$ .  $l$  and  $f$  can be 0. For input sequence  $G$ , the output sequence begins with the same  $l$  batches of size  $S1$  but the last batch has size  $f + 1 \leq S1$ . As for both sequences, the sizes of queue 2 are equal at time  $t_1$  and we observe the same sequence of inputs at queue 2 up to time  $t_1 + l$ , we have the same number of losses:  $\mathbf{SP}_F^H(t) = \mathbf{SP}_F^G(t)$ ,  $\forall t \in t_1 + 1, \dots, t_1 + l$
- At time instant  $t_1 + l + 1$  we have one extra arrival at queue 2. We must consider two cases according to the acceptance or the rejection by the queue.
  - Rejection: we lose one customer more with sequence  $G$  than with sequence  $H$ . But, after this loss, both queues  $\mathbf{N}_F^H$  et  $\mathbf{N}_F^G$  are equal and the following sequences of inputs are also equal. Thus,

$$\mathbf{SP}_F^G(t) = \mathbf{SP}_F^H(t) + 1, \quad \forall t > t_1 + l.$$

- Acceptation: thus we have:  $\mathbf{N}_F^G(t + l + 1) = \mathbf{N}_F^H(t + l + 1) + 1$ . And the arrival sequence are equal in the future. We still have two cases to consider according to the first event which occurs: loss or empty queue.

If with sequence  $G$ , a loss occurs before queue 2 becomes empty, then  $\mathbf{N}_F^G$  and  $\mathbf{N}_F^H$  will be equal in the remaining part of the sample-paths. Let  $t_2$  be this time instant. We have:

$$\mathbf{SP}_F^G(t) = \mathbf{SP}_F^H(t), \quad t_1 < t \leq t_2 - 1, \quad \text{and} \quad \mathbf{SP}_F^G(t) = \mathbf{SP}_F^H(t) + 1, \quad \forall t \geq t_2.$$

Otherwise, queue 2 becomes empty with input sequence  $G$  before the first loss. Thus, it will also be empty for sequence  $H$ . Let  $t_3$  this time instant. As there is not loss between  $t_1 + l$  and  $t_3$  we have:

$$\mathbf{SP}_F^G(t) = \mathbf{SP}_F^H(t), \quad t_1 + 1 + l \leq t \leq t_3.$$

As the arrivals will be the same since  $t_0 + 1$ , we have for all  $t \geq t_3$ :

$$\mathbf{N}_F^G(t) = \mathbf{N}_F^H(t), \quad \text{et} \quad \mathbf{SP}_F^G(t) = \mathbf{SP}_F^H(t).$$

And the proof is complete.  $\square$

**Lemma 4** *For any sequence of arrivals  $H$  between 0 and  $T - 1$ , if  $\mathbf{SP}_I^H(t) = 0$ , then  $\mathbf{SP}_F^H(t) = 0$  for all  $t, 0 \leq t \leq T$ .*

*Proof* By induction on  $t$ , we state that for all time instant  $t$  before the first loss in model with queue 1, we have:

$$\mathbf{SP}_I^H(t) = \mathbf{SP}_F^H(t) = 0, \tag{1}$$

$$\mathbf{SS}_I^H(t) = \mathbf{SS}_F^H(t), \tag{2}$$

$$\mathbf{X}^H(t) + \mathbf{K}_F^H(t) + \mathbf{N}_F^H(t) = \mathbf{K}_I^H(t) + \mathbf{N}_I^H(t). \tag{3}$$

Note that Eq. 3 is a simple consequence of Eqs. 1 and 2 and Lemma 1. Furthermore if Eq. 3 holds, we have, due to Lemma 2,  $\mathbf{N}_F^H(t) \leq \mathbf{N}_I^H(t)$  for all time instant  $t$  smaller than the instant with the first loss in the second model.

Clearly, Eqs. 1–3 holds for  $t = 0$ . Suppose now they also holds at an arbitrary time  $t < T$ . According to 3, and as  $\mathbf{X}^H(t) \geq 0$ , we get:

$$\mathbf{K}_F^H(t) + \mathbf{N}_F^H(t) \leq \mathbf{K}_I^H(t) + \mathbf{N}_I^H(t). \tag{4}$$

Subtracting S2 to both sides, we obtain:

$$\mathbf{K}_F^H(t) + \mathbf{N}_F^H(t) - S2 \leq \mathbf{K}_I^H(t) + \mathbf{N}_I^H(t) - S2. \tag{5}$$

Because of the assumption on the time instant, we have  $\mathbf{K}_I^H(t) + \mathbf{N}_I^H(t) - S2 \leq B2$ . By transitivity, we obtain:  $\mathbf{K}_F^H(t) + \mathbf{N}_F^H(t) \leq B2$  and we do not lose packet at time  $t$  in the first model. As  $\mathbf{SP}_I^H(t + 1) = 0$ , Eq. 1 holds at time  $t + 1$ .

We now prove that Eq. 2 also holds at time  $t + 1$ . Remember that  $\mathbf{K}_I^H(t) = H(t - 1)$ .

- If  $H(t - 1) \geq S1$ ,  $\mathbf{K}_F^H(t) = S1$  as queue 1 is work conserving. Thus  $S2$  packets are leaving queue 2 in both models as  $S1$  packets arrive in that queue and  $S1 > S2$ . By induction, we get  $\mathbf{SS}_I^H(t) = \mathbf{SS}_F^H(t)$ , and finally:

$$\mathbf{SS}_I^H(t + 1) = S2 + \mathbf{SS}_I^H(t) = S2 + \mathbf{SS}_F^H(t) = \mathbf{SS}_F^H(t + 1)$$

- Otherwise  $H(t - 1) < S1$  and we still have two cases:
  - If  $\mathbf{K}_F^H(t) < S1$ , then  $\mathbf{X}^H(t) = 0$ , as the queue is work conserving. Thus,  $\mathbf{K}_F^H(t) + \mathbf{N}_F^H(t) = \mathbf{K}_I^H(t) + \mathbf{N}_I^H(t)$  And:

$$\min(S2, \mathbf{K}_F^H(t) + \mathbf{N}_F^H(t)) = \min(S2, \mathbf{K}_I^H(t) + \mathbf{N}_I^H(t)). \quad (6)$$

By induction  $\mathbf{SS}_I^H(t) = \mathbf{SS}_F^H(t)$ . And the previous equation states that both models have the same outputs at time instant  $t + 1$ . Thus,

$$\mathbf{SS}_I^H(t + 1) = \mathbf{SS}_F^H(t + 1).$$

- Otherwise  $\mathbf{K}_F^H(t) = S1$ . In model 1, the output queue will output  $S2$  packets because  $S1$  packets enter the queue and  $S1 > S2$ . In model 2,  $\mathbf{K}_I^H(t) < S1$  packets enter the queue. Equation 3 still holds. As  $\mathbf{K}_F^H(t) = S1$ , we have  $\mathbf{K}_I^H(t) + \mathbf{N}_I^H(t) \geq S1 > S2$ . Thus queue 2 in model 2 also outputs  $S2$  packets. With the same argument as in the previous case, we get:

$$\mathbf{SS}_I^H(t + 1) = \mathbf{SS}_F^H(t + 1).$$

As Eqs. 1 and 2 are satisfied at  $t + 1$ , Eq. 3 also holds.  $\square$

**Theorem 1** *For all  $t$ , we have:  $\mathbf{SP}_F^H(t) \leq \mathbf{SP}_I^H(t)$ . The sample-path of the number of losses is always smaller when the first queue has a finite capacity which is sufficient for stability.*

*Proof* We consider an arbitrary sequence  $H(0), \dots, H(t - 1)$ . The proof is based on the four following steps.

- Step 1 : We build the sample-path for model 2 (i.e. with infinite  $S1$ ) with input sequence  $H$ . We get  $\mathbf{SP}_I^H(t) = g$ . If  $g = 0$ , then according to Lemma 4, we have  $\mathbf{SP}_F^H(t) = 0$  and the proof is complete. We now suppose that  $g > 0$  and we mark all the packets lost during this sample-path.
- Step 2 : We remove from sequence  $H$  all the packets which have been lost during Step 1. We obtain a new sequence on inputs, called  $G$ . This sequence is a real sequence of inputs. Indeed, in model 2 (i.e. with infinite  $S1$ ), when a packet is lost, all the packets after it in the same batch of fresh arrivals are also lost. Thus the rejection in model 2 removes the tail of some batches. The remaining part of the

batches (i.e. the heads until the first loss for each batch) is a valid input sequence for both models.

- Step 3: We build the sample-paths for both models with input sequence  $G$ . Clearly, we have:

$$\mathbf{N}_I^G(t) = \mathbf{N}_I^H(t), \quad \mathbf{SS}_I^G(t) = \mathbf{SS}_I^H(t), \quad \mathbf{SP}_I^G(t) = 0.$$

And Lemma 4 implies that  $\mathbf{SP}_F^G(t) = 0$ .

- Step 4: We now add the  $g$  packets lost during Step 1 to sequence  $G$  to obtain sequence  $H$ . The packets are added one by one. After each modification, we can apply Lemma 3. Finally:

$$\mathbf{SP}_F^H(t) \leq \mathbf{SP}_F^G(t) + g \leq 0 + \mathbf{SP}_I^H(t) = \mathbf{SP}_I^H(t).$$

And the proof is complete.  $\square$

To define the strong stochastic ordering ( $\leq_{st}$ ), we use the following theorem which gives several characterizations (see [7] for more details).

**Theorem 2** *Let  $X$  and  $Y$  be two random variables defined on a totally ordered space.  $X \leq_{st} Y$  denotes that  $X$  is less or equal to  $Y$  in the sense of the  $\leq_{st}$  order and  $X =_{st} Y$  denotes that they have the same distribution ( $F_X(a) = F_Y(a)$ ,  $\forall a$ ). The following statements are equivalent:*

1.  $X \leq_{st} Y$ ,
2.  $\forall a, F_X(a) \geq F_Y(a)$ ,
3.  $E[f(X)] \leq E[f(Y)]$ , for all increasing (non decreasing functions  $f$ , whenever the expectations exist,
4. there exist two random variables such that  $\widehat{X} =_{st} X$ ,  $\widehat{Y} =_{st} Y$ , and  $\text{Prob}(\widehat{X} \leq \widehat{Y}) = 1$ .

**Corollary 1** *For all  $t$ , the number of losses in model 1 with finite input queue is less or equal in the sense of the  $\leq_{st}$  order than the the number of losses in model 2 with infinite input queue:*

$$SP_F^H(t) \leq_{st} SP_I^H(t), \quad \forall t > 0.$$

*Proof* It follows from the sample-path property of the  $\leq_{st}$  order given in the theorem (last statement).  $\square$

## 4 Conclusion

We now try to adapt our results to tandem networks with finite queues and prove some comparison results when we reorder the queue in the tandem. Our aim is to generalize the well-known result by Friedman [5] about the interchangeability of queues in a tandem. Indeed, for infinite buffers, the end to end delay in the tandem does not depend of the order of the queue in the network.

## References

1. Aït-Salaht, F., Castel-Taleb, H., Fourneau, J., Pekergin, N.: Stochastic bounds and histograms for network performance analysis. In: *Computer Performance Engineering—10th European Workshop, EPEW Italy*, volume 8168 of *Lecture Notes in Computer Science*, pp. 13–27. Springer, New York (2013)
2. Chang, C.-S.: Smoothing point processes as a means to increase throughput. *Oper. Res.* **43**(1), 117–129 (1995)
3. Dupuis, A., Hébuterne, G.: Dimensioning the continuous state leaky bucket for geometric arrivals. In: *Data Communications and their Performance, Proceedings of the Sixth IFIP WG6.3 Conference on Performance of Computer Networks, Istanbul, Turkey, (1995)*, volume 36 of *IFIP Conference Proceedings*, pp. 289–301. Chapman and Hall (1995)
4. Fourneau, J.-M., Le Coz, M., Pekergin, N., Quessette, F.: An open tool to compute stochastic bounds on steady-state distributions and rewards. In: *11th International Workshop on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS: Orlando)*. IEEE Computer Society, FL (2003)
5. Friedman, H.D.: Reduction methods for tandem queuing systems. *Oper. Res.* **13**(1), 121–131 (1965)
6. Hernández-Orallo, E., Vila-Carbó, J.: Network queue and loss analysis using histogram-based traffic models. *Comput. Commun.* **33**(2), 190–201 (2010)
7. Muller, A., Stoyan, D.: *Comparison Methods for Stochastic Models and Risks*. Wiley, New York (2002)

# Network-Based Job Dispatching in the Cloud

Byungseok Kang

**Abstract** We describe and evaluate an Inter Cloud Manager (ICM) that dispatches jobs in Cloud environment. ICM consists of two main parts, Cloud monitoring and decision making. For Cloud monitoring, ICM use Cloud Probes (CPs) that observe and collect data, and decision making is based on both the measured execution time and network delay in forwarding the jobs and receiving back the result of the execution. ICM checks each Cloud hosts current number of waiting jobs and average execution time, and also makes use of delay and loss information regarding the network. Measurements are used to compare ICM with a Round Robin (RR) allocation of jobs between Clouds which spreads the workload equitably, and with a and Honeybee Foraging Algorithm (HFA). We see that under heavy load, ICM is better at avoiding system saturation than HFA and RR.

**Keywords** Cloud · Inter Cloud Manager · ICM · Job dispatching

## 1 Introduction

Cloud load balancing has always been a hot issue since emergence of distributed systems. Load balancing algorithms can be classified into sub categories from various perspectives. From one view point, they can be classified into static and dynamic algorithms. Static algorithms are suitable for homogeneous and stable environments. However, static algorithms are often not flexible enough to match dynamic changes in Cloud resources during execution time. On the other hand, dynamic algorithms are more flexible and take into consideration different types of system resources both prior and during run-time. Dynamic algorithms can adapt to changes and provide

---

B. Kang (✉)

Imperial College London, Department of Electrical and Electronic Engineering,  
Intelligent Systems and Networks Group, London, UK  
e-mail: bskang@imperial.ac.uk



better results in heterogeneous and dynamic environments. However, when the physical distribution of system resources becomes complex and dynamic, such algorithms can cause overhead, resulting in an overall degradation of QoS.

A dynamic approach can provide not only the software, hardware, and infrastructure redundancy necessary to optimize fault tolerance, but also steer work from different users to the best or fastest parts of the system. Resilience is particularly important for applications that rely on the Cloud for non-stop service such as emergency management [4], but even in such applications, diversity may need to be exploited for different aspects. For example, a certain system may handle large numbers of requests per unit time requiring small data transfers on average such as reporting the location of the people in a large building or sports arena, but another one may perform better for smaller numbers of requests per unit time involving large real-time jobs such as augmented reality simulations, or searching in large spaces for the best possible evacuation route [9, 10]. Due to this variety of jobs and servers, task scheduling mechanisms for single systems are not directly applicable to distributed environments [6, 11, 13, 14]. Furthermore, mobile devices are becoming one of the major sources of the workload for Clouds in order to save energy at the mobile device itself and avoid moving bulky data over wireless networks between the mobile devices and data repositories.

Thus our work is motivated by the need to investigate job dispatching in distributed system environments where the job requests must first travel over a network (mobile and wired) incurring some delay, then they are served by some host, and the results are returned back to the mobile device for review by the end user or for some service that was requested by the mobile device. We therefore investigate the operation of a job dispatcher that would reside at some network entry point in order to attempt to make judicious assignments to different Cloud hosts. To this effect we propose the Inter Cloud Manager (ICM) which monitors a set of Cloud nodes among which it may choose to send a job. The ICM uses a best-fit type technique to make the assignment based on jobs information and Cloud nodes status. We then run experiments on a small scale laboratory test-bed to evaluate ICM's performance, and compare it with other algorithms such as the static Round Robin approach (RR) [17] and the Honey-Bee Foraging which were previously proposed [18]. Measurements focus in particular on the observed total average response time including network delay in congested environments.

In the following section, we first present the core concept of ICM and then explain its details. Experiments and measurement results are provided in Sect. 3. Finally, we conclude the paper in Sect. 4.

## 2 The Inter Cloud Manager (ICM)

We developed an ICM which performs two tasks: Cloud monitoring and decision-making. The first part monitors and maintains the current status of Cloud information by using *Cloud Probes* (CPs) which are sent out by ICM to monitor each of the Clouds, collect response time data and bring it back to ICM. The basic concept of CPs was

inspired from the *Smart Packets* (SPs) [7] that exploit reinforcement learning based on the Random Neural Network [5, 16]. The rate at which this sampling takes place is varied between some maximum value, say (1 CP/s) to some low value (e.g. 0.1 CP/s), which increases when ICM's job queue contains more jobs. The second part of ICM is the decision-making process that uses a *Best-Fit* decision-making rule based on data that has been collected by the CP.

## 2.1 Cloud Monitoring

It is important to have a good knowledge about the condition of the current Cloud [3] in order to appropriately allocate the job. In order to be able to meet demands and provide satisfactory QoS [2, 8, 12, 15], individual monitoring mechanisms are needed and can lead to collection and processing of a large amount of runtime data. To monitor the Clouds continuously, we use CPs to collect system's conditions (e.g. num. of waiting jobs and average job execution time) of the Cloud hosts and the end-to-end delay from the client to the host.

Algorithm 1 shows basic Cloud monitoring process of ICM. ICM increases its sending rate of CP when meets the certain threshold value of its job waiting queue size. This threshold value can be altered depending on conditions in the network. ICM normally sends 0.1 CPs per second in random Cloud host, but in a busy situation (e.g. many clients send a job request at the same time), ICM changes sending rate of CP to 1.0 per second to promptly recognize the current conditions of the Cloud. In an environment where job requests are not as frequent (under the threshold), ICM sends 0.1 CP per second which minimizes network overhead. In return, this saves up the energy that is required to run the ICM.

The ICM uses the Current Status of Cloud (CSC) table. Every interconnected Cloud information is periodically updated by CP packet through ACK packet that is being sent by the destination (Cloud host). The CSC Table contains C\_ID, time, delay, loss rate, average job execution time, and the number of waiting jobs. C\_ID is the identification code for each cloud; time is the time that is used to identify the information from each cloud. Delay and loss rate have a particular function in this case to find out how much network delay is being made from client (source) to Cloud host (destination). Furthermore, we can get the average job execution time and the number of waiting jobs from the CP. These two sets of information are the key things that enable ICM to recognize the current condition of Cloud hosts and networks. All of this information is being used in the decision-making process. ICM uses the history from each Cloud in its decision-making.

## 2.2 Decision-Making

Decision-making is a core function of ICM and maintains best efficiency in dispatching job requests from a client in the ideal Cloud location. Algorithm 2 shows the

details of ICM's decision-making process. There are two steps that decision-making follows.

The first step is selecting one request among the job list that ICM holds. All job requests are saved in a "Linked List" containing its first time that it was originally entered as job request and a basic description. One reason to use the linked list is to provide primary benefit of limiting memory waste as insertion and removal of data constantly takes place. ICM processes job requests in order, from head to tail which means in our case that we used a First-In-First-Out (FIFO) job selection process. We will explain details about experimental system settings in Sect. 3.

---

**Algorithm 1** Cloud monitoring of ICM

---

```

1: forever
2: while ICM's job input queue length  $q > 0$  do
3:   if Receive ACK from destination[ $C\_ID$ ] then
4:     adjust CSC table[ $C\_ID$ ]
5:   end if
6: end while
7: /*adjust sending rate of CP*/
8: SET sending rate of  $CP = 0.1$  CP/sec
9: if ( $\geq$ ) threshold( $q$ ) then
10:   SET sending rate of  $CP = 1.0$  CP/sec
11: end if
12: end forever

```

---

The second step is the job dispatching process that was initially placed by the client and its task is to make sure that the dispatching process is well undertaken to the most ideal Cloud host. ICM uses a *Best-Fit* approach that dispatches a job to the shortest Average Response Time (ART) among operating Clouds. Firstly, ICM checks the ART value. If all the hosts exceed the threshold of ART, ICM do not send and hold a job until a host's ART is under the threshold. Secondly, ICM checks the status of hosts to see whether they are idle or not. If a host does not have any waiting jobs, ICM sends a job to that host. Finally, ICM sends a job to the host with the minimum value of ART.

As we discussed, ICM uses history base decision-making. In this case, ART is calculated by summing between Expected Network Delay (ENT) and Expected Job Transfer Time (EJTT) from the cloud host. ENT is calculated by using formula of (avg. network delay  $\times$  num. of packets for current job) over  $(1 - \text{loss rate})$ . From a networks point of view, a job (application) consists of several numbers of data packets. In our case, avg. network delay included processing, queuing, transmission, and propagation delays. EJTT is calculated by avg. job execution time  $\times$  (num. of waiting jobs + 1). Average job execution time can be calculated as shown in expression (1) where  $W_{new}$  stands for current average job execution time and  $W_{old}$  is previous value,  $W_{last}$  is last job execution time.  $a$  is the value ( $0 < a < 1$ ) that is worthwhile to notice as it is the usage of system memory from the last job within

the Cloud host. For instance, suppose the latest job used 30% of system memory. Therefore, value of  $a$  becomes 0.3 in the simplest term.

$$W_{new} \leftarrow W_{old}(1 - a) + W_{last} \times a \quad (1)$$

---

**Algorithm 2** Decision-making of ICM
 

---

```

1: Select job( $j$ )
2: Compute each Cloud's observed Average Response Time (ART)
3: /*no appropriate Cloud at current time*/
4: if Minimum ART  $\geq$  threshold(ART) then
5:   Do not Dispatch Job( $j$ )
6:   /*prevent starvation of the Cloud*/
7: else if destination[ $C\_ID$ ] has no waiting jobs then
8:   Allocate job( $j$ ) to the destination[ $C\_ID$ ]
9: else
10:  Allocate job( $j$ ) to the appropriate Cloud
11: end if

```

---

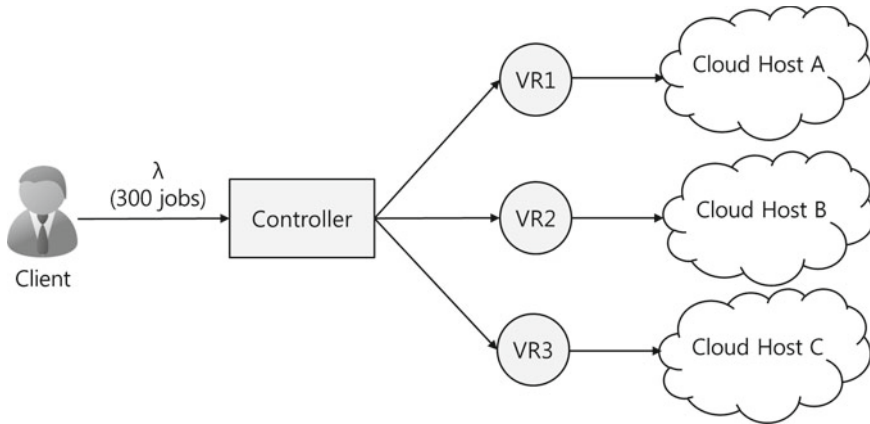
### 3 Experiment Results

In the real environment, congestion [1] can occur in any intermediate node, often due to limitation in resources, when data packets are being transmitted from the client to the destination. Congestion will lead to high packet loss, long delay and a waste of resource utilization time. Therefore, for this measurement, we compare three job dispatching algorithms with a congested environment.

#### 3.1 System Settings

Figure 1 shows a experimental model for congested Cloud networks. Each Cloud host has different computing machines. Host server A has Intel Xeon 2.4 GHz dual core CPUs and size of RAM is 0.5 GB where host B has Xeon 3.0 GHz dual core CPUs with RAM size of 2.0 GB. Finally, host C has Xeon 2.8 GHz dual core CPUs with RAM size of 1.0 GB. The systems run Linux (Ubuntu) with CPU throttling enabled with the *on demand governor*, which dynamically adjusts the cores frequencies depending on load. In addition, each server processes jobs based on First-Come-First-Serve (FCFS) scheduling policy.

First of all, we use the types of jobs for delay-sensitive rather than processing intensive jobs. All jobs (data packets) are transferred and executed through the network session between client and host server. We use our VPN network for this measurement. The long job is YouTube real-time video streaming with a 10 MB medium quality file; the medium job is uploading a Skype voice file (5 MB sensitive



**Fig. 1** Congested Cloud networks

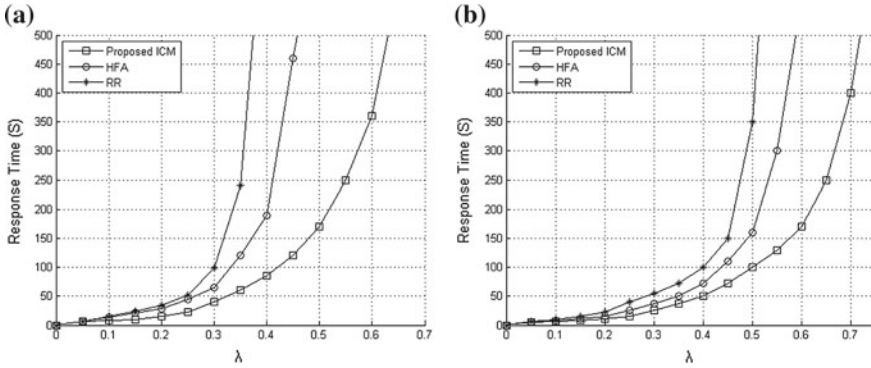
traffic) which is regarded as a normal workload. Finally, the short job is sending email (best-effort traffic). A job (application) consists of several data packets. For instance, a long job (10 MB YouTube video streaming) creates more than 150 wireless TCP data packets as it is transferred to its destination. If the client successfully received all the YouTube video data packets, we can count a job as finished and disconnect the network session. In this environment, network delay is much more important than computation delay. We used total 300 jobs (100 short, 100 medium, and 100 long) for all the tests.

Secondly, we use the controller. The controller plays an important role not only in interconnecting the client with cloud hosts but in dispatching jobs to the ideal Cloud host. The controller uses three different algorithms for dispatching a job, such as proposed ICM, HFA, and RR. These algorithms received a job without knowing if jobs are long, medium, or short work load. Therefore we use a single queue rather than a multi queue for the controller. In addition, the controller uses FIFO scheduling. That is the commonly used and simplest way to develop a single waiting queue.

Finally, we added software Virtual Router (VR); every VR is located in a nearby Cloud host. By using VR, we can manually adjust each network's QoS settings such as delay, loss rate, bit rate, etc., and in this measurement, we just changed the network propagation delay settings from 10 to 4000 ms. Except propagation delay, other settings depend on the real-network environment.

### 3.2 Performance Comparison of Three Algorithms

In order to carry out performance evaluation of ICM, comparative measurement was undertaken with well-known HFA and round robin algorithms. HFA is the state of the art in Cloud environmen and different from ICM, which uses a *First-Fit* job dispatching approach. We measured each algorithm's job response time under the



**Fig. 2** Measured response time for congested network. **a** Long propagation delay network, **b** short propagation delay network

congested environment. We fixed the threshold value of ART for 10min at every measurement. If every system has reached that value, ICM does not dispatch a job to the Cloud hosts anymore, because the systems are normally saturated at that point.

Figure 2 shows the measurement results under the long (1000–4000 ms) and the short (10–999 ms) propagation delay networks, respectively. The reported values were obtained by averaging the measurements corresponding to 300 jobs. We changed job arrival rate  $\lambda$  from 0.05 to 0.70 and measured 30 times at each  $\lambda$  point. In order to capture the sudden changes in response time at high traffic rates, we measured response time for every 0.05 point after  $\lambda = 0.4$ . In Fig. 2a for the long-latency network, the saturation rate of ICM was on average 0.631, HFA was on average 0.464 and RR was 0.371. In the congested network, ICM shows contrasting performance in that it effectively dispatched a job to the ideal Cloud host to find best place. In addition, ICM has smooth curves while others are rapidly increased for high traffic rates. This result caused by ICM increases the CP sending rate to immediately recognize the current status of Cloud hosts in a busy situation. In contrast, HFA and RR cannot efficiently keep track of the current status of host servers.

As we can see from the measurement results, network delay is more important than computation delay in congested networks. To monitor the Clouds, ICM sends CP periodically to collect Cloud information; this adds network overhead and is deemed to be the remaining issue. In general, it is important to focus on both performance and energy while the algorithm is evaluating what is used in the Cloud.

## 4 Conclusion

In this paper, we suggest a dynamic job dispatching algorithm that is better known as ICM, designed to be suitable for a congested environment. Comparative measurement is carried out to compare the performance of ICM, HFA and RR while increasing the job sending rate. After evaluation, average response time from ICM

shows outstanding performance than the other two algorithms, but ICM generated additional control (dummy) packets continuously into the entire network. To use these experimental results, we can estimate the expected saturation point in congested environments. Furthermore, we can improve the performance of the job allocation algorithm. Our future work will extend job dispatching to large-scale and energy-aware Cloud environments.

## References

1. Apostolopoulou, D., Gross, G., Guler, T.: Optimized fit portfolio construction based on the identification of congested network elements. *IEEE Trans. Power Syst.* **28**(4), 4968–4978 (2013)
2. Berl, A., Gelenbe, E., Di Girolamo, M., Giuliani, G., De Meer, H., Dang, M.Q., Pentikousis, K.: Energy-efficient cloud computing. *Comput. J.* **53**(7), 1045–1051 (2010)
3. Clay, R.W., Wild, N.R., Bird, D.J., Dawson, B.R., Johnston, M., Patrick, R., Sewell, A.: A cloud monitoring system for remote sites. *Publ. Astronom. Soc. Aust.* **15**(03), 332–335 (1998)
4. Filippopolitis, A., Gelenbe, E.: A distributed decision support system for building evacuation. In: 2nd Conference on Human System Interactions, 2009, HSI'09, pp. 323–330 (2009)
5. Gelenbe, E.: The first decade of g-networks. *Eur. J. Operat. Res.* **126**(2), 231–232 (2000)
6. Gelenbe, E.: Sensible decisions based on QoS. *Comput. Manag. Sci.* **1**(1), 1–14 (2003)
7. Gelenbe, E.: Steps toward self-aware networks. *Commun. ACM* **52**(7), 66–75 (2009)
8. Gelenbe, E.: Energy packet networks: smart electricity storage to meet surges in demand. In: Proceedings of the 5th International ICST Conference on Simulation Tools and Techniques, pp. 1–7. ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering) (2012)
9. Gelenbe, E., Cao, Y.: Autonomous search for mines. *Eur. J. Operat. Res.* **108**(2), 319–333 (1998)
10. Gelenbe, E., Hussain, K., Kaptan, V.: Simulating autonomous agents in augmented reality. *J. Syst. Softw.* **74**(3), 255–268 (2005)
11. Gelenbe, E., Lent, R.: Trade-offs between energy and quality of service. In: Sustainable Internet and ICT for Sustainability, SustainIT 2012, 4–5 Oct 2012, Pisa, Italy, Sponsored by the IFIP TC6 WG 6.3 “Performance of Communication Systems”, pp. 1–5. IEEE (2012), <http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=6388014>
12. Gelenbe, E., Lent, R.: Trade-offs between energy and quality of service. In: Sustainable Internet and ICT for Sustainability (SustainIT), 2012, pp. 1–5. IEEE (2012)
13. Gelenbe, E., Lent, R.: Energy-qos trade-offs in mobile service selection. *Future Internet* **5**(2), 128–139 (2013)
14. Gelenbe, E., Lent, R., Douratsos, M.: Choosing a local or remote cloud. In: 2012 Second Symposium on Network Cloud Computing and Applications (NCCA), pp. 25–30. IEEE (2012)
15. Gelenbe, E., Morfopoulou, C.: A framework for energy-aware routing in packet networks. *Comput. J.* **54**(6), 850–859 (2011)
16. Gelenbe, E., Timotheou, S.: Random neural networks with synchronized interactions. *Neural Comput.* **20**(9), 2308–2324 (2008)
17. Sakata, M., Noguchi, S., Oizumi, J.: An analysis of the m/g/1 queue under round-robin scheduling. *Operat. Res.* **19**(2), 371–385 (1971)
18. Venkata Krishna, P.: Honey bee behavior inspired load balancing of tasks in cloud computing environments. *Appl. Soft Comput.* **13**(5), 2292–2303 (2013)

# Discrete Time Stochastic Automata Network with Steady-State Product Form Distribution

J.M. Fourné

**Abstract** We present some sufficient conditions for a discrete time SAN to have a steady-state distribution which has a multiplicative form. The proofs are based on algebraic properties of the tensor operations associated with SAN. Some examples are given.

## 1 Introduction

We consider a discrete-time model of a system with multiple components which evolve according to some global events. The description of the model is based on the stochastic automata network (SAN) formalism and on the generalized tensor product introduced by Plateau [9]. Assuming that all components are associated with Markov transition (say  $M_i^{(e)}$  for effect of event  $e$  on component  $i$ ), we obtain a global Markov chain which models the whole system. Let  $N$  be the number of automata and  $Pr(e)$  the probability for global event  $e$ .  $Pr(e)$  does not depend on the state of the automata. The transition probability matrix of the chain is

$$\left(\sum_e Pr(e) \bigotimes_{i=1}^N M_i^{(e)}\right).$$

A distribution of probability  $\pi$  for a multi component chain with dimension  $N$  has a product form expression if there exist  $N$  marginal distributions  $\pi_e$  such that  $\pi = \bigotimes_{e=1}^N \pi_e$ . Thus, we have a strong relation between product form and tensor, and one may ask if it is possible to prove algebraically that:

$$\left(\bigotimes_{e=1}^N \pi_e\right) \left(\sum_e Pr(e) \bigotimes_{i=1}^N M_i^{(e)}\right) = \bigotimes_{e=1}^N \pi_e,$$

---

J.M. Fourné (✉)  
PRISM, UVSQ, CNRS, Versailles, France  
e-mail: Jmf@prism.uvsq.fr



i.e. the steady-state solution of  $(\sum_e Pr(e) \otimes_{i=1}^N M^{(e)})$  is  $(\otimes_{e=1}^N \pi_e)$  (i.e. it has a multiplicative form). Note that each component is modeled by several stochastic matrices (one for each global event). Thus, the automata are synchronized and they are not independent in general, as shown in the examples. We prove some sufficient conditions for such a result to hold. We extend the results obtained in [2] to discrete time SAN without functional dependencies. Some sufficient conditions for continuous time SANs associated with continuous time Markov chains have been introduced in [10] and generalized in [4, 5]. The discrete time version of the SAN is more complex due to the synchronizations between automata.

The technical part of the paper is as follows. In Sect. 2, we give some algebraic properties of the tensor algebra. Section 3 is devoted to the main theoretical results while in Sect. 4, we present some examples to illustrate the result. For technical results on SAN, one can read [9, 10] and the references therein, for an example of modeling with SAN, one can refer to [8].

## 2 Some Results About Tensor

We first define the usual tensor and we give some of its algebraic properties (see [11] for more details). We give some algebraic properties of these operators and we prove a sufficient condition for product form.

**Definition 1** The tensor product  $C = A \otimes B$  is defined by assigning the element of  $C$  that is in the  $(i_2, j_2)$  position of block  $(i_1, j_1)$ , the value  $a_{i_1 j_1} b_{i_2 j_2}$ . We shall write this as

$$c_{\{(i_1, j_1); (i_2, j_2)\}} = a_{i_1 j_1} b_{i_2 j_2}.$$

**Property 1** (Basic properties of Tensor) *Let  $A, B, C$  and  $D$  be four arbitrary stochastic matrices, and  $\lambda \in R$  the following properties hold (see [1] for a proof):*

- *Associativity:*

$$(A \otimes B) \otimes C = A \otimes (B \otimes C).$$

- *Associativity of the scalar product:*

$$(\lambda A) \otimes B = A \otimes (\lambda B).$$

- *Distributivity over Addition:*

$$\begin{aligned} (A + B) \otimes (C + D) &= A \otimes C \\ &+ A \otimes D \\ &+ B \otimes C \\ &+ B \otimes D. \end{aligned}$$

**Property 2** *Let  $A$  and  $B$  two stochastic matrices, then  $A \otimes B$  is a stochastic matrix.*

These properties are also true for generalized tensor product. But the most important property for our problem is the compatibility with the ordinary product.

**Property 3** (Compatibility with product) *Let  $A, B, C$  and  $D$  four arbitrary matrices with respective size  $n \times m, m \times p, q \times r$ , and  $r \times s$ . We have:*

$$(A \otimes B) \cdot (C \otimes D) = (A \cdot C) \otimes (B \cdot D),$$

where ‘ $\cdot$ ’ is the ordinary matrix product.

Now suppose that we consider vectors instead of matrix, we clearly have if  $v$  (resp.  $w$ ) is a vector with size  $m$ , (resp.  $r$ )

$$(v \otimes w) \cdot (C \otimes D) = (v \cdot C) \otimes (w \cdot D),$$

**Property 4** (Eigenvector) *if  $v$  is an eigenvector of  $C$  with eigenvalue  $\lambda$  and  $w$  is an eigenvector of  $D$  with eigenvalue  $\mu$ , then  $v \otimes w$  is an eigenvector of  $C \otimes D$  with eigenvalue  $\lambda\mu$ .*

*Proof*

$$(v \otimes w) \cdot (C \otimes D) = (v \cdot C) \otimes (w \cdot D) = (\lambda C) \otimes (\mu D) = \lambda\mu (C \otimes D).$$

□

Finally note that  $M^t$  is the transpose of  $M$  and  $Id$  is the Identity matrix with the appropriate size.

### 3 Model and Theorems

Assume that the discrete time SAN allows the following decomposition of its stochastic matrix:

$$P = \sum_{e \in \mathcal{E}} Pr(e) \bigotimes_{i=1}^N M_i^{(e)},$$

where:

- $N$  is the number of Automata
- $e$  is a global event.  $\mathcal{E}$  is the set of global events.
- $Pr(e)$  is the probability of event  $e$ . Of course  $\sum_{e \in \mathcal{E}} Pr(e) = 1$ .
- $M_i^{(e)}$  is the stochastic matrix describing the effect of global event  $e$  on Automaton  $i$ .

**Property 5** Matrix  $\bigotimes_{i=1}^N M_i^{(e)}$  is stochastic. By the convex composition of stochastic matrices,  $P$  is also a stochastic matrix

The proofs are well-known and they appear in many papers on tensor product.

*Example 1* We consider 3 events with probability  $1/2$ ,  $1/3$ ,  $1/6$  and 2 automata. We assume that both automata have state space  $\{a, b\}$ . Remember that  $M_1^{(e)}$  is the description of event  $e$  on automata 1. Let these matrices be:

$$M_1^{(1)} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \quad M_1^{(2)} = \begin{pmatrix} 1/3 & 2/3 \\ 2/3 & 1/3 \end{pmatrix} \quad M_1^{(3)} = \begin{pmatrix} 0.1 & 0.9 \\ 0.9 & 0.1 \end{pmatrix},$$

$$M_2^{(1)} = \begin{pmatrix} 1/2 & 1/2 \\ 1/4 & 3/4 \end{pmatrix} \quad M_2^{(2)} = \begin{pmatrix} 3/4 & 1/4 \\ 1/8 & 7/8 \end{pmatrix} \quad M_2^{(3)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

The transition matrix associated with the SAN is:

$$M = 1/2 (M_1^{(1)} \otimes M_2^{(1)}) + 1/3 (M_1^{(2)} \otimes M_2^{(2)}) + 1/6 (M_1^{(3)} \otimes M_2^{(3)}).$$

Remember also the following definition of a product form steady-state distribution:

**Definition 2**  $\pi$  is the steady state for the ergodic DTMC associated to stochastic matrix  $P$  and it is product form if

- $\pi P = \pi$
- $\pi = \bigotimes_{i=1}^N \pi_i$ , where  $\pi_i$  is the solutions of  $\pi_i = \pi_i B_i$  for some stochastic matrix  $B_i$  defined on the state space of automaton  $A_i$ . Note that  $B_i$  is not in general equal to one of the  $M_i^{(e)}$  for some  $e$ .

**Theorem 1** Assume that the Markov chain is ergodic. Assume also that there exist an automaton index  $k$  such that the two following conditions hold:

- for all  $i \neq k$ ,  $\pi_i M_i^{(e)} = \pi_i$  for all event  $e$ ,
- $\pi_k (\sum_{e \in \mathcal{E}} Pr(e) M_k^{(e)}) = \pi_k$

then the steady-state distribution has product form:  $\pi = \bigotimes_{i=1}^N \pi_i$ , where the  $\pi_i$  are obtained from the conditions of the theorem.

*Proof* We write the Chapman-Kolmogorov balance equation assuming that the steady-state has a product form.

$$\bigotimes_{i=1}^N \pi_i \left[ \sum_{e \in \mathcal{E}} Pr(e) \bigotimes_{i=1}^N M_i^{(e)} \right] = \bigotimes_{i=1}^N \pi_i.$$

Let  $LHS = \bigotimes_{i=1}^N \pi_i \left[ \sum_{e \in \mathcal{E}} Pr(e) \bigotimes_{i=1}^N M_i^{(e)} \right]$ .

We want to prove that  $LHS = \bigotimes_{i=1}^N \pi_i$ . First, we apply the distributivity of the tensor product on the ordinary sum.

$$LHS = \sum_{e \in \mathcal{E}} Pr(e) \left[ \bigotimes_{i=1}^N \pi_i \right] \cdot \left[ \bigotimes_{i=1}^N M_i^{(e)} \right] \quad (1)$$

where “.” is the ordinary product. Let us now use the key property for this tensor approach: the compatibility between tensor product and ordinary product.

$$LHS = \sum_{e \in \mathcal{E}} Pr(e) \left[ \bigotimes_{i=1}^N \left[ \pi_i \cdot M_i^{(e)} \right] \right], \quad (2)$$

Let us now take into account the assumptions of the theorem. First we use transposition matrix  $R_{(k)}$  defined by Plateau . Remember that it is defined as:

$$\bigotimes_{i=1}^N A = R_{(k)} \left[ A_k \bigotimes_{i=1, i \neq k}^N A_i \right] (R_{(k)})^t \quad (3)$$

Indeed the tensor product is not commutative. Now substitute Eq. 3 into Eq. 2. After some factorization, we get:

$$LHS = \sum_{e \in \mathcal{E}} Pr(e) R_{(k)} \left[ \pi_k M_k^{(e)} \bigotimes_{i=1, i \neq k}^N \left[ \pi_i \cdot M_i^{(e)} \right] \right] (R_{(k)})^t, \quad (4)$$

Now remember the first part of the assumptions: for all index  $i$  not equal to  $k$  we have, for all  $e$ ,  $\pi_i M_i^{(e)} = \pi_i$ . Thus,

$$LHS = \sum_{e \in \mathcal{E}} Pr(e) R_{(k)} \left[ \pi_k M_k^{(e)} \bigotimes_{i=1, i \neq k}^N \pi_i \right] (R_{(k)})^t \quad (5)$$

Using again distributivity, we get:

$$LHS = R_{(k)} \left[ \pi_k \left( \sum_{e \in \mathcal{E}} Pr(e) M_k^{(e)} \right) \bigotimes_{i=1, i \neq k}^N \pi_i \right] (R_{(k)})^t, \quad (6)$$

Now we use the second assumption on the theorem:  $\pi_k \left( \sum_{e \in \mathcal{E}} Pr(e) M_k^{(e)} \right) = \pi_k$ .

$$R_{(k)} \left[ \pi_k \bigotimes_{i=1, i \neq k}^N \pi_i \right] (R_{(k)})^t = \bigotimes_{i=1}^N \pi_i, \tag{7}$$

which is obviously true by definition of matrix  $R_{(k)}$ . □

We now extend the theorem using a decomposition of the set of events.

**Theorem 2** *Assume that the Markov chain is ergodic. Assume also that there exist a partition of the set of event  $\mathcal{E}$  into  $N$  subsets:  $\mathcal{E}_1, \dots, \mathcal{E}_N$  such that for all subset index  $k$  we have:*

- *if  $\mathcal{E}_k$  is not empty.*
  - *for all  $i \neq k, \pi_i M_i^{(e)} = \pi_i$  for all event  $e$  in  $\mathcal{E}_k$*
  - *$\pi_k (\sum_{e \in \mathcal{E}_k} Pr(e) M_k^{(e)}) = \pi_k \sum_{e \in \mathcal{E}_k} Pr(e)$ ,*

*then the steady-state distribution has product form:  $\pi = \bigotimes_{i=1}^N \pi_i$ , where the  $\pi_i$  are obtained from the conditions of the theorem.*

*Proof* First remark that we do not ask that the partition is a proper partition; therefore we may have that for some  $k, \mathcal{E}_k = \emptyset$ . Each subset is associated to an automaton. Therefore  $k$  is also an automaton index. We use the same technique as in the previous result. We begin with the Chapman Kolmogorov equation in tensor form:

$$\bigotimes_{i=1}^N \pi_i \left[ \sum_{e \in \mathcal{E}} Pr(e) \bigotimes_{i=1}^N M_i^{(e)} \right] = \bigotimes_{i=1}^N \pi_i,$$

and we readily obtain:

$$LHS = \sum_{e \in \mathcal{E}} Pr(e) \left[ \bigotimes_{i=1}^N \left[ \pi_i \cdot M_i^{(e)} \right] \right]. \tag{8}$$

We now decompose among the subsets of the partition.

$$LHS = \sum_{k=1}^N \sum_{e \in \mathcal{E}_k} Pr(e) \left[ \bigotimes_{i=1}^N \left[ \pi_i \cdot M_i^{(e)} \right] \right]. \tag{9}$$

Consider now the subset  $\mathcal{E}_k$  and the associated automaton with the same index. First assume that  $\mathcal{E}_k$  is not empty. We can reorder the tensors product with the help of matrices  $R_{(k)}$ .

$$\sum_{e \in \mathcal{E}_k} Pr(e) \left[ \bigotimes_{i=1}^N \left[ \pi_i \cdot M_i^{(e)} \right] \right] = R_{(k)} \pi_k \left( \sum_{e \in \mathcal{E}_k} Pr(e) M_k^{(e)} \right) \left[ \bigotimes_{i=1, i \neq k}^N \left[ \pi_i \cdot M_i^{(e)} \right] \right] (R_{(k)})^t \tag{10}$$

Due to the assumptions of the theorem, we get after simplifications:

$$\sum_{e \in \mathcal{E}_k} Pr(e) \left[ \bigotimes_{i=1}^N \left[ \pi_i \cdot M_i^{(e)} \right] \right] = R_{(k)} \pi_k \sum_{e \in \mathcal{E}_k} Pr(e) \left[ \bigotimes_{i=1, i \neq k}^N \pi_i \right] (R_{(k)})^t. \quad (11)$$

Reordering the tensor with matrix  $R_{(k)}$  we finally get;

$$\sum_{e \in \mathcal{E}_k} Pr(e) \left[ \bigotimes_{i=1}^N \left[ \pi_i \cdot M_i^{(e)} \right] \right] = \sum_{e \in \mathcal{E}_k} Pr(e) \left[ \bigotimes_{i=1}^N \pi_i \right]. \quad (12)$$

Let now consider that  $\mathcal{E}_k = \emptyset$ . Then  $\sum_{e \in \mathcal{E}_k} Pr(e) = 0$  and this part of the equation cancelled. We finally get:

$$\sum_{k=1}^N \sum_{e \in \mathcal{E}_k} Pr(e) \left[ \bigotimes_{i=1}^N \left[ \pi_i \cdot M_i^{(e)} \right] \right] = \left[ \sum_{k=1}^N \sum_{e \in \mathcal{E}_k} Pr(e) \right] \cdot \left[ \bigotimes_{i=1}^N \pi_i \right]. \quad (13)$$

As  $[\sum_{k=1}^N \sum_{e \in \mathcal{E}_k} Pr(e)] = 1$  because the subsets form a partition, we finally obtain:

$$LHS = \bigotimes_{i=1}^N \pi_i,$$

□

**Corollary 1** Assume that the Markov chain is ergodic. Assume also that for all index  $i$  there exist a probability distribution  $\pi_i$  such that for all event  $e$  in  $\mathcal{E}$ , we have:

- $\pi_i M_i^{(e)} = \pi_i$ ,

then the steady-state distribution has product form:  $\pi = \bigotimes_{i=1}^N \pi_i$ .

## 4 Shoals of Markovian Automata

The examples are based on the three following properties which give examples of stochastic matrices which all have the same dominant left eigenvector. We then show how we can use these properties to find groups of Markov chains which collectively react to global events and have a product-form steady-state distribution.

**Property 6** Let  $M$  be an arbitrary stochastic matrix associated with an ergodic DTMC. For all  $k \geq 0$ ,  $M^k$  has the same dominant eigenvector than  $M$ .

*Proof* Assume that  $\pi M = \pi$ . Then,  $\pi M^k = (\pi M)(M^{k-1}) = \pi M^{k-1} = \pi M = \pi$ . The first step is by associativity of the product. The second by induction on  $k$ .

**Property 7** Let  $M1$  and  $M2$  be two arbitrary stochastic matrices associated with ergodic DTMCs and such that they share the same dominant eigenvector. For all  $v$  such that  $1 \geq v \geq 0$ ,  $vM1 + (1-v)M2$  has the same dominant eigenvector than  $M1$ .

*Proof* Assume that  $\pi M1 = \pi$  and  $\pi M2 = \pi$ . Then, by distributivity,  $\pi(vM1 + (1-v)M2) = \pi vM1 + \pi(1-v)M2 = v\pi + (1-v)\pi = \pi$ .

**Property 8** Let  $M1$  be an arbitrary stochastic matrix associated with an ergodic DTMC. Let  $Id$  be the identity matrix with a consistent size. For all  $v$  such that  $1 \geq v \geq 0$ ,  $vM1 + (1-v)Id$  has the same dominant eigenvector than  $M1$ .

*Proof* This is a simple corollary of Property 7. Note that the ergodicity of  $M2$  is not necessary to prove Property 7.

## 4.1 Global Synchronizations and Number of Steps

For the sake of readability we only consider two automata and two events. Thus we have to describe 4 stochastic matrices. We assume that  $M_1^{(1)}$  and  $M_2^{(1)}$  are arbitrary stochastic matrices associated with two ergodic Markov chains. For some illustrative reason we also assume that the diagonal entries of  $M_1^{(1)}$  are zero. Assume that:

$$M_1^{(2)} = (M_1^{(1)})^0 = Id, \text{ and } M_2^{(2)} = (M_2^{(1)})^3.$$

Assume also that  $Pr(event1) = a$  and  $Pr(event2) = 1 - a$  for an arbitrary  $a$ , such that  $0 < a < 1$ . Such a model may represent, in a very abstract way, a system with two modes for consuming energy. At each time slot, the system may use 3 units of energy (or Energy Packets like in [6, 7]). Component 1 uses 2 EP to move while component 2 only uses 1 EP. It chooses, probabilistically, one of the two possible modes. In Mode 1 (with probability  $a$ ) both components make one move according respectively to matrices  $M_1^{(1)}$  and  $M_2^{(1)}$ . In Mode 2, Component 1 does not move (the transition matrix is  $Id$ ) et Component 2 performs 3 steps with transition matrix  $M_2^{(1)}$  (the transition matrix is  $(M_2^{(1)})^3$ ).

Clearly the two components are not independent as they obey some common rules associated to global events. When Component 1 does not move, it implies that event 2 occurs (remember that  $M_1^{(1)}[i, i] = 0$  for all  $i$ ). Thus, we know that Component 2 makes three transitions with probability 1. However the assumptions of the corollary hold because the matrices are built to satisfy Property 6. Thus, such a system has a product form steady-state distribution. However as the components are not independent, the product form we prove is not based on independence and to the best of our knowledge, such a result is original.

## 4.2 Global Synchronizations and Local Speed

We now present a slightly different model. The events are now used to fix the speed of the transitions of the components. We consider two automata and a set of events of size  $E$ . Here we have to describe 4 stochastic matrices and two sets of parameters. We will build these matrices to satisfy Property 8. We assume that  $M_1^{(1)}$  and  $M_2^{(1)}$  are arbitrary stochastic matrices associated with two ergodic Markov chains. Assume that:

$$M_1^{(e)} = v1_e Id + (1 - v1_e)M_1^{(1)}, \text{ and } M_2^{(e)} = v2_e Id + (1 - v2_e)M_2^{(1)},$$

Of course  $0 \leq v1_e \leq 1$  and  $0 \leq v2_e \leq 1$  for all event  $e$ . We consider an arbitrary distribution of probability for the events. Again the two components are not independent as their speeds are fixed by some global events. Nevertheless, the assumptions of the corollary hold because Property 8 is satisfied by the matrices we have considered. Thus, the steady-state has a multiplicative form.

## 4.3 Global Set of Synchronized Movements

Consider now a network with  $N$  automata. Each automaton is associated with a stochastic matrix (say  $P_i$  for Automaton  $i$ ). We consider a set of events  $\mathcal{E}$  of an arbitrary size. An event is characterized by a subset of the set of automata: this subset contains the automata which participate to the event. We assume that:

- if Automaton  $i$  belongs to the subset describing event  $e$ , then  $M_i^{(e)}$  is equal to  $P_i$ ,
- else it is equal to the identity matrix.

We consider an arbitrary distribution of probability for the events. Then, the model satisfies the assumptions of the theorem and this steady-state distribution (if it exists) has a product-form solution. Note however that the ergodicity of the chain cannot be proved without more detailed assumptions on matrices  $P_i$  and the subsets defining the events.

## 5 Conclusions

Clearly, the sufficient conditions we found allow to obtain new descriptions of multiple components systems the steady-state distribution of which (in a Markovian framework) has a multiplicative structure. We hope that these results will help to derive methods and tools to analyze discrete time systems. A numerical method will be added into the Xborne tool [3] in the near future.

**Acknowledgments** The author is supported by grant ANR-12-MONU-00019 Marmote.



## References

1. Fernandes, P., Plateau, B., Stewart, W.J.: Efficient descriptor-vector multiplications in Stochastic Automata Networks. *J. Acm* **45**(3), 381–414 (1998)
2. Fourneau, J.-M.: Collaboration of discrete-time markov chains: Tensor and product form. *Perform. Eval.* **67**(9), 779–796 (2010)
3. Fourneau, J.-M., Le Coz, M., Pekergin, N., Quessette, F.: An open tool to compute stochastic bounds on steady-state distributions and rewards. In: 11th International Workshop on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS 2003). IEEE Computer Society, Orlando (2003)
4. Fourneau, J.-M., Plateau, B., Stewart, W.: Product form for stochastic automata networks. In: Glynn, P.W. (ed.) Proceedings of the 2nd International Conference on Performance Evaluation Methodologies and Tools, VALUETOOLS 2007. Icsst, Nantes (2007)
5. Fourneau, J.-M., Plateau, B., Stewart, W.: An algebraic condition for product form in Stochastic Automata Networks without synchronizations. *Perform. Eval.* **85**, 854–868 (2008)
6. Gelenbe, E.: Energy packet networks: smart electricity storage to meet surges in demand. In: International ICST Conference on Simulation Tools and Techniques, SIMUTOOLS '12, ICST/ACM, 2012, pp. 1–7. Sirmione-Desenzano, Italy (2012). March 19–23, 2012,
7. Gelenbe, E., Marin, A.: Interconnected wireless sensors with energy harvesting. In: Analytical and Stochastic Modelling Techniques and Applications—22nd International Conference, ASMTA 2015, Albena, Bulgaria, volume 9081 of Lecture Notes in Computer Science, pp. 87–99. Springer, New York (2015)
8. Mokdad, L., Ben-Othman, J.: Admission control mechanism and performance analysis based on stochastic automata networks formalism. *J. Parallel Distrib. Comput.* **71**(4), 594–602 (2011)
9. Plateau, B., Fourneau, J., Lee, K.: PEPS: A package for solving complex Markov models of parallel systems. In: Proceedings of the 4th International Conference on Modeling Techniques and Tools for Computer Performance Evaluation, pp. 341–360. Majorca (1988)
10. Plateau, B., Stewart, W.: Advances in computational probability, chapter stochastic automata networks. In: Grassmann, W. (Ed.) Chapter Stochastic Automata Network. International Series in Operations Research and Management Science, vol. 24, Kluwer Academic Publishers, New York (1999)
11. Van Loan, C.: The ubiquitous Kronecker product. *J. Comput. Appl. Math.* **123**, 85–100 (2000)

# Modelling Dynamics of TCP Flows in Very Large Network Topologies

Monika Nycz, Tomasz Nycz and Tadeusz Czachórski

**Abstract** The article presents an approach to numerical modeling of dynamics of flows in wide area TCP/IP computer networks with the use of SAP HANA in-memory database. The aim is to explore the possibility of transforming a broadly known modeling method—fluid-flow approximation—into database language, overcoming this way the need to develop dedicated solutions and to transmit the results from the application to the database. We implemented the model logic into SQL procedures and performed mathematical calculations for an exemplary vast topology. The experiments show that the database engine may be used to perform all model computations.

**Keywords** Internet topology · Fluid-flow approximation · SAP HANA · SQL · Numerical calculation · Databases · Computer networks · Opte project · Queueing models · Big data

## 1 Introduction

The analysis of transient phenomena occurring in computer networks is performed by simulation or analytical models. Simulation models are time-consuming and require considerable computing power. Therefore their application is limited to small topologies. Analytical models are based on solved numerically Markov chains, e.g. [11], diffusion approximation [2, 3], and fluid-flow approximation [5, 8]. Each approach

---

M. Nycz · T. Nycz  
Institute of Informatics, Silesian University of Technology, ul. Akademicka 16,  
44-100 Gliwice, Poland  
e-mail: monika.nycz@polsl.pl

T. Nycz  
e-mail: tomasz.nycz@polsl.pl

T. Czachórski (✉)  
Institute of Theoretical and Applied Informatics, Polish Academy of Sciences,  
Baltycka 5, 44-100 Gliwice, Poland  
e-mail: tadek@iitis.pl

has its advantages and drawbacks, but it is the fluid-flow approximation that is widely used for modeling states in wide area networks, including the Internet.

The fluid approximation uses first-order ordinary linear differential equations to determine the dynamics of the average length of node queues and the dynamics of TCP congestion windows in a modeled network. The changes of a queue length  $dq_j(t)/dt$  at a station  $j$ , Eq. (1), are defined as the intensity of the input stream, i.e. the sum of all flows  $i = 1, \dots, K$  traversing a particular node minus the constant intensity of output flow  $C_j$ , i.e. the number of packets sent further in a time unit:

$$\frac{dq_j(t)}{dt} = \sum_{i=1}^K \frac{W_i(t)}{R_i(\mathbf{q}(t))} - \mathbf{1}(q_j(t) > 0) C_j. \quad (1)$$

A router allows reception of traffic from  $K$  TCP flows ( $K \leq N$ ), where  $N$  is the entire number of flows in the network. Following the TCP congestion avoidance principles, each flow  $i$  ( $i = 1, \dots, N$ ) is determined by its time varying congestion window size  $W_i$  which gives the number of packets to be sent without waiting for the acknowledgment of previously sent and received packets. This flow is expressed here by  $W_i(t)/R_i(\mathbf{q}(t))$ .

The window size, Eq. (2), increases by one at each RTT (round trip time, denoted here by  $R_i$ ) in the absence of packet losses (first term on the right side of Eq. (2)) and decreases by half of its current value after every packet loss occurring in nodes on the flow path (the latter decision is taken after the time  $\tau$ ). The loss for the entire TCP connection is defined as flow intensity multiplied by total drop probability—the probability which specifies that the loss occurs in nodes on the route. It is based on a matrix  $\mathbf{B}$  that stores drop probabilities in each router in all flows in the network,

$$\frac{dW_i(t)}{dt} = \frac{1}{R_i(\mathbf{q}(t))} - \frac{W_i(t)}{2} \frac{W_i(t - \tau)}{R_i(\mathbf{q}(t - \tau))} \left( 1 - \prod_{j \in V_i} (1 - B_{ij}) \right). \quad (2)$$

The values  $B_{ij}$  specify the drop probability at node  $j$  for packets of connection  $i$ ;  $V_i$  is the set of nodes belonging to this connection, and  $\mathbf{q}(t)$  is the vector of queues at these nodes. Delay  $R_i$  in these formulas determine the time needed for the information on congestion and packet loss to propagate through the network back to the sender of a flow  $i$ . It consists of queue delays at all nodes  $j$ , defined as  $q_j(t)/C_j$  along this connection and the propagation delay  $Tp_i$ :

$$R_i(\mathbf{q}(t)) = \sum_{j \in V_i} \frac{q_j(t)}{C_j} + Tp_i. \quad (3)$$

The drop probability  $p_j(x_j)$ , Eq. (4) in a single node is determined according to RED mechanism, e.g. [8] as a function of the moving average queue length  $x_j(t)$

which is the sum of current queue  $q_j(t)$  taken with a weight parameter  $w$  and previous average queue taken with  $(1 - w)$  weight,

$$p_j(x_j) = \begin{cases} 0, & 0 \leq x_j < t_{min_j} \\ \frac{x_j - t_{min_j}}{t_{max_j} - t_{min_j}} p_{max_j}, & t_{min_j} \leq x_j < t_{max_j} \\ 1, & t_{max_j} \leq x_j. \end{cases} \quad (4)$$

In general, the fluid-flow differential equations are solved numerically with 4th order Runge-Kutty method. In our current database approach, the Euler method has been chosen to reduce the complexity of calculations, however we have already tested that we are able to use RK4 as well.

## 2 Modeling Real Internet Topologies

The real Internet topologies form dependent structures with varying degrees of complexity. This study aims to analyze selected Internet network topologies with a large number of nodes and flows with the use of fluid-flow approximation. As a base topology serves here the Internet map developed within The Opte Project, [6]. The project was appointed to gather data from the global network and create an image of the Internet. During the initial phases of the project, data were collected using the traceroutes. The collected data on routing are stored in the database and then they are exported to files in the Large Graph Layout (LGL) format. Text files are used as the data source for the visualization tools, i.e. LGL, Gephi to create the final image.

The Opte Project provides currently two images of the Internet taken at 2003 and at 2014. However, the only 2003 data are publicly available and they are used in this study.

The project provides three types of source files:

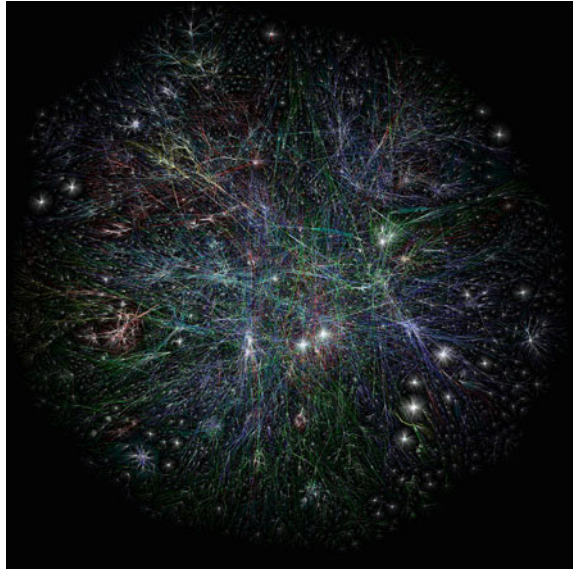
- LGL (Large Graph Layout), including undirected graph, with nodes and all its neighbors;
- 2D Coords, containing nodes together with the coordinates in the 2D space;
- Edge Coords, consisting of edges with assigned territory colours.

We focused on the LGL file, as it contains the network topology.

## 3 The Use of In-Memory Database

Our previous experiences with fluid-flow approximation applied to large topologies, e.g. [1] learned us that the main problem lies in storing and processing large amount of output data. The numerical solution of differential equations (1), (2) was possible only when at least one of two dimensions (number of flows, number of routers) was relatively small or the historical results were not collected. Otherwise the available

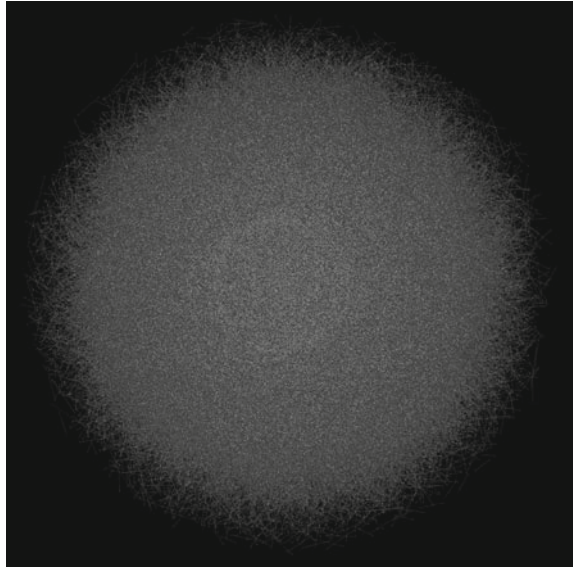
**Fig. 1** The map of the Internet from 2003 generated by The Opte Project, [7]



amount of memory in standard PC was insufficient to store the intermediate results. A standard solution is to create a dedicated software structure for storing and analyzing the obtained data or to use a database as a storage. However, we wanted to check the possibility of transferring the modeling application logic into database engine to overcome this way the need of development of dedicated solution and transmission of large amount of data from the application to the database. If we consider a thousand- and million-node topologies, the calculations generate a large amount of results (millions of values). To eliminate the necessity of customizing the resulting structures and waiting times for sending the results to storage, we transferred the application logic into database layer. Such solution requires a specialized hardware infrastructure. We focused on SAP HANA as an in-memory database solution. We used the HPI Future SOC Lab HP DL980 G7 server having i. a. 4 x Xeon (Nehalem EX) X7560 and 1 TB RAM. The calculations were performed by running the SQL procedures and scripts within SAP HANA Studio on SAP HANA 1.00.091 version (Fig. 1).

The task consists of three phases: preprocessing, implementation and modeling, visualization. In the first phase we selected a whole exemplary real network topology, [6] with 134,023 nodes and generated 50,000 router pairs for flows. Next, we generated the link propagation delays (tens to hundreds of milliseconds) based on which the flows paths were determined using Dijkstra algorithm. Last step concerned the preparation of the initial configuration data as the CSV files which were imported into SAP HANA database (Fig. 2).

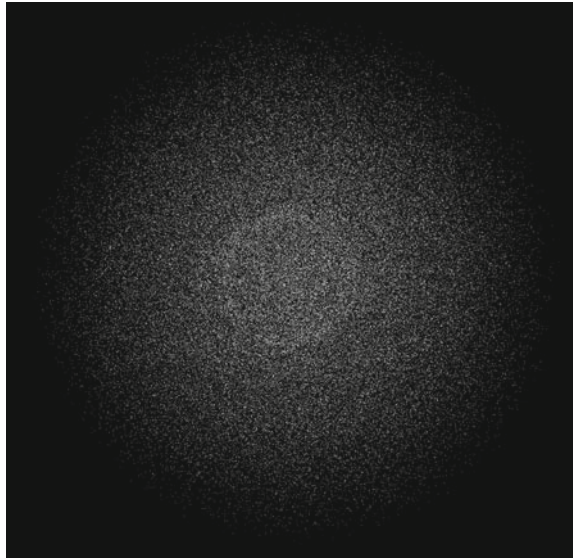
**Fig. 2** The Opte Project 2003 Internet map visualized in Gephi—nodes with edges



The application modeling logic assumed the update of all network parameters at each step and stored them as historical data. The values of particular step were used in subsequent modeling step. The main project phase involved the translating of the numerical fluid-flow approximation logic into the database language. In this case we created database tables, loaded them with initial data and tested several algorithms based on combination of updates, upserts and inserts. SAP HANA [9] provides two types of data storage: row store and column store. Both were used to determine the most effective way to perform modeling inside the database engine. The logic was processed within the SQL procedure using loop or as the SQL script. Despite the fact that the modeling logic calculated aggregates on both routers and flows, the preliminary results indicated the update-insert on row store as the best solution. However, the project period turned out to be too short to check all necessary optimization possibilities which could change the best option for the benefit of insert-only approach on column store. Our next publication will present details of the final implementation of this approach (Fig. 3).

Unfortunately, at the moment we cannot compare the performance of the database fluid-flow algorithm with any other solution. There is not any other dedicated solution using standard PC computer or any other technology. To compare the implementation with dedicated solution, it is necessary to “move” it to multiprocessor server which the SAP HANA database operate on. Only with the same hardware one can make a conclusive comparison.

**Fig. 3** The Opte Project 2003 Internet map visualized in Gephi—only nodes (edges removed for greater readability)

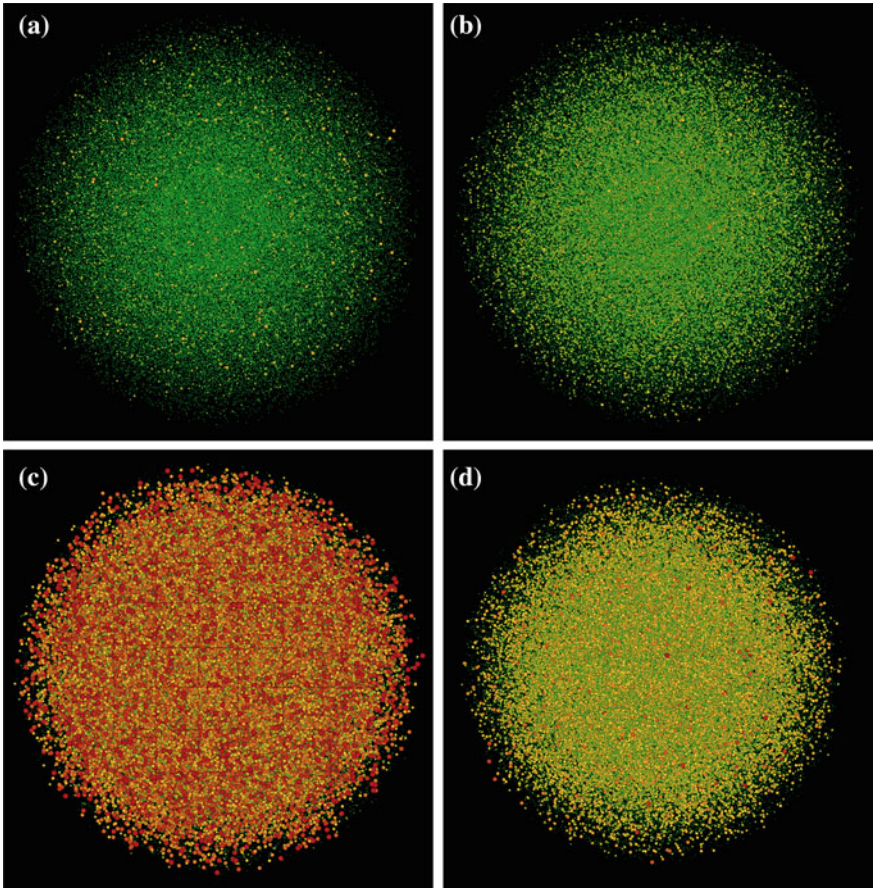


The final phase involved the visualization of calculated results. Therefore we extracted the values, such as the queue length, for particular steps from SAP HANA database as CSV files. Then we imported the basic topology into Gephi tool (The Open Graph Viz Platform, [4]) and updated it with files obtained from database. The process ends with visualization which presents the changes of queues in the whole network.

## 4 Numerical Results

The loads of network nodes (percentage) are defined in our visualization by colours: 0% (green), 50% (yellow), 100% (red). At the beginning ( $t = 1$  s), their queues are mainly empty, Fig. 4a.

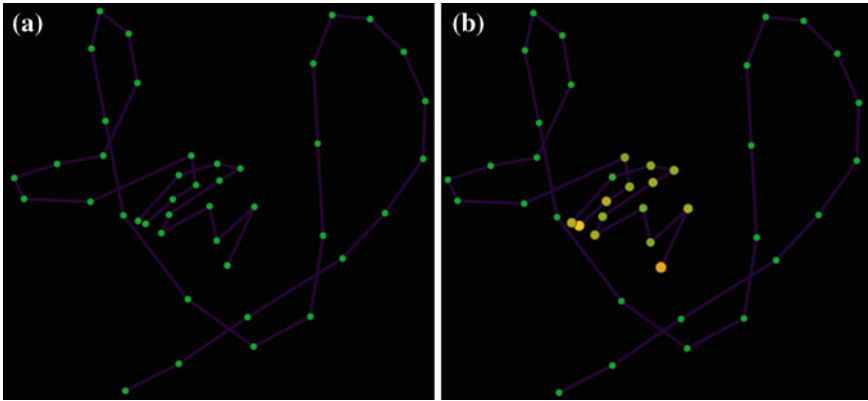
With the increase of time window sizes increase, thus the network congestion increases, Fig. 4b (the moment when the first congested routers occurs in the network) reaching the critical point of total congestion at ( $t = 26$  s), Fig. 4c. Then the increasing congestion activates the mechanisms to reduce window sizes and consequently the flows (the senders decrease their transmissions), a counter-measure against the overflows in nodes allowing to balance the network loads, Fig. 4d. In the rest of the analyzed time period the total congestion is no longer present.



**Fig. 4** The loads of nodes in the modeled network for time = 1 (a), 13 (b), 26 (c) and 50 (d) sec

We may also apply the filtering on the particular routers or flows and observe the situation from a detailed point of view. In our example, Figs. 5a, b show the behavior of the longest flow (42 nodes) in the whole network. As it turned out, it was not congested at the critical point (time = 26 s).





**Fig. 5** The loads of nodes of the longest flow in the modeled network for time = 1 (a) and 26 (b) sec

## 5 Conclusions and Next Steps

The studies, intending to analyze the final behavior of the fluid-flow approximation model for large networks, indicate that it is possible to solve first-order differential equations with the use of database engine. Our proposed and examined method concentrated on pure SQL code as an implementation language. We used the code to solve the model for real Internet topology and the characteristics of obtained results, concerning the network congestion changes, were presented with the use of visualization tool. The preliminary results confirm that certain analytical models can be implemented within SQL language, thus the databases can be used not only to store the results but to calculate them as well.

The time needed for the load of the initial configuration data is surprisingly small as well as the time for extracting the results in appropriate format acceptable by Gephi (several seconds). Thus, once the data are generated, we have wide spectrum of tools and algorithms to mine the knowledge about the network loads. Generally, the main advantages of the our solution are its scalability, portability and universality. However, further studies on this subject are still necessary.

The next step of the investigations will focus on further optimization of the developed SQL algorithms, based on a set of time performance tests, both on column and row store, with the use of inserts, updates and information models in case of aggregates. We would like to model other topologies also based on CAIDA measurements. Additionally, we plan to compare the performance of our approach with competitive OLTP engines, graph databases solutions and a multiprocessor native implementation.

**Acknowledgments** This work was supported by a grant *Modelling wide area networks using SAP HANA in-memory database* of Hasso-Plattner-Institute (HPI) in Potsdam, Germany.

## References

1. Czachórski, T., Nycz, M., Nycz, T., Pekergin, F.: Analytical and numerical means to model transient states in computer networks. In: 20th International Conference, CN 2013, Lwówek Śląski, Poland, June 17–21, 2013. Springer Proceedings Series: Communications in Computer and Information Science, Vol. 370, pp. 426–435, ISBN:978-3-642-38864-4
2. Gelenbe, E.: On approximate computer systems models. *J. ACM* **22**(2), 261–269 (1975)
3. Gelenbe, E., Pujolle, G.: The behaviour of a single queue in a generalqueueing network. *Acta Informatica* **7**(2), 123–136 (1976)
4. Gephi: The open graph viz platform. <http://gephi.github.io/>
5. Liu, Y., Presti, F.L., Misra, V., Towsley, D., Gu, Y.: Fluid models and solutions for large-scale ip networks. *ACM/SigMetrics* (2003)
6. Lyon, B.: The opte project. <http://www.opte.org/>
7. Lyon, B.: The opte project—internet 2003. <http://blyon.com/blyon-cdn/opte/maps/static/1069646562.LGL.2D.4000x4000.png>
8. Misra, V., Gong, W.-B., Towsley, D.: A fluid-based analysis of a network of aqm routers supporting tcp flows with an application to red. In: Proceedings of the Conference on Applications, Technologies, Architectures and Protocols for Computer Communication (SIGCOMM 2000), pp. 151–160 (2000)
9. SAP HANA. <http://hana.sap.com/>
10. SAP HANA developer guide. [http://help.sap.com/hana/sap\\_hana\\_developer\\_guide\\_en.pdf](http://help.sap.com/hana/sap_hana_developer_guide_en.pdf)
11. Wiczorek, B., Polonski, M., Pecka, P., Deorowicz, S.: An effective way of storing and accessing very large transition matrices using multi-core CPU and GPU architectures. In: Man-Machine Interactions, Springer, Seria Advances in Intelligent and Soft Computing (2014)

# Numerically Efficient Analysis of a One-Dimensional Stochastic Lac Operon Model

Neslihan Avcu, Nihal Pekergin, Ferhan Pekergin and Cüneyt Güzeliş

**Abstract** Gene expression models and their analysis play a key role to understand gene regulation mechanisms. The lac operon mechanism has been largely studied to analyze its bistable behavior. In this paper a stochastic quasi steady-state lac operon model which is indeed a one dimensional birth-death process is considered. Nevertheless the well known closed-form solutions, due to the nonlinearity of parameters, the intermediate computed values become out of the representation range with the increase of the state space size. An aggregation-based two step algorithm is proposed to compute the steady-state distribution efficiently. The results of the stochastic model give the same parameter range for the bistable behavior as with the deterministic ODE model.

## 1 Introduction

In gene regulatory networks, the regulation of the gene expression by the products of other genes, proteins, constitutes complex interaction pathways between network elements and the analysis of these network dynamics is one of the most important steps to understand and to control of these interactions. For the network

---

N. Avcu (✉)  
Department of Electrical-Electronics Engineering,  
Dokuz Eylül University, İzmir, Turkey  
e-mail: neslihan.avcu@deu.edu.tr

N. Pekergin  
Université Paris-Est Créteil, LACL, Créteil, France  
e-mail: nihai.pekerkin@u-pec.fr

F. Pekergin  
Université Paris Nord, Sorbonne Paris Cité, LIPN, CNRS, Villetaneuse, France  
e-mail: ferhan.pekerkin@lipn.univ-paris13.fr

C. Güzeliş  
Department of Electrical-Electronics Engineering, Izmir University of Economics,  
İzmir, Turkey  
e-mail: cuneyt.guzelis@ieu.edu.tr

interaction comprehension and its behavior prediction, lots of modeling and analysis approaches based on different mathematical methods are developed in the literature. These approaches change from qualitative to quantitative, continuous to discrete and deterministic to stochastic frameworks.

The growing interest in stochastic models for biochemical reaction systems has lead to several works for the design of efficient solution methods. Most of the stochastic models are in discrete-state and continuous time as proposed by Gillespie [4]. The usual way of analyzing these models is simulation, i.e. the generation of large numbers of sample trajectories and then the estimation of probabilistic behaviors by statistical methods. The main drawback of this method is the high number of simulation runs to be able to retrieve estimations within acceptable confidence intervals. Moreover the simulation time of a trajectory may be very long in the case of rare events. Although the simulation is largely applied to the stochastic analysis in systems biology, the obtained results must be carefully interpreted and obtaining good estimations is in general very time-consuming.

The numerical analysis of Markov chains has been largely studied [9]. Despite of lots of works, the large state space sizes and the difference in the magnitude of parameters are still the main problems to compute the exact numerical solutions. The infinite state space cases may be considered for special structures. Several approximate methods to make numerical analysis tractable and/or to provide efficient numerical analysis have been proposed in previous studies for performance and reliability evaluation and more recently for biological system analysis [3, 7, 11]. The main idea of these studies is to take into account not the whole state space but only a portion of it. Since the probabilities are concentrated in some regions of the space these approximations provide useful insights on the underlying models [7, 11]. Even bounding models provide approximative results, they also guarantee the accuracy of results [3].

The lac operon is the most studied gene regulatory network in experimental and theoretical works due to its bistable behavior [5]. This bistability phenomenon which defines biological switching between two stable states based upon the external inputs, i.e. external glucose and lactose concentrations, is seen frequently in gene regulatory networks. Such switching mechanisms are mainly due to positive feedback creating an unstable state which actually separates two stable states caused by saturated reaction rates. To explain the bistable behavior of lac operon, different mathematical models and analysis tools have been developed since the experimental discovery in 1960s [5]. A one dimensional ODE model of lac operon has been given under quasi steady-state assumption [1, 2]. The stochastic version of this model built by considering exponentially distributed reaction times is indeed a birth-death type continuous time Markov chain. Due to the simple structure of the model, the closed-form equations of the steady-state distribution are known. However the nonlinearity of parameters, and the large state space in terms of molecule numbers result in some numerical problems such as the computed intermediate values go out of the range of representation of standard programming languages.

In this paper, an aggregation-based two steps calculation is proposed to overcome the numerical problems. As in Aggregate/Disaggregate methods [9], the state space is divided into subsystems and each subsystem is analyzed in isolation for the

conditional steady-state distributions. The subsystems are chosen in the analysis to be with the same size, however the partition of the state-space may also be made in a nonuniform way. In the second step, we consider each subsystem as an individual state, and the steady-state probabilities to be in each subsystem are calculated. The results of the original system is derived by combining the results of these two steps. The considered measure is the steady-state probability distribution to detect the bimodal behavior of the model. The calculated results are the exact values for the finite molecule number case while they provide approximative results when the infinite state space is truncated. The comparison of the results obtained from the deterministic and the stochastic one-dimensional lac operon models leads to the same parameter range for the bistable behavior.

This paper is organized as: One dimensional lac operon model is presented in Sect. 2. The calculation of the steady-state probabilities is given in Sect. 3. Section 4 is devoted to the numerical results. The conclusions are given in Sect. 5.

## 2 One Dimensional Lac Operon Model

The lac operon of *E. coli* is a gene regulatory network comprised three genes *lacZ*, *lacY* and *lacA* which respectively encode  $\beta$ -galactosidase, galactoside permease and thiogalactoside transacetylase enzymes. These three enzymes are functionary in lactose metabolism in the absence of glucose which is actually the preferable energy source. When there is no glucose in the medium, *E.coli* switches to a different living type for metabolizing lactose as the energy source. This switching mechanism, so called bistability, is due to the positive feedback in the induction of enzyme products: Under the absence of glucose, external lactose is transferred into cell via permease enzyme and then hydrolyzed by  $\beta$ -galactosidase into allolactose or glucose and galactose. The allolactose inhibits the repressor protein and hence enables to induce the production of permease enzyme in higher level of concentration. The successive increase in the permease and consequently internal lactose concentrations leads eventually saturation. The induced (ON) and uninduced (OFF) stable states of lac operon are observed to coexist at the same external lactose concentration depending on the level of external glucose. This bistability phenomenon of lac operon is widely studied in the literature both theoretically and experimentally to understand the underlying mechanism of biological switches. In experimental studies, two artificial lactose analogs, namely methyl-1-thio- $\beta$ -D-galactoside (TMG) or isopropyl- $\beta$ -D-thio-galactoside (IPTG) are used since they are non-metabolizable, making the experiment easy.

In the previous studies, an Ordinary Differential Model (ODE) with three state variables for mRNA, permease and TMG is constructed to analyze these behavior of lac operon and some critical analyses such as boundedness and local stability are employed to show that this model is sufficient and explains the dynamics of lac operon [1, 2]. This TMG induced lac operon model includes the catabolite repression and inducer exclusion effects of the extracellular glucose and the positive effect of TMG on the transcription of mRNA. The translational and transcriptional delays are

ignored in the lac operon mechanism. Also, a reduced version of ODE model in (1) is obtained under the quasi steady-state assumption by setting differential equations of mRNA and permease to zero [1, 2]. The bistability ranges for model parameters are calculated with the discriminant based bistability analysis method and root locus based bistability analysis method for the following reduced lac operon model [1, 2].

$$\frac{1}{\tilde{\gamma}_T} \frac{dT}{dt} = pf_{M,T}(T) - T \quad (1)$$

where  $p$  is a combination of the ODE model parameters such as degradation and production constants, the effects of external glucose and TMG in [1] to simplify the analysis of the model and  $f_{M,T}(T) = (1 + K_1 T^n)/(K + K_1 T^n)$  is the Hill type production function of mRNA under TMG concentration. It is an allosteric interaction [12] and  $n$  is the required number of TMG to inactivate a repressor protein,  $K_1$  is the equilibrium constant of TMG-repressor protein interaction, and  $K^{-1}$  is the basal level of mRNA transcription in *E.coli* [8, 12]. To inactivate the repressor protein, at least two TMG molecules have to bind the repressor, thus  $n$  is taken as 2.

In the present study, a stochastic model is developed to compare the results of bistability analyses. The quasi steady-state model in (1) is then a one-dimensional birth-death process. The change rate of the TMG concentration can be defined with two terms. The birth term explaining the increase of the concentration by the transition to the next state,  $\lambda_T = \tilde{\gamma}_T p \frac{1+K_1 T^2}{K+K_1 T^2}$ . The death term expressing the decrease in the concentration level by the transition to the previous state,  $\mu_T = \tilde{\gamma}_T T$ . In the steady-state analysis, the loss constant,  $\tilde{\gamma}_T$ , having only an effect in transient response of the model is neglected.

The stochastic 1-D birth-death lac operon model derived from reduced deterministic one is in *Chemical Master Equation* form as follows.

$$\frac{dPr(T, t)}{dt} = f(T-1)Pr(T-1, t) - f(T)Pr(T, t) + g(T+1)Pr(T+1, t) - g(T)Pr(T, t) \quad (2)$$

where,  $f(T) = \lambda_T$  defines the birth rate and  $g(T) = \mu_T$  defines the death rate, the state variable  $T$  represents the number of TMG molecules and is a discrete random variable taking values in  $T \in \{0, 1, \dots, m\}$ .  $m$  is the maximum number of  $T$  molecules.  $dPr(T, t)/dt$  is the change rate for the probability of TMG molecule number.

A basic transformation is applied to the deterministic model parameters into stochastic model. In deterministic model, the nominal value of  $p$  parameter is estimated as  $p = 74.59 \mu\text{M}$  from the mathematical model in [8] by taking external TMG concentration  $T_e = 20 \mu\text{M}$  and external glucose concentration  $G_e = 0 \mu\text{M}$ . It is reported that the bistable behavior of lac operon is observed for these external input concentrations in the experimental study [8]. To convert the unit into *the number of molecules*, its value is multiplied by  $N \cdot V$ , where  $N$  is Avogadro constant and  $V$  is average cell volume taken as  $1 \times 10^{-18} \text{m}^3$  similar in [6]. The nominal value of  $p$  parameter in stochastic model is calculated as 44,754 molecules. A similar

transformation is applied to  $K_1$  parameter and the value in the deterministic model,  $1 (\mu M)^{-2}$ , is multiplied by  $1/(N \cdot V)^2$  thus  $K_1$  is  $2.7594 \times 10^{-6} (\text{molecule})^{-2}$  in the stochastic one. The dimensionless  $K$  parameter is the same in both models and its value is equal to 167.1. The nominal values of  $K$  and  $K_1$  parameters in deterministic model are taken from the mathematical model in [8]. The equations of 1-D birth-death model in (2) can be written in matrix form as follows.

$$\frac{d}{dt} \begin{bmatrix} Pr(0, t) \\ Pr(1, t) \\ \vdots \\ Pr(m, t) \end{bmatrix} = \begin{bmatrix} -\lambda_0 & \mu_1 & 0 & \cdots & 0 \\ \lambda_0 & -\lambda_1 - \mu_1 & \mu_2 & \cdots & 0 \\ 0 & \lambda_1 & -\lambda_2 - \mu_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \mu_m \\ 0 & 0 & 0 & \cdots & -\mu_m \end{bmatrix} \begin{bmatrix} Pr(0, t) \\ Pr(1, t) \\ \vdots \\ Pr(m, t) \end{bmatrix}$$

There exists a unique steady-state probability distribution,  $\pi$ , for this time homogeneous, finite, irreducible Continuous Time Markov Chain (CTMC) [10]. The steady-state probability distribution is calculated by taking their time derivatives equal to 0, which lead to the following closed-form equations:

$$\pi_i = \pi_0 \prod_{j=0}^{i-1} \frac{\lambda_j}{\mu_{j+1}} \quad \text{with} \quad \pi_0 = \frac{1}{1 + \sum_{i=1}^m \prod_{j=0}^{i-1} \frac{\lambda_j}{\mu_{j+1}}} \tag{3}$$

### 2.1 Numerical Problems

Although the closed form solution in (3) has a simple structure, the nonlinear nature of parameters together with the large state space size results in numerical problems. In expression of  $\pi_i$  in (3),  $\pi_0$  must be known with a good accuracy. However, in large Markov chains with multi-modal distributions where probability concentrations are far from state 0,  $\pi_0$  may be very small. In the case of unimodal distribution, this problem can be avoided by defining the steady-state probabilities in terms of the  $\pi_M$  where  $M$  is the state having the largest probability.  $M$  can be found by checking the sign of the variations. Then the state space is divided in two regions:

$$\pi_i = \pi_M \prod_{j=M-1}^i \frac{\mu_{j+1}}{\lambda_j} \quad \text{for } i = 0, \dots, M - 1 \tag{4}$$

$$\pi_i = \pi_M \prod_{j=M+1}^{j=i} \frac{\lambda_j}{\mu_{j+1}} \quad \text{for } i = M + 1, \dots, m \tag{5}$$

where  $\pi_M$  is deduced from the normalization equation as  $\pi_M = \frac{1}{1 + \sum_{j \neq M} \pi_j}$ .

However this computation scheme which reduces the errors can not be extended to multimodal distributions if some intermediate states probabilities become so small to be represented with the used real number format. To consider null the probability values when they come close to the minimum number in the representation scale, breaks the link between the probability masses accumulated around the mode (peak) states. Thereafter, there is no more common expression in the formula of the probabilities in the different modes. Hence, the normalization is insufficient, we must have some extra relations between the probability masses standing in each mode subrange. A computing environment capable to represent much smaller values for the probabilities may overcome this kind of problems. Indeed there is some computing packages which permit to use larger format to represent very big numbers, more precisely higher than  $10^{+306}$ , and very small numbers, lower than  $10^{-306}$ . Gnu Multiple Precision Library environment (GMP) is such an open source software without any limitation for the representation of the numbers.

### 3 Calculation of Steady-State Probabilities by Aggregation

First, the notation and some explanations are given. Figure 1 shows the aggregation of the state space.

- The state space of the original birth-death process is  $\sigma = \{0, 1, \dots, m\}$  where  $m$  is the maximum state number. This system will be called also as System  $\sigma$ . The birth and death rates of State  $i$  are represented as  $\lambda_i$  and  $\mu_i$ , respectively.
- The unique transition from State 0 is to State 1 with rate  $\lambda_0$ . Similarly, there is only one transition from State  $m$  to the State  $m - 1$  with rate  $\mu_m$ .
- A partition of  $\sigma$  is defined, such that the number of subsystems is  $n$  and the number of elements of each subsystem is  $k = \lfloor (m + 1)/n \rfloor$ . Then, each subsystem contains  $k$  states:  $\sigma_0 = \{0, 1, \dots, k - 1\}$ ,  $\sigma_1 = \{k, \dots, 2k - 1\}$ ,  $\dots$ ,  $\sigma_{n-1} = \{(n - 1)k \dots m\}$ . It should be note that the aggregation method is by no means restricted to such a uniform partition. Based on some a priori information about the probability changes in the steady-state distribution, a nonuniform partition may be preferred as assigning large size macro states for flat regions while small size macro states for steep ones.

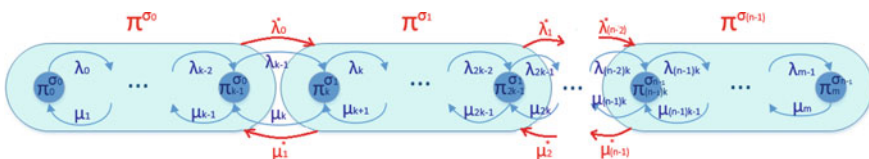


Fig. 1 The aggregated 1-D birth-death process



- Each subsystem,  $\sigma_i$ , is only connected to the previous and the following ones with the death and birth rates,  $\mu_i^*$  and  $\lambda_i^*$ , respectively. The connection between two subsystem is through the last and the first states in the consecutive subsystems.
- The ratio of the transition rates between two consecutive states  $i$  and  $i + 1$  in System  $\sigma$  and two consecutive subsystems,  $\sigma_j$  and  $\sigma_{j+1}$ , are respectively defined as  $\rho_i = \frac{\lambda_i}{\mu_{i+1}}$  for all  $i \in \sigma$  and  $\rho_j^* = \frac{\lambda_j^*}{\mu_{j+1}^*}$  for all  $j \in \{0, 1, \dots, n - 1\}$ .
- The steady-state probability to be in state  $i \in \{0, 1, \dots, m\}$  in the original system, System  $\sigma$ , is denoted by  $\pi_i^\sigma$ . The summation of all the states is unity.

$$\sum_{i=0}^m \pi_i^\sigma = 1 \tag{6}$$

It is assumed that the probabilities of the other states,  $j \notin \{0, 1, \dots, m\}$ , are equal to zero.

- Also, the steady-state of the conditional probability to be in state  $i \in \sigma_j$  is represented as  $\pi_i^{\sigma_j} = Prob(X(\infty) = i | X(\infty) \in \sigma_j)$ , where  $X(t)$  is the state of the underlying birth-death process at time  $t$ . The conditional steady-state probabilities in each subsystem are calculated by isolating the related subsystem from the System  $\sigma$  and assuming that the summation of the probabilities of all the states in  $\sigma_j$  is equal to 1.

$$\sum_{i=j \cdot k}^{(j+1) \cdot k - 1} \pi_i^{\sigma_j} = 1 \text{ for } j \in \{0, 1, \dots, n - 1\} \tag{7}$$

- Let  $\pi^{\sigma_j}$ ,  $j \in \{0, \dots, n - 1\}$  be the steady-state probability that the system is in one of the states of System  $\sigma_j$ , which is equal to the summation of all state probabilities in  $\sigma_j$ :

$$\pi^{\sigma_j} = \sum_{i=j \cdot k}^{(j+1) \cdot k - 1} \pi_i^{\sigma_j} \tag{8}$$

- All the subsystems cover the whole states  $i \in \sigma$  and satisfy the given equality condition in (9).

$$\sum_{j=0}^{n-1} \pi^{\sigma_j} = 1 \tag{9}$$

- The steady-state probability  $\pi_i^\sigma$  of State  $i$  for  $i \in \sigma_j$  can be written in terms of the steady-state probability of subsystem  $\sigma_j$  and the conditional steady-state probability of State  $i$  known to be in subsystem  $\sigma_j$  for  $j \in \{0, 1, \dots, n - 1\}$

$$\pi_i^\sigma = \pi^{\sigma_j} \cdot \pi_i^{\sigma_j} \tag{10}$$

### 3.1 Algorithm

In the first step, the conditional steady-state probability  $\pi_i^{\sigma_j}$  of each state  $i$  for every subsystem is calculated. In the second step, the conditional probability  $\pi^{\sigma_j}$  of each subsystem is computed and then, the steady-state probability  $\pi_i^\sigma$  of the state  $i$  in system  $\sigma$  is obtained in terms of these two probabilities,  $\pi_i^{\sigma_j}$  and  $\pi^{\sigma_j}$ .

*Step I* The conditional steady-state probabilities of each state in subsystem  $\sigma_j$  for  $j = \{0, \dots, n - 1\}$  are computed by considering in isolation subsystem  $\sigma_j$  according to the local balance equations. In a subsystem  $\sigma_j$ , the local balance equation between two consecutive states,  $i$  and  $i + 1$ , is written as in (11).

$$\pi_{i+1}^{\sigma_j} = \rho_i \cdot \pi_i^{\sigma_j} \quad \text{for } i \in \{j \cdot k, \dots, (j + 1) \cdot k - 2\} \quad (11)$$

- The summation of the conditional probability is equal to 1.

$$\sum_{i \in \sigma_j} \pi_i^{\sigma_j} = 1 \quad (12)$$

- By using (12), the conditional steady-state probability  $\pi_{j \cdot k}^{\sigma_j}$  of the first states in each subsystem  $\sigma_j$  is determined as follows:

$$\pi_{j \cdot k}^{\sigma_j} = \frac{1}{1 + \rho_{j \cdot k} + \rho_{j \cdot k} \cdot \rho_{j \cdot k + 1} + \dots + \rho_{j \cdot k} \dots \rho_{(j+1) \cdot k - 2}} \quad (13)$$

- After calculation of the conditional steady-state probability of the first state, the other conditional probabilities can be calculated iteratively as in (11).

*Step II* Similar calculations in Step I are performed to compute the steady-state probability  $\pi^{\sigma_j}$  of each subsystem  $\sigma_j$ . By assuming each subsystem as an individual state, the local balance equation between each consecutive subsystems,  $\sigma_j$  and  $\sigma_{j+1}$ , is defined as follows for  $j \in \{0, 1, \dots, n - 2\}$ :

$$\pi^{\sigma_{j+1}} = \rho_j^* \cdot \pi^{\sigma_j} \quad (14)$$

- When the derived steady-state probability of subsystem  $\sigma_j$  from (10) is substituted in (14), transition rates  $\mu_j^*$  and  $\lambda_j^*$  of the subsystem  $\sigma_j$  become functions of the conditional steady-state probabilities of the first and the last states in these subsystems, for  $j \in \{0, 1, \dots, n - 1\}$ .

$$\mu_j^* = \mu_{j \cdot k} \cdot \pi_{j \cdot k}^{\sigma_j} \quad (15)$$

$$\lambda_j^* = \lambda_{(j+1) \cdot k - 1} \cdot \pi_{(j+1) \cdot k - 1}^{\sigma_j} \quad (16)$$

- Again by using (9), the steady-state probability of the first subsystem is calculated.

$$\pi^{\sigma_0} = \frac{1}{1 + \rho_0^* + \rho_0^* \cdot \rho_1^* + \dots + \rho_0^* \dots \rho_{n-2}^*} \tag{17}$$

- The steady-state probability of each subsystem  $\sigma_j$  can be computed iteratively by substituting the steady-state probability of the first subsystem,  $\sigma_0$ , into (14).

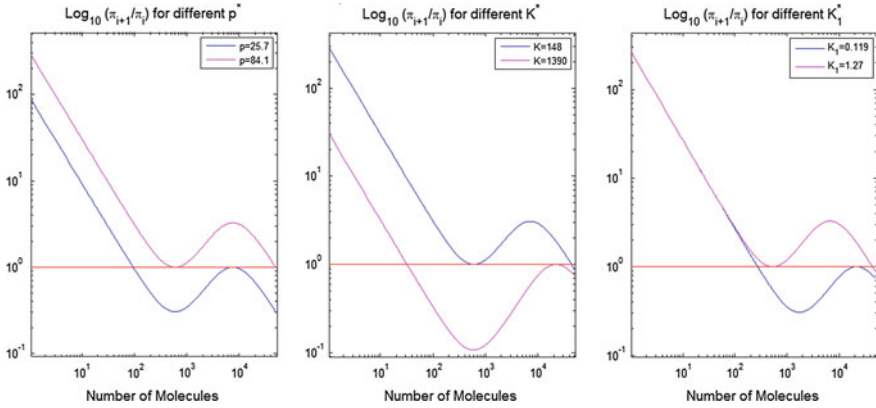
Finally, the steady-state probability distribution  $\pi_i^\sigma$  of system  $\sigma$  is obtained according to (10) after calculations in these two steps.

The above algorithm attempts to tackle the numerical problems due to large state space sizes for nonlinear birth-death processes in a two-level scheme which works if the number of meta-states and also the sizes of the meta-states are not too large. For too large state space sizes, the proposed algorithm can be implemented in a scheme employing more than two levels to obtain numerically tractable number and sizes of meta-states in each level. Such a multi-level approach is seen to solve the mentioned numerical problems due to large state space sizes in a definite way.

### 4 Comparison of Deterministic and Stochastic Models

The bistability parameter ranges are determined by discriminant based and also by root locus analysis for the deterministic 1-D reduced lac operon model in [1, 2]. When the calculation is made for one parameter, the other two parameters are taken in nominal values. The bistability parameter range is found as  $p \in (25.7, 84.1 \mu\text{M})$ ,  $K \in (148, 1390)$  and  $K_1 \in (0.119 (\mu\text{M})^{-2}, 1.27 (\mu\text{M})^{-2})$  in the deterministic model [1, 2]. These ranges are determined by capturing the boundary limits either with a discriminant analysis [1] specifying the changes of the roots of the equilibrium equation or with a root-locus analysis providing the break in/away points [2].

In the stochastic model, the units of parameters are transformed into a suitable form and then the aggregation method is applied to determine the bistability ranges of model parameters one by one. In the analysis of  $p$  parameter range, the nominal values of  $K = 167.1$  and  $K_1 = 1 (\mu\text{M})^{-2}$  are considered. The bistability range of  $p$  parameter is computed as  $p \in (15,420, 50,460 \text{ molecules})$  in the stochastic model. This range is obtained by checking the bimodality of the steady-state distribution for a sufficiently large set of  $p$  parameter values. When the molecule numbers are converted into molar concentration, it is seen that the bistability range of  $p$  parameter obtained in the stochastic model is the same with the deterministic ones of [1, 2]. The bistability ranges for  $K$  and  $K_1$  parameters are computed in a similar way. The bistability range of  $K$  parameter is obtained as  $K \in (148, 1390)$ . The  $K$  parameter is unitless and its bistability range is found the same for the deterministic and stochastic models. The obtained bistability range of  $K_1$  parameter is  $K_1 \in (3.056 \times 10^{-7} (\text{molecules})^{-2}, 3.5278 \times 10^{-7} (\text{molecules})^{-2})$ . After the conversion, it is observed that the bistability range of  $K_1$  parameter in both models



**Fig. 2** Bistability ranges for model parameters

is computed as the same. Due to the small values of the steady-state probabilities, these values are given in the logarithmic scale. In Fig. 2, the ratios of the steady-state probabilities of the consecutive states are plotted in the logarithmic scale. For the limit values of parameters, the steady-state distributions of the model have unimodal shapes (Fig. 2). In the bistability ranges of parameters, the steady-state probability distributions become bimodal as in the deterministic model.

## 5 Conclusion

Some numerical problems are encountered in the numerical analysis of large, one dimensional birth-death processes with nonlinear transition rates such as the reduced lac operon stochastic model. An aggregation-based two step algorithm to compute steady-state probabilities is proposed to make possible the numerical analysis of such systems with the standart programming languages lacking of multiple precision feature. The comparison of deterministic and stochastic lac operon models yields to the same parameter ranges for the bistable behavior. This approach may be extended to other special structures such as multi-dimensional birth-death processes.

**Acknowledgments** This work was supported in part by the Turkish Scientific and Technological Research Council and the Centre National des Recherches Scientifiques in the framework of the Bosphorus PIA Program as a joint research project with grant number 111E082.

## References

1. Avcu, N., Demir, G.K., Pekergin, F., Alyürük, H., Çavaş, L., Güzeliş, C.: Discriminant based bistability analysis of a TMG induced lac operon model supported with boundedness and local stability results. *Turkish J. Elec. Eng. Comput. Sci.* 1–11 (2013). doi:[10.3906/elk-1305-264](https://doi.org/10.3906/elk-1305-264)
2. Avcu, N., Alyürük, H., Demir, G.K., Pekergin, F., Cavas, L., Güzeliş, C.: Determining the bistability parameter ranges of artificially induced lac operon using the root locus method. *Comput. Biol. Med.* **61**, 75–91 (2015). doi:[10.1016/j.combiomed.2015.03.009](https://doi.org/10.1016/j.combiomed.2015.03.009)
3. Dayar, T., Hermanns, H., Spieler, D., Wolf, V.: Bounding the equilibrium distribution of markov population models. *Numer. Linear Algebra Appl.* **18**, 931–946 (2011)
4. Gillespie, D.T.: Stochastic simulation of chemical kinetics. *Annu. Rev. Phys. Chem.* **58**, 35–55 (2007)
5. Jacob, F., Perrin, D., Sanchez, C., Monod, J.: Lac Operon: groupe de gène à expression par un operator. *C. R. Acad. Sci.* **250**, 1727–1729 (1960)
6. Lestas, I., Paulsson, J., Ross, N.E., Vinnicombe, G.: Noise in gene regulatory networks. *IEEE Trans. Autom. Control* **53**, 189–200 (2008)
7. Munsy, B., Khammash, M.: The finite state projection algorithm for the solution of the chemical master equation. *J. Chem. Phys.* **124**, 201–214 (2006)
8. Özbudak, E.M., Thattai, M., Lim, H.N., Shraiman, B.I., van Oudenaarden, A.: Multistability in the lactose utilization network of *Escherichia coli*. *Nature* **427**, 737–740 (2004)
9. Stewart, W.J.: *Introduction to the numerical Solution of Markov Chains*. Princeton University Press, New Jersey (1994)
10. Trivedi, K.S.: *Probability and statistics with reliability, queuing and computer science applications*. Wiley, Chichester (2001)
11. Wolf, V., Goel, R., Mateescu, M., Henzinger, T.A.: Solving the chemical master equation using sliding windows. *BMC Syst. Biol. J.* **4**, 42 (2010)
12. Yagil, G., Yagil, E.: On the relation between effector concentration and the rate of induced enzyme synthesis. *Biophys. J.* **11**, 11–27 (1971)

**Part VI**  
**Image Processing**  
**and Computer Vision**

# Brain MR Image Denoising for Rician Noise Using Intrinsic Geometrical Information

Hamit Soyel, Kamil Yurtkan, Hasan Demirel and Peter W. McOwan

**Abstract** A new image denoising algorithm based on nonsubsampling contourlet transform is presented. Magnetic Resonance (MR) images corrupted by Rician noise are transformed into multi-scale and multi-directional contour information, where a nonlinear mapping function is used to modify the contour coefficients at each level. The denoising is achieved by improving edge sharpness and inhibiting the background noise. Experiments show the proposed algorithm preserves the intrinsic geometrical information of the noised MR image and can be effectively applied to T1-, T2-, and PD-weighted MR images without any parameter tuning under diverse noise levels.

## 1 Introduction

By analysing magnetic resonance (MR) images of the brain, surgeons can make an appropriate decision for diagnosing many neurological diseases, such as Parkinsons disease, Alzheimers disease, brain tumors, and stroke. However, MR images are affected by several artifacts and noise sources. One of them is the random fluctuation of the MR signal which is mainly due to thermal noise. Such a noise seriously degrades the acquisition of any quantitative measurements from the MR images, such as registration, segmentation, classification, and visualization. To obtain reliable

---

H. Soyel (✉) · P.W. McOwan  
Queen Mary University of London, London, UK  
e-mail: h.soyel@qmul.ac.uk

P.W. McOwan  
e-mail: p.mcowan@qmul.ac.uk

K. Yurtkan  
Cyprus International University, Lefkoşa, Mersin Turkey  
e-mail: kyurtkan@ciu.edu.tr

H. Demirel  
Eastern Mediterranean University, Gazimağusa, Mersin Turkey  
e-mail: hasan.demirel@emu.edu.tr

analysis results, denoising MR image is vital before further analysis steps can be conducted. MR images are computed from both real and imaginary images, which are assumed to contain Gaussian distribution noises with zero means. Thus, noise is image dependent, which follows a Rician distribution, and makes removing noises difficult [12].

Postprocessing filtering techniques have the advantage of not increasing the acquisition time and, hence have been extensively used in MR image denoising. Most denoising methods are based on the signal averaging principle by using the spatial pattern redundancy in the image. However, there are other filtering techniques that use other statistical estimates such as the Median [10], Gaussian [1] and Anisotropic Diffusion [16] that have received considerable attention due to their simplicity. Although these denoising methods can significantly remove noise, they tend to cause blurring and erase small features if there are high peaks in the histogram, often resulting a mask effect in uniform regions and other artifacts in the output image.

One way to approaching this problem is to use multi-scale image decomposition, that is, processing images in each scale independently and recombining each processed image to obtain the final image. Advances in wavelet theory combined with multi-scale analysis applied to image contrast enhancement can achieve promising results. The decomposition of images into different frequency ranges permits the isolation of the frequency components introduced by intrinsic deformations or extrinsic factors in certain subbands [18].

In [13], the undecimated wavelet transform is employed to provide effective representation of the noisy coefficients information. However, the 2D wavelet transform used is a separable extension of the 1D wavelet transform, which does not work well in capturing the geometry of image edges [14].

We propose a new image enhancement method based on the nonsubsampled contourlet transform (NSCT) [6]. The proposed algorithm enhances the MR image while amplifying weak edges and suppressing noise by modifying the NSCT coefficients using a nonlinear mapping function in each directional subband.

The rest of this paper is organized as follows. In Sect. 2, we explain the intrinsic geometrical information based methodology for image denoising. A variety of experimental results are presented in Sect. 3. Finally, the conclusions are given in Sect. 4.

## 2 Methodology

Existing image-denoising methods amplify noise when they amplify weak edges since they cannot distinguish noise from weak edges. In the frequency domain, both weak edges and noise produce low-magnitude coefficients. Since weak edges are geometric structures and noise is not, we can use the NSCT to distinguish them [17].

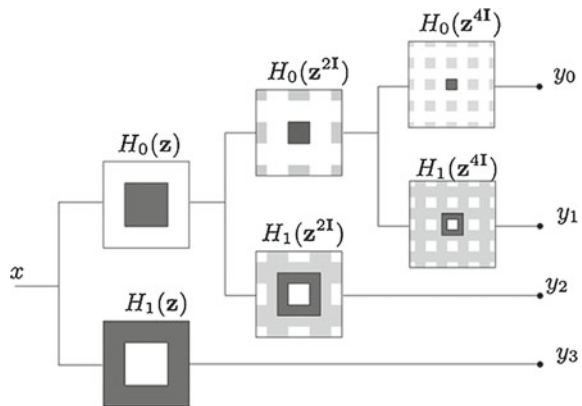


### 2.1 Nonsampled Countourlet Transform

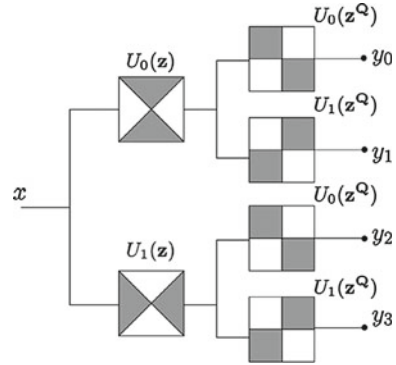
In [7], Do and Vetterli proposed the contourlet transform (CT), which is a multiscale directional representation constructed in the discrete grid by combining the Laplacian pyramid and the directional filter bank. Due to its directionality and anisotropy CT can represent curve more sparsely. NSCT, based on the theory of CT, is a kind of multi-scale, multi-directional computation framework of discrete images. The main difference lies in that, in the course of decomposition and reconstruction, traditional upsamplers and downsamplers in CT do not exist in NSCT any more, so that the NSCT is a fully multi-scale, multi-directional, good time-frequency property and shift invariant expansion.

NSCT falls into two phases, including nonsampled pyramid (NSP) and non-sampled directional filter bank (NSDFB). The former phase ensures the multi-scale property by using two-channel nonsampled filter bank, and a low-pass image with a band-pass one can be produced at each NSP decomposition stage. The subsequent NSP decomposition stages are carried out to decompose the low-pass component available iteratively to capture the singularities in the image. As a result, NSP can result in  $k + 1$  subband images including one low-pass image and  $k$  band-pass images, whose sizes are all the same as that of the source image, where  $k$  denotes the number of decomposition stages. Figure 1 gives the NSP decomposition with  $k = 3$  stages with a 2-D low-pass filter is represented by its  $z$ -transform  $H_0(z)$  where  $z = [z_1, z_2]^T$  and  $H_1(z) = 1 - H_0(z)$ . The filters for subsequent stages are obtained by upsampling the filters of the first stage. This gives the multi-scale property without the need for additional filter design. The NSDFB, constructed by combining the directional fan filter banks devised by Bamberger and Smith [2], is two-channel nonsampled filter banks. NSDFB allows the direction decomposition with  $l$  stages in band-pass images from NSP at each scale and produces  $2l$  directional subband images which have the same size as the source image. Thus, the NSDFB endows the NSCT with the multi-direction property and we can benefit a lot from

**Fig. 1** Three-stage NSP decomposition



**Fig. 2** Four-channel nonsampled DFB



the NSDFB because it provides us with more precise directional detail information. A four-channel NSDFB, constructed with two-channel fan filter banks and parallelogram filters without downsamplers and upsamplers, is illustrated in Fig. 2. Note that in the second level, the upsampled fan filters  $U_i(z^Q), i = 0, 1$  have checker-board frequency support, and when combined with the filters in the first level give the four directional frequency decomposition. We use the ‘maxat’ filters and ‘dmaxat7’ filters for NSP and NSDFB, respectively. The concrete filter banks construction methods and more NSCT details can be found in [6].

### 2.2 Proposed Image Denoising Using Intrinsic Geometrical Information

NSCT differs from other multi-scale analysis methods in that contourlet transform allows for different and flexible number of directions at each scale. According the direction information contours are obtained by directional filter bank concatenated the neighbouring singular points into local contours in the frequency domain. By combination of NSP and NSDFB, NSCT is constructed as a fully shift invariant, multi-scale, and multi-direction expansion that has better directional frequency localization and a fast implementation. It is worth noting that shift invariance is very important. Being shift invariant, each pixel of the transform subbands corresponds to that of the original image in the same spatial location. Therefore, we gather the geometrical information pixel by pixel from the NSCT coefficients. All directional contour subbands can be expressed as:

$$\{C_{m,d}\}, m = 1, 2, \dots, k, d = 1, 2, \dots, l_m, k \in (1, 2, \dots, N), l_m = 2^N. \tag{1}$$

where  $m$  and  $d$  are the scale and direction of the decomposition respectively,  $k$  is the number of contour decomposition scale,  $l_m$  is the number of contour decomposition directions of  $m$ th scale and  $\{C_{m,d}\}$  is the coefficient at the  $d$ th directional subband of the  $m$ th scale. We observe that in the NSCT domain, the Rician noise corresponds to those pixels with small magnitude coefficients in all directional subbands at a specific pyramidal level. Based on this observation, we can classify pixels into two categories by analysing the distribution of their coefficients in different subbands. One simple way is to compute the mean (denoted by  $mean$ ) and the maximum (denoted by  $max$ ) magnitude of the coefficients for each pixel across directional subbands, and then classify it by:

$$pixel = \begin{cases} noise, & \text{if}((mean < \alpha\sigma\gamma) \text{ and } (max < \alpha\sigma\gamma)) \\ edge, & \text{otherwise} \end{cases} \tag{2}$$

where  $\alpha$  is the amplifying gain of the subbands at a specific pyramidal level (3),  $\sigma$  is the noise variance of the input MR image and  $\gamma$  is the noise standard deviation of the subbands at a specific pyramidal level.

$$\alpha = \frac{\log\left(\frac{mean(|C_{m,d}|)}{\max(|C_{m,d}|)}\right)}{\log\left(\sin\left(\frac{\pi}{2} \frac{mean(|C_{m,d}|)}{\max(|C_{m,d}|)}\right)\right)} \tag{3}$$

We first estimate the noise variance of the input MR image,  $\hat{\sigma}$ , with the robust median operator defined in [4] and then compute the noise variance of each subband,  $\hat{\gamma}$ , by using [18]. The core requirement for image denoising by multi-scale transforms is to suppressing noise while amplifying weak edges. To this end we modify the NSCT coefficients,  $\tilde{C}_{m,d}$ , by a nonlinear mapping function in each scale and direction where  $m > 1$ :

$$\tilde{C}_{m,d} = \begin{cases} \left(\frac{\max(|C_{m,d}|)}{\text{sgn}(C_{m,d})}\right) \sin\left(\frac{\pi}{2} \sin\left(\frac{\pi}{2} \frac{|C_{m,d}|}{\max(|C_{m,d}|)}\right)^\alpha\right)^{\sqrt{\alpha}}, & \text{edge} \\ 0, & \text{noise.} \end{cases} \tag{4}$$

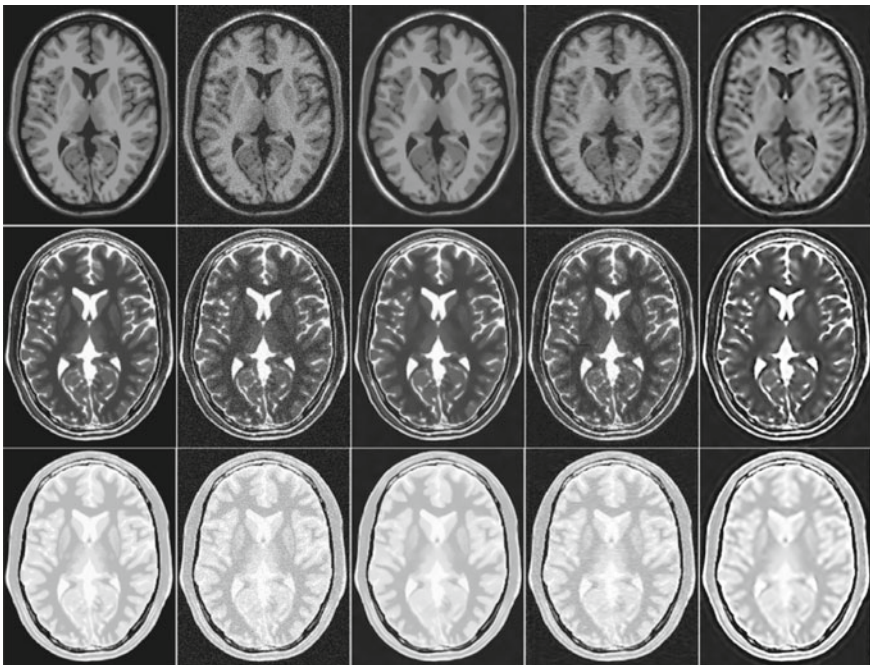
We summarize our denoising method in the following algorithm.

1. Compute the NSCT of the input MR image for  $k$  levels according to (1).
2. Estimate the noise standard deviation ( $\hat{\sigma}$ ) of the input image.

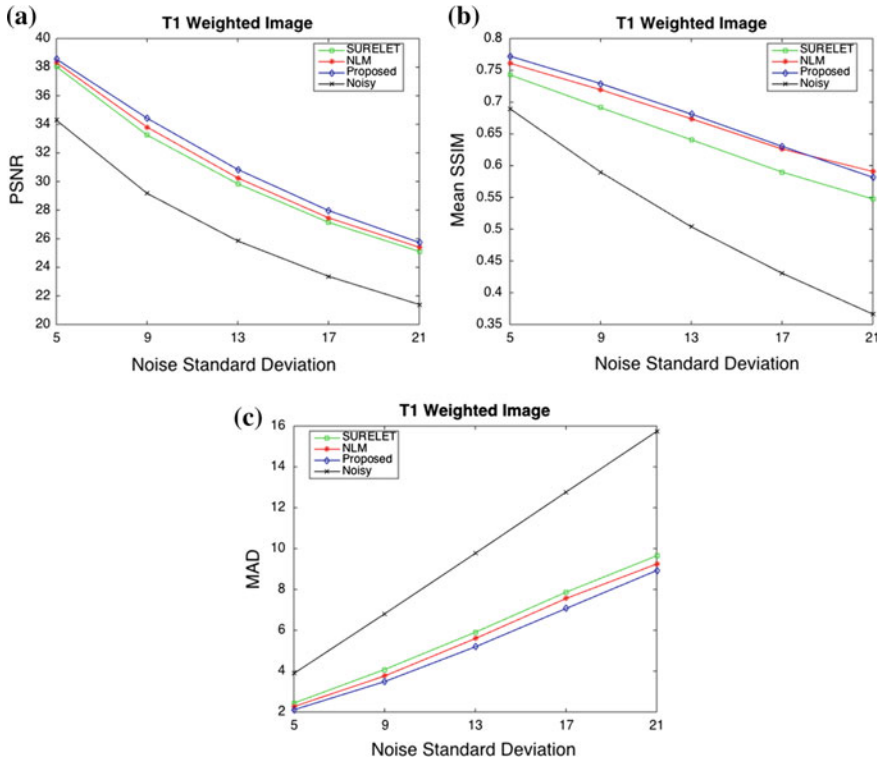
3. For each level of the pyramid where  $m > 1$ :
  - a. Estimate the noise variance ( $\hat{\gamma}$ ).
  - b. At each pixel location, compute the mean and the maximum magnitude of the corresponding coefficients in all directional subbands at this level, and classify each pixel according to (2) and (3) into “edges” or “noise”.
  - c. For each directional subband, use the nonlinear mapping function given in (4) to modify the NSCT coefficients according to the classification.
4. Reconstruct the denoised MR image from the modified NSCT coefficients.

### 3 Experimental Results

We have compared the performance of our proposed image denoising using intrinsic geometrical information with two state-of-the-art denoising algorithms, the non-local mean (NLM) algorithm [3] and a modern wavelet-based denoising algorithm (SURELET) [11] using the 3D-simulated MR images downloaded from the



**Fig. 3** Visual quality comparison: *Top row*; from *left to right*: **a** Ground truth T1-weighted image. **b** Ground truth corrupted with Rician noise of 13%. Denoised with **c** NLM, **d** SURELET and **e** the proposed method. *Middle row*; and *bottom row*: same results for a simulated T2- and PD-weighted images

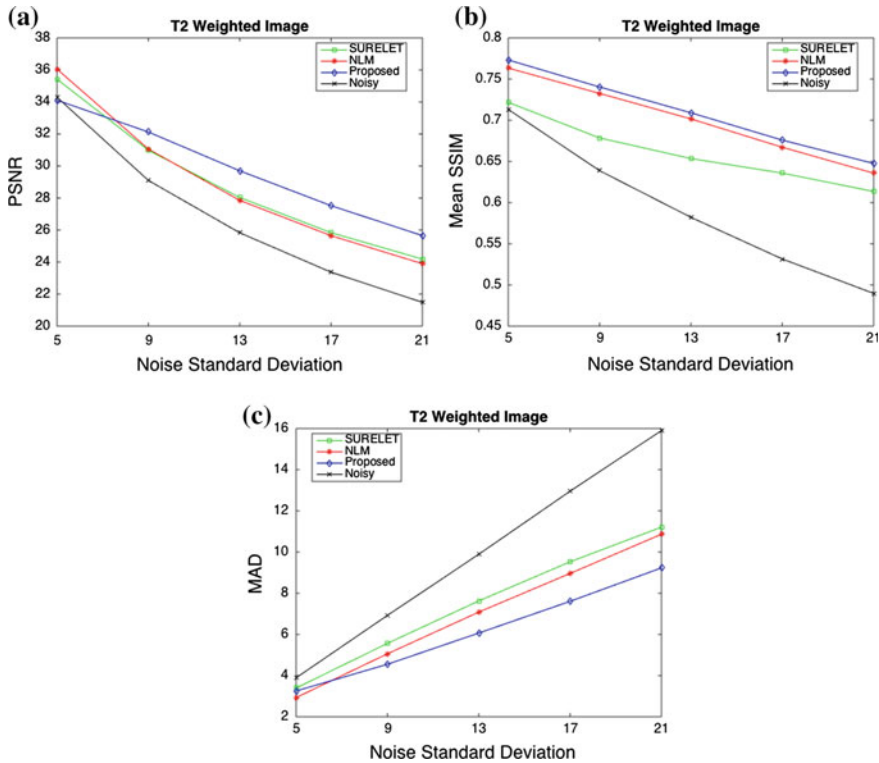


**Fig. 4** Quantitative comparison of the proposed NSTC based denoising method with NLM and SURELET methods based on **a** PSNR, **b** SSIM and **c** MAD for simulated T1 images corrupted with Rician noise varying from 5 to 21

BrainWeb database [5]. The simulated T1-, T2-, and PD-weighted synthetic noise-free MR images are downloaded where the size of each image is  $181 \times 217 \times 181$  and the intensity is 256 bins.

Visual quality comparison of the MR images denoised with the NLM, the SURELET and the proposed method was conducted on the simulated T1-, T2-, and PD-weighted synthetic noise-free MR images after corrupting the images by Rician noise [8] with 13%. It can be observed from Fig. 3 that the images denoised with the proposed method are more closer to the ground truth than the images denoised with the other approaches.

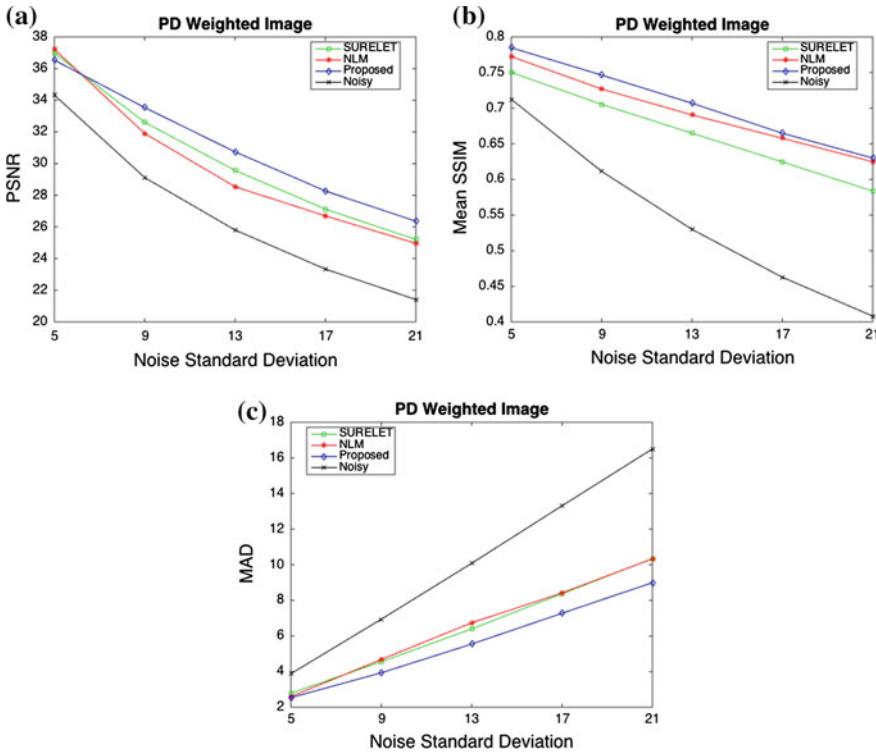
For a quantitative analysis, MR images were degraded with Rician noise for a wide range of noise levels and the denoising efficiency of the algorithms was evaluated based on the Peak Signal to Noise Ratio (PSNR) [9], the mean Structural Similarity Index Matrix (SSIM) [19] and the Mean Absolute Difference (MAD) [15]. A higher



**Fig. 5** Quantitative comparison of the proposed NSTC based denoising method with NLM and SURELET methods based on **a** PSNR, **b** SSIM and **c** MAD for simulated T2 images corrupted with Rician noise varying from 5 to 21

PSNR, MSSIM and a lower MAD correspond to a higher performance. Figures 4, 5 and 6 show the quantitative analysis of the denoising methods in terms of PSNR, mean SSIM and MAD. It can be observed from the plots that the proposed method performs better than the NLM and SURELET based on the quality metrics used for the evaluation. Furthermore, proposed methodology does not require any parameter selection and on a PC with 2.5 GHz Intel Core i5 CPU and 8 GB RAM, NSCT costs on average of 5.23 s to decompose a  $181 \times 217$  MR image into directional subbands and to reconstruct denoised MR image using Matlab<sup>1</sup> hence, can be easily applied for a clinical application.

<sup>1</sup>The nonsubsamped contourlet toolbox used in this paper can be downloaded at <http://www.mathworks.com/matlabcentral/fileexchange/10049-nonsubsamped-contourlet-toolbox>.



**Fig. 6** Quantitative comparison of the proposed NSTC based denoising method with NLM and SURELET methods based on **a** PSNR, **b** SSIM and **c** MAD for simulated PD images corrupted with Rician noise varying from 5 to 21

### 4 Conclusion

We have proposed a novel multi-scale image denoising algorithm to remove Rician noise from MR images based on the NSCT. We have shown that the new algorithm can sufficiently remove Rician noise, while simultaneously preserving edges and fine structures in a given noisy MR image. Experimental results on the BrainWeb database verified that the intrinsic geometrical information extracted from directional contour coefficients is quite efficient, speedy and can be effectively applied to T1-, T2-, and PD-weighted MR images without any parameter tuning under diverse noise levels.

## References

1. Ashburner, J., Friston, K.J.: Voxel-based morphometry—the methods. *NeuroImage* **11**(6), 805–821 (2000)
2. Bamberg, R.H., Smith, M.J.T.: A filter bank for the directional decomposition of images: theory and design. *IEEE Trans. Sig. Process.* **40**(4), 882–893 (1992)
3. Buades, A., Coll, B., Morel, J.M.: A non-local algorithm for image denoising. *Computer Vision and Pattern Recognition. CVPR 2005. IEEE Computer Society Conference on* 2, 60–65 (2005)
4. Chang, S., Yu, B., Vetterli, M.: Spatially adaptive wavelet thresholding with context modeling for image denoising. *IEEE Trans. Image Process.* **9**(9), 1522–1531 (2000)
5. Cocosco, C.A., Kollokian, V., Kwan, Evans, A.C.: BrainWeb: Online interface to a 3D MRI simulated brain database. *NeuroImage* **5**(4) (1997)
6. da Cunha, A., Zhou, J., Do, M.: The nonsubsampling contourlet transform: Theory, design, and applications. *IEEE Trans. Image Process.* **15**(10), 3089–3101 (2006)
7. Do, M., Vetterli, M.: The contourlet transform: An efficient directional multiresolution image representation. *IEEE Trans. Image Process.* **14**(12), 2091–2106 (2005)
8. Gudbjartsson, H., Patz, S.: The Rician distribution of noisy MRI data. *Magn. Reson. Med.* **34**, 910–914 (1995)
9. Huynh-Thu, Q., Ghanbari, M.: Scope of validity of PSNR in image/video quality assessment. *Electron. Lett.* **44**(13), 800–801 (2008)
10. Livin, M., Luthon, F., Keeve, E.: Entropic estimation of noise for medical volume restoration. *Proceedings of the 16th International Conference on Pattern Recognition*, **16**(3), 871–874 (2002)
11. Luisier, F., Blu, T., Unser, M.: A new SURE approach to image denoising: Interscale orthonormal wavelet thresholding. *IEEE Trans. Image Process.* **16**(3), 593–606 (2007)
12. Manjri, J.V., Carbonell-Caballero, J., Lull, J.J., Garcia-Mart, G., Mart-Bonmat, L., Robles, M.: MRI denoising using non-local means. *Med. Image Anal.* **12**(4), 514–523 (2008)
13. Pizurica, A., Philips, W., Lemahieu, I., Acheroy, M.: A versatile wavelet domain noise filtration technique for medical imaging. *IEEE Trans. Med. Imaging* **22**(3), 323–331 (2003)
14. Po, D.Y., Do, M.: Directional multiscale modeling of images using the contourlet transform. *IEEE Trans. Image Process.* **15**(6), 1610–1620 (2006)
15. Rajan, J., Jeurissen, B., Verhoye, M., Audekerke, J.V., Sijbers, J.: Maximum likelihood estimation-based denoising of magnetic resonance images using restricted local neighborhoods. *Phys. Med. Biol.* **56**(16), 5221 (2011)
16. Samsonov, A., Johnson, C.: Noise-adaptive nonlinear diffusion filtering of MR images with spatially varying noise levels. *Magn. Reson. Med.* **52**(4), 798–806 (2004)
17. Soyel, H., McOwan, P.: Automatic image enhancement using intrinsic geometrical information. *Electron. Lett.* **48**(15), 917–919 (2012)
18. Starck, J.L., Candes, E., Donoho, D.: The curvelet transform for image denoising. *IEEE Trans. Image Process.* **11**(6), 670–684 (2002)
19. Wang, Z., Bovik, A., Sheikh, H., Simoncelli, E.: Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**(4), 600–612 (2004)



# Image Analysis in a Parameter-Free Setting

Yu Zhu and Thomas Zeugmann

**Abstract** The paper proposes a new method to approximate the normalized information distance by a compression method that is particularly suited for image data. The new method is based on a video compressor. The new method is used to compute the distance matrix of all the images in the data sets considered. Moreover, the hierarchical clustering method from the R package is used to cluster the distance matrix obtained. Two different datasets are considered to demonstrate the usefulness of our new image analysis method. The results are very promising and show that one can obtain a very good clustering of the image data.

## 1 Introduction

To measure the similarity between objects is a fundamental notion in everyday life. For many data mining and machine learning algorithms the task of measuring the similarity between objects is also fundamental. Usually, the similarity between objects is measured by a domain-specific measure based on features of the objects. For example, the distance between pieces of music can be measured by using features like rhythm, pitch or melody; i.e., these features do not make sense in any other domain. For defining the right domain-specific distance measure one needs special knowledge about the application domain for extracting the relevant features beforehand. By using these parameters, one can then control the algorithms' sensitivity to certain features. Determining how relevant particular features are is often difficult and may require a certain amount of guessing. Expressing this differently, one has to tune the algorithms, which is requiring domain knowledge and a larger amount of experience. Such an approach does not only cause difficulties, but includes a certain

---

Y. Zhu (✉) · T. Zeugmann  
Division of Computer Science, Hokkaido University, N-14, W-9,  
Sapporo 060-0814, Japan  
e-mail: zhuyu07@ist.hokudai.ac.jp

T. Zeugmann  
e-mail: thomas@ist.hokudai.ac.jp

danger or risk of being biased. Furthermore, it may be expensive, error prone, and time consuming to arrive at a suitable tuning.

However, as a radically different approach; i.e., the paradigm of *parameter-free data mining*, has emerged (cf. Keogh et al. [4]). The main idea of parameter-free data mining is the design of algorithms that have no parameters and that are universally applicable in all areas.

The problem is whether or not such an approach can be realized at all. It is only natural to ask how an algorithm can perform well if it is not based on extracting the important features of the data, and if we are not allowed to adjust its parameters until it is doing the right thing. As expressed by Vitányi et al. [12], *if we a priori know the features, how to extract them, and how to combine them into exactly the distance measure we want, we should do just that. For example, if we have a list of cars with their color, motor rating, etc. and want to cluster them by color, we can easily do that in a straightforward way.*

So the approach of parameter-free data mining is aiming at scenarios where we are not interested in a certain similarity measure but in the similarity between the objects themselves. The most promising approach to this paradigm is based on Kolmogorov complexity theory. The intuitive idea can be described as follows: If two objects  $x$  and  $y$  are similar then it should be possible to obtain a short description of how to transform object  $x$  into  $y$  and object  $y$  into object  $x$ . Conversely, if two objects have (almost) nothing in common, then obtaining  $x$  from  $y$  and  $y$  from  $x$  is (almost) as complex as describing  $x$  and  $y$ , respectively, from scratch. Note that we need both directions here. For example, if we are given a blue image and a beautiful flower image (of the same size) the one can easily obtain the blue image from the flower image by assigning to each pixel the color blue. But the converse is not true; i.e., obtaining the flower from the blue image is as complex as describing the flower from the empty image.

The key ingredient to this approach is the so-called *normalized information distance* (NID) which was developed by various researchers during the past decades in a series of steps (cf., e.g., [1, 2, 5]).

More formally the *normalized information distance* between two strings  $x$  and  $y$  is defined as

$$NID(x, y) = \frac{\max\{K(x|y), K(y|x)\}}{\max\{K(x), K(y)\}}, \quad (1)$$

where  $K(x|y)$  is the length of the shortest program that outputs  $x$  on input  $y$ , and  $K(x)$  is the length of the shortest program that outputs  $x$  on the empty input. It is beyond the scope of the present paper to discuss the technical details of the definition of the NID. We refer the reader to Vitányi et al. [12].

The NID has nice theoretical properties, the most important of which is universality. The NID is called *universal*, since it accounts for the dominant difference between two objects (cf. Li et al. [5] and Vitányi et al. [12] and the references therein).

In a sense, the NID captures all computational ways in which the features needed in the traditional approach could be defined. Since its definition involves the

Kolmogorov complexity  $K(\cdot)$ , the NID cannot be computed. Therefore, to apply this idea to real-world data mining tasks, standard compression algorithms, such as `gzip`, `bzip2`, or `PPMZ`, have been used to approximate the Kolmogorov complexity and resulted in the *normalized compression distance* (NCD). We modify this approach by using a video compressor. This yields the *normalized MPEG distance* (NMD) as an approximation of the NID (cf. Definition 1). In [3] we have shown that the NCD based method is useful for text strings, such as influenza virus datasets.

But the NCD method could not find the real similarity in image datasets, even if we get a good compression ratio for images by the standard compression algorithms.

In this paper, we focus our attention to image datasets and provide the NMD to calculate the distance between images. We show our method to create an efficient and robust image distance measure. Using the NMD we compute a symmetric matrix  $D$  such that  $d_{ij}$  is the NMD between the data entries  $i$  and  $j$  (henceforth called distance matrix). The next step is the clustering. Here of course the variety of possible algorithms is large. We have decided to try the *hierarchical clustering* algorithm from the R package (called `hclust`) with the average option. In this way we obtain a rooted tree showing the relations among the input data. The results obtained are generally very promising. For the butterflies dataset, we obtained a perfect clustering result, which nicely coincides with a clustering obtained by human intuition. For the spider dataset, we still could get accuracy of 83.59 % without any occasional human intervention.

## 2 Background and Theory

### 2.1 Background

As explained in the Introduction, the theoretical basis for computing the distance matrix is deeply based in Kolmogorov complexity theory. Since the definition of the NID depends on the function  $K$  and since  $K$  is *uncomputable*, the NID is *uncomputable*, too. Thus it must be approximated. Experience showed that universal compression algorithms yield good approximations of the NID.

In our previous work we also demonstrated that these compressors gave unexpected good results for text based datasets such as virus datasets (cf. [3]).

However for the images, these “universal” compressors could not find the real similarity between images but only reduced the size of the data, even if they are lossless compressors. The main reason is that image data are usually large and thus the standard compressors are *not normal* when applied to image data. For example, if  $C$  is compressor then it should satisfy the condition  $C(xx) = C(x)$ , it should be symmetric and distributive and obey a monotonicity condition (cf. [3] for the formal definition).

The usual way to deal with images is to extract their features and then to compare the features extracted. Various methods have been proposed to perform this feature extraction.

Here we aim at an approach that maintains the benefits of parameter-free data mining. The idea is to use a video compressor to approximate the NID in an appropriate manner. It is known that widely used video compressors work as a tool to reduce the amount of data needed for subsequent frames. This means if the two images are similar, then the compressor can reduce more data and returns a file of smaller size. If the two images have almost nothing in common, then the video compressor basically just concatenates the two images and returns a file of bigger size.

However, using a video compressor poses some technical difficulties, since it requires the input of at least two images. Therefore, the original definition of the normalized compression distance has to be modified appropriately. This is done in the following subsection.

## 2.2 Our New Modified Measure Method

Looking at the definition of the NID, we see that it takes two inputs and then determines the length of the shortest program to produce  $x$  and  $y$  (expressed by  $K(x)$  and  $K(y)$ , respectively). Furthermore, determining  $K(x|y)$  means to figure out how much information about  $x$  is already contained in  $y$  and similarly for  $K(y|x)$ .

Since the video compressor  $m$  requires at least two images, this poses the problem how to relate  $m(xy)$  and  $m(yx)$  which maybe considered as approximations of  $K(y|x)$  and  $K(x|y)$ , respectively, to  $m(xx)$  and  $m(yy)$ . The result is provided in Definition 1 below; i.e., we compare the resulting compressions  $m(xy) - m(xx)$  and  $m(yx) - m(yy)$  and normalize it by dividing by the maximum of  $m(xx)$  and  $m(yy)$ .

We choose to realize this idea by using the *MPEG* encoder provided by MathWorks in Matlab for its simplicity and availability [9]. Of course, all of the input images should be in the same format, it could be *jpg*, *jpeg* or *png* any uniform files. The first step will transform the two individual images to two frames, and then transform them to one movie. The only thing we care about here is the size of the movie, as this indicate how similar the two images are. We perform a pairwise transformation for all of the images in the dataset. Then we can calculate the distance between each pair of images as defined in our measure method. More precisely, we have the following.

**Definition 1** The distance between two images  $x$  and  $y$  is defined as

$$MD(x, y) = \frac{\max\{m(xy) - m(xx), m(yx) - m(yy)\}}{\max\{m(xx), m(yy)\}}$$

where  $m$  is the given video compressor.

Having this definition we can put all images in a list  $X(x_1, \dots, x_n)$  and compute the distance matrix  $MD = (md(x, y))_{x, y \in X}$ . Here, the  $MD$  stands for “MPEG Distance.” Note that the distance matrix  $MD(x, y)$  returned is positive and symmetric.

Next, we turn our attention to clustering. We shortly outline the hierarchical clustering as provided by the R package, i.e., by the program `hclust` (cf. [10]). The input is the  $(n \times n)$  distance matrix  $MD$ . The program uses a measure of dissimilarity for the objects to be clustered. Initially, each object is assigned to its own cluster and the program proceeds iteratively. In each iteration the two most similar clusters are joint, and the process is repeated until only a single cluster is left. Furthermore, in every iteration the distances between clusters are recomputed by using the Lance–Williams dissimilarity update formula for the particular method used.

The methods differ in the way in which the distances between clusters are recomputed. Provided are the *complete linkage method*, the *single linkage method*, and the *average linkage clustering*. In the first case, the distance between any two clusters is equal to the greatest similarity from any member of one cluster to any member of the other cluster. This method works well for compact clusters but causes sensitivity to outliers. The second method pays attention solely to the area where the two clusters come closest to one another. The more distant parts of the clusters and the overall structure of the clusters is not taken into account. If the total number of clusters is large, a messy clustering may result.

The *average linkage clustering* defines the distance between any two clusters to be the average of distances between all pairs of objects from any member of one cluster to any member of the other cluster. As a result, the average pairwise distance within the newly formed cluster, is minimum.

Heuristically, the average linkage clustering should give the best results in our setting, and thus we have chosen it (see also Manning et al. [6] for a thorough exposition).

Hierarchical clustering builds clusters within clusters, and does not require a pre-specified number of clusters like  $k$ -means and  $k$ -medoids do. A hierarchical clustering can be thought of as a tree and displayed as a dendrogram; at the top there is just one cluster consisting of all the observations, and at the bottom each observation is an entire cluster. In between are varying levels of clustering.

As shown below, our algorithm works in this way.

**Algorithm** MPEG distance clustering for a data list

*Input:* image data list  $X = (x_1, x_2, \dots, x_n)$ ;

*Output:* clustering tree displayed as a dendrogram;

*Step 1.* for  $x, y \in X$ , use MPEG compressor to get the video size  $m(xy)$  for pairwise images;

*Step 2.* compute the distance matrix  $MD = (md(x, y))_{x, y \in X}$ ;

*Step 3.* cluster the matrix  $MD$  by using hierarchical clustering from the R package (cf. [10]).

### 3 Experiments and Results

In this section we describe the two datasets used and the results obtained. The clusterings obtained provide evidence for the usefulness of our new method for image classification.

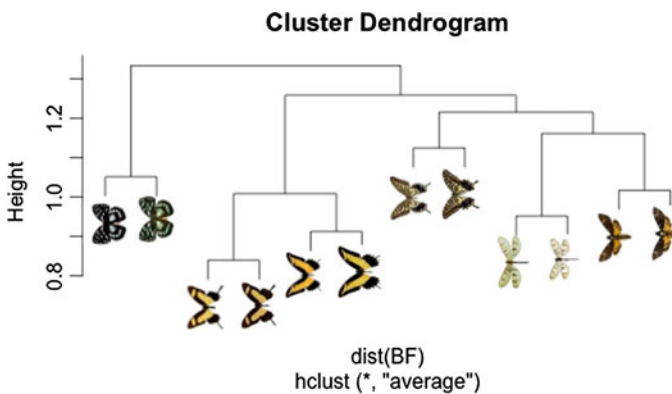
#### 3.1 Butterflies Dataset

First, we applied our new method to a set of butterfly images in order to judge its effectiveness intuitively. These images were extracted from various websites devoted to entomology. The result obtained is displayed in Fig. 1. As we see the butterflies belong to six clusters based on their color and shapes.

In order to obtain this clustering, first we performed an alignment of the images; i.e., we ensured that all images have the same size. Here we kept the resolution ratio as its original ratio and added a frame to ensure that all images have the same size. This preprocessing is necessary, since in a movie all images are assumed to be of the same size. If an image did not fit in the frame, a white background was added to fill the blank part.

After this step, we calculated the symmetrical pairwise distance matrix. Then we applied the hierarchical clustering as described above to judge the *similarity* between these images. As Fig. 1 shows, the obtained result is really satisfying.

Note that we clustered this set of images, because it was used to test the utility of a color and shape distance measures in [13]. We found that the parameter-free method's result is even better than the one obtained by the carefully tuned color or shape measure.



**Fig. 1** The butterflies dataset clustered with our new distance measure method (hierarchical clustering method as provided by R package)

### 3.2 Spider Dataset

To show that the MD measure method could be widely used, we chose the Australasia ground spiders of the family Trochanteriidae as our dataset. This selection was made primarily because all species were already available and the size of the family is 121 species in 14 genera, this seems a reasonable dataset and we can use it to judge the effectiveness of our method [11].

Species discrimination in spiders is based primarily on the shape of the male and female genitalia. If we try to identify the species, or to systematically describing new species, we need to examine these structures. Epigyna is the reproductive structures for female spider. The epigynum is found on the ventral side of an adult female and is visible without dissection (cf. Figs. 2 and 3).

Fig. 2 Spider ventral view

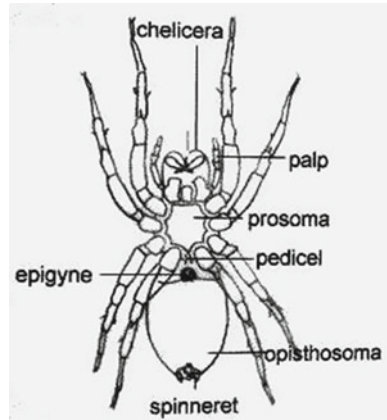
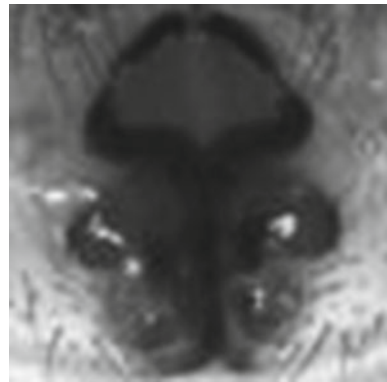
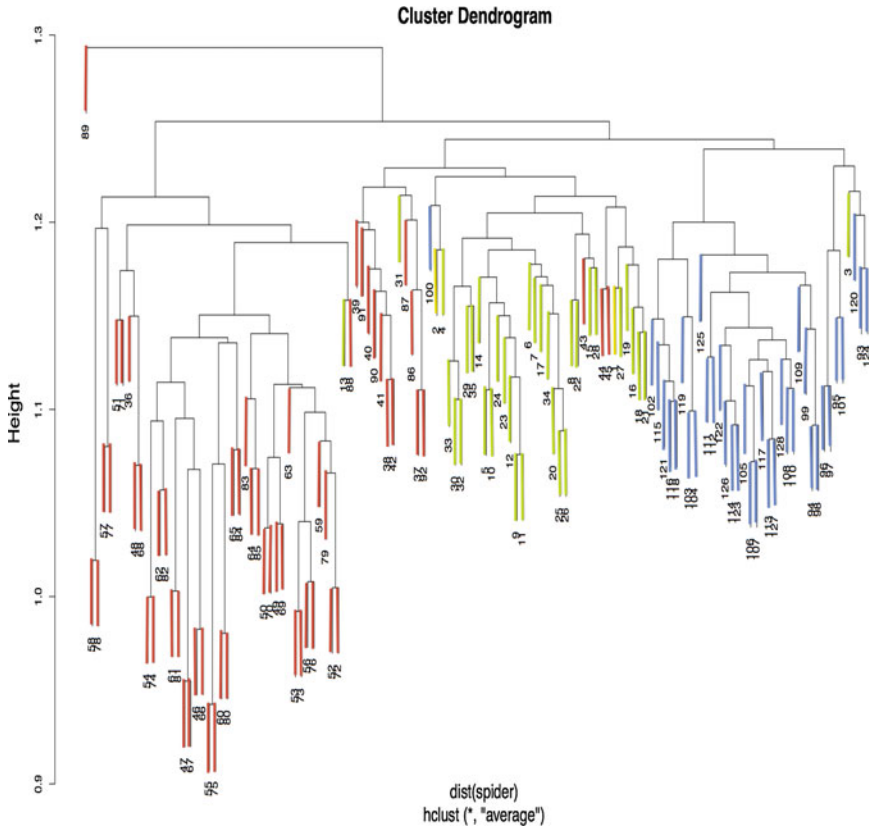


Fig. 3 One example of Epigynum for spider





**Fig. 4** The spider dataset clustered with our new distance measure method (hierarchical clustering method as provided by R package)

In order to allow the reader to check the results obtained, we shortly describe the data set used. We chose 128 epigyna images from four species; i.e., the first two belong to the *Boothana* species (displayed in yellow), the third to the 35th belong to the *Fissarena* species (shown in green), the 36th–92nd are *Hewicloeina* species (displayed in red) and the remaining ones (93rd–128th) are *Longrita* species (drawn in blue) (cf. Fig. 4).

All experiments were performed under Mac OS X Lion. We calculated the distance matrix by using MPEG in Matlab. From the 128 spider images 107 are correctly clustered. Thus, the accuracy of the cluster for the spider data is 83.59 %.

By performing our experiments we demonstrated the usefulness of our new measure method for image data. We could easily justify the clustering result for butterflies even we do not have background knowledge in this area. For the spider dataset, results of informal surveys of archaeologists suggest that acceptable cutoffs for accuracy vary widely and often depend on the background of the respondents. Systematists



or taxonomic specialists demand on the highest accuracy level—95 % minimum for such a system to be useful for them. Ecologist and conservationists would be happy with 80–90 percent if it meant they could have a species list to work with [8]. Our method gave an acceptable result without any trained personal who are able to identify known species correctly. Especially for the *Fissarena* species (in green) and *Longrita* species (in blue) we only got 7 individual mixed in the clustering result, this part gave an 91.67 % accuracy result. We have not reported the speed of the *MPEG* distance method, one of the reasons is we wanted to show the utility first. Another reason is that optimizing speed may be irrelevant in many domains. As in some medical application it may take over an hour to produce an image, and wait another hour to find matches in a database [7].

## 4 Conclusion and Future Work

The usefulness of our new measure method, *MPEG* distance, for clustering the butterflies images and spider's epigyna images has been demonstrated. For the butterflies dataset, we got a perfect clustering result. And for the spider subset, we got an accuracy of 83.59 %. We are not claiming the *MPEG* distance is the best measure possible for image analysis problems. We have not reported the running time here, which it is still acceptable, since the process of identification and description of new species usually takes months or even years. For specialized application areas, there may be better measures, which include domain specific constraints and features. However, the *MPEG* distance measure offers a useful simple way when we do not have so much background about the application domain.

For the future work, we also would like to combine our measure method with the multisets to find the similarity between a pair of finite objects based on compression. Exploiting some other video compressors maybe increase the speed of our method. Of course, to apply the *MPEG* distance to other fields is the first task we will pursue next.

**Acknowledgments** We would like to thank to the program committee and the anonymous referees for their valuable comments.

## References

1. Bennett, C.H., Gács, P., Li, M., Paul M.B.V., Zurek, W.H.: Information distance. *IEEE Trans. Inf. Theor.* **44**(4), 1407–1423 (1998)
2. Cilibrasi, R., Vitányi, P.M.B.: Clustering by compression. *IEEE Trans. Inf. Theor.* **51**(4), 1523–1545 (2005)
3. Ito, K., Zeugmann, T., Zhu, Y.: Clustering the normalized compression distance for influenza virus data. In: *Algorithms and Applications*, volume 6060 of *Lecture Notes in Computer Science*, pp. 130–146. Springer, New York (2010)

4. Keogh, E., Lonardi, S., Ann, C.: Ratanamahatana. Towards parameter-free data mining. In: KDD '04: Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 206–215. ACM Press, New York (2004)
5. Li, M., Chen, X., Li, X., Ma, B., Vitányi, P.M.B.: The similarity metric. *IEEE Trans. Inf. Theor.* **50**(12), 3250–3264 (2004)
6. Manning, C.D., Raghavan, P., Schütze, H.: *Introduction to Information Retrieval*. Cambridge University Press, Cambridge (2008)
7. Pavlidis, T.: Limitations of content-based image retrieval, 2008. unpublished manuscript: <http://www.theopavlidis.com/technology/CBIR/PaperB/vers3.htm>
8. Russell, K.N., Do, M.T., Huff, J.C., Platnick, N.I.: Introducing spida-web: Wavelets, neural networks and internet accessibility in an image-based automated identification system. In: MacLeod, N. (eds.) *Automated Taxon Identification in Systematics: Theory, Approaches and Applications*, pp. 131–152. CRC Press, New York (2007)
9. Sumathi, S., Paneerselvam, S.: *Computational Intelligence Paradigms Theory and Applications using MATLAB*. CRC Press, New York (2010)
10. The R project for statistical computing. <http://www.r-project.org/>
11. Ticay-Rivas, J.R., del Pozo-Baños, M., Eberhard, W.G., Alonso, J.B., Travieso, C.M.: Spider specie identification and verification based on pattern recognition of it cobweb. *Expert Syst. Appl.* **40**(10), 4213–4225 (2013)
12. Paul M.B.V., Frank J.B., Rudi L.C., Li, M.: Normalized information distance. In: *Information Theory and Statistical Learning*, pp. 45–82. Springer, New York (2008)
13. Wang, X., Ye, L., Keogh, E., Shelton, C.: Annotating historical archives of images. In: *Joint Conference on Digital Libraries*, pp. 341–350 (2008)

# Age Estimation Based on Hybrid Features of Facial Images

Asuman Günay and Vasif V. Nabiyev

**Abstract** This paper proposes a new age estimation method using hybrid features produced by fusing the global and local features of facial images at decision level. The global facial features are extracted using active appearance models (AAM) which contains both the shape and appearance of a human face. In the local feature extraction phase, the wrinkle features are extracted using a set of Gabor filters, capable of extracting deep and fine wrinkles in different directions and the skin features are extracted using local binary patterns (LBP), capable of extracting the detailed textures of skin. After the feature extraction module, three aging functions are modeled separately with multiple linear regression. Then decision level fusion is performed to combine the results of these aging functions to make a final decision. Experimental results showed that the performance of the proposed method was superior to that of the previous methods when using the FG-NET and PAL aging databases.

**Keywords** Age estimation · Hybrid features · Active appearance models · Local binary patterns · Gabor filters

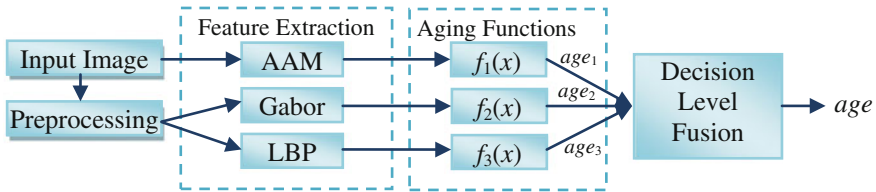
## 1 Introduction

With the increasing need of automatic recognition systems, the researches on facial image processing have received considerable interest in recent decades. Face detection, face recognition, gender classification and facial expression recognition are the research topics that have been studied by many researchers in this area. Facial age estimation is a relatively new topic and the interest in this topic has significantly increased because it has many real world applications. For example, an automatic

---

A. Günay (✉) · V.V. Nabiyev  
Department of Computer Engineering, Karadeniz Technical University,  
61080 Trabzon, Turkey  
e-mail: gunaya@ktu.edu.tr

V.V. Nabiyev  
e-mail: vasif@ktu.edu.tr



**Fig. 1** Proposed age estimation method

age estimation system can prevent under ages from accessing cigarettes, alcohol or obscene contents on websites. In addition, age specific target advertising, face recognition systems robust to age progression and age prediction systems for finding the lost children and criminals are important age estimation applications.

Facial age estimation is a multi-class classification problem because an age label can be seen as an individual class. This makes age estimation much harder than other facial image processing problems such as gender classification, face detection, etc. Besides, real world age progression displayed on faces is varied and personalized. Aging process of a person is affected by the genetics, race, living styles, eating and drinking habits, climate, etc. Exposure to sunlight, smoking, extreme weight loss, emotional stress, extent and frequency of facial expressions, usage of anti-aging products, and plastic surgery also affect the person's facial appearance. Therefore, it is very difficult to determine the type of facial features directly represents the age. Moreover, the accuracy of age estimation systems are insufficient, even the human skills about age estimation are limited. The lack of proper large data set including the chronological image series of individuals is another drawback in these systems.

In this paper we propose a new age estimation method based on decision level fusion of global and local features of facial images as shown in Fig. 1. The global facial features are extracted using active appearance models (AAM) which contains both the shape and appearance of a human face. The local facial features are extracted using Gabor filters and local binary patterns (LBP). A set of Gabor filters capable of extracting deep and fine wrinkles in different directions are used to extract wrinkle features and LBP is used to extract the detailed skin texture features of facial images. Then dimensionality reduction is performed using principal component analysis (PCA) for each feature vector separately. After finding the lower dimensional subspaces three distinct aging functions are modeled using global features, wrinkle features and skin features of facial images. Finally the results of these aging functions are combined to estimate the age of the subject.

This paper is organized as follows. A survey on age estimation methods is given in the next section. Section 3 introduces the proposed method including preprocessing, feature extraction, regression and decision level fusion modules. In Sect. 4, experimental results are given and Sect. 5 concludes the paper.

## 2 Related Work

The existing age estimation systems typically consist of an age image representation and an age estimation module. Age image representation techniques rely often on shape-based and texture-based features that were extracted from facial images. They can be grouped under the topics of Anthropometric Models, AAM, AGing pattErn Subspace (AGES), Age Manifold and Appearance Models. Then, age estimation can be performed with age group classification or regression methods.

The anthropometric model based representations only consider the facial geometry. The earliest paper published in the area of age classification relying on facial geometry was the work by Kwon and Lobo [1]. They computed six ratios of distances on frontal images to separate babies from adults. They also used wrinkle information extracted by snakelets in several regions of human face to separate young adults from senior adults.

AAM based approaches incorporate shape and appearance information together rather than just the facial geometry as in the anthropometric model based methods. Using the AAMs for age estimation was initially proposed by Lanitis et al. [2]. The relationship between the age of individuals and the parametric description of face images was defined with an aging function  $age = f(b)$ . Later on various works are presented using AAM features for age estimation [3–5]. In some studies AAM features are extracted as global facial features and fused with local facial features for efficient age estimation [6].

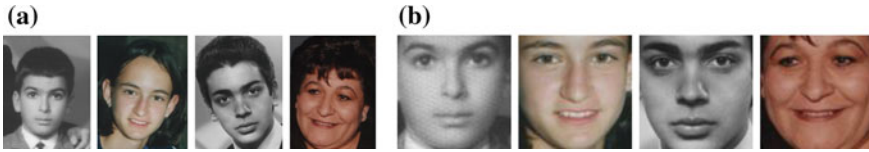
The idea of using the sequence of an individual's aging face images arranged in chronological order, to model the aging process is explored by Geng et al. [7]. The features of face images in the aging pattern are extracted with AAM. Then a specific aging subspace is learned with PCA and EM-like iterative algorithm for each individual. The age of a test image is estimated by searching a suitable aging pattern and a proper age position in that aging pattern for it.

Age manifold methods aims to learn a common aging pattern from the images of different individuals at different ages. Age manifold utilizes a manifold embedding technique to discover the aging trend in a low dimensional domain. Thus, the mapping from the image space to the manifold space can be done either by linear or by nonlinear functions [8–11].

Appearance models are mainly focused on the aging-related facial feature extraction. Both global and local features extracted with edge detection methods, Gabor filters, LBP were used in existing age estimation systems [6, 12, 13].

## 3 Proposed Method

This paper presents a new age estimation method using hybrid features which are the combination of global and local features of facial images. Global features are extracted using AAM. The local features are extracted using Gabor filters and LBP.



**Fig. 2** Image preprocessing **a** Original images. **b** Preprocessed images

Then three aging functions are modeled with these features using multiple linear regression. Finally the results of these aging functions are combined to make a final decision of the accurate age. These modules are explained in the following sections.

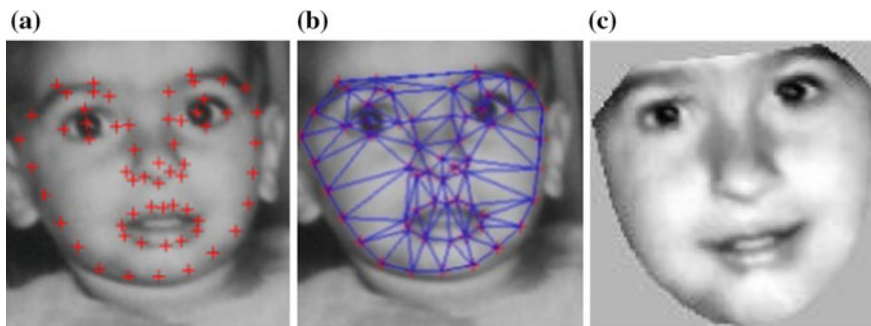
### 3.1 Preprocessing

The orientation and the size of original images are different from each other. Also they have unnecessary features such as background, cloth and hair which are not related to the face and can affect the performance of the algorithm. Therefore, an image preprocessing step is performed to extract only the facial regions and to adjust the orientation and size of the faces. In the preprocessing module, the facial images are cropped, scaled and transformed to the size of  $88 \times 88$ , based on the eye center locations as shown in Fig. 2.

### 3.2 Feature Extraction

In this module hybrid features which are the combination of global and local features are extracted from facial images. Global features are extracted using AAM. The local features are extracted using Gabor filters and LBP. The details are given below.

**Active Appearance Models.** AAM is a statistical shape and appearance model of facial images [14]. These models are generated by combining a model of shape variations with a model of the appearance variations in a shape-normalized frame. A statistical shape model can be generated with a training set of face images labeled with landmark points as shown in Fig. 3a. The mean shape is produced with taking the mean of the landmark points in the training set. Then, PCA is applied to the data to extract the main principal components along which the training set varies from the mean shape. To build a statistical appearance model, each image in the training set has to be normalized, so that its control points match the mean shape. For this purpose all images are warped to the mean shape using affine transformation and Delaunay triangulation (Fig. 3b). Then, the gray-level intensities within a pre-specified image region of the shape normalized images (Fig. 3c) are used to train an intensity model. By applying PCA to the gray level intensities, a linear model is obtained. Since there



**Fig. 3** **a** Example of face image labeled with 68 landmark points. **b** Result of the Delaunay triangulation used in warping process. **c** Shape normalized image

may be correlations between the shape and gray-level variations, a further PCA is applied to them and, finally, the combined shape and appearance parameters are obtained.

In this study the statistical shape model is generated with a training set of face images labeled with 68 landmark points. For the intensity model, approximately 7000 gray-level intensities in the facial region of the corresponding shape-free image are used to represent the training samples. The resulting combined shape and intensity model requires 362 model parameters for FG-NET database and 418 model parameters for PAL database to explain 99% of the variance in the training set. These model parameters are used as a global descriptor of facial images.

**Gabor Filters.** A Gabor filter is obtained by modulating a sinusoid with a Gaussian. This filter will therefore respond to some frequency, but only in a localized part of the signal. Let  $g(x, y, \theta, \phi)$  be the function defining a Gabor filter centered at the origin with  $\theta$  as the spatial frequency and  $\phi$  as the orientation. We can view Gabor filters as:

$$g(x, y, \theta, \phi) = \exp(-(x^2 + y^2)/\sigma^2) \exp(2\pi\theta(x \cos \phi + y \sin \phi)) \quad (1)$$

It has been shown that  $\sigma^2$ , the standard deviation of the Gaussian kernel depends upon the spatial frequency to measured, i.e.  $\theta$ . The response of a Gabor filter to an image is obtained by a 2D convolution operation. Let  $I(x, y)$  denote the image and  $G(x, y, \theta, \phi)$  denote the response of a Gabor filter with frequency  $\theta$  and orientation  $\phi$  to an image at point  $(x, y)$  on the image plane.  $G(\cdot)$  is obtained as,

$$G(x, y, \theta, \phi) = \int \int I(p, q)g(x - p, y - q, \theta, \phi)dpdq \quad (2)$$

In the study the fine and deep wrinkles at different orientations are extracted using a Gabor filter set with 4 scales and 6 orientations as shown is Fig. 4. The responses of these filters for an image are also given in the figure.

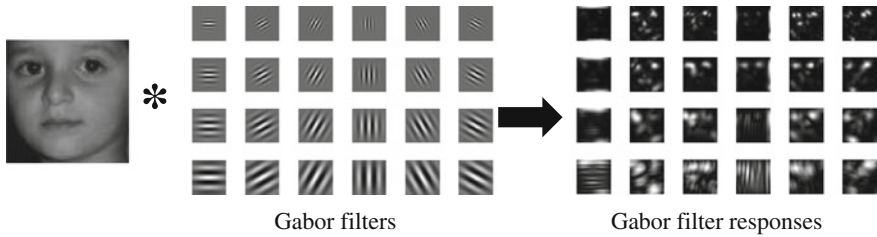


Fig. 4 Gabor filter process

**Local Binary Patterns.** Local binary patterns are fundamental properties of local image texture and their occurrence histogram is a powerful texture descriptor [15]. Original LBP operator assigns a label for every pixel of the image, by means of thresholding the center pixel with its  $3 \times 3$  neighborhood and considering the results as a binary number. Then the occurrence histogram of these labels is used as a texture feature. Every pixel of the image is labeled with the following equation,

$$LBP_{P,R}(x_c) = \sum_{p=0}^{P-1} u(x_p - x_c)2^p \quad u(y) = \begin{cases} 1, & y \geq 0 \\ 0, & y < 0 \end{cases} \quad (3)$$

where  $x_c$  is center pixel,  $x_p$  represents one of his  $P$  neighbors and  $R$  is the radius. In texture description all LBP codes are not used. The uniform patterns which contains at most two bitwise transitions from 0 to 1 or vice versa when the binary pattern is considered circular, account for a bit less than 90% of all patterns when using (8,1) neighborhood [15]. But holistic descriptions of facial images are not reasonable since the texture descriptors tend to average over the image area [16]. However retaining the information about spatial relations is important for facial images. Furthermore local representations are more robust against variations in pose or illuminations than holistic representations. As a result, spatial LBP histograms are extracted for an efficient representation of facial images. For this purpose image is divided into  $m$  regions and for every region spatial histograms are produced as follows,

$$H_{i,j} = \sum_{x_c \in R_j} f\{LBP_{P,R}(x_c) = U(i)\} \quad (4)$$

$$f(y) = \begin{cases} 1, & \text{if } y \text{ is true} \\ 0, & \text{otherwise} \end{cases} \quad i = 0, 1, \dots, n - 1, j = 0, 1, \dots, m - 1$$

where  $H_{i,j}$  is the  $i$ . value of the LBP histogram of  $j$ . region and  $U(i)$  is the vector keeps uniform patterns. These regional histograms are concatenated to build a global description of the image.



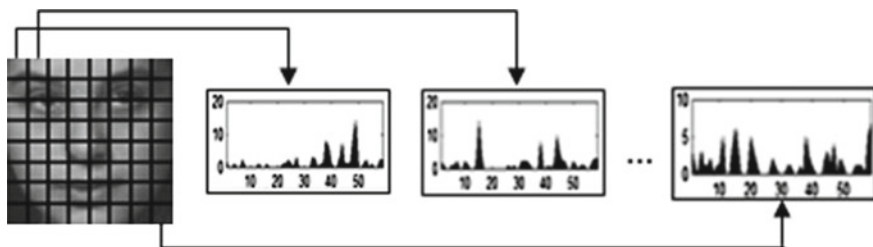


Fig. 5 Regional LBP histograms

In this work the detailed skin textures of facial images are extracted using spatial LBP histograms. For this purpose the facial images are divided into  $8 \times 8$  sub-regions from which LBP histograms are produced and concatenated into a single vector as shown in Fig. 5.

After the feature extraction module, PCA is performed on feature vectors to find a lower dimensional subspace which carries significant information for age estimation.

### 3.3 Aging Function Modeling

After finding the low dimensional representation of facial features, we define the age estimation problem as a multiple linear regression problem as  $age = f(M) \Leftrightarrow \hat{L} = \hat{f}(Y)$  where  $\hat{L}$  denotes the estimated age label,  $f(\cdot)$  the unknown regression function, and  $\hat{f}(\cdot)$  is the estimated regression function. The aging function used in this study is a linear function given in  $\hat{l} = \hat{\beta}_0 + \hat{\beta}_1^T y$ , where  $\hat{l}$  is the estimate of age,  $\hat{\beta}_0$  is the offset,  $\hat{\beta}_1$  is the weight vector and  $y$  is the extracted feature vector.

### 3.4 Decision Level Fusion

In the previous module three aging functions are modeled separately using global features, wrinkle features and skin features for age estimation. The results of these aging functions are combined to make a final decision for the age of the test sample as follows:

$$age = \sum_{i=1}^N age_i / N \tag{5}$$

where  $age$  is the final age of the test sample,  $age_i$  is the age estimated by the  $i$ th aging function and  $N$  is the total number of aging functions.

## 4 Experiments and Results

To evaluate the performance of the proposed method FG-NET and PAL databases are used in experiments. The FG-NET aging database [17] comprises of 1,002 images of 82 subjects (6–18 images per subject) in the age range 0–69 years. Since the images were retrieved from real-life albums of different subjects, aspects such as illumination, head pose, facial expressions etc. are uncontrolled in this dataset. The PAL aging database was developed to be more representative of age groups across the lifespan [18]. It contains 580 individual faces ranging from ages 18 to 93. The images were captured under natural lighting conditions using a digital camera. The database includes various expressions such as smiling, sadness, anger or neutral faces. The data distributions of FG-NET and PAL aging databases are shown in Fig. 6.

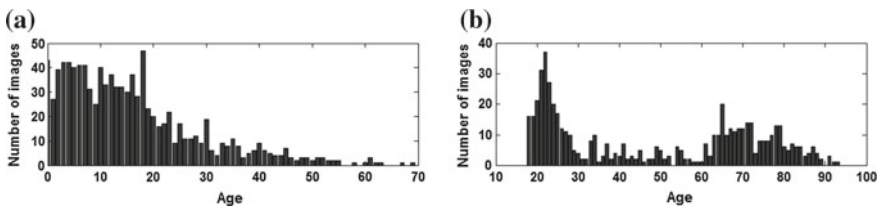
The evaluation framework is Leave-One-Person-Out mode for FG-NET aging database. That is in each fold the images of one person are used as test set and those of the others are used as the training set. After 82 folds, each subject has been used as test set once, and the final results are calculated based on all the estimations. In the experiments we also use 3-fold cross validation mode for PAL database in which the 1/3 of the images are selected randomly as test set and the rest are used as training set. After 3 folds the mean of all estimations is determined as estimation performance of the system

For the performance comparison, the Mean Absolute Error (MAE) metric is used. MAE is the mean of the absolute error between the estimated age labels and the ground truth labels as follows:

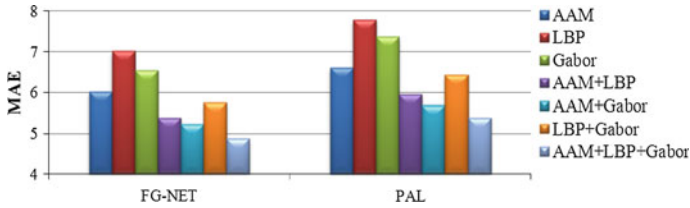
$$MAE = \frac{1}{N_t} \sum_{i=1}^{N_t} |\hat{y}_i - y_i| \quad (6)$$

where  $\hat{y}_i$  is the estimated age for the  $i$ th testing sample,  $y_i$  is the corresponding ground truth, and  $N_t$  is the total number of the test samples.

The estimation results of the proposed method for FG-NET and PAL databases are shown in Fig. 7. We can see from the figure that age estimation performance of the AAM features is better than Gabor and LBP features on FG-NET and PAL databases as the AAM features both include the shape and appearance information



**Fig. 6** The data distributions of **a** FG-NET and **b** PAL aging databases



**Fig. 7** The estimation results of feature vectors on FG-NET and PAL database

**Table 1** Comparison of the proposed methods to the previous works on FG-NET database

Method	WAS [7]	AGES [7]	AGES <sub>lda</sub> [7]	LARR [9]	PFA [10]	Proposed
MAE	8.06	6.77	6.22	5.07	4.97	<b>4.87</b>

of facial images. But when they are fused with local features including wrinkle and skin texture information at decision level, age estimation performance noticeably increased. As a result the proposed method achieves the MAE of 4.87 on FG-NET database and MAE of 5.38 on PAL database when using the hybrid features of facial images.

As the FG-NET database is the most common database used in age estimation works, the performance of the proposed method is compared with the previous works in which the performances are reported on the FG-NET aging database. The MAEs of the previous methods and the proposed method are tabulated in Table 1. Our method has an MAE of 4.87 which is lower than all the previous methods using only AAM features listed in Table 1. This result also shows that the combination of the global and local features at decision level improves the age estimation performance.

Also, when compared our age estimation performance with the result reported in [6], our 4.87 years MAE is better than their 5.32 years MAE when using a single level age estimation scheme.

## 5 Conclusion

In this paper we have presented an age estimation method using the global and local features of facial images. The global features are extracted using AAM while the local features are extracted using Gabor filters and LBP. Three aging functions are modeled with these features separately and the age estimation results of these aging functions are combined in decision level to make a final decision for the age label. Experimental results on the FG-NET and PAL aging databases have shown that proposed method is better than most conventional methods.

## References

1. Kwon, Y.H., Lobo, N.V.: Age classification from facial images. *Comput. Vis. Image Underst.* **74**(1), 1–21 (1999)
2. Lanitis, A., Taylor, C., Cootes, T.: Toward automatic simulation of aging effects on face images. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(4), 442–455 (2002)
3. Wang, C.-C., Su, Y.-C., Hsu, C.-T., Lin, C.-W., Liao, H.Y.: Bayesian age estimation on face images. In: *IEEE International Conference on Multimedia and Expo, ICME'09*, pp. 282–285. (2009)
4. Kohli, S., Prakash, S., Gupta, P.: Hierarchical age estimation with dissimilarity-based classification. *Neurocomputing* **120**, 164–176 (2013)
5. Chao, W.-L., Liu, J.-Z., Ding, J.-J.: Facial age estimation based on label-sensitive learning and age oriented regression. *Pattern Recogn.* **43**, 628–641 (2013)
6. Choi, S.E., Lee, Y.J., Lee, S.J., Park, K.R.: Age estimation using a hierarchical classifier based on global and local facial features. *Pattern Recogn.* **44**(6), 1262–1281 (2011)
7. Geng, X., Zhou, Z.H., Miles, K.S.: Automatic age estimation based on facial aging patterns. *IEEE Trans. on Pattern Anal. Mach. Intell.* **29**(12), 2234–2240 (2007)
8. Fu, Y., Huang, T.S.: Human age estimation with regression on discriminative aging manifold. *IEEE Trans. Multimedia* **10**(4), 578–584 (2008)
9. Guo, G., Fu, Y., Dyer, C.R., Huang, T.S.: Image-based human age estimation by manifold learning and locally adjusted robust regression. *IEEE Trans. Image Process.* **17**(7), 1178–1188 (2008)
10. Guo, G., Fu, Y., Dyer, C.R., Huang, T.S.: A probabilistic fusion approach to human age prediction. In: *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW'08)*, pp. 1–6 (2008)
11. Lu, J., Tan, Y.-P.: Ordinary preserving manifold analysis for human age and head pose estimation. *IEEE Trans. Hum.-Mach. Syst.* **43**(2), 249–258 (2013)
12. Gao, F., Ai, H.: Face age classification on consumer images with gabor feature and fuzzy LDA method. In: *Proceedings of 3rd International Conference on Advances in Biometrics. LNCS*, vol. 5558, pp. 132–141. Springer, Heidelberg (2009)
13. Ma, Y., Liu, J., Yang, Y., Zheng, N.: Double layer multiple task learning for age estimation with insufficient training samples. *Neurocomputing* **147**, 380–386 (2015)
14. Cootes, T., Edwards, G., Taylor, C.: Active appearance models. *IEEE Trans. Pattern Anal. Mach. Intell.* **23**(6), 681–685 (2001)
15. Ojala, T., Pietikainen, M., Maenpaa, T.: Multi-resolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* **24**(7), 971–987 (2002)
16. Ahonen, T., Hadid, A., ve Pietikainen, M.: Face description with local binary patterns: application to face recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* **28**(12), 2037–2041 (2006)
17. FG-Net aging database. <http://sting.cyccollege.ac.cy/~alanitis/fgnetaging>
18. Minear, M., Park, D.C.: A lifespan database of adult stimuli. *Behav. Res. Methods Instrum. Comput.* **36**(4), 630–633 (2004)

**Part VII**  
**Algorithm Design for Biological  
and Chemical Systems**

# Proposal of a New Method for de Novo DNA Sequence Assembly Using de Bruijn Graphs

Adriano Donato Couto, Fabio Ribeiro Cerqueira,  
Ricardo dos Santos Ferreira and Alcione de Paiva Oliveira

**Abstract** In this work, we propose a method where all main concerns in the construction of a DNA sequence assembler is encompassed in a single computational approach. The graph that represents sequences and their relationships is broken into simpler components through a matching approach on bipartite graphs, which we show to be very suitable for parallelism. Next, a simple heuristic is used to combine the components into long paths in the graph, which correspond to contigs in the genome. Experiments with five datasets have shown high quality assemblies.

**Keywords** DNA sequence assembly · de Bruijn graphs ·  $k$ -mer graphs · Graph algorithms · Bioinformatics

## 1 Introduction

A major difference between sequencing technologies concerning how to handle the assembly procedure is the directed graph that is used to represent reads and their relationships. For the Sanger method, the most common structure is the overlap graph, where each read is a vertex and an arc connects vertex  $a$  to vertex  $b$  if a suffix of  $a$  can align to a prefix of  $b$  [4, 5, 8, 9]. Due to the enormous input in the case of data produced by the so-called next-generation sequencing (NGS) [1], on the other hand, the assemblers usually use a *de Bruijn* graph that is also called  $k$ -mer graph in the context of genomics. This structure is preferred because it eliminates the necessity of performing a huge amount of pairwise sequence alignment that is needed in the case of overlap graphs.

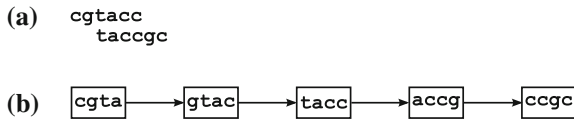
To build a  $k$ -mer graph, instead of representing reads in their original form, each read of length  $l$  is broken into a set of subsequences of  $k$  letters ( $k$ -mers), where

---

Couto and Cerqueira contributed equally to this work.

---

A.D. Couto · F.R. Cerqueira (✉) · R. dos Santos Ferreira · A. de Paiva Oliveira  
Department of Informatics, Universidade Federal de Viçosa, Viçosa,  
Minas Gerais 36570-900, Brazil  
e-mail: fabio.cerqueira@ufv.br; frcerqueira@gmail.com



**Fig. 1** A  $k$ -mer graph built from two overlapping reads. **a** Two reads with an overlap of 4 bases. **b** A  $k$ -mer graph representing the reads in (a) for  $k = 4$ . The alignment is naturally obtained by a Hamiltonian path in the graph

$0 < k < l$ . As a result, each pair of consecutive  $k$ -mers of a given read has an overlap of  $k - 1$  bases. All  $k$ -mers resultant of the breakage of input reads are considered vertices of the graph. Arcs are included for consecutive  $k$ -mers for which the  $k - 1$  exact overlap occurs naturally. Figure 1 illustrates this process. During the graph construction, every time a read gives rise to a  $k$ -mer that is already present in the graph, no new vertex is created. Instead, the vertex count for the number of occurrences of that  $k$ -mer sequence is increased by one [4, 5, 7–9].

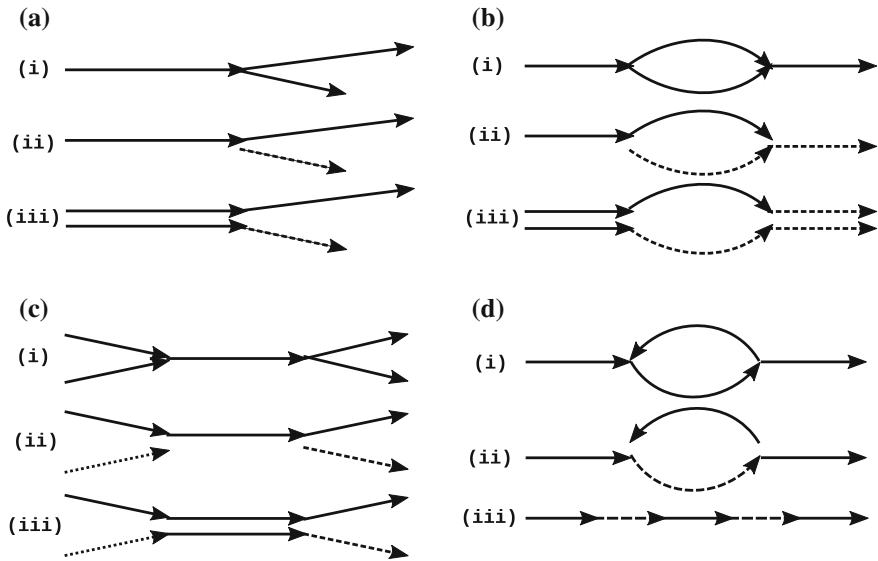
Despite the use of different strategies, the main purpose of DNA assemblers is to find a path (or a set of paths when no single path is possible) covering the vertices to establish an order in which the reads have to be taken to form the consensus sequence of the genome. As  $k$ -mer graphs, by definition, consider only perfect alignments, the graph complexity is sensitive to differences in the sequence such as sequencing errors and single-nucleotide polymorphisms (SNPs). The complexity is also affected by ambiguities caused by repeated areas of the genome. Such situations bring the necessity to run extensive pre and post-processes on the dataset.

In order to overcome the above issues related to the graph complexity, but without extra procedures, we propose an unified process based on a previous work for overlap graphs [3]. In this work, we present a new method for assembling reads represented by  $k$ -mer graphs. Our approach belongs to the category of de novo methods, i.e., it does not use any extra information besides the reads themselves. In this first stage of our work, that is reported in this paper, our goal is to propose and validate a method that can be easily parallelized. As can be seen in Sect. 3, the proposed strategy succeeded in assembling the genomes we used in our experiments and is ready to pass through the next stage of our effort that concerns the use of parallelism to increase performance in both aspects: CPU and memory usage.

## 2 Methods

### 2.1 Problematic Topologies of $k$ -mer Graphs

The complexity of  $k$ -mer graphs is highly dependent on differences in the read sequences and on ambiguities that are natural in genomes, such as the occurrence of repeats. According to Miller et al. [5], there are basically four problematic topologies

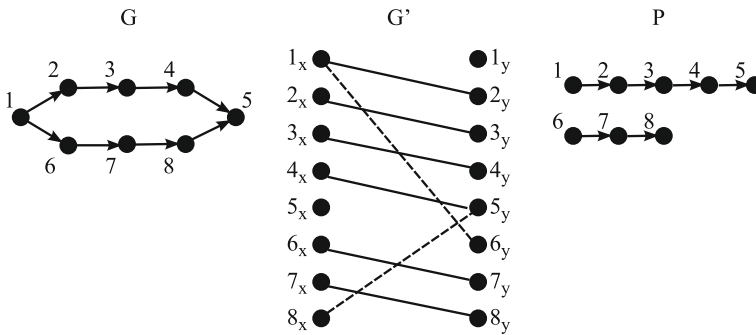


**Fig. 2** Treating problematic topologies. **a** Handling spurs: **a-i** The spur topology. **a-ii** A possible set of components returned by the MWMCM procedure (*solid* and *dashed lines* are to distinguish paths returned by this procedure). **a-iii** Ideal result, keeping both valid alternatives. **b** Coping with a bubble caused by natural genomic variations: **b-i** The bubble topology. **b-ii** A possible set of components returned by the MWMCM procedure. **b-iii** Ideal result, keeping both valid alternatives. **c** Handling frayed rope: **c-i** The frayed rope topology. **c-ii** A possible solution found by the MWMCM procedure. **c-iii** A possible combination of components in (**c-ii**). **d** Coping with cycles: **d-i** Cycle due to repeats. **d-ii** A maximum matching can induce either only paths or paths plus cycles. In both cases, in order to cover all vertices, the combination of components have to cover some vertices of the cycle more than once. Here, we want to minimize cover repetition so that a final solution (**d-iii**) is as short as possible. We see, then, that a solution is not a path but a walk

that can be found in a  $k$ -mer graph: Spur—short dead-end divergence (Fig. 2a-i). This topology is caused by nucleotide differences present at the end of similar reads; Bubble—divergence of a path followed by a convergence (Fig. 2b-i). Also caused by differences in the sequences, but in the middle of similar reads; Frayed rope—convergence of paths followed by divergence (Fig. 2c-i). This topology arises as a consequence of the presence of repeats in the genome; Cycle—also a consequence of repeats (Fig. 2d-i).

Our work proposes a single approach to handle all situations described above, instead of applying intensive pre/post-processing procedures that are performed by previous DNA assemblers with high computational costs [2, 10].





**Fig. 3** Maximum matching applied to a bubble.  $G$  is the original graph,  $G'$  is the bipartite version (*double*) of  $G$ , and  $P$  is a set of paths in  $G$  induced by a maximum matching in  $G'$ . Edges in  $G'$  that are part of the matching are solid. Otherwise, they are *dashed*. Notice that there are other three possibilities for a maximum matching. One of them, for instance, can be obtained by replacing edge  $1_x 2_y$  to edge  $1_x 6_y$ .

### 2.2 Breaking the Graph into Simpler Components

When possible, DNA sequence assemblers based on graph models aim at finding a path that covers all vertices in the graph, which represents the original genome sequence. This scenario, however, is hard to be achieved due to several issues such as sequencing errors, SNPs, lack of coverage, and repeats. Therefore, the objective becomes to find a minimum set of paths (or walks in the case of repeats) that cover the vertices. Such paths (and possibly walks) will represent contigs, i.e., contiguous regions of the genome. A small set of paths is preferable because few and long contigs make more sense biologically speaking.

The initial part of the method proposed here is based on a previous work for overlap graphs [3]. Adaptations and increments are presented here for the case of  $k$ -mer graphs. In a first stage, we break the graph into components that are easier to cope with. This breakage uses the idea of maximum-weight maximum-cardinality matching (MWMCM) over a bipartite version of  $k$ -mer graphs. This version is called the double of the  $k$ -mer graph, as stated below.

Let  $G = (V, A)$  be a directed graph. The *double* of  $G$  is an undirected bipartite graph  $G' = (V1, V2, E)$ , where  $V1 = V2 = V$  and  $E$  is defined as follows. For each arc  $(u, v) \in A$ , there is an edge  $uv \in E$ , where  $u \in V1$  and  $v \in V2$ . Therefore,  $|A| = |E|$ . Figure 3 shows an example. Based on this definition, the following theorem holds (see complete proof in the work of Cerqueira and Meidanis [3]).

**Theorem 1** *Let  $G$  be a directed graph and  $G'$  its double. Every matching  $M$  in  $G'$  corresponds to a set  $P$  of disjoint paths, and possibly a set of cycles, that cover all the vertices in  $G$  such that  $|M| + |P| = |V|$ .*

Figure 3 shows the paths (set  $P$ ) in a graph  $G$ , which are inferred from a matching in a double graph  $G'$ . According to the theorem, considering that  $V$  is constant, as we

maximize the number of edges of a matching, the number of corresponding paths is minimized, which is our goal here. Note that this step is enough to obtain a solution (obtain good contigs) for the DNA assembly problem when overlapping graphs are used. In the case of  $k$ -mer graphs, however, this is just a first stage to produce simpler components of the graph. These components pass through a second phase for the generation of final contigs, as we describe throughout the text.

In our work, we extend the matching approach, by building a weighted graph and applying the idea of MWCM, so that the best matching can be taken. Notice that a graph can present more than one possibility for a maximum matching. In the case of a bubble, for instance, if the divergence occurred due to sequencing errors, the algorithm will produce a matching that leads to two paths, one long and one short (see Fig. 3). The long one presents a high weight, while the short one has a low weight. It facilitates the elimination of misleading paths corresponding to sequences that contain errors. The weight of an edge is defined according to its occurrence during the graph construction. In this way, an edge that connects a pair of  $k$ -mer with sequencing errors has a much lower frequency than those connecting good  $k$ -mers because the number of correct reads is much higher than the number of wrong reads. As a consequence, a matching automatically isolates paths that cover such wrong  $k$ -mers.

On the other hand, problematic topologies may appear also by natural biological variation, such as SNPs. For the case of bubbles produced by such natural variations, for instance, any possible matching produces paths with high weights. In this case, we do not want to discard but combine them to generate two long genuine paths. This combination procedure comprises our second stage. Figure 2 depicts some scenarios where combinations of components are made for the described topologies. However, we use an heuristic that encompasses all the cases shown in the figures, instead of treating each case separately, as described in the next section. Notice that for the case of frayed ropes and cycles, that are caused by repeats, our method provides one possible solution. To pursue the ideal solution, it is necessary to perform sequencing with the so-called paired-end reads, where the distance of pair of fragments in the genome is known. Currently, our method treats only the case of single-end reads.

### ***2.3 Combining Graph Components to Generate Contigs***

After breaking the graph into simpler components, and eliminating the low-weight ones (sequencing error treatment), the second part of our procedure combines the resulting components to generate long paths (or walks in the case of cycles—Fig. 2d) corresponding to contigs of the genome. However, in the case of bubbles, for instance, which is illustrated in Fig. 2b, there are four possibilities for the MWCM procedure output. One possible set of components is shown in the part 2b-ii. Each possibility requires different steps to combine the components for generating the answer shown in Fig. 2b-iii.

In order to use a single procedure to combine components from any of the described topology, covering all possible MWMCM solution for each case, we further break the components into even simpler elements called unipaths. The term unipath was first coined by Butler et al. for the ALLPATHS assembler [2]. A unipath is a subgraph that represents a maximal unbranched sequence in a genome, relative to a minimum overlap  $k$ . To obtain such subgraphs from a  $k$ -mer graph  $G$ , we inspect the MWMCM components to remove the arcs that are incident to vertices with indegree and/or outdegree greater than 1 in  $G$ . The graph in Fig. 4 has its unipaths labeled  $a$ ,  $b$ ,  $c$ ,  $d$ , and  $e$ . All arcs removed to obtain unipaths and also the arcs not selected by the matching procedure are used later to combine unipaths in order to produce long paths/walks in the graph. These arcs used for the combination procedure are called outarcs (labeled with numbers in Fig. 4). Algorithm 1 sums up our method.

---

### Algorithm 1 DNA sequence assembler

---

```

1: procedure DNAASSEMBLER(reads)
2: input: a list of reads
3: output: a set of contigs
4:  $G \leftarrow \text{BUILDKMERGRAPH}(\text{reads})$ 
5:  $\text{comps} \leftarrow \text{GETCONNECTEDCOMPONENTS}(G)$ 
6:  $\text{matchings} \leftarrow \text{MWMCM}(\text{comps})$ 
7:  $H \leftarrow \text{JOINMATCHINGS}(\text{matchings})$ 
8:  $NLP \leftarrow \text{GENERATEUNIPATHS}(H, G)$ 
9:  $ScLP \leftarrow -1$ 
10:  $ScNLP \leftarrow \text{CALCLPSCORE}(NLP)$ 
11: while  $\text{SIZE}(NLP) > 1$  and  $ScNLP > ScLP$  do
12:    $LP \leftarrow NLP$ 
13:    $ScLP \leftarrow ScNLP$ 
14:   for all outarc  $a$  do
15:      $NLP \leftarrow NLP \cup \text{GENERATENEWPATHS}(LP, a)$ 
16:   end for
17:    $NLP \leftarrow \text{REMOVEREDUNDANCY}(NLP)$ 
18:    $ScNLP \leftarrow \text{CALCLPSCORE}(NLP)$ 
19: end while
20: if  $\text{SIZE}(NLP) = 1$  then
21:    $LP \leftarrow NLP$ 
22: end if
23:  $\text{contigs} \leftarrow \text{GENERATECONTIGS}(LP)$ 
24: return  $\text{contigs}$ 
25: end procedure

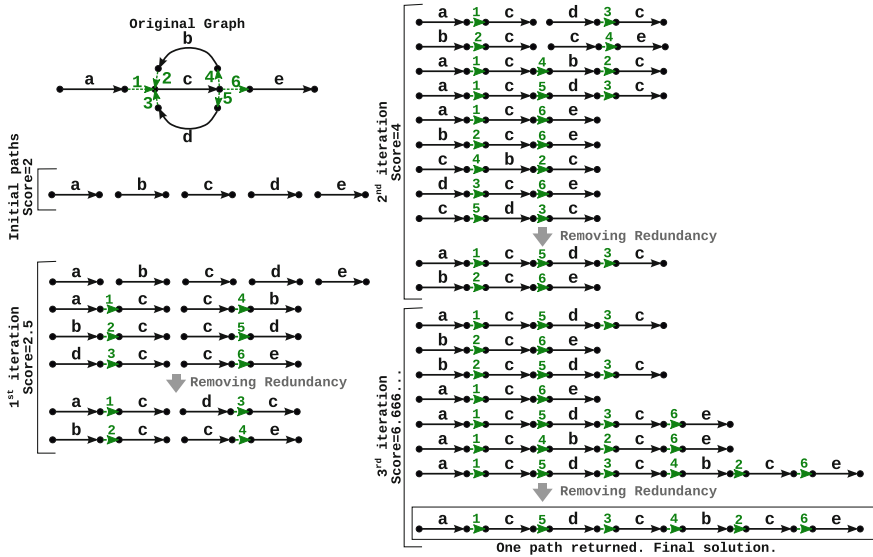
```

▷ Finds an MWMCM for each component in *comps*  
 ▷  $H$  is the union of all matchings obtained in Line 6  
 ▷  $NLP$ : list of unipaths (initial set of paths)  
 ▷ Score of the current list of paths/walks  
 ▷ Score of the new list of paths/walks  
 ▷  $LP$ : the current list of paths/walks  
 ▷  $LP$  gets the single path in  $NLP$

---

After building the  $k$ -mer graph, the first part of Algorithm 1 (lines 5–8) concerns the breakage of the graph into unipaths through the MWMCM procedure plus the elimination of arcs incident to branching vertices. Notice that the double of the  $k$ -mer graph is not explicitly constructed because it is trivial to work directly in the  $k$ -mer graph simulating its double.

Another important thing to highlight in our approach is the efficiency in solving the MWMCM problem in a divide-and-conquer fashion. Although this problem is solvable by the Kuhn-Munkres algorithm, which is  $O(n^3)$  for a complete bipartite graph  $K_{n,n}$  [6], the enormous size of  $k$ -mer graphs would result in a very large instance of the problem. Nevertheless, the highest value of the in/outdegree of any vertice is bounded to the size of the alphabet  $\Sigma = \{A, C, T, G\}$ . As a consequence, in the worst



**Fig. 4** An example applying Algorithm 1 for cycles. Numbers represent outarcs and letters represent unipaths

case, a connected component of a double graph is a bipartite subgraph of the form  $K_{|\Sigma_1|, |\Sigma_2|}$ . Thus, in place of trying to solve a huge instance (the entire double graph) of the MWCM problem, we solve multiple small instances, corresponding to each connected component of the double graph (algorithm’s line 5 and 6). It provides a two-fold advantage. First, the Kuhn-Munkres algorithm delivers a quick answer to such small instances. Second, this solution is very suitable for parallelization because each small instance (each connected component) can be solved independently. The merge procedure (algorithm’s line 7) of sub-solutions into a global solution is trivial. This is the union of all edges comprising each connected component matching.

In a second stage (algorithm’s lines 9–19), after creating a first pool of paths, the algorithm enters a repetition, where in each iteration a new pool of paths (and, eventually, walks) is generated. The new pool is constructed by using outarcs to combine elements of the previous pool (lines 14–16), so that previous paths can be elongated. Note that the crossing of paths using an outarc might lead to a redundant path, i.e., a path that covers unipaths that have already been covered in other paths. Remember that we need a minimum number of long contigs because it makes more sense in biological terms. For these reason, a procedure was created to eliminate redundant paths (line 17). This procedure basically removes paths whose covered unipaths appear also in longer distinct paths. This process of producing new pools of paths ends when the newly created pool contains exactly one path or could not be improved when compared to the path set of the previous iteration.

In order to be able of assessing each pool, we created a scoring system to paths and pools. The score of a path (or walk) is given by:

$$S_{path} = \frac{c}{1 + \sum_{i=1}^c r_i} + c, \quad (1)$$

where  $c$  is the number of distinct unipaths covered by the path, and  $r_i$  is the number of repetitions of unipath  $i$  in the current pool of paths. Notice that we do not want to cover a unipath more than once, unless it is part of a cycle. Even in this case, we have a parsimony criterion, i.e., we want to use such a unipath as few as possible. As a result, a path score is maximum if the path covers all unipaths without repeating them. The “+ $c$ ” part of Eq. 1 is to favor paths that cover more unipaths in case of a tie concerning the fractional value. The score of an entire pool is given simply by the mean of the scores of the constituent paths.

Still concerning the possibility of parallelization, the combination of paths and the redundancy elimination (lines 15 and 17 of Algorithm 1) are adequate procedures to parallelize either. In the case of path combination, for each outarc  $a$  used to combine a pair of paths, a different process might be started. The procedure of redundancy elimination of a pool, in turn, might be also divided into independent processes, one for each path, where unipaths are classified as redundant or not by consulting a common data structure that indicates the paths that cover each unipath.

After obtaining the best possible pool of paths, the algorithm extracts and returns the consensus sequences of the resulting contigs, based on the  $k$ -mer order imposed by the paths (lines 23 and 24). Figure 4 shows an example for the case of a cycle to illustrate how the algorithm works.

### 3 Results and Discussion

Algorithm 1 was implemented using the Java programming language. We performed experiments with three bacterial genomes as well as two stretches extracted from two distinct eukaryote genomes (all sequences were downloaded from NCBI): *Carsonella ruddii* pv (159662 nt); *Nasuia deltocephalinicola* (112091 nt); *Sulcia muelleri* (190733 nt); *Drosophila melanogaster*—a stretch of 10000 nt of chromosome 3R starting at position 8728682; *Homo sapiens* – a stretch of 10000 nt of chromosome X starting at position 122951623. For each of these sequences, we used the program simLibrary<sup>1</sup> to produce simulated reads of 100 nt according to the characteristics of the Illumina NGS technology, and considering a genome coverage = 80. The sets of resultant reads produced for these genomes contain, respectively, 127730, 89673, 152586, 8000, and 8000 reads.

---

<sup>1</sup><http://www.ebi.ac.uk/goldman-srv/simNGS/simNGS/man/simLibrary.1.html>.

**Table 1** Results for  $k = 41$  (a), 51 (b), and 61 (c)

		Consensus (nt)	Identity (nt) <sup>a</sup>	Gaps	Strand	t(s)	mem. (GB)
(a)	<i>C. ruddi</i>	159662	159652/159652	0	←	113	22
		159661	159661/159661	0	→		
	<i>N. deltocephalinicola</i>	112088	112081/112088	0	←	75	6
		112088	112081/112088	0	→		
	<i>S. muelleri</i>	190656	190635/190714	79	→	144	13
		190678	178150/178181	30	←		
	<i>D. melanogaster</i>	9991	9991/9991	0	←	6	2
		9996	9996/9996	0	→		
	<i>H. sapiens</i>	6486	4538/4551	0	→	5	1
4578		4434/4447	0	→			
9998		9998/9998	0	←			
(b)	<i>C. ruddi</i>	159652	159652/159652	0	←	121	13
		159661	159661/159661	0	→		
	<i>N. deltocephalinicola</i>	112088	112081/112088	0	→	80	10
		112088	112081/112088	0	←		
	<i>S. muelleri</i>	190769	90790/90797	3	←	132	31
		190680	190674/190709	35	→		
	<i>D. melanogaster</i>	9991	9991/9991	0	←	6	1
		9996	9996/9996	0	→		
	<i>H. sapiens</i>	9964	9964/9964	0	←	5	1
9975		9975/9975	0	→			
(c)	<i>C. ruddi</i>	159661	159661/159661	0	→	136	21
		159617	159617/159617	0	←		
	<i>N. deltocephalinicola</i>	112088	112081/112088	0	→	97	25
		112088	112081/112088	0	←		
	<i>S. muelleri</i>	190731	190726/190731	5	←	161	8
		190703	190703/190713	10	→		
	<i>D. melanogaster</i>	9996	9996/9996	0	→	7	2
		9991	9991/9991	0	←		
	<i>H. sapiens</i>	9985	9985/9985	0	→	6	2
9974		9974/9974	0	←			

<sup>a</sup> Identity is the number of nucleotide matches within the region of alignment

The assembler was run in a computer of 32 processors Intel® Xeon® CPU E7-8837, 2.67 GHz, 8 CPU cores, and 270 GB of shared memory. For each set of reads, we have run our program using three different values for  $k$ : 41, 51, and 61. Next, we compared the resulting assembly to the original genome sequence by using the online sequence alignment program Blast. Table 1 shows the results for all species.

Table 1 demonstrates for all organisms, and for the three values of  $k$ , that the algorithm could find at least one consensus in some direction (sense or antisense)

corresponding to the reference genome sequence with a very good alignment, i.e., a very high percentage of coverage in both aligned sequences, a very high percentage of identity, and a very low percentage of gaps (mostly zero). In most cases, both consensus sequences (in both directions) found by the algorithm agreed very well with the reference genome. Especially in the case where we used  $k = 61$ , all organisms could have both sense and antisense consensus sequences successfully assembled. It is a normal practice to test many values for  $k$  in the process of genome assembly because methods based on  $k$ -mer graphs are very sensitive to this parameter. In our case,  $k = 61$  was the best value for  $k$ , although we could obtain a very good assembly for all organisms with other values either.

## 4 Conclusions

Our first challenge to create a new DNA sequence assembler was to formulate the method described so far. The second effort was to implement and validate the method. After the accomplishment of these two stages, a third and last phase will be to improve the performance of our solution in terms of CPU and memory usage by the application of parallel execution, for instance, by the use of graphics processing units (GPUs).

Concerning the method to obtain unipaths, we could use a depth-first search. However, the MWMCM procedure, as already mentioned, has two very good advantages. First, the resultant paths (and possibly cycles) presenting low weight are easily detected and eliminated from the rest of the execution. Therefore, the identification of uninteresting paths formed due to sequencing errors is simple and there is no necessity of including a distinct error-detection procedure. The second advantage is that the bipartite version of the  $k$ -mer graph is composed of a huge number of very small connected components for which the MWMCM problem is solved individually and the overall matching is trivially obtained by the union of individual solutions. This is the perfect scenario for the use of parallel solutions such as GPUs.

Other important issues that we intend to cover in the future is to include the possibility of accepting paired-end reads and also a special treatment in the algorithm for distinguishing paths corresponding to each DNA strand. We have seen in the experiments that the lack of this treatment was not a problem possibly because paths produced by the MWMCM approach tend to be long, i.e., short misleading paths are rare to occur. However, if some regions in one strand of the genome are repeats of regions in the opposite strand, it might result in chimeric contigs.

The experiments have shown a promising DNA assembly tool that is very suitable for parallel approaches. A new version based on a parallel implementation will be able to provide fast and high quality assemblies for complete eukaryotic genomes, which will allow us to compete with current assemblers. This future endeavor will also bring the possibility to handle metagenomic data.

**Acknowledgments** This work is supported by FAPEMIG, CAPES, and CNPq.

## References

1. Bahassi, E.M., Stambrook, P.J.: Next-generation sequencing technologies: breaking the sound barrier of human genetics. *Mutagenesis* **29**(5), 303–310 (2014)
2. Butler, J., MacCallum, I., Kleber, M., Shlyakhter, I.A., Belmonte, M.K., Lander, E.S., Nusbaum, C., Jaffe, D.B.: Allpaths: De novo assembly of whole-genome shotgun microreads. *Genome Research* **18**(5), 810–820 (2008)
3. Cerqueira, F.R.: DNA fragment assembly (in Portuguese). Master's thesis, Institute of Computing, Unicamp, São Paulo (2000)
4. Conway, T.C., Bromage, A.J.: Succinct data structures for assembling large genomes. *Bioinformatics* **27**(4), 479–486 (2011)
5. Miller, J.R., Koren, S., Sutton, G.: Assembly algorithms for next-generation sequencing data. *Genomics* **95**(6), 315–327 (2010)
6. Munkres, J.: Algorithms for the assignment and transportation problems. *Journal of the Society for Industrial and Applied Mathematics* **5**(1), 32–38 (1957)
7. Nagarajan, N., Pop, M.: Sequence assembly demystified. *Nat Rev Genet* **14**(3), 157–167 (2013)
8. Pop, M.: Genome assembly reborn: Recent computational challenges. *Briefings in Bioinformatics* **10**(4), 354–366 (2009)
9. Schatz, M.C., Delcher, A.L., Salzberg, S.L.: Assembly of large genomes using second-generation sequencing. *Genome Research* **20**(9), 1165–1173 (2010)
10. Zerbino, D.R., Birney, E.: Velvet: Algorithms for de novo short read assembly using de Bruijn graphs. *Genome Research* **18**(5), 821–829 (2008)



# A New Graph Algorithm for the Analysis of Conformational Dynamics of Molecules

D. Barth, S. Bougueroua, M.-P. Gageot, F. Quessette, R. Spezia and S. Vial

**Abstract** In molecular dynamics, there can be conformational dynamics of molecules. To identify the conformational space, chemists take a series of snapshots of molecule at regular time then they try to analyze them by their eyes. When a serie contains thousands of snapshots or when the molecule is very big, this work is not possible. We present a new algorithm which aims to analyze the evolution of molecules under specific environmental conditions and characterize and quantify the isomers (conformations that could take a molecule). The algorithm is based on graph theory, it uses the distance geometry and some chemical properties (covalent bonds, H-bonds, proton transfer...). Optimizations were done to the algorithm to let it faster an efficient and to have a good time complexity. It can handle series with thousands of snapshots of molecules in Gas phase and analyze them in a few milliseconds.

## 1 Introduction

In the context of molecular dynamics simulations, different structures of the same molecule can be obtained due to changes of covalent bonds, proton transfers or appearances/disappearances of hydrogen bonds. The behavior of a molecule can change according to the environmental conditions (temperature, solvation...) [1]. These structures are called *conformations*. One of the basic challenges of computational chemistry is to identify the conformational dynamics of these molecules along trajectories and to be able to understand how conformations are related to one another. Currently, there are some softwares like VMD [4], Avogadro [3], GaussView [2] that allow a dynamic visualization of the whole molecule along the trajectory. However the dynamic analysis of different structures obtained is done by chemists using their

---

D. Barth · S. Bougueroua · F. Quessette · S. Vial (✉)  
PRISM FRE 3709 CNRS—University of Versailles Saint-Quentin-en-Yvelines,  
Versailles, France  
e-mail: sandrine.vial@uvsq.fr

M.-P. Gageot · R. Spezia  
LAMBE UMR 8587 CNRS CEA—University of Evry Val D'Essonne, Évry, France

eyes. This analysis is possible for short trajectories, however, it takes a lot of time up to some months for only one trajectory containing less than hundred snapshots. In addition, when it comes to big trajectories containing thousands of molecules, the manual analysis becomes impossible and an automatic method is crucial.

Few years ago, Logan et al. have proposed new methods for performing topological analyzes of chemical networks based on graph theory [5, 6]. Unfortunately, these methods are not very useful because they are restricted to solvatations environment and they use an external database to confirm the results.

In this paper, we present a new systematic algorithm for analyzing the evolution of molecules under specific environmental conditions and for characterizing and quantifying the conformations found along the trajectories in a fast and efficient way.

In the next section we define chemical basis for the conformation definition and discuss how they are used in our algorithm. We then describe the algorithm and the main outputs related to the conformations of molecule. In the last section, we show some experimental results using trajectories of molecules in Gas phase.

## 2 Materials and Methods

### 2.1 Material

In this work, we use the atomic positions in trajectory as input to have the list of conformations. Before describing our algorithm, below we give some definitions related to the context used:

**Definition 1** (*Trajectory*) A *trajectory* is a series of snapshots of molecule during a period of time taken at regular intervals. Each snapshot contains cartesian coordinates of molecular's atoms and their types. In one trajectory, atoms are numbered in the same order in all snapshots and therefore atoms can be easily identified.

**Definition 2** (*Conformation*) The spatial arrangements of atoms and chemical bonds between them give the molecule a specific structure which is called *conformation*.

Over the time, atoms move and bonds between them change. This dynamics leads to change of the structure/conformation of this molecule. In a trajectory many conformations can be identified. Chemists need to get this list of conformations over the time and how these conformations change.

In this work, we propose an algorithm for analyzing molecular trajectory, identifying the different conformations that can take the molecule and describing how conformations are related to each other.

In order to identify molecular's conformation in one snapshot, we are interested only in the bonds between atoms. There exist several types of bonds:

**Covalent bond** it is the strongest link that can be created between atoms. It is a chemical bond in which two atoms share some of their electrons.

**H-Bond** it is weaker than the covalent bond. It is a type of chemical bond in which a hydrogen atom that has a covalent link with an electronegative atoms called *donor* forms an electrostatic link with another electronegative atom called *acceptor*.

In the following, we modelize these bonds according to the following definition:

**Definition 3** (*Covalent bond*) Two atoms are related with a covalent bond if they are at a distance less or equal to a threshold distance  $D_c$ .

**Definition 4** (*H-bond*) H-bond is created if two conditions are satisfied: distance between hydrogen and acceptor is less or equal to the threshold distance  $D_h$  and the angle between the three atoms is in the intervalle to  $[\pi - \alpha_h, \pi + \alpha_h]$ .

In trajectories, bonds can be broken and new ones can be created, especially for H-bonds. A special case for H-bonds deletion and creation is, when the distance between the donor and the hydrogen is greater than the distance between the acceptor and the hydrogen, hydrogen atom is more attracted by the electronegativity of the acceptor. Therefore, the covalent bond is broken and then formed again with the acceptor and an H-bond is made between the hydrogen and the former donor. This case is called a *proton transfer*.

## 2.2 Methods

To analyze all trajectories and identify the different conformations, we are interested in atom's positions in order to identify all types of bonds. In our case, co-valent bonds are considered as fixed over time. Covalents bonds are only computed in the first snapshot of a trajectory. In order to identify the conformation for every snapshot, we only need to compute H-bonds and proton transfers.

To identify H-bond for an hydrogen atom, we should compare this atom with all atoms of the molecule and check if the H-bond conditions related to distance and angle are satisfied. However, this is very expensive. If  $N$  is the number of atoms and  $S$  the number of snapshots, the complexity of computing bonds is in  $O(N^2 \times S)$ . However, bonds between hydrogen and atoms are done in the neighborhood of Hydrogen. Moreover, we have noticed that between two consecutive snapshots, atoms move slowly. So instead of computing bonds between hydrogen and all the other atoms, we define a neighborhood of each hydrogen and compute H-bonds in this neighborhood. The neighborhood vary very slowly over time, so it can be set for consecutive snapshots and re-compute only when necessary. The snapshots where we re-compute neighborhood are called *references*. The first snapshot in a trajectory is the first reference and we will change it if the following sufficient condition is fulfilled: in a snapshot find the two atoms that have had the largest displacement (distance between their actual position and their position in the reference). These distances are called  $max_1$  and  $max_2$ . We compare the sum of  $max_1$  and  $max_2$  to a threshold distance  $D_r$ , if  $max_1 + max_2 \geq D_r$  the current snapshot is the new reference.

The neighborhood or orbit of an atom is define as follows:

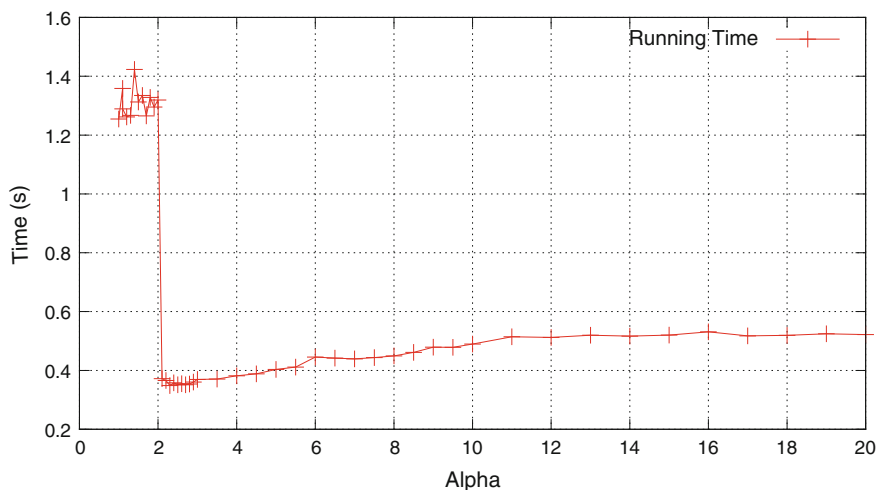
**Definition 5 (Orbit)** An orbit of an atom is a set of atoms that are at a distance less than a threshold distance  $D_r$  from this atom.

Using orbits not only reduces the number of comparisons but reduces also the running time of the algorithm. Note that orbits are fixed between two references.

We set  $D_r = \alpha \times D_c$ , where  $D_c$  is the threshold distance used for covalent bonds and  $\alpha$  is a constant. To determine the *optimum* of  $\alpha$ , we test the algorithm with several values of  $\alpha$  on many trajectories and we compare the size of orbits and the number of references and then we compare the running time. We have noticed that the different trajectories gives similar curves, Fig. 1 shows the results obtained by testing the algorithm on  $\text{H}^+ - \text{Ala}_2 - \text{COOH}^*$  trajectory. It contains 10,200 snapshots with 24 atoms/molecule.

From the Fig. 1, we notice that when  $\alpha < 2$  running time is about 1.3 s. This running time is due to the frequent computation of the orbits and reference snapshots. However, when  $\alpha \geq 2$ , this time is divided by three. The optimum interval is [2, 4]. For our tests we set  $\alpha$  to 3.

To summarize, our algorithm sets the covalent bonds and the first reference by reading the first snapshot. At each following step, the algorithm read a new snapshot, tests if this new snapshot is the new reference. If it is the new reference, it computes new orbits. In each case, algorithm computes H-Bonds and proton transfer using the orbits. The following section gives the outputs of our algorithm.



**Fig. 1** Running time according to the coefficient  $\alpha$ . We note that the running time is very high when the coefficient tends to zero, this is due to the frequent computation of the orbits and reference snapshots. We note that the optimum value of  $\alpha$  is on the interval [2, 4]

## 2.3 Outputs

After analyzing all the trajectory's snapshots, algorithm gives many output files as:

- **List of changes:** a list that contains all changes happened along the trajectory. In this file we find three types of information: changes of reference, creation or break of an H-Bond and proton transfer.
- **Evolution of H-bonds throughout the trajectory:** for each H-bond found during the trajectory, the figure indicates where it appears/disappears and if there is a proton transfer. Y-axis presents H-bonds and X-axis presents time (number of snapshots). Evolution is shown with lines, when a line is bold it means that the H-bonds is present, else it is absent. In addition to color codification is used to distinguish the proton transfer with blue color and a normal H-bond with red color.
- **Graph of conformations:** It is a weighted directed graph. It contains the different conformations found and how they are related to each other. Each state corresponds to a specific conformation, it is composed of  $n$  bits where  $n$  is the number of H-bonds found along trajectory, every bit corresponds to one H-bond, it can take the 1 value or 0 that show respectively if the H-bond is present or absent and a color codification is added to show if there is a proton transfer (blue color) or not (red color). The weights on the graph edges represent the frequencies of transition from state to another.

## 2.4 Complexity

Finally, we can give the complexity of this algorithm as follows: for a trajectory of  $S$  snapshots with  $N$  atoms/molecule, reading the input file has a time complexity of  $O(S \times N)$ . Computing covalent bonds is quadratic relative to number of atoms, the time complexity is  $O(N^2)$  and finally getting orbits and H-bonds has a time complexity of  $O(n_h \times N + n_h \times \frac{N}{10})$  where  $n_h$  is the number of Hydrogen atoms and  $\frac{N}{10}$  maximum number of atoms in an orbit. Then the whole time complexity becomes:  $O(S \times N + N^2 + k \times n_h \times N + S \times n_h \times \frac{N}{10})$  where  $k$  is the number of references snapshots (for each reference we recompute orbits). But the number of atoms  $N$  is small comparing to  $S$ , the complexity becomes:  $O(S \times N)$ .

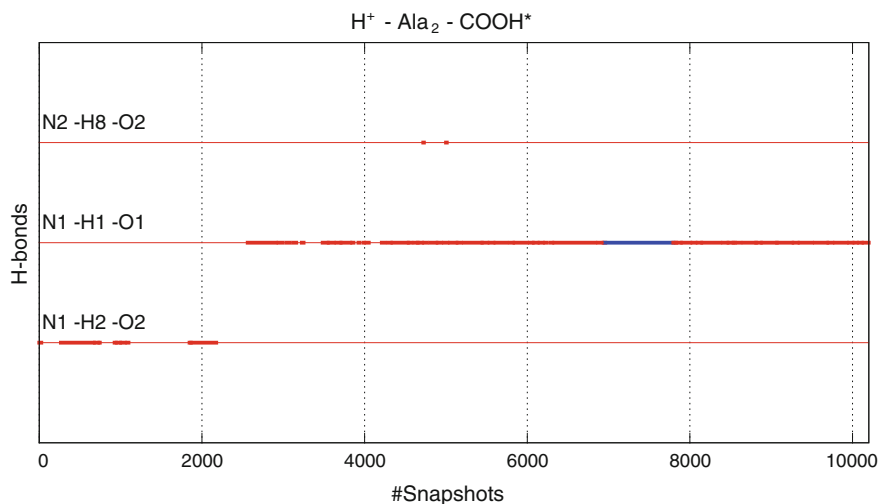
## 3 Results and Discussion

We implemented the algorithm in C environment, and we test it on many trajectories of molecules in Gas phase. We show some results, the Table 1 presents the running time (RT) in seconds for some trajectories of  $S$  snapshots an  $N$  atoms/molecule:

Results obtained by testing program on  $H^+ - Ala_2 - COOH^*$  trajectory. It contains 10,200 snapshots with 24 atoms/molecule:

**Table 1** Running time values for some trajectories

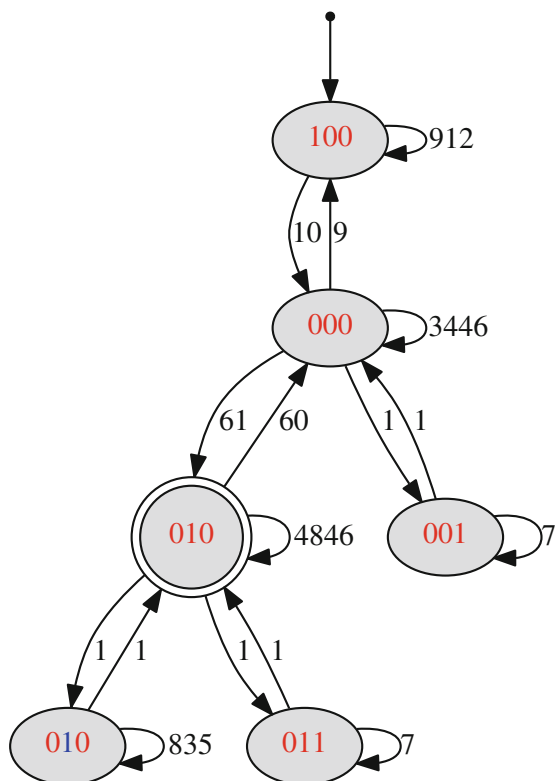
Trajectory	Number of atoms by molecule (N)	Trajectory size (S)	RT (s)
H <sup>+</sup> -Ala <sub>3</sub> -COOH	34	2101	0.50
H <sup>+</sup> -Ala <sub>3</sub> -COOH*	34	6599	0.60
H <sup>+</sup> -Ala <sub>2</sub> -COOH	24	27,000	1.00
H <sup>+</sup> -Ala <sub>2</sub> -COOH*	24	10,200	0.60
H <sup>+</sup> -Ala <sub>7</sub> -COOH	74	26,601	2.50



**Fig. 2** Evolution of hydrogen bonds according to time for H<sup>+</sup>-Ala<sub>2</sub>-COOH\* trajectory. The presence of an H-Bond is shown in *red bold line*. The presence of a proton transfer is shown in *blue bold line*. In all the other cases the line is *thin*

- Evolution of H-bonds throughout the trajectory:** in this trajectory we found 3 H-Bonds: *N1-H2-O2*, *N1-H1-O1* and *N2-H8-O2*. As shown on the Fig. 2, we notice that *N1-H2-O2* and *N1-H1-O1* do not appear at the same time. First, we have *N1-H2-O2* then after the snapshot 2000, this H-bond is broken and *N1-H1-O1* is created. We notice also that there is a proton transfer between snapshots 7000 and 7500.
- Graph of conformations:** Figure 3 presents the graph of conformations of this trajectory. There are 3 H-bonds, each state of the graph is composed of 3 bits, the most left bit represents the H-bond *N1-H2-O2* and the most right is *N2-H8-O2* and the remaining one is *N1-H1-O1*. After analyzing the trajectory, the algorithm identify 6 conformations for H<sup>+</sup>-Ala<sub>2</sub>-COOH\* molecule. They are represented in a graph as shown by the Fig. 3. On this figure, we notice that the state 010 is the most pertinent and there is an alternation between this state and 000. This means that *N1-H2-O2* and *N2-H8-O2* are almost absent and H-bond *N1-H1-O1* is created and broken all time.

**Fig. 3** Example of graph of conformations of the molecule  $\text{H}^+-\text{Ala}_2-\text{COOH}^*$  containing three hydrogen bonds



Evolution of H-Bonds and the graph of conformations give the same information. For the previous example, we noticed from the evolution of H-Bonds that  $N2-H8-O2$  rarely appears. This is confirmed by the graph, the corresponding states are 011 and 001 have very low frequency (7 times for every state).

All our results are available at [www.hydrochronographe.prism.uvsq.fr](http://www.hydrochronographe.prism.uvsq.fr).

## 4 Conclusion and Future Works

We have introduced in this paper a new algorithm to identify the conformational dynamics of a molecule along trajectory. Using the chemical properties and the geometry distances and angles, the algorithm gives good results about the conformational dynamics of molecules in Gas phase. We introduced some optimizations on the algorithm using the behaviour of atoms through the trajectory.

Further work will concern the integration of more criteria to identify conformations (dynamic of covalent bonds) and integration of the hierarchical clustering and the isomorphism methods in order to simplify the graph of conformations and allow to chemists to compare many trajectories simultaneously and analyze the behaviour of the molecule.

## References

1. Cimas, A., Vaden, T., De Boer, T., Snoek, L., Gaigeot, M.-P.: Vibrational spectra of small protonated peptides from finite temperature md simulations and irmpd spectroscopy. *J. Chem. Theory Comput.* **5**(4), 1068–1078 (2009)
2. Dennington, R., Keith, T., Millam, J., Eppinnett, K., Hovell, W.L., Gilliland, R.: Gaussview. Semichem Inc., Shawnee Mission (2009)
3. Hanwell, M.D., Curtis, D.E., Lonie, D.C., Vandermeersch, T., Zurek, E., Hutchison, G.R.: Avogadro: an advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminform.* **4**, 17 (2012)
4. Humphrey, W., Dalke, A., Schulten, K.: Vmd: visual molecular dynamics. *J. Mol. Graph.* **14**(1), 33–38 (1996)
5. Mooney, B.L., Corrales, L.R., Clark, A.E.: MoleculaRnetworks: an integrated graph theoretic and data mining tool to explore solvent organization in molecular simulation. *J. Comput. Chem.* **33**(8), 853–860 (2012)
6. Ozkanlar, A., Clark, A.E.: ChemNetworks: a complex network analysis tool for chemical systems. *J. Comput. Chem.* **35**(6), 495–505 (2014)



**Part VIII**  
**Natural Language Processing**  
**and Language Design**

# Two-Stage Feature Selection for Text Classification

Levent Özgür and Tunga Güngör

**Abstract** In this paper, we focus on feature coverage policies used for feature selection in the text classification domain. Two alternative policies are discussed and compared: corpus-based and class-based selection of features. We make a detailed analysis of pruning and keyword selection by varying the parameters of the policies and obtain the optimal usage patterns. In addition, by combining the optimal forms of these methods, we propose a novel two-stage feature selection approach. The experiments on three independent datasets showed that the proposed method results in a statistically significant increase over the traditional methods in the success rates of the classifier.

## 1 Introduction

Text classification, which is a sub-domain of classification and has been subject to active research for many years, is a learning task where pre-defined category labels are assigned to documents based on the likelihood suggested by a training set of labeled documents.

There exist several research topics related to text classification that have been extensively studied in the literature, such as the machine learning scheme used for classification, feature representation, generating new feature types (syntactic or semantic features), feature selection, performance measures, etc. In this paper, by using the well-known and the state-of-the-art methods in most of these topics, we mainly focus on the feature selection and feature filtering process and propose a novel two-stage feature extraction approach. Basically feature selection aims at eliminating unimportant and uninformative features using some statistical ranking techniques in order to reach more scalable and accurate solutions.

---

L. Özgür (✉) · T. Güngör  
Boğaziçi University Computer Engineering Department, Bebek, 34342 Istanbul, Turkey  
e-mail: ozgurlev@boun.edu.tr

T. Güngör  
e-mail: gungort@boun.edu.tr

In traditional studies, all available words in the document set were used as features instead of limiting to a set of keywords and satisfactory results were obtained [1]. Some studies even stated that using all the words leads to the best performance and using keywords may be unsuccessful without optimal parameter tuning [2, 3]. On the other hand, some studies reveal that feature selection may improve the performance in terms of accuracy and scalability with a significant cut in the solution vector size [4]. There are different types of feature selection implementations: Filter methods determine a ranking among all features with respect to some statistical metrics, wrapper methods use classical artificial intelligence techniques (e.g. greedy hill climbing) with cross validation, and embedded methods employ a linear prediction model for optimization [4]. Among them, filter methods are the simplest (in terms of implementation) and the most scalable ones for text classification problems having large feature spaces.

There are various types of feature selection metrics used in the text classification domain, such as chi-square, information gain, tf-idf, odds ratio, pruning, probability ratio, document frequency, and bi-normal separation. Concerning these metrics, there exist many studies analyzing and comparing their performances [3, 5], combining them based on specific measurements [6], and proposing supervised and unsupervised selection algorithms [7, 8].

The main concern of this paper is not the analysis or extension of these methods and metrics which has already been discussed in many recent studies. Instead, we deal with the coverage policy employed during the feature selection process: corpus-based and class-based feature selection approaches are analyzed using the appropriate metrics. The corpus-based approach uses the same feature vector for the discrimination of all the classes by selecting terms from the whole corpus as global keywords and thus favors the prevailing classes. On the other hand, the class-based approach uses a distinct feature vector for each class by considering the document set of each class separately, so that rare classes are represented equally well as the prevailing classes. In this work, we use two alternative selection approaches within these coverage policies. The first one is corpus-based pruning that takes into account the total frequencies of the terms in the whole dataset and filters the less frequent ones. The second one is class-based tf-idf (term frequency—inverse document frequency) metric that focuses on the frequencies of the terms in the documents of a class and favors those terms that do not commonly occur in other classes.

Corpus-based feature selection is the traditional approach used in classification problems: filtering the rare features that occur less than a threshold value is a classical usage of corpus-based selection [9]. As an alternative to the corpus-based approach, class-based feature selection aims at identifying important features for each class separately. A related study covers several feature selection metrics for text classification using support vector machines (SVM) [3]. While this study makes extensive use of class-based features, it also does not include an explicit comparison of the two approaches. A direct comparison between these approaches was performed with the Reuters dataset by using the tf-idf metric [10]. In that work, optimal

results were obtained around 2000 terms and the class-based approach yielded significantly more successful results than the corpus-based approach, especially with the macro-averaged F-measure. Reuters is a highly skewed dataset, so it is an expected result for macro-averaged performance to be much more affected by the class-based coverage of the terms. Singh et al. [5] proposes a new metric named as within class popularity for class-based feature selection. They aim at taking two issues about feature selection into consideration, which are the skewness of a dataset and the global importance of a term. Experiments on three datasets showed more successful results than other classifiers for class-based feature selection. In another study, a scalable architecture was proposed and class-based results were given on the Reuters dataset [11].

In this paper, we compare the class-based and corpus-based feature selection approaches using three datasets having different characteristics. The main motivation in the paper is not only making a comparison of these policies, but also analyzing their optimal usage patterns and combining these patterns to obtain higher classification performances. In this respect, we propose a two-stage feature selection approach that combines corpus-based pruning and class-based tf-idf filtering.

## 2 Proposed System

### 2.1 Datasets

In this work, we use three well-known datasets from the UCI Machine Learning Repository: Reuters-21578 (Reuters), National Science Foundation Research Award Abstracts (NSF), and Mini 20 Newsgroups (MiniNg20) [12]. These datasets have different characteristics which may be critical for the classification performance. Skewness is one of the key properties of a dataset that is defined as the distribution of the number of documents over classes. A dataset having a low skewness factor indicates that it is a balanced dataset with approximately the same number of document samples for each class. Allowance of multiple classes for documents (indicating that a document may belong to more than one topic), document length (e.g. short abstracts or long news articles), split proportions (training and test sets), formality level (e.g. formal journal documents or informal internet forum messages) are other properties of datasets.

In the experiments, we use the standard splits of the Reuters and MiniNg20 datasets. For NSF, data related with year 2001 was selected randomly and five sections (four sections for training and one section for test) were picked out from this year. We form five different splits, repeat all the tests with these five cross folds, and take their average as the final result.

## ***2.2 Preprocessing, Document Representation, and Term Weighting***

For the preprocessing of the documents, we perform all the standard preprocessing operations such as removal of non-alphabetic characters and mark-up tags, case folding, elimination of stopwords, and stemming. We use the Smart system stoplist for the removal of stopwords (<ftp://ftp.cs.cornell.edu/pub/smart>) and the widely-used Porter stemmer for extracting the root words.

The bag-of-words (bow) form is accepted as the simplest and the most successful approach for document representation in text classification problems. In this standard approach, only the words in the documents are considered as features in the machine learning algorithm used for classification. Using a machine learning algorithm with these basic features with training and test data is the direct, fundamental and conventional architecture for text classification problems [13].

As the term weighting approach, we use the tf-idf metric which is a simple measure that takes the term frequencies into account and that decreases the importance of terms common to the entire dataset by using the document frequencies [13]. For the optimized tf-idf calculation, each document vector is normalized so that it is of unit length to account for documents of different lengths.

## ***2.3 Machine Learning Algorithm***

Several studies have compared the performances of different machine learning approaches and in general SVM with linear kernel was shown to yield successful results [3, 10, 14]. For the fundamental challenges in the text classification domain (high dimensionality, sparse instances, separability of classes), SVM provides efficient solutions by being more immune to the overfitting problem, using an additive algorithm with an inductive bias that suits problems with dense concepts and sparse instances, and employing a basic linear separation model that fits the discrimination of most of the classes [15]. Based on these positive aspects and its success in previous studies, we decided to use SVM with linear kernel as the machine learning module in this work.

## ***2.4 Feature Selection***

As stated in the first section, the main motivation in this paper is focusing on both the corpus-based and class-based feature selection approaches and combining them in such a way that will increase the classifier's performance. On the one hand, for corpus-based feature selection, we apply pruning (filtering low-frequency terms) to the whole dataset and perform an analysis for the optimal pruning level using seven

different levels between 2 and 30. In the literature, usually an arbitrarily selected and small value (e.g. 2 or 3) has been used for this purpose. On the other hand, we examine class-based feature selection based on tf-idf (which is superior to corpus-based tf-idf as mentioned previously) by extracting a number of the most informative keywords in each class. We experiment with five different number of keywords between 250 and 4000 to determine the optimal number of keywords. In the rest of the paper, we will refer to these two steps as (corpus-based) pruning and (class-based) keyword selection, respectively. Finally, by analyzing the results of these filtering and selection processes and by extracting parameters corresponding to the optimal performances in these experiments, we derive an additional stage that combines the corpus-based pruning with the class-based keyword selection.

## 2.5 Methods

We basically implement four main approaches in this work: all words (AW), all words with corpus-based pruning (AWP), all words with class-based keyword selection (AWK), and two-stage feature selection with both pruning and keyword selection (AWPK).

The AW method is the baseline method that uses the standard bow approach with all the words in the feature vector. AWP considers all the words in the document collection, but filters them by the pruning process. In this method, the terms that occur less than a certain threshold value in the whole training set are filtered. We name this threshold value as the *pruning level* (PL).  $PL = n$  ( $n \geq 1$ ) indicates that terms occurring at least  $n$  times in the training set are used in the solution vector while the others are ignored. Note that  $PL = 1$  corresponds to the AW method (i.e. no pruning). We perform parameter tuning by analyzing different values for each dataset to reach the optimal PL values for the AWP method. We conduct experiments with different pruning levels between 2 and 30: 2, 3, 5, 8, 13, 20, and 30.

In the AWK method, distinct keywords are selected for each class. This approach gives equal weight to each class in the keyword selection phase. We experiment with five different number of keywords (250, 500, 1000, 2000, and 4000) and compare the results with AW that includes all the words as features in the solution vector. The AWPK method is designed as the optimal combination of AWP and AWK by varying the pruning level and the number of keywords parameters. The parameter values that yield the best results in the underlying methods are used for the AWPK experiments.

## 3 Experiments and Results

Based on the approaches discussed in the previous section, in this section we determine the optimal parameter values (pruning level and number of keywords) for the methods in all the datasets. The experiments were evaluated and the methods were

**Table 1** AWP Success Rates (Optimal Results Shown in Bold)

Method, Parameter	Reuters			NSF			MiniNg20		
	Feature#	MicroF	MacroF	Feature#	MicroF	MacroF	Feature#	MicroF	MacroF
AW	20292	85.58	43.83	13424	64.46	46.11	30970	46.42	43.44
AWP,2	12959	85.55	43.84	8492	64.41	46.21	13102	49.73	47.13
AWP,3	9971	85.52	43.93	6328	64.62	46.42	9092	49.64	47.19
AWP,5	7168	85.51	44.56	4528	64.86	46.49	6000	51.26	48.52
AWP,8	5268	85.73	44.91	3376	64.66	46.38	4169	52.48	49.90
<b>AWP,13</b>	<b>3976</b>	<b>85.84</b>	<b>44.85</b>	<b>2478</b>	<b>64.58</b>	<b>46.49</b>	<b>2863</b>	<b>53.62</b>	<b>51.02</b>
AWP,20	3046	86.02	44.55	1875	64.23	46.67	2025	53.78	51.02
AWP,30	2237	81.29	43.59	1419	63.84	46.21	1384	52.89	50.46

compared with respect to micro-averaged F-measure (MicroF), which is an average of the success rates of the documents, and macro-averaged F-measure (MacroF), which is an average of the success rates of the categories [13].

### 3.1 Pruning Level Analysis—AWP

In this experiment, the AWP method was implemented with several PL values (PL=1 corresponds to AW) for the three datasets. Table 1 shows the feature number and the MicroF and MacroF success rates for each pruning level. The first column of the table indicates the method and the value of the PL parameter, separated by comma. As can be seen, the pruning process improves the success rate of the classifier and the best results (high accuracies with low feature numbers) are obtained around PL = 13 consistently in all the three datasets with two different performance measures. By following the generalization that words occurring less than 10–15 times in a dataset are most probably not a good indicator for the classification of texts [9], we set PL = 13 in the pruning-based experiments. This result indicates that the usual belief in the literature that a pruning level of 2–3 suffices to eliminate uninformative terms does not hold.

### 3.2 Class-Based Keyword Selection Analysis—AWK

In this experiment, the performance of the AWK method was analyzed using different keyword (feature) number parameters. The results are shown in Table 2. The success rates for AW are also included in the table for comparison.

In general, the AWK method with number of keywords between 2000 and 4000 increases the success rates in all the datasets compared to the AW method. Therefore, we can conclude that using a specific set of keywords for each class gives more successful results than using all the words in the feature vector.

**Table 2** AWK Success Rates (Optimal Results Shown in Bold)

Method, Parameter	Reuters		NSF		MiniNg20	
	MicroF	MacroF	MicroF	MacroF	MicroF	MacroF
AWK,250	83.69	51.15	62.04	49.51	56.65	55.72
AWK,500	84.71	50.92	62.92	49.31	56.16	55.01
AWK,1000	85.16	51.72	64.69	49.33	53.68	52.17
<b>AWK,2000</b>	<b>85.58</b>	<b>52.03</b>	<b>65.19</b>	<b>49.31</b>	<b>54.04</b>	<b>52.10</b>
<b>AWK,4000</b>	<b>85.84</b>	<b>52.10</b>	<b>65.71</b>	<b>49.35</b>	<b>55.25</b>	<b>53.73</b>
AW	85.58	43.83	64.46	46.11	46.42	43.44

When we analyze the results of AWP and AWK together, we see that the improvement of AWP over AW is explicit in the balanced dataset (MiniNg20) while there is less improvement in the skewed datasets (Reuters and NSF). On the other hand, the improvement of AWK over AW is more significant than that of AWP in all the datasets. This performance increment is more explicit in the MacroF measure. In corpus-based approaches documents of rare classes tend to be more misclassified since the words of prevailing classes dominate the feature vector. The MacroF measure gives equal weight to each class in determining the success rate of the classifier. Thus, especially for highly skewed datasets, when the rare classes are not represented well with the selected features, average of correct classifications for rare classes drops dramatically. This is the case for both AW and AWP in skewed datasets that use a common set of features for all the classes. However, with class-based keyword selection, since each class has its own keywords during classification, rare classes are characterized in a more successful way. So, we observe a significant success rate (MacroF) increase with the AWK method in skewed datasets.

### 3.3 Two-Stage Feature Selection Analysis—AWPK

The AWPK method combines the optimal usage patterns of the AWP and AWK approaches. Therefore, the parameters in the method are the pruning level and the number of keywords. In this experiment, we use the optimal values of these parameters determined during the previous analyses for each dataset: pruning level 13 and number of keywords 2000 and 4000. The results are given in Table 3. The table also shows the best performances of AW, AWP, and AWK for comparison.

As can be seen in the table, the two-stage feature selection approach outperforms the previous approaches. Selecting the best 2000–4000 keywords for each class with an initial pruning step significantly improves the best performances of AWP (with  $PL = 13$ ) and AWK (with 2000–4000 keywords) in all the three datasets. So, we can conclude that the incremental effect of corpus-based pruning continues when it is combined with the class-based tf-idf keyword selection metric. As a result, the method proposed in this work, AWPK, yields the best performance.



**Table 3** AWPk Success Rates (Optimal Results Shown in Bold)

Method, Parameters	Reuters		NSF		MiniNg20	
	MicroF	MacroF	MicroF	MacroF	MicroF	MacroF
<b>AWPK,13,2000</b>	<b>86.40</b>	<b>53.95</b>	<b>66.06</b>	<b>50.11</b>	<b>57.43</b>	<b>55.66</b>
<b>AWPK,13,4000</b>	<b>86.70</b>	<b>53.98</b>	<b>66.10</b>	<b>50.12</b>	<b>57.43</b>	<b>55.66</b>
AW	85.58	43.83	64.46	46.11	46.42	43.44
AWP,13	85.84	44.85	64.58	46.49	53.62	51.02
AWK,2000	85.58	52.03	65.19	49.31	54.04	52.10
AWK,4000	85.84	52.10	65.71	49.35	55.25	53.73

The significance of the results for the three methods were measured using the statistical sign test. We observed that, in general, each method significantly outperforms its predecessor method. In this sense, AWP and AWK are significantly better than the standard benchmark method AW, and AWPk is significantly better than both AWP and AWK. So, the most advanced method in this study (AWPK) is the optimal method with its two-stage feature selection analysis.

## 4 Conclusions

In this paper, we focused on feature coverage policies (corpus-based or class-based selection of features) used in the text classification domain. First, we analyzed the performances of corpus-based pruning (AWP) and class-based keyword selection with tf-idf (AWK) separately. Then, determining the optimal parameter values for each method, we formed the AWPk method which is a combination of these two approaches. To the best of our knowledge, this is the first work that combines class-based and corpus-based feature selection in the text classification domain.

A possible future work is applying the two-stage feature selection approach to more semantically-oriented text classification methods, such as those using language models, linguistic features, or lexical dependencies. Integrating the concepts of pruning and keyword selection into those methods as two consecutive steps may lead to a higher classification performance.

**Acknowledgments** This work was supported by the Boğaziçi University Research Fund under the grant number 05A103D and the Turkish State Planning Organization (DPT) under the TAM Project, number 200K120610.

## References

1. Joachims, T.: Text categorization with support vector machines: learning with many relevant features. In: Proceedings of the European Conference on Machine Learning (ECML), 137–142, Springer (1998)
2. Aizawa, A., Linguistic techniques to improve the performance of automatic text categorization. In: Proceedings of 6th Natural Language Processing Pacific Rim Symposium, 307–314, Tokyo (2001)
3. Forman, G.: An extensive empirical study of feature selection metrics for text classification. *J. Mach. Learn. Res.* **3**, 1289–1305 (2003)
4. Forman, G.: Feature selection for text classification, in Computational methods of feature selection, ed. Liu, H., Hiroshi, M.: Chapman and Hall/CRC Press (2007)
5. Singh, S.R., Murthy, H.A., Gonsalves, T.A.: Feature selection for text classification based on Gini coefficient of inequality. In: Proceedings of the 4th International Workshop on Feature Selection in Data Mining, 76–85, India (2010)
6. Shoushan, L., Rui, X., Chengqing, Z., Huang, C.R.: A framework of feature selection methods for text categorization. In: Proceedings of the 47th Annual Meeting of the ACL, 692–700, Singapore (2009)
7. Dasgupta, A., Drineas, P., Harb, B., Josifovski, V., Mahoney, M.W.: Feature selection methods for text classification, In: Proceedings of 13th International Conference on Knowledge Discovery and Data Mining, 230–239, ACM, San Jose (2007)
8. Shang, W., Huang, H., Zhu, H., Lin, Y., Qu, Y., Zhihai, W.: A novel feature selection algorithm for text categorization. *Expert Syst. Appl.* **33**, 1–5 (2007)
9. Özgür, L., Güngör, T.: Text classification with the support of pruned dependency patterns. *Pattern Recogn. Lett.* **31**, 1598–1607 (2010)
10. Özgür, Arzucan, Özgür, Levent, Güngör, Tunga: Text Categorization with Class-Based and Corpus-Based Keyword Selection. In: Yolum, pInar, Güngör, Tunga, Gürgen, Fikret, Özturan, Can (eds.) Computer and Information Sciences - ISCIS 2005. Lecture Notes in Computer Science, vol. 3733, pp. 606–615. Springer, Heidelberg (2005)
11. Ghiassi, M., Olschimke, M., Moon, B., Arnaudo, P.: Automated text classification using a dynamic artificial neural network model. *Expert Syst. Appl.* **39**(12), 10967–10976 (2012)
12. Frank, A., Asuncion, A.: UCI machine learning repository, University of California, School of Information and Computer Science, Irvine, CA (2010). <http://archive.ics.uci.edu/ml>
13. Manning, C., Raghavan, P., Schütze, H.: Introduction to information retrieval, Cambridge University Press (2008)
14. Gao, Y., Sun, S.: An empirical evaluation of linear and nonlinear kernels for text classification using support vector machines. In: Proceedings of the 7th International Conference on Fuzzy Systems and Knowledge Discovery (FSKD), 1502–1505, China (2010)
15. Joachims, T.: Advances in kernel methods: support vector learning, MIT Press (1999)

# Constructing a Turkish Constituency Parse TreeBank

Olcay Taner Yıldız, Ercan Solak, Şemsinur Çandır,  
Razieh Ehsani and Onur Görgün

**Abstract** In this paper, we describe our initial efforts for creating a Turkish constituency parse treebank by utilizing the English Penn Treebank. We employ a semi-automated approach for annotation. In our previous work [18], the English parse trees were manually translated to Turkish. In this paper, the words are semi-automatically annotated morphologically. As a second step, a rule-based approach is used for refining the parse trees based on the morphological analyses of the words. We generated Turkish phrase structure trees for 5143 sentences from Penn Treebank that contain fewer than 15 tokens. The annotated corpus can be used in statistical natural language processing studies for developing tools such as constituency parsers and statistical machine translation systems for Turkish.

## 1 Introduction

Treebanks annotated for the syntactic or semantic structures of the sentences are essential for developing state-of-the-art statistical natural language processing (NLP) systems including part-of-speech-taggers, syntactic parsers, and machine translation systems. There are two main groups of syntactic treebanks, namely treebanks annotated for constituency (phrase structure) and the ones that are annotated for dependency structure. The first large-scale treebank, the Penn Treebank, was developed for English and annotated for constituency structures of sentences [13]. The development of other treebanks followed for a wide variety of languages including German [3], French [1], Arabic [12], Chinese [17], Hungarian [7], and Finnish [10].

In this study, we report our preliminary efforts for constructing a Turkish constituency parse treebank corpus. Turkish is morphologically rich language with a highly agglutinative nature. Sentences in general have an SOV order. However,

---

O.T. Yıldız (✉) · E. Solak · Ş. Çandır · R. Ehsani · O. Görgün  
Işık University, Istanbul, Turkey  
e-mail: olcaytaner@isikun.edu.tr

O. Görgün  
Alcatel Lucent Teletaş Telekomünikasyon A.Ş., Istanbul, Turkey

constituents can also be reordered for emphasizing and topicalizing certain elements based on the discourse. In addition, case markings (e.g. locative, dative) play crucial roles in identify the syntactic functions of the constituents [11].

We utilize a subset of the English Penn Treebank and introduce a semi-automatic annotation approach consisting of three phases. In our previous work [18], the parse trees in Penn Treebank were manually translated to Turkish. In this paper, the words are automatically annotated morphologically, and refined manually. As a second step, a rule-based method is developed to refine the parse trees based on the morphological annotations of the words. To the best of our knowledge, this is the first Turkish treebank annotated for phrase structure. We believe such a resource will promote statistical NLP research and applications for Turkish.

The paper is organized as follows: In Sect. 2, we give the literature review for treebank construction efforts in Turkish. We give the details of our corpus construction strategy in Sect. 3. Finally, we conclude in Sect. 4.

## 2 Related Work

There are only a handful of studies for creating Turkish treebank corpora. The METU-Sabancı Dependency Treebank is the first Turkish treebank [2]. It contains 7262 sentences manually annotated morphologically and syntactically. The syntactic annotation consists of head-dependent relations and functional categories. The METU-Sabancı Treebank has been used in several Turkish NLP studies [8, 9, 15, 16, 19].

There have also been some efforts for transforming the dependency-structure representation of METU-Sabancı Treebank into different syntactic representations. Cakici [4] extracted a Combinatory Categorical Grammar (CCG) from the METU-Sabancı Treebank with annotation of lexical categories. A finite state machine based approach was developed for generating a Lexical Grammar Formalism (LGF) for Turkish by using the sub-lexical units that reveal the internal structures of the words in [5, 6].

To our knowledge, this study is the first effort for creating a Turkish treebank corpus annotated for phrase-structures. The sentences in the corpus are selected from the English Penn Treebank. Therefore, besides other statistical NLP studies, the corpus can also be used for English-Turkish machine translation studies as a parallel treebank.

## 3 Corpus Construction Strategy

In this initial phase of our annotation efforts, we selected sentences that have at most 15 tokens, including punctuations, from the Penn Treebank II [13]. This choice reduced the total number of trees in the Penn Treebank to 9560 including 8660 trees

from the training set, 540 trees from the test set and 360 trees from the development set. We translated all of the test and development sets and nearly half of trees in the training set and obtained a total of 5143 Turkish trees.

Our corpus construction strategy is semi-automatic consisting of three stages. In the first stage, we manually translated the selected trees from the Penn Treebank. The last two stages are mostly automatic. In the second stage, we first automatically annotated words morphologically and then human annotators manually selected the correct morphological annotation for words having more than one possible morphological analyses. In the final stage, we refined the constituency parse trees by using the morphological analyses from stage two and following a set manually designed rules. In this paper, we emphasize on the last two steps, but for the sake of completeness, we also explain the first stage, which was completed in our previous paper [18].

### 3.1 Translation of Trees

We built a tool to facilitate translators' task. The tool both visualizes trees and makes the tree manipulation task much easier and faster. Additionally, the tool recommends glosses to translators based on the statistics of previously translated trees. Hence, as the number of translated trees increases, the translation task gets easier.

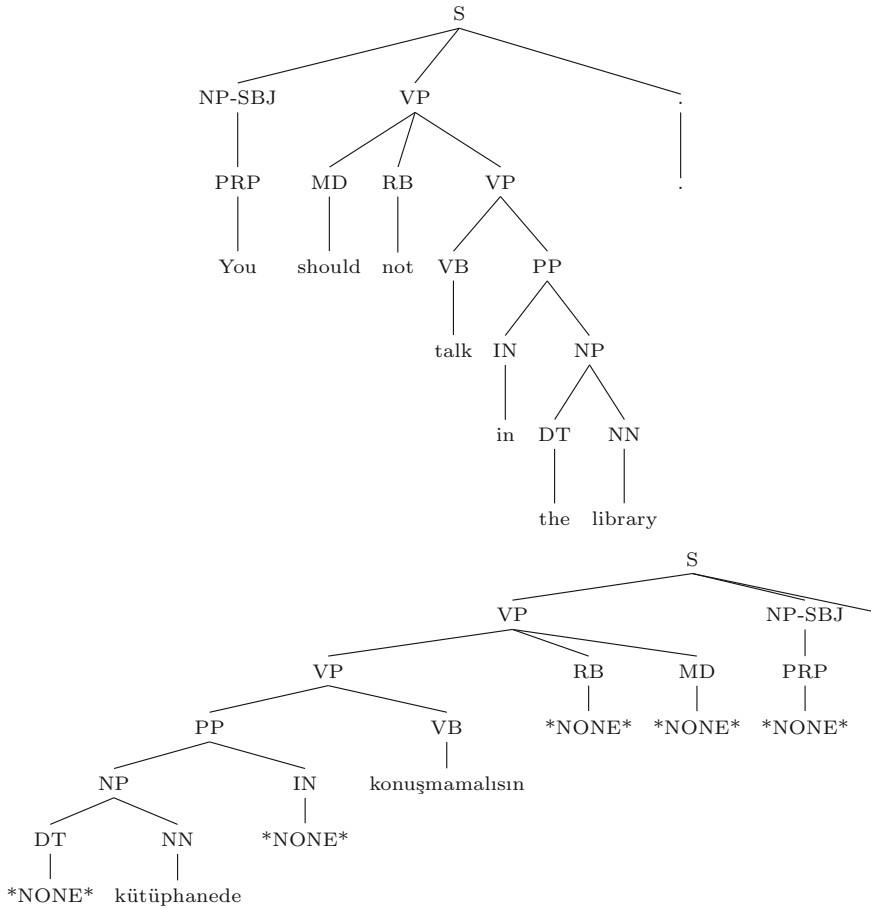
We constrain the translations of trees to two operations. We can only permute the children of a node and replace the leaf nodes with translated glosses. Since Turkish is an agglutinative language, it is often the case that we embed an English constituent in the morphemes of a Turkish stem. In such cases, we replace the English constituent leaf with \*NONE\*.

Turkish sentences generally have the SOV order. When translating English trees, subtrees are permuted to follow that order. Additionally, Turkish morphotactics determine the order of constituents.

We obtain the translated tree in Fig. 1 after following the above rules. Note that (VB talk), (RB not), (MD should), and (PRP you) are embedded in the morphological analysis "konuş-NEG-NECES-2SG" of the verb "konuşmamalısın".

### 3.2 Morphological Analysis

In general, functional words in English correspond to specific morphemes attached to the word stem in Turkish. For instance, "I/PRP will/MD not/RB do/VB" is translated to Turkish as "yap-ma-yacağ-ım" which corresponds to "do-NEG-FUT-1SG" in English. Hence, even though annotating trees syntactically is sufficient for English, morphological analysis is necessary for Turkish because of its highly agglutinative structure.



**Fig. 1** Original and translated trees, kütüphane-de konuş-ma-malı-sın (library-LOC talk-NEG-NECES-2SG)

This stage consists of two substages. First, words are morphologically analyzed using an analyzer implementation based on Oflazer’s [14]. Second, human annotators manually select the correct analysis using a graphical user interface. Figure 2 shows the morphological analysis of the translated tree in Fig. 1.

### 3.3 Morphological Annotation

This final stage gets translated trees with morphological analyses as input. We perform morphological annotation to translated trees, in addition to the syntactic annotation. We follow the set of rules described below to obtain the final Turkish constituency parse trees.

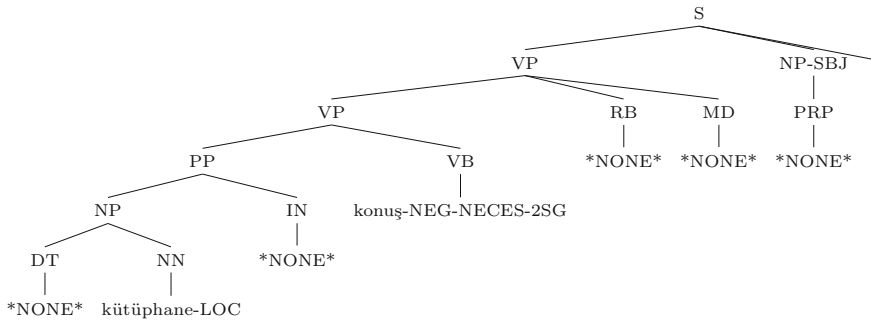
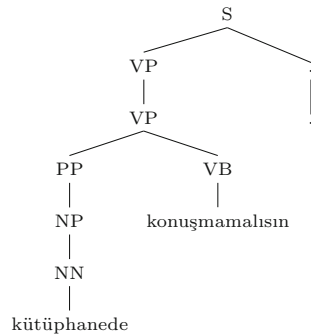


Fig. 2 Morphological analysis of the translated tree in Fig. 1

Fig. 3 After removal of \*NONE\* leaves

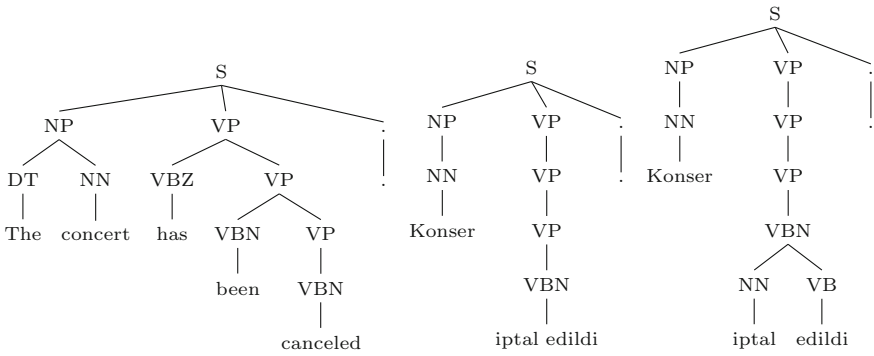


### 3.3.1 Removal of \*NONE\* Leafs

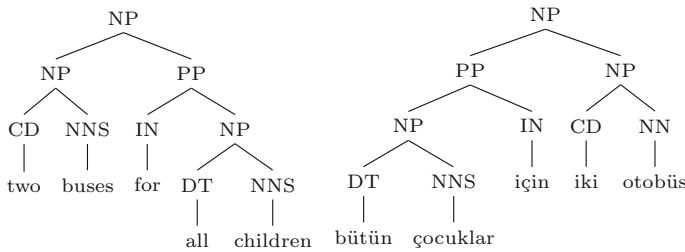
The translated trees contain \*NONE\* leafs vacated by English constituents embedded in the morphemes of Turkish stems. After morphological annotation, the semantic aspects of those English constituents will be represented in the morpheme leaves. Hence, we remove all \*NONE\* leaves and all their ancestors until we reach an ancestor that has more than one child. When we apply this rule to the translated tree in Fig. 1, we end up with the tree in Fig. 3.

### 3.3.2 Branching Multiword Leaves

A single English word may be translated to Turkish as a multiword expression. In such cases, we branch the multiword leaf into multiple leaves. We assign tags to the parents of the new leaf nodes according to their morphological analyses. For instance, in Fig. 4, the translation of “cancel” is a two word expression “iptal et”. Since “iptal” is a noun and “et” is a verb, their new parents are tagged as NN and VB respectively.



**Fig. 4** Original tree, translated tree without \*NONE\* and tree after multiword branching konser iptal ed-il-di (concert cancel-PASS-PAST-A3SG)



**Fig. 5** Original tree and translated tree after fixing NNS tags

### 3.3.3 Fixing Plural Nouns

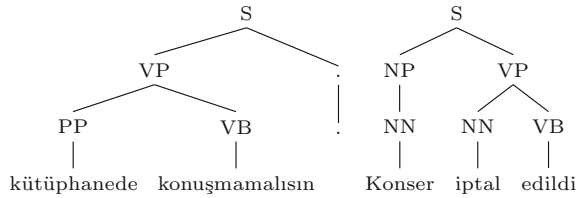
Plural nouns tagged as NNS in an English tree are sometimes translated as singular to Turkish. For example, while English nouns next to a cardinality are attached the plurality suffix, Turkish nouns are not. In such cases, we check the morphological analyses of the nouns to detect whether they have the plurality suffix “-lAr” which is equivalent to the “-s” plurality suffix in English. Since we rely on Turkish morphological analysis, irregular plural nouns of English are also tagged correctly following this rule. Figure 5 shows examples for these cases. While the translation of “children”, “çocuk-lAr”, contains a plurality suffix, the translation of “buses”, “otobüs”, does not. Therefore, we convert the NNS tag above “otobüs” to NN.

### 3.3.4 Removal of Unnecessary Ancestors

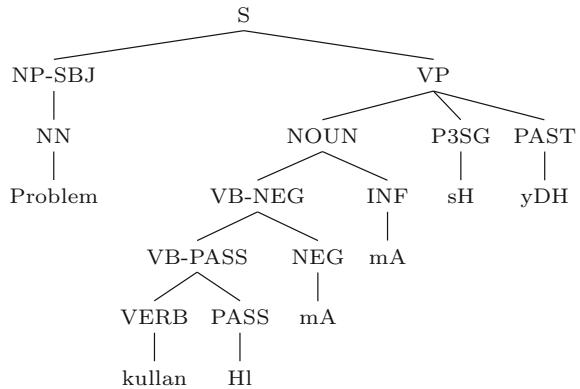
After removal of \*NONE\* leaves, we generally end up with trees that have unnecessary ancestors. For example, in Fig. 3 we have PP–NP–NN sequence and in Fig. 4 we have VP–VP–VP–VBN sequence. In the former tree NP–NN sequence and in the latter tree VP–VP–VBN sequence is unnecessary and can be removed. For each



**Fig. 6** Translated trees in Figs. 3 and 4 after ancestor removal



**Fig. 7** Example noun suffixes problem kullan-il-ma-ma-sı-ydı (problem use-PASS-NEG-INF-P3SG-PAST)



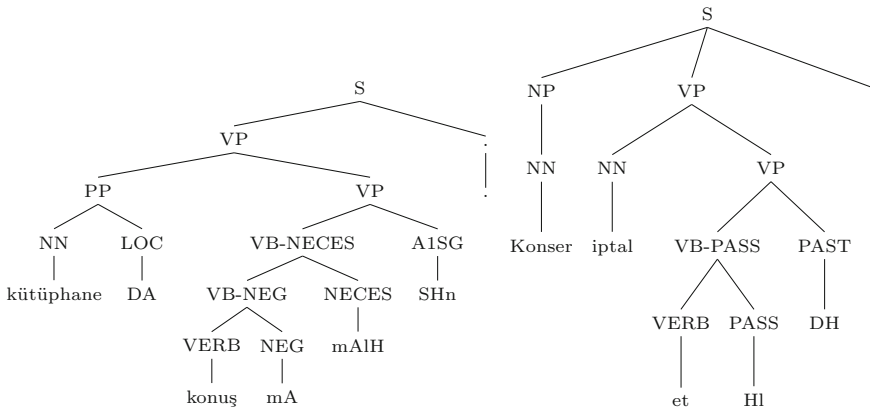
node, we remove all its ancestors until we reach an ancestor that has more than one child. However, if the leaf node does not contain any suffixes, we do not remove the immediate ancestor of that node. For instance, the leaf node “Konser” in Fig. 4 does not contain any suffixes. Hence, we do not remove its parent, i.e. the NN tag, from the tree. After applying this rule, we end up with the trees in Fig. 6.

### 3.3.5 Branching Morphemes

The final and probably the most important operation is to branch the morphemes. We need to exploit the morphological analysis to compensate the information loss that occurred during the removal operations in the previous stages.

We treat the suffixes attached to nouns and verbs differently. While all suffixes attached to a noun will be siblings of that noun, each suffix attached to a verb will generate a new parent node and the next suffix will be a sibling of that new node. Figure 7 shows examples of both noun and verb suffixes and Fig. 8 shows examples of verb suffixes. The sentence in Fig. 7, “Problem kullanılmamasıydı” is the translation of “The problem was not to be used.” As it can be seen, the suffixes attached to the nouns are siblings (i.e. P3SG and PAST) of that noun, whereas the suffixes attached to verbs create new tags (i.e. VB-PASS, VB-NEG).

In Turkish, suffixes may convert nouns to verbs or vice versa. In such cases, the corresponding node will be treated according to its final form and any additional



**Fig. 8** Final forms of the trees in Fig. 6 after morpheme branching

suffixes will be attached following the above rule. Figure 7 shows an example of converting a verb to a noun.

In this stage, we introduced new tags for morphological annotation. Verb and noun roots are tagged as VERB and NOUN respectively. Suffixes attached to verbs determine the new tag that will be created. We simply concatenate their morphological annotation to VB-. For example, as shown in Fig. 8, VB-NECES is used when the NECES suffix is attached to a verb. Figure 8 shows final forms of the trees in Fig. 6 after morpheme branching.

## 4 Conclusion

We reported our preliminary efforts for building a constituency parse tree corpus for Turkish by translating and transforming a subset of the parse trees of the English Penn Treebank. We covered over 50% of the sentences with a maximum length of 15 tokens including punctuation. We believe that this will be a useful resource for Turkish statistical NLP studies. The corpus will enable the development of statistical phrase-structure parsers for Turkish. It can also be used in English-Turkish statistical machine translation studies as a parallel treebank. As a future work, we plan to expand the dataset to include all the Penn Treebank sentences.

## References

1. Abeillé, A., Clément, L., Toussanel, F.: Building a treebank for French. In: Treebanks, pp. 165–187. Springer (2003)
2. Atalay, N.B., Oflazer, K., Say, B.: The annotation process in the Turkish treebank. In: 4th International Workshop on Linguistically Interpreted Corpora (2003)

3. Brants, S., Dipper, S., Hansen, S., Lezius, W., Smith, G.: The tiger treebank. In: Proceedings of the workshop on treebanks and linguistic theories, vol. 168 (2002)
4. Cakici, R.: Automatic induction of a ccg grammar for Turkish. In: ACL Student Research Workshop (2005)
5. Cetinoglu, O., Oflazer, K.: Morphology-syntax interface for Turkish lfg. In: Computational Linguistics and Annual Meeting of the Association (2006)
6. Cetinoglu, O., Oflazer, K.: Integrating derivational morphology into syntax. In: Recent Advances in Natural Language Processing V (2009)
7. Csendes, D., Csirik, J., Gyimóthy, T., Kocsor, A.: The szeged treebank. In: Text, Speech and Dialogue, pp. 123–131. Springer (2005)
8. Eryigit, G., Nivre, J., Oflazer, K.: Dependency parsing of Turkish. *Comput. Linguist.* (2008)
9. Eryigit, G., Oflazer, K.: Statistical dependency parsing for Turkish. In: 11th Conference of the European Chapter of the Association for Computational Linguistics (2006)
10. Haverinen, K., Nyblom, J., Viljanen, T., Laippala, V., Kohonen, S., Missilä, A., Ojala, S., Salakoski, T., Ginter, F.: Building the essential resources for Finnish: The Turku dependency treebank. *Lang. Resour. Eval.* 1–39 (2013)
11. Kornfilt, J.: Turkish. Routledge (1997)
12. Maamouri, M., Bies, A., Buckwalter, T., Mekki, W.: The penn Arabic treebank: Building a large-scale annotated Arabic corpus. In: NEMLAR Conference on Arabic Language Resources and Tools, pp. 102–109 (2004)
13. Marcus, M.P., Marcinkiewicz, M.A., Santorini, B.: Building a large annotated corpus of english: The penn treebank. *Comput. Linguist.* **19**(2), 313–330 (1993)
14. Oflazer, K.: Two-level description of Turkish morphology. *Literary Linguist. Comput.* **9**(2), 137–148 (1994)
15. Riedel, S., Cakici, R., Meza-Ruiz, I.: Multi-lingual dependency parsing with incremental integer linear programming (2006)
16. Ruket, C., Baldrige, J.: Projective and non-projective Turkish parsing. In: Fifth International Workshop on Treebanks and Linguistic Theories (2006)
17. Xue, N., Xia, F., Chiou, F.D., Palmer, M.: The penn Chinese treebank: Phrase structure annotation of a large corpus. *Nat. Lang. Eng.* **11**(2), 207–238 (2005)
18. Yıldız, O.T., Solak, E., Görgün, O., Ehsani, R.: Constructing a Turkish-English parallel treebank. In: Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, pp. 112–117. Association for Computational Linguistics, Baltimore, Maryland (2014)
19. Yuret, D.: Dependency parsing as a classification problem. In: Proceedings of the Tenth Conference on Computational Natural Language Learning (2006)

# A TV Content Augmentation System Exploiting Rule Based Named Entity Recognition Method

Yunus Emre Işıklar and Nihan Çiçekli

**Abstract** This paper presents a TV content augmentation system that enhances the contents of TV programs by retrieving context related data and presenting them to the viewers without the necessity of another device. The paper presents both the conceptual description of the system and a prototype implementation. The implementation utilizes program descriptions crawled from web resources in order to extract named entities such as person names, locations, organizations, etc. For this purpose, a rule based Named Entity Recognition (NER) algorithm is developed for Turkish texts. Information about the extracted entities is retrieved from Wikipedia with the help of semantic disambiguation and its summarized form is presented to the users. A set of experiments have been conducted on two different data sets in order to evaluate the performance of the rule based NER algorithm and the behavior of the TV content augmentation system.

**Keywords** Content augmentation · Connected TV · EPG (Electronic Program Guide) · Named Entity Recognition (NER) · Semantic disambiguation

## 1 Introduction

Nowadays, television is the main entertainment device in our living rooms and there are two cognitive modes of TV watching; experiential and reflective [1]. In the experiential mode, users usually watch TV in order to relax and entertain themselves. On the other hand, in the reflective mode, they want to learn more about TV content or program. When they want to learn detailed or specific information related to TV program while watching experience, they usually search specific keywords on the Internet by using mobile devices. However users may miss the content while

---

Y.E. Işıklar (✉) · N. Çiçekli  
Dept. of Computer Engineering METU, Ankara, Turkey  
e-mail: emre.isiklar@metu.edu.tr

N. Çiçekli  
e-mail: nihan@metu.edu.tr

searching. Even if the user doesn't care about missing the original content, it may not be easy to reach related results of TV program due to the excessive results. So what we need is a TV content augmentation system which retrieves additional information automatically and presents them to the viewers by preserving user experience quality.

Our main motivation in this study is to design a TV content augmentation system retrieving additional information and presenting them to the viewers by preserving user experience quality. Enhancement of the content meaning and intelligibility by providing additional information is known to be content augmentation. We design a TV content augmentation system by using information extraction methods.

In order to construct such a system, we first interpret meaning of TV content by extracting keywords in program descriptions. The most challenging part of this study is keyword extraction i.e. named entity recognition (NER). We have implemented a rule based NER method for Turkish texts to overcome this issue by utilizing language morphological structure and lexical resources. After keyword extraction we retrieve relevant data about the content by collecting detailed information of extracted keywords from Wikipedia web site. While retrieving the description of a keyword we also deal with ambiguous named entities. During the presentation of the relevant information to the users, we summarize additional information by constructing an information box including the most important data pieces. In order to provide the users with the additional information quickly whenever they request it, we collect and prepare additional relevant information as a summary before the broadcasting of TV programs. Towards our goal, we also evaluate our NER algorithm and TV augmentation system performance.

The rest of the paper is organized as follows. Section 2 presents the related work on TV content augmentation systems and named entity recognition on Turkish. Section 3 describes the details of the TV content augmentation system that we propose. Section 4 discusses experimental results of our rule based NER implementation. Finally, Sect. 5 presents the conclusion and future work.

## 2 Related Work

There are some TV content augmentation systems which use image and audio processing methods to interpret the meaning of the content [2, 3]. However, televisions do not have powerful processors and sufficient resources to implement image processing techniques. Exploiting social media is another method to extract content meaning and relevant data [4]. Nevertheless, in this technique there are limited resources for extra-content. Also, additional information may not be accurate and may not fulfill the requirements of the users. Some TV content augmentation systems use multi devices while presenting additional data to the users [4–6]. Presenting summarized relevant content on a second screen different from TV is a considerable functionality among similar systems. However, in such a system, viewers have to manage second device while watching TV and may miss the original content on TV screen. On the contrary, our system uses predefined TV program descriptions as a

sufficient resource and present additional data in the *overlay* mode without any other device.

Since one of the most important parts of our TV content augmentation system is NER and the system is designed for Turkish, here we summarize NER studies in Turkish particularly. Some Turkish NER systems utilize rule based approach [7]. In these systems, the information sources which are used to recognize named entities are lexical and pattern based resources. While determining candidate entities, a morphological analyzer is exploited for noun inflections, in case named entities are inflected. We have also dealt with inflected named entities similarly in our NER system. Rule based systems can be enhanced by constructing a hybrid recognizer [8] that learns from available annotated data through the agency of rote learning [9]. Another study on Turkish NER literature is given in [10] where authors utilize supervised learning strategy to learn rules automatically from the annotated data and employ rule filtering and rule refinement to increase the accuracy. Although supervised learning systems are more powerful than rule based systems, rule based approach is more suitable for our augmentation system. Therefore we exploit rule based techniques in our system.

### 3 TV Content Augmentation System

The proposed system generates and presents relevant data about current content on TV so that viewers can get specific information about the program content. The major contribution of the proposed system is using textual description of a TV broadcast program and providing additional data in real time. General structure of this system is shown in Fig. 1.

In order to show the feasibility of the proposed system, a prototype system is implemented according to its conceptual design except the actual smart TV interface part.

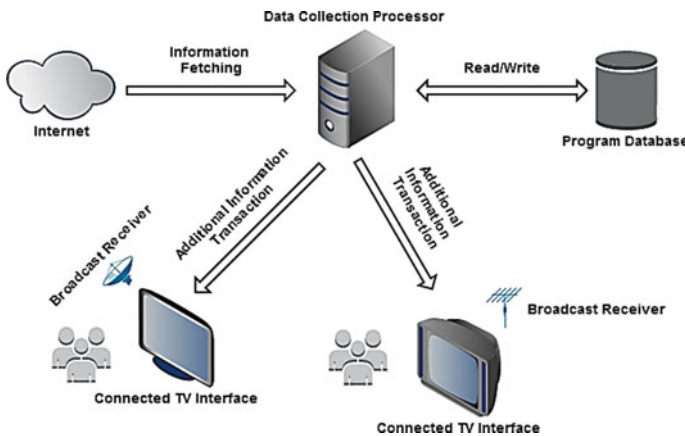


Fig. 1 General structure of the proposed system

Server side and client side are integrated in the prototype system and it is implemented as a java application running on PC. The prototype mainly consists of a TV guide collection module, Turkish NER module for keyword extraction, information retrieval module including semantic disambiguation and summarization, and a java SWT desktop application. The main issues in the implementation have been developing a NER algorithm for Turkish texts and constructing a TV content augmentation application for Turkish TV programs.

**TV Guide Collection:** In our implementation, we have used Radikal<sup>1</sup> website, which is a news site publishing daily news in Turkish to collect descriptions of TV programs. It periodically publishes three days program schedule of thirty nine channels in TV guide page of the web site. A program is described with some features like the channel name, day and time of the program, type, director, cast information, summary and long description. We have parsed html page for each channel to extract the necessary information.

**Turkish NER:** Program descriptions in Turkish are processed to recognize named entities. Since our prototype system works with Turkish TV programs and text documents, we have implemented our NER algorithm for Turkish language. We set our goal to recognize person names, organizations, locations including cities and countries, and abbreviations. Due to the limited number of entity types to be recognized, we have chosen rule based approach among other NER methods. We have defined lexical and pattern base resources in order to tag named entities. Lexical resources consist of entity type dictionaries: person names, cities and countries. Unlike other rule based systems, the number of dictionaries in our system is few. Thus, we have tried to recognize named entities by using pattern base resources. Pattern base resources include organization and location patterns which are determined by analyzing Turkish texts. There are more than 20 patterns in these resources and the following patterns are examples where X represents the entities before the specified word in each pattern:

- X Üniversitesi/Derneği/Hastanesi/Takımı/Ligi/...
- X University/Association/Hospital/Team/League/...
- X Meydanı/Caddesi/Köyü/Yolu/...
- X Square/Street/Town/Road/...

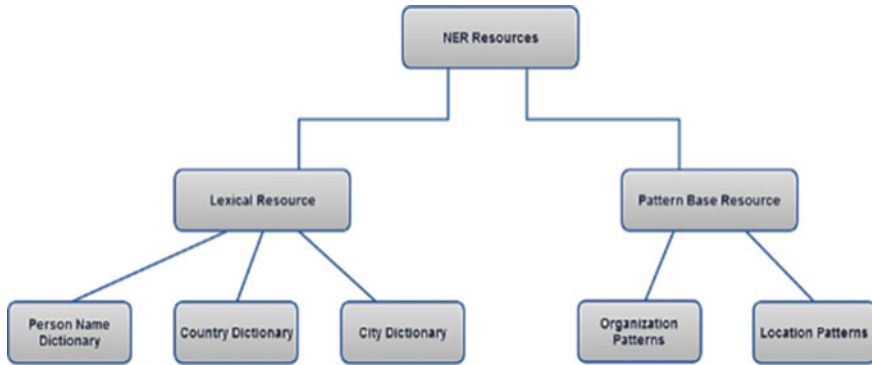
The classification of these resources is presented in Fig. 2.

We have developed our algorithm by considering the resources and determined its steps as follows:

- Determine the boundary of sentences in a given text by using sentence boundary detection method of Zemberek [11].
- Extract each word as a token from sentences and put it to the word list of the corresponding sentence.
- Run abbreviation, location, organization and person name entity finder on word lists of each sentence respectively.
- If a word or word group is tagged as an entity, add it into the annotated word list.

---

<sup>1</sup><http://www.radikal.com.tr/tvrehberi/>.



**Fig. 2** Classification of resources in our NER approach

An entity finder basically takes a word list as input and gives named entities as output. There are four types of entity finders:

*Abbreviation Entity Finder:* This entity finder simply processes each word by checking whether all letters of them are capitalized or not, after affixes are discarded. If all letters of a word are capitalized, then this word is tagged as an abbreviation. If there are dots between letters such as *S.W.A.T*, this is also considered as an abbreviation.

*Location Entity Finder:* Location named entity finder first takes entities from location dictionaries and searches them on the given sentence without addressing each word separately. If a matching case occurs, then the matched word or word group is tagged as a location entity. After using the lexical resource, finder exploits location pattern base resource. In order to use patterns, noun inflections of the word are determined by using the morphological parser of Zemberek. Noun inflections of each word are searched on the patterns list whether there is a match or not. If there is a match, it is tagged as a possible location and previous words are analyzed iteratively until finding a word that does not begin with a capital letter.

*Organization Entity Finder:* This entity finder performs the same operations as the location entity finder with the only difference that the organization entity finder does not use any lexical resource to find entities.

*Person Name Entity Finder:* Person names are recognized by utilizing only lexical resource which is a Turkish person name dictionary. The entity finder searches words on person name list after discarding affixes. If there is a match, it is tagged as person name and next words are analyzed iteratively until finding a word that does not begin with a capital letter to recognize middle name and surname. Note that, if a word is already tagged as an abbreviation, organization or location; it is ignored by this entity finder.

**Wikipedia Information Retrieval:** The user may request some more information about an extracted named entity. Our content augmentation system provides additional information for the named entity by searching it as a keyword in Wikipedia and bringing a summarized form of the information found. We retrieve wiki page



of keyword as a textual data including the information box. If there are different pages in Wikipedia for the same keyword, we retrieve all possible descriptions of the keyword from Wikipedia as a document. In this case, we carry out our disambiguation process. We use cosine similarity method which is a measure of similarity between two vectors of an inner product space [12] to measure similarity scores of each document. The vector for the document from which the keyword is extracted is compared to other documents and the one with the highest similarity score is chosen as the actual meaning of the given keyword. Retrieved Wikipedia data is too long and complicated to read while watching TV. Thus, we summarize Wikipedia texts so that users understand the related additional data at a glance. We use short descriptions of the related wiki page as a definition of keyword. Since infobox of the retrieved Wikipedia page may be exhaustive for learning important features of the keyword, we reduce the number of infobox features with respect to Wiki type of the keyword such as person or city.

**TV Scenario Application:** We have implemented this desktop application in accordance with the conceptual description of the proposed system. Since users cannot select or switch channels on a desktop application as in TV, we provide the users with the channel and program list, gathered from Radikal TV guide. The user interface of TV scenario application has four main parts as seen in Fig. 3; channel buttons on the leftmost area, program combobox with a representative image of the selected program in the middle, keyword/actor selection list at the bottom of program image, and infobox/additional information of selected keyword/actor on the rightmost area.

The usage of the application is straightforward. When the user selects a channel, the corresponding program list is presented to the user. When the user selects a program, keywords and actors about selected program are extracted by NER module and presented in the actors and keywords list. The selection of each keyword or actor triggers loading of infobox and relevant additional information area.

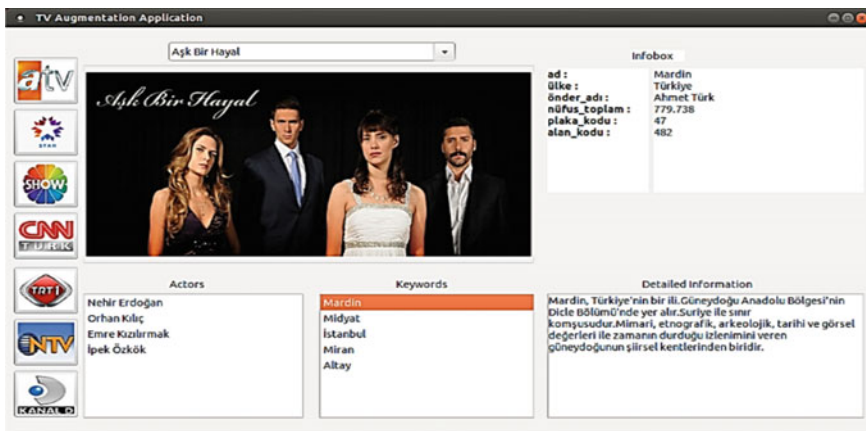


Fig. 3 A screenshot of TV augmentation application

## 4 Experimental Results

The quantitative results for each entity type and overall scores, which are calculated by considering partial matching in the data set are shown at Table 1.

The fact that F-scores for person, location and organization entities are higher than 80 % shows that our system is successful for all named entity categories. High recall and relatively low precision values for person and location entities indicate that our system tags some entities as a location or person although they are actually not. Such cases decrease the performance of our algorithm for location and person entities. On the other hand, F-score is 85.97 % for organization type entities, which is a more challenging type, indicating that our system is more successful in finding organization names than person and location entities. The overall performance F-score = 84.90 % can be considered as a successful result among similar rule based systems.

Since learning techniques in NER systems are superior to rule based methods, the systems exploiting learning techniques with lexical and morphological features produce better results than our rule based system. As a future work we might consider enhancing our rule-based system with learning methods.

In order to evaluate our TV content augmentation system, we analyzed the system by considering augmentation performance with NER operation on program description data. We utilized three-month period program description data gathered from Radical TV guide web site. Gathered data includes 28 channels and about 1700 *prime time* programs. We have considered the performance scores with overlapping boundaries presented at Table 1, while commenting on NER results of program descriptions. We have also determined a measurement technique for evaluating the performance of augmentation in our system as follows:

$$\text{Augmented Named Entity} = \text{The entity has a Wikipedia page}$$

$$\text{Augmentation Rate} = \frac{\# \text{ of Augmented Named Entity}}{\# \text{ of Entities Recognized by the system}}$$

Table 2 shows the statistical results of the experiments on the data set performed with the help of our NER and information retrieval modules.

We believe that programs including sufficient number of named entities and having high augmentation rate can increase the augmentation performance in a content augmentation system. Therefore we extracted top channels according to the number of named entities after processing 100 programs for each channel.

**Table 1** Evaluation results of the system considering partial matches

NE Category	Precision (%)	Recall (%)	F-score (%)
Person	82.43	86.52	84.42
Location	77.54	90.12	83.35
Organization	86.20	85.75	85.97
Overall	81.57	88.52	84.90

**Table 2** Augmentation performance of the system

	Person names (%)	Location names (%)	Organization names (%)	All named entities (%)
RoA (Rate of Augmentation)	34.56	83.93	61.85	49.52

**Table 3** Top channels according to number of named entities and augmentation rate

Channel name	Number of named entities	Augmentation rate (%)
Fox TV	481	34.71
A Haber	67	94.02

Table 3 shows the channels which have higher number of named entities and augmentation rate. We have observed that news channels and channels usually broadcasting series contain more satisfactory content for augmentation purposes. We exploited this inference while constructing the demo of the TV content augmentation system on PC.

## 5 Conclusion & Future Work

In this study, we have proposed a TV content augmentation system exploiting named entity recognition methods for Turkish language. We have implemented a prototype by considering user desires for such an augmentation system. Our NER system produces better results than similar systems using rule based and local grammar approach, with an overall f-score of 84.90. We have also evaluated TV content augmentation system by processing 2700 TV program descriptions. According to the experimental results, implementing such a TV content augmentation system for news channels and channels usually broadcasting series is more appropriate. Also we observed that our content augmentation system can easily be integrated to smart TVs of Turkish companies.

As a future work, we are planning to extend our NER module with different languages and supervised learning methods. Moreover, we will exploit different web resources besides Wikipedia and Radikal. Integrating our TV content augmentation system with smart TVs as an application instead of PC prototype is another future work.

**Acknowledgments** This work is supported by the Scientific and Technical Council of Turkey Grant TUBITAK EEEAG-112E111

## References

1. Taylor, A., Harper, R.: Analysis of Routine TV Watching Habits and Their Implications for Electronic Program Guide Design, pp. 1–12 (2002)
2. Dimitrova, N., Zimmerman, J., Janevski, A., Agnihotri, L., Haas, N., Bolle, R.: Content augmentation aspects of personalized entertainment experience. In: Proceedings of Third workshop on Personalization in Future TV (2003)
3. Chattopadhyay, T., Pal, A., Garain, U.: Mash up of breaking news and contextual web information: a novel service for connected television. In: Proceedings of International Conference on Computer and Communication Networks, ICCCN (2010)
4. Martin R., Holtzman, H.: Newstream: A Multi-Device, Cross-Medium, and Socially Aware Approach to News Content, pp. 83–90. (2010)
5. Hemsley, R., Ducao, A., Toledano, E., Holtzman, H.: ContextController: Augmenting broadcast TV with Real-Time Contextual Information, pp. 833–836 (2013)
6. Prata, A., Chambel, T.: Going Beyond iTV: designing flexible video-based crossmedia interactive services as informal learning contexts. In: Proceedings of EuroITV'11—Ubiquitous TV Conference, pp. 65–74 (2011)
7. Dilek, K.: Named Entity Recognition Experiments on Turkish Texts, pp. 524–535 (2009)
8. Küçük, D., Yazıcı, A.: A hybrid named entity recognizer for Turkish. *Expert Syst. Appl.* **39**(3), 2733–2742 (2012)
9. Freitag, D.: Machine learning for information extraction in informal domains. *Mach. Learn.* **39**, 169–202 (2000)
10. Tatar, S., Cicekli, I.: Automatic rule learning exploiting morphological features for named entity recognition in Turkish. *J. Inf. Sci.* **37**(2), 137–151 (2011)
11. Akın, A.A., Akın, M.D.: Zemberek, an open source nlp framework for Turkic languages. *Structure* (2007). <http://zemberek.googlecode.com/>
12. Baeza-Yates, R., Ribeiro-Neto, B.: *Modern information retrieval*, vol. 9. ACM press, New York (1999)

# A Comparison Study on Ensemble Strategies and Feature Sets for Sentiment Analysis

Deniz Aldogan and Yusuf Yaslan

**Abstract** This paper is devoted to the comparison of different common base and ensemble classifiers for sentiment classification of reviews. It is also aimed to generate different feature sets and to observe their contribution to the classification accuracy. In detail, these feature sets are formed in an hierarchical manner, which is accomplished by first forming part-of-speech (POS) based word groups and then utilizing feature frequencies, SentiWordNet scores and their combination to obtain feature sets. In addition, several common base classifiers, namely Multinomial Naive Bayes (MNB), Support Vector Machine (SVM), Voted Perceptron (VP), K-Nearest Neighbor (k-NN), as well as common ensemble strategies, Random Forests (RFs), Stacking and Random Subspace (RSS) are each tested on the generated feature sets. Also, the Behavior-Knowledge Space (BKS) method has been derived to be applied on the set of outcomes for different algorithm and feature set combinations. Furthermore, a probability based meta-classifier technique has been tested on this set of outcomes. Finally, Information Gain (IG) feature selection technique has been applied to reduce the feature spaces. The experiments are conducted on a widely used movie review dataset and an equally common multi-domain review dataset. The results indicate that the probabilistic ensemble method generally gives comparatively better results than the other algorithms tested on the chosen datasets and that IG method can be utilized to save computational time while maintaining allowable accuracy.

**Keywords** Sentiment analysis · Ensemble algorithms · Text classification · Machine learning

---

D. Aldogan (✉)

B3Lab, Department of Cloud Computing and Big Data Analysis Systems,  
Information Technologies Institute, Bilgem, Tubitak  
e-mail: deniz.aldogan@hotmail.com; deniz.aldogan@tubitak.gov.tr

Y. Yaslan

Faculty of Computer and Informatics, Department of Computer Engineering,  
Istanbul Technical University, Istanbul, Turkey

## 1 Introduction

Owing to the massive increase in the quantity and variety of data shared on social media, the analysis of such data has become crucial for most product owners, companies, universities, business entrepreneurs, etc. One of the most important reasons for these analyses is to retrieve feed-back about people, products, services, events and so on. Therefore, the sentiment analysis, which is basically a natural language processing (NLP) application for identifying text sentiment as positive, negative, neutral, etc., has become a hot research topic as mentioned in [1] or in [2]. In [2], sentiment prediction or classification is further shown as the simplest form of an opinion summary, where opinion summarization attempts to create a summary of a large number of opinions on social media. Sentiment analysis can be an extremely difficult task to achieve very high accuracies. One of the main reasons for this can be traced to the fact mentioned in [3], where it is stated that NLP's unresolved problems such as negation handling, word-sense disambiguation or anaphora resolution also exist in sentiment classification.

Several machine learning algorithms have been applied for sentiment analysis on social media data such as website texts ([4]), reviews ([5–7]) and micro-blogs ([8–11]). Another study [12] discusses sentence-level sentiment analysis. The authors claim that unigrams/bigrams combined with sentiment lexicon features result in good classification performance for subjectivity. As for polarity classification, unigrams/bigrams along with Sentiment Rhetorical Structure Theory (RST) features help creation of more robust classifiers.

There are many studies where ensemble algorithms have been used for sentiment analysis. To illustrate, the studies [13] where the authors apply ensemble methods to integrate both different feature sets and different classification algorithms or [14] where a novel boosting algorithm SharedBoost is introduced to implement transfer learning can be given. In other studies such as [15] or [16], the Behavior Knowledge Space (BKS) ensemble strategy is utilized for sentiment analysis. In their discussion of NLP techniques for feature discovery, the authors of [2] point out that shallow NLP approaches like POS tagging and parsing may not be enough to discover all the features since not all the features are explicitly specified in the text. This fact necessitates the requirement for some domain knowledge or help from some ontological word dictionary. In this study, this idea has been realized by the employment of SentiWordNet, which is a lexical resource for opinion mining and which assigns sentiment scores to words [17].

In this paper, it is aimed to utilize meta-classifier approach on hierarchically formed feature sets for a popular movie review dataset. Firstly, part-of-speech (POS) based features are obtained. Then, these groups are further divided into 3 different features, namely feature frequency based features, SentiWordNet based features and the combination of both of these features. Furthermore, the base and ensemble classifiers are tested on each of these feature sets. Moreover, the outcomes of each algorithm feature set combination are used to test a derived version of BKS algorithm and a probability based meta-classifier strategy, which benefits from the outcomes

of both base and ensemble strategies. In addition, the multi-domain review dataset is also experimented with to test the base and ensemble methods as well as a widely used feature selection approach.

The remainder of this paper is organized as follows: Sects. 2 and 3 summarize base and ensemble classification algorithms respectively. Sections 4 and 5 discuss the proposed feature sets and the ensemble models. In Sect. 6, experimental results are both introduced and discussed. Section 7 is comprised of the conclusions drawn and future work.

## 2 Base Classification Algorithms

Multinomial Naive Bayes (MNB), which is a widely used classification algorithm in sentiment analysis, is a specialized version of Naive Bayes. This version is intended to be used for solving classification problems about text documents. While Naive Bayes (NB) represents a text by constructing a model via the presence and absence of particular words, MNB models by capturing the number of times a word occurs in the document [18]. The Support Vector Machine (SVM) technique is an equally popular supervised learning technique for sentiment analysis. It has a theoretical foundation and performs classification more accurately than most other algorithms in many applications. As a supervised classification approach, SVM aims to maximize the distance to the closest point from each class so as to obtain better generalization/classification performance on the test data. The Sequential Minimal Optimization (SMO) algorithm is an algorithm for solving the quadratic programming (QP) problem that arises during the training of support vector machines [19]. The k-Nearest Neighbors algorithm (k-NN) is a non-parametric lazy learning algorithm used for both classification and regression [20]. In both cases, the input consists of the k closest training examples. The Voted Perceptron (VP) method stems from the classical Perceptron algorithm. The algorithm takes advantage of data that are linearly separable with large margins. It can also be used in very high dimensional spaces using kernel functions.

## 3 Ensemble Classifiers

We can define a classifier ensemble as a set of classifiers whose individual decisions are combined in some way to classify new instances. This resembles the case where the president of a committee (the ensemble classifier) decides on the polarity of a given document by considering the opinions of the members (the base classifiers) and the document itself. Therefore, a president who has successfully learnt when to trust each of the members can improve overall performance [21]. Research in ensemble methods has largely revolved around designing ensembles consisting of competent yet complementary models [22]. A necessary and sufficient condition for an ensemble of classifiers to be more accurate than any of its individual members is

that the base classifiers must be accurate and diverse [23]. In addition, the accuracy rate of each base classifier should not be below 0.5 and their errors should be at least somewhat uncorrelated.

In the Random Subspace (RSS) method, the algorithm randomly chooses subsets of training features rather than choosing subsets of training data. Random forest (RF) is an ensemble learning method which operates by constructing a multitude of decision trees. The RFs are chosen to correct for decision trees' habit of overfitting to the training set. The BKS Method is known to efficiently aggregate decisions of individual classifiers to derive better results [24]. It contains two stages. Firstly, it extracts knowledge from the former behavior of classifiers and constructs a K-dimensional behavior knowledge space, where K is the number of base classifiers. Then, an operation stage is performed for each test sample to combine decisions from individual classifiers. The intersection of the decisions of classifiers occupies one unit of the BKS, i.e. the focal unit (FU). In each such unit, the method accumulates three values; the total number of incoming samples, the number of incoming samples for each class and the best representative class.

## 4 Feature Engineering

In text classification, bag of words (BOW) model is widely used for feature formation. Another equally common method is using POS information. In this study, the words in the reviews are first divided into 3 groups, which are adjectives and adverbs (POS1), nouns (POS2) and verbs (POS3). Moreover, another group of words (JointPOS) is also formed by combining the 3 previous groups. In this way, 4 different vocabularies are obtained to be used by the BOW models for the same dataset. This categorization has been obtained from the study [13].

Having 4 different word groups at hand, 3 different methods are applied to represent a word in the feature set. Firstly, a feature set can be constructed by using the number of occurrences, i.e. frequencies, of the words in each of the POS based groups. Secondly, SentiWordNet can be aided to build more feature sets. By being normalized between 0 and 1, SentiWordNet scores can be utilized to represent words for each of the POS based word groups. Finally, frequency and SentiWordNet scores are both used respectively to generate a feature corresponding to a word in the BOW model. As a result, totally 12 different representations are created for the very same review in the dataset. The neutral SentiWordNet scores are discarded in order to tighten the feature space.

In this study, it is aimed to observe the contribution of each feature set to the accuracy of the classification algorithms. Another equally important reason for the creation of varied feature sets is given in the next section. Since feature space is excessively huge due to a massive number of words in the vocabulary, it becomes necessary to reduce the number of features by a proper feature selection algorithm. Information Gain (IG) has been selected as it is widely used for sentiment analysis. The number of attributes to be chosen by IG has also been experimented with.



## 5 The Proposed Ensemble Model

There are 2 different ensemble models that are proposed in this study. Firstly, the BKS model is derived by using base and ensemble classifiers as well as the instance sets. In the standard BKS algorithm, a  $C^K$  space is allocated, where  $C$  is the number of classes and  $K$  is the number of classifiers to store the outcomes of each algorithm for the training data. In this study,  $C$  equals to 2 since a review can be either positive or negative, while  $K$  is 84 as there are 12 different outcomes for each of the 7 algorithms (4 base classifiers and 3 ensemble classifiers). This is due to applying all classifiers to all the different feature sets.

These 84 different outcomes for a test instance (out of totally 7 algorithms and 12 feature sets) all possess a probability value. The outcome whose probability is highest is chosen as the resulting class in the second derived meta algorithm, namely the Probabilistic Ensemble (PE) technique.

## 6 Experimental Study

Experiments have been carried out in two steps. In the first step, different feature sets have been extracted from a widely used dataset, namely the document-level polarity dataset v2.0 of the Cornell movie-review corpora, which has been introduced in [5]. The dataset contains 1000 positive and 1000 negative processed reviews. The dataset is evenly divided into 10 folds. Each fold contains 1800 instances of training and 200 instances of test data. In the second step of experiments, another dataset, the multi-domain sentiment dataset presented in [25], has been utilized. This dataset is comprised of product reviews taken from Amazon.com. Four product types, i.e. Book, DVD, Electronics and Kitchen, have been considered to form the dataset. In the second part of the experiments for this study, the preprocessed versions of these datasets that contain 1000 positive and 1000 negative reviews for each of the four domains have been worked with. Again, the dataset is evenly divided into 10 folds so that each fold contains 1800 instances of training and 200 instances of test data as in the previous step.

### *6.1 Experiment I: Experimentation with Feature Sets and Ensemble Algorithms on Movie Review Dataset*

As a preprocessing tool for POS-tagging, POS-Tagger tool of the Stanford Core NLP was utilized. During the generation of feature sets, unigrams have been used throughout the study. Each of the four category of words, namely POS1, POS2, POS3 and JointPOS, has been further divided into other 3 new deeper subcategories in correspondence with what has been used to represent a word in the feature vector.

The first of such subcategories uses the feature frequency of a word in the review. The second one uses the normalized SentiWordNet score of a word, whereas the third one uses both of these data for each word. If a word lacks a non-zero sentiment score, it is discarded from the feature set. Eventually, 12 kinds of feature sets are formed.

MNB, SMO, k-NN and VP have been chosen as the individual classifiers, while RF, Stacking that uses RF as meta-classifier and MNB, SMO, k-NN and VP as base classifiers and Random Subspace strategy with SMO have been chosen as the three ensemble methods. WEKA framework has been used to realize all these algorithms. Instead of using merely different algorithms as base classifiers for the BKS ensemble algorithm, different algorithms on different feature sets have been inputted to the BKS application. In other words, 7 algorithms over 12 feature sets cause totally 84 different classifier outcomes to be considered for the meta-classifier. Since a BKS table of  $2^{84}$  entries is required for the algorithm to run, a hash function along with a hash table has been produced to keep these data. The results of 48 different base algorithm and feature set combinations as well as 36 different meta-classifier and feature set combinations make totally 84 outcomes for each test instance. The outcome whose probability is highest is chosen for each test instance for the PE algorithm. The accuracies obtained for each of the algorithms on each feature set are given in Table 1. The accuracies for the derived ensemble algorithms are given in Table 2. The k-NN algorithm with  $k = 1$  has been used to provide the ensemble classifiers with its contribution. In order to enhance its low performance on accuracy results, different  $k$  values can be experimented with or different stronger features other than unigrams can be applied. However, boosting k-NN's success is not considered in the scope of this study. The accuracies for JointPOS and POS1 are higher than the accuracies for

**Table 1** Accuracies of classification algorithms on different feature sets. FF, SS and both refer to feature frequency, SentiWordNet scores and usage of both methods respectively

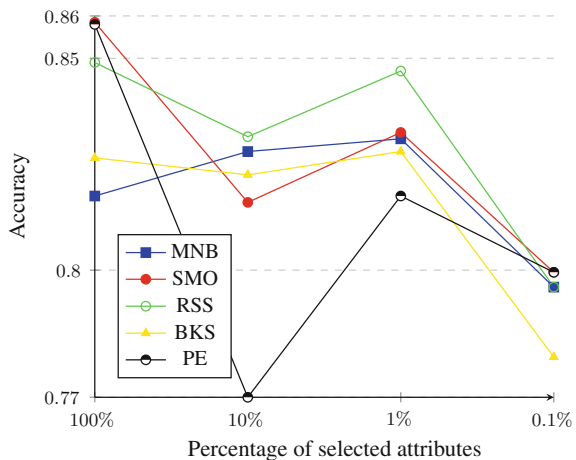
Feature set versus Algorithm	MNB	SMO	RF	k-NN	VP	Stacking	RSS
JointPOS-FF	81.75	<b>85.85</b>	63.95	50.40	80.05	80.70	84.90
JointPOS-SS	73.05	80.60	65.95	51.45	65.35	76.00	<b>80.75</b>
JointPOS-Both	76.20	80.60	66.25	51.45	71.80	77.45	<b>80.65</b>
POS1-FF	<b>84.75</b>	81.95	67.60	51.85	79.90	80.75	82.30
POS1-SS	72.75	78.30	66.80	53.15	74.60	75.65	<b>80.00</b>
POS1-Both	76.15	78.60	67.75	53.15	77.80	75.25	<b>78.75</b>
POS2-FF	<b>74.70</b>	70.50	58.10	50.40	67.10	68.15	69.90
POS2-SS	56.30	60.30	57.05	52.35	58.55	53.05	<b>60.70</b>
POS2-Both	56.60	58.85	56.40	52.35	58.70	53.00	<b>59.50</b>
POS3-FF	75.60	<b>76.60</b>	60.10	50.55	70.95	68.75	75.00
POS3-SS	65.90	67.30	60.80	51.60	50.60	63.80	<b>70.75</b>
POS3-Both	68.90	67.35	61.85	51.60	52	65.90	<b>69.00</b>

**Table 2** Accuracies of derived BKS and PE algorithms on all feature sets combined

BKS	PE
82.65	<b>85.80</b>

POS2 and POS3. Therefore, using all the words generally causes better performance than using a specific word group. POS1, the group of adjectives and adverbs, is better than the 2 other specific word groups. The worst group of words in terms of resulting accuracies is POS2, namely the group of nouns. These results for 4 different word groups are similar to those produced in [13]. As for the base classifiers, MNB and SMO are the top 2 algorithms, which are followed by VP. Among the three ensemble methods, RSS performs the best. Using feature frequencies is significantly better than using merely SentiWordNet scores or the combination of both. The derived BKS and PE algorithms obtain good accuracies, in particular PE gives the overall second best accuracy, 85.80. By observing the log files for the results, it has been inferred that the correct class is always predicted by at least one of the algorithms for all the test instances in all folds. However, this outcome is not always chosen by the PE meta-classifier since its probability may not be the highest. In order to perform an attempt to reduce the feature space to increase the accuracy, feature selection has been applied. IG technique has been applied to all of the 12 different feature sets. Afterwards, all the algorithms have been experimented with these reduced feature sets. Figure 1 shows the results for the top 5 algorithms. It has been prepared by using the average frequency results of JointPOS word group for base classifiers MNB and SMO along with RSS and the 2 derived ensemble classifiers, BKS and PE. In Fig. 1, it can be observed that the usage of IG doesn't improve accuracy except for BKS and MNB. These 2 algorithms produce their best accuracies with one hundredth of the feature set. SMO, RSS and PE algorithms undergo a performance decrease with one

**Fig. 1** Accuracy versus percentage of selected attributes for the five best classifiers on the movie review dataset



**Table 3** Accuracies of classification algorithms on multi-domain feature sets

Alg.	MNB	SMO	RF	k-NN	VP	Stacking	RSS	BKS	PE
Books	77.50	78.60	76.25	56.35	73.75	74.10	78.95	78.45	<b>79.55</b>
DVD	78.70	78.70	79.25	52.85	74.70	76.65	79.55	78.90	<b>80.10</b>
Electronics	80.70	83.15	81.70	55.75	78.75	78.10	83.20	82.55	<b>83.40</b>
Kitchen	83.55	84.60	83.40	55.40	80.60	82.25	85.30	84.05	<b>85.50</b>

tenth of the attributes, while the accuracy improves again with one hundredth of the feature set. Dividing the attribute number by 1000 again causes a much more obvious fall in accuracy. However, it can be deduced that the application of IG to keep only the one hundredth of the attributes can be a reasonable choice since it reduces the computational complexity while maintaining still acceptable accuracy results.

We can observe a very similar fall and rise pattern for the accuracy results with decreasing number of attributes in the study [26], where SVM is used with IG on the same movie review dataset. Their accuracy results versus the number of chosen attributes show close resemblance with the currently produced results. It should be noted that not only the SMO here or their SVM algorithm but also the other algorithms used in this study exhibit a similar behavior in terms of accuracy results.

## ***6.2 Experiment II: Experimentation with Base and Ensemble Algorithms on Multi-Domain Review Dataset***

In the second step of the experiments, both base and ensemble algorithms previously used have been applied to the second multi-domain dataset collection. The dataset has been preprocessed so that each word or 2 subsequent words paired together appear with their frequency in the review. Instead of using these word pairs, each word in the review has been assigned as a specific attribute. In this way, the number of attributes in the vocabulary, i.e. the feature space, has been drastically reduced.

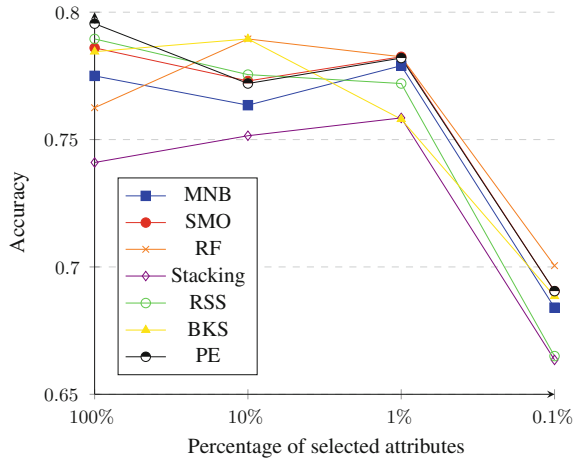
Table 3 summarizes the results of accuracies for each algorithm and review domain. From the table, it can be inferred that SMO is approximately the best performing base classifier. Unlike the previous dataset, the accuracies of the RF algorithm get very close to those of MNB. However, MNB classifier is still the second best classifier among the base classifiers.

We can observe that PE algorithm slightly exceeds the performance of all other algorithms except for the movie review dataset. The number of attributes in this dataset is much higher than the number of attributes in other datasets. Therefore, we can deduce that using the PE algorithm is more advantageous when we have less features to process at hand. The PE algorithm makes use of the results for all other algorithms, which compensates for having fewer attributes. The RSS algorithm is the second best algorithm for the datasets other than the movie review dataset. As the number of attributes decrease, the RSS algorithm becomes a better choice

**Table 4** Ensemble algorithms with the best accuracy results for the five datasets in the study [13]

Movie	Strategy 3 with meta-classifier combination
Books	Strategy 3 with weighted combination
DVD	Strategy 3 with meta-classifier combination
Electronics	Strategy 3 with weighted combination
Kitchen	Strategy 2 with weighted combination

than the SMO algorithm. In order to compare the results with previous studies on ensemble strategies, the study [13] can be referenced. However, it should be kept in mind that the results produced by using word-relation (WR) based features should not be taken into consideration since WR based features outperform unigrams due to their high computational complexity. In the paper, the results of three different strategies, namely ensemble of feature sets, ensemble of classification algorithms and finally ensemble of feature sets and classification algorithms, are shown. For each of these strategies, three meta-classifier combination approaches, fixed combination, weighted combination and meta-classifier combination, are used. Therefore, nine different outcomes for accuracy are displayed for all the five datasets. The strategy and combination pair that performs the best for each dataset is given in Table 4. In Table 4, we can see that no strategy combination pair generally performs better than the rest of the algorithms. However, in this study, it can be stated that the PE algorithm performs slightly better than all the other methods except for only one dataset. There is another aspect that needs to be considered here. The benchmark study uses feature presence approach, which has been widely accepted to perform better than feature frequency approach [13]. Also, the study [3] puts forward that feature presence forms a more effective basis to review polarity classification. It is stated that although recurrent keywords indicate a topic, repeated terms might not reflect the overall sentiment. Since the feature frequency approach has been experimented with in this study, it is believed that more promising results will be obtained once the latter approach is utilized. In order to measure the effects of feature selection on the second dataset, IG has been applied. As in the previous step, one tenth, one hundredth and one thousandth of the number of original attributes have been chosen by the IG algorithm for each of the four review domains. Figure 2 shows the results of the top 7 algorithms. As seen in Fig. 2, the accuracies for MNB, SMO and PE algorithms decrease slightly with one tenth of the attribute set, while their accuracies increase with one hundredth of the features. MNB even achieves better results with this configuration. RF, BKS and RSS algorithms output increased accuracies with one tenth of the attributes. Like the previous experimentation step, it can be inferred that using one hundredth of the attributes can be an appealing choice due to the reduction in computational time and still acceptable accuracies.



**Fig. 2** Accuracy versus percentage of selected attributes for the seven best classifiers on the books review dataset

## 7 Conclusion and Future Work

It can be inferred from the experiments that the derived PE ensemble strategy is slightly better than the other algorithms tested. Its hierarchical structure enables the utilization of ensemble algorithms over other ensemble algorithms. By combining the outcomes of base classifiers as well as the outcomes of both base and other ensemble classifiers, PE can obtain comparatively better results for 4 of the 5 different datasets.

Experimenting with different feature sets gives insight for determining the usefulness of distinct word groups. It has been concluded that usage of nouns, verbs, adjectives and adverbs altogether causes better performance, whereas the group of adjectives and adverbs display a bit less average accuracy. Furthermore, the utilization of SentiWordNet scores for forming feature sets enable to observe its performance with respect to using feature frequencies. Combining both approaches while representing each single word in the vocabulary is also attempted.

In this study, the IG technique has been applied to the movie review dataset and a multi-domain review dataset so as to decide if computational time could be saved without a significant drop in accuracy. The results display that one hundredth of the attributes can be utilized to achieve this goal. As a next step, it is aimed to work more on the PE algorithm along with feature engineering while adding more base and ensemble classifiers in the experimentation process.

## References

1. Pang, B., Lee, L.: Opinion mining and sentiment analysis. *Found. Trends Inf. Retrieval* **2**(1–2), 1–135 (2008)
2. Kim, H., Ganesan, K., Sondhi, P., Zhai, C.: Comprehensive review of opinion summarization (survey) (2011)
3. Cambria, E., Schuller, B., Xia, Y., Havasi, C.: New avenues in opinion mining and sentiment analysis. *IEEE Intell. Syst.* **28**(2), 15–21 (2013)
4. Li, S., Zhang, H., Xu, W., Chen, G., Guo, J.: Exploiting combined multi-level model for document sentiment analysis. In: International Conference on Pattern Recognition, IEEE Computer Society Washington, pp. 4141–4144 (2010)
5. Pang, B., Lee, L.: A sentimental education: Sentiment analysis using subjectivity summarization based on minimum cuts. In: Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics. ACL '04, Stroudsburg, Association for Computational Linguistics (2004)
6. Koncz, P., Paralic, J.: An approach to feature selection for sentiment analysis. In: International Conference on Intelligent Engineering Systems, Poprad, Slovakia (2011). June 23–25, 2011
7. Varma, S.: Cross-product sentiment analysis via ensemble svm classifiers. In: International Conference on Advancements in Information Technology (2011). Dec 17–18, 2011
8. Jiang, L., Yu, M., Zhou, M., Liu, X., Zhao, T.: Target-dependent twitter sentiment classification. In: Association for Computational Linguistics (ACL) (2011)
9. Bermingham, A., Smeaton, A.: Classifying sentiment in microblogs: is brevity an advantage? In: International Conference on Information and Knowledge Management, pp. 1833–1836 (2010)
10. Claster, W.B., Hung, D.Q., Shanmuganathan, S.: Unsupervised artificial neural nets for modeling movie sentiment. In: Second International Conference on Computational Intelligence (2010)
11. Li, G., Hoi, S.C.H., Chang, K., Jain, R.: Micro-blogging sentiment detection by collaborative online learning. In: IEEE International Conference on Data Mining (2010)
12. Chenlo, J., Losada, D.E.: An empirical study of sentence features for subjectivity and polarity classification. *Inf. Sci.* **280**, 275–288 (2014)
13. Xia, R., Zong, C., Li, S.: Ensemble of feature sets and classification algorithms for sentiment classification. *Inf. Sci.* **181**(6), 1138–1152 (2011)
14. Huang, P., Wang, G., Qin, S.: Boosting for transfer learning from multiple data sources. *Pattern Recognit. Lett.* **33**(5), 568–579 (2012)
15. Li, W., Wang, W., Chen, Y.: Heterogeneous ensemble learning for chinese sentiment classification. *J. Inf. Comput. Sci.* **9**(15), 4551–4558 (2012)
16. Zhang, Z., Miao, D., Wei, Z., Wang, L.: Document-level sentiment classification based on behavior-knowledge space method. *Adv. Data Min. Appl. Lect. Notes Comput. Sci.* **7713**(15), 330–339 (2012)
17. Baccianella, S., Esuli, A., Sebastiani, F.: Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In: Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10), Valletta, Malta, European Language Resources Association (ELRA) (2010)
18. McCallum, A., Nigam, K.: A comparison of event models for naive Bayes text classification. In: Learning for Text Categorization: Papers from the 1998 AAAI Workshop, pp. 41–48 (1998)
19. Platt, J.C.: Fast training of support vector machines using sequential minimal optimization. In: Advances in Kernel Methods—Support Vector Learning, MIT Press (1998)
20. Altman, N.S.: An introduction to kernel and nearest-neighbor nonparametric regression. *Am. Stat.* **46**(3), 175–185 (1992)
21. Su, Y., Zhang, Y., Ji, D., Wang, Y., Wu, H.: Ensemble learning for sentiment classification. *Lect. Notes Comput. Sci.* **7717**, 84–93 (2013)
22. Oza, N.C.: Online Ensemble Learning. Ph.D. thesis, The University of California, Berkeley (2001)

23. Dietterich, T.G.: Ensemble methods in machine learning. In: Proceedings of the First International Workshop on Multiple Classifier Systems. MCS '00, Springer, London, pp. 1–15 (2000)
24. Huang, Y.S., Suen, C.Y.: A method of combining multiple experts for the recognition of unconstrained handwritten numerals. *IEEE Trans. Pattern Anal. Mach. Intell.* **17**(1), 90–94 (1995)
25. Blitzer, J., Dredze, M., Pereira, F.: Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification. In: Proceedings of the Association for Computational Linguistics (ACL) (2007)
26. Li, S., Xia, R., Zong, C., Huang, C.: A framework of feature selection methods for text categorization. In: Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP: Volume 2- Volume 2. ACL '09, Association for Computational Linguistics, Stroudsburg, pp. 692–700 (2009)



# Feature Selection for Enhanced Author Identification of Turkish Text

Yasemin Bay and Erbuğ Çelebi

**Abstract** The rapid growth of the Internet and the increasing availability of electronic documents poses some problems, such as identification of an anonymous text and plagiarism. This study aims to determine the author of a given document among the set of text documents whose author is known. Despite the excess number of researches conducted in English language for author identification in the last century, Turkish and other languages are gaining interest only in the last decade. Therefore, this study deals with the Author Identification problem using two different Turkish datasets, collected from two different Turkish newspapers. The datasets comprises 850 columns written by 17 columnists as a total, 50 columns from each columnist. 4 different Machine Learning algorithms (Naive Bayes, Support Vector Machine, the K-Nearest Neighbor and Decision Tree) have been employed and 99.7% accuracy is achieved with K-Nearest Neighbor algorithm. The classification fully recognized with Chi-square feature selection method by reducing the features from 20 to 17.

**Keywords** Author identification · Text classification · Machine learning · Feature selection

## 1 Introduction

With the rapid growth of the Internet and the increase on the online documents it becomes one of the challenging studies to identify the author of an anonymous text. Authorship attribution and identification is one of the popular research topics gaining increased interest in text classification (TC). One of the dominating study on

---

Y. Bay (✉)

Management Information Systems, Cyprus International University,  
Lefkosa, Mersin 10, Turkey  
e-mail: yaseminbay@ciu.edu.tr

E. Çelebi

Computer Engineering Department, Cyprus International University,  
Lefkosa, Mersin 10, Turkey  
e-mail: ecelebi@ciu.edu.tr

© Springer International Publishing Switzerland 2016  
O.H. Abdelrahman et al. (eds.), *Information Sciences and Systems 2015*,  
Lecture Notes in Electrical Engineering 363,  
DOI 10.1007/978-3-319-22635-4\_34

computer-based author identification was carried out by [1] in 1964 and has inspired many researches since then. There are many applications concerning different fields in author identification, such as author verification—mining email content for author identification [2, 3], plagiarism detection [4], forensic investigations [5, 6], author profiling (i.e. gender of the author [7]), identifying authors from their source code [6, 8], etc.

Every person has their own distinct style of talking and writing, which distinguishes them from others, acts as a fingerprint and this made possible measuring features in written text. Even they use the same words as others, they tend to use them in a different way, with different combinations and patterns unique among speakers/writers, therefore, there are large number of researches [2, 9, 10] have paid increasing attention on idiolects, linguistics and stylometry during analysis of author characterization for their identification.

In 2007 [11] performed their author attribution experiments on Turkish text with function words, stylometric features and bag of words as a feature set using 500 articles from 18 different writers. The classifiers they have used are Bayes classifier, Support Vector Machines, Histogram, K-Nearest Neighbor Method and K-Means Clustering and obtained 95 % success rate with SVM.

In [12], the authors conducted their experiments using two corpuses; a training and a test sets each having 20 different authors. For identifying the authors, at the beginning 35 of style markers has been used. Researchers have applied several Machine Learning Methods in their author identification study on Turkish and obtained 80 % success rate with Naive Bayes Multinomial method using CFS (Correlation Feature Selection) Subset Evaluator for feature selection.

Another investigation of author attribution in Turkish was performed successfully by [13] using 3 different datasets in combination with 5 different feature vectors; function words, lexical, statistical, grammatical and n-gram. They have applied 5 classification methods that are Naive Bayes, Random Forest, Multilayer Perceptron, Support Vector Machine and k-Nearest Neighbour and obtained 96.9 % success rate with Multilayer Perceptron.

Despite the excess number of researches conducted in English language for author identification in the last century, Turkish and other languages are gaining interest only in the last decades. Therefore, in this paper the effectiveness of Machine Learning algorithms has been examined in Turkish text. Without considering the genre of the content two different datasets have been collected randomly, from 2 different daily newspapers broadcasting in Turkish, *Milliyet* and *Kıbrıs Gazetesi*, using a system developed in Java. The datasets consist of 850 columns written by 17 authors as a total(50 columns of each author).

The remainder of the paper is organized as follows. In Sect. 2, the Author Identification is described. Section 3 gives details of the feature vector obtained. Section 4 briefly defines the Machine Learning algorithms used in this study. The experimental results are presented in Sect. 5 and the final section covers the conclusion and related work.

## 2 Author Identification

Authorship identification is the task of predicting the most likely author of a text given a predefined set of candidate authors and a number of text samples per author of undisputed authorship [14]. From a machine learning point of view, this task can be seen as a single-label multi-class text categorization problem [15]. The basic steps of author identification includes:

1. Extract the characteristics of the textual documents of known authors,
2. extract the characteristics of the textual document to be classified,
3. compare 1 and 2,
4. classify the author based on the result of the comparison.

However, there are some possible limitations on author identification that a research may encounter. These are limited or excess number of documents to be analyzed, limited number of text data in documents, excess number of known authors to be characterized, and distribution of the training corpus over the authors (balanced or imbalanced) [16–18]. With limited number of authors the classification process will be relatively high, gaining 95 % and over success rate and this will lead the system to lose its objectivity especially if the candidate authors (classes) having distinguishable features. [19] has proven this theory by 40 % drop on the performance by increasing the number of authors from 2 to 20.

The author identification work in this study starts with collecting the documents of known authors from websites of Milliyet and Kıbrıs Gazetesi newspapers randomly. All the HTML tags are cleaned as a pre-processing step with the system developed in Java and plain text data are saved in MySQL. The database consists of 850 columns written by 17 columnists as a total, 50 columns from each columnist. The obtained datasets are; Dataset I: Milliyet, has 10 authors from [www.milliyet.com.tr](http://www.milliyet.com.tr) with 500 documents and have an average of 18,800 sentences and 305,000 words, Dataset II: Kıbrıs, has 7 authors from [www.kibrisgazetes.com](http://www.kibrisgazetes.com) with 350 documents and have an average of 265 sentences and 5,960 words, 50 columns per writer.

## 3 Feature Extraction

Feature extraction is one of the most important stages of author identification since it finds distinctive features that exhibit the writing style of each author. From the list of stylometric features such as, lexical, character, syntactic and semantic features, this study focuses on the lexical features. A significant advantage of such features is that they can be applied to any language and any corpus with no additional requirements [16].

After obtaining the corpus, 20 lexical features (style markers) were used to extract the feature sets which are presented in Table 1.

**Table 1** Style Markers

No	Style marker	No	Style marker
1	# of sentences	11	# of question mark
2	# of words	12	# of punctuation marks
3	Average # of words	13	Average # of dots
4	# of words after stopword removal	14	Average # of commas
5	Average # of words after stopword removal	15	Average # of semicolon
6	# of dots	16	Average # of colon
7	# of commas	17	Average # of exclamation
8	# of semicolon	18	Average # of question mark
9	# of colon	19	# of Non-Turkish words
10	# of exclamation	20	Average # of non-Turkish words

The system used to construct the feature vectors was developed in Java and Zemberek [20] library was used for stopword removal and Non-Turkish word identification which are the crucial style markers for Turkish Language. Zemberek is an open source, Natural Language Processing library for Turkish Languages.

## 4 Classifiers

There are variety of powerful machine learning algorithms applied in TC systems. As indicated by [16] in his deep survey about authorship attribution, these are Support Vector Machines, Decision Tree, Naive Bayes, Neural Networks, Random Forest, Genetic Algorithm, Maximum Entropy and K-Nearest Neighbors. In this study 4 of the most popularly used classifiers in Text Classification were implemented using WEKA tool.

### 4.1 Naive Bayes

This method is popularly used in TC due to its simplicity and computational efficiency. It computes the probability of instance  $d_j$  being in class  $c_i$  as shown in Eq. 1. Hence,  $P(c_i)$  is the probability of the occurrence of class  $c_i$ ,  $P(d_j/c_i)$  is the probability of generating instance  $d$  given class  $c_i$  and  $P(d_j)$  is the probability of instance  $d$  occurring.

$$P(c_i/d_j) = \frac{P(c_i)P(d_j/c_i)}{p(d_j)} \quad (1)$$

Based on this theorem, Naive Bayes classifier estimates the probabilities of classes (authors) given a document (features). The classifier chooses the highest probability among all authors as the author of the anonymous text.

## 4.2 Support Vector Machine (SVM)

SVM was introduced by Vapnik [21] as a machine learning algorithm, followed by many researchers. In SVM model text documents are represented as vectors and the model is built based on the extracted features from the training data set. The main idea of SVM is to divide the  $n$  dimensional space into two classes (positive and negative) and maximize the distance between the boundaries of each class. A classifier with large margin increases the certainty of the classification decision. New samples (test data) are then mapped into that space to predict the category it can fall into. Given training data  $x_i$ :

$$wTx_i + b \geq 1 \text{ for all } x_i \in P \text{ (for positive samples)} \quad (2)$$

$$wTx_i + b \leq -1 \text{ for all } x_i \in N \text{ (for negative samples)} \quad (3)$$

where  $w$  is the weight vector and  $b$  is a bias value (an average over marginal support vectors).

## 4.3 K-Nearest Neighbor (KNN)

The aim of KNN algorithm is to assign the instance (text to be classified) based on  $k$  nearest neighbors among the pre-defined classes (authors) using associated distance function. The performance of this algorithm greatly depends on an appropriate value of the parameter  $k$  where  $k$  is a user defined constant. When  $k=1$ , the algorithm simply assigns the instance to the class of its nearest neighbor which might work well if there are only few classes.  $k$  is usually set to be an odd number to make ties less likely [22]. When the value chosen for  $k$  is large the effect of noise on the classification is reduced, but this makes the boundaries between classes less distinct. In this study we used  $k=3$ . The pseudo-code of simple KNN algorithm is presented in Fig. 1.

## 4.4 Decision Tree

Decision Tree is another widely used classification algorithm in TC for inductive inference. It is a decision-modelling tool used to classify an instance by starting

```
Read the training data from the file <x,f(x)>
Read the testing data from the file <y,f(y)>
Assign value for parameter k
For i=1 to test size
    Get a new sample y to be classified
    Calculate the distance of y among its k neighbors in training set x
    Assign the corresponding class label based on max. similarity to sample y
End for.
```

Fig. 1 Pseudo-code of KNN

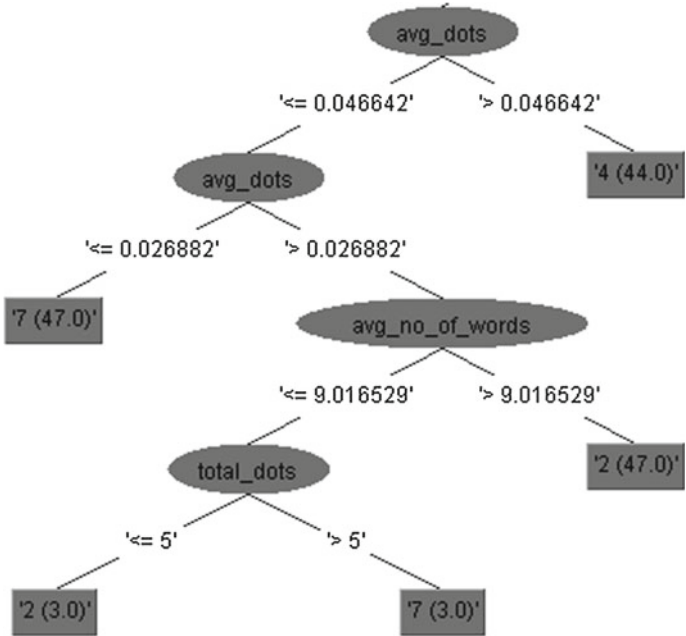
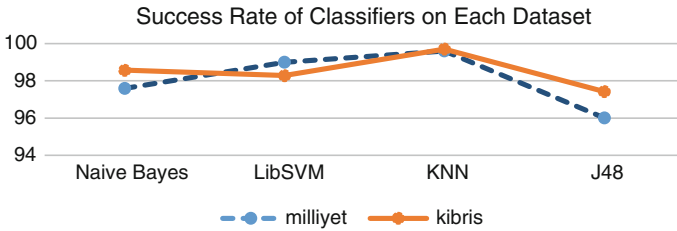


Fig. 2 A section from a decision tree generated by WEKA

at the root of the tree and moving through it until a leaf node, which provides the classification of the instance is reached. The classifier adopts a top-down approach, constructing the tree using if-then rules. Each path that starts from the root of a decision tree and ends at one of its leaf represents a rule. In this study the datasets were tested using the J48 decision tree algorithm which is a WEKA implementation of C4.5. Figure 2 illustrates a section from a decision tree generated by WEKA.

**Table 2** Correctly classified instances in %

Classification method	Milliyet (%)	Kıbris (%)
Naïve Bayes	97.6	98.28
SVM	98.8	98.57
KNN	<b>99.6</b>	<b>99.71</b>
Decision Tree	96.2	97.42



**Fig. 3** Comparison of 4 classifiers over 2 datasets

## 5 Experimental Results

For each feature set, all classifiers are trained and tested by applying 10-fold cross validation where it estimates the performance of a classifier by breaking the dataset into 10 partitions and then the model is trained on 9 datasets and tested on 1. The mean accuracy is obtained by repeating the process 10 times.

Table 2 shows the correctly classified instances in percentage for both datasets. Comparison of 4 classifiers over 2 datasets can be seen graphically in Fig. 2. It can be clearly observed in both representations that the success rate of all classifiers are relatively high however, KNN performs better among all 4 classifiers (Fig. 3).

## 6 Feature Selection

There could be many potential features available in the written text. Hereby, Feature Selection (FS) methods are used to utilize for more promising features and achieve better results. By identifying the dependent features, the FS method aims to eliminate the useless features. To maximize the success of the author identification system and reduce the size of the vector space we used Chi-Square function in WEKA. We reduced our feature set from 20 to 17 by eliminating the worst 3 features based on the ranking process of Chi-square (number of words, number of commas and number of words after stopword removal) which were common to both feature vectors. Improvements in the author identification rates after applying FS procedure can be observed in Table 3.

**Table 3** Correctly classified instances in % before and after feature selection

Classification method	Recognition rate (%)			
	Milliyet before FS	After FS	Kıbrıs before FS	After FS
Naïve Bayes	97.6	97.6	98.57	98.55
SVM	98.8	<b>99</b>	98.28	98.28
KNN	99.6	99.6	99.71	<b>100</b>
Decision Tree	96.2	96	97.42	97.42

## 7 Conclusion

The motivation of this study was to experiment an Author Identification task in Turkish with feature selection method using classifiers. The empirical evaluations carried on two datasets, Dataset I with 10 authors, Dataset II with 7 authors, 50 columns per author. 4 machine learning classifiers have been used; Naive Bayes, Support Vector Machines, K-Nearest Neighbors and Decision Tree. Performance results show that KNN performs the best score compared with the other 3 classifiers (Naive Bayes, SVM and Decision Tree) with 99.7 % accuracy and classification fully recognized by reducing the features from 20 to 17 with Chi-square feature selection method. To measure the robustness of the developed system it can be tested on more datasets as a part of future work. Since the features used for author identification task are mostly language independent, it is also possible to experiment with other languages as well.

## References

1. Mosteller, F., Wallace, D.: Inference and disputed authorship: the federalist. Adison Wesley (1964)
2. Mikros, G.K., Perifanos, K.: Authorship identification in large email collections: experiments using features that belong to different linguistic levels, CLEF (2011)
3. Hill, S., Provost, F.: The myth of the double-blind review?author identification using only citations. ACM SIGKDD Explor. Newsl. **5**(2), 179–184 (2003)
4. Zhao, J., Zhan, G., Feng, J.: Disputed authorship in C program code after detection of plagiarism. Int. Conf. Comput. Sci. Softw. Eng. **1**, 86–89 (2008)
5. de Vel, O., Anderson, A., Corney, M., Mohay, G.: Mining e-mail content for author identification forensics. SIGMOD Rec. **30**(4), 55–64 (2001)
6. Gray, A., Sallis, P., MacDonnel, S.: Software forensics: extending authorship analysis techniques to computer programs. In: Biannual Conference of the International Association of Forensic Linguists (IAFL'97), pp. 1–8 (1997)
7. Cheng, N., Chen, X., Chandramouli, R., Subbalakshmi, K.P.: Gender identification from E-mails. In: IEEE Symposium on Computational Intelligence and Data Mining, CIDM '09, pp. 154–158 (2009)
8. Bandara, U., Wijayarathna, G.: Source code author identification with unsupervised feature learning. Pattern Recogn. Lett. **34**(3), 330–334



9. Coulthard, M.: Author identification. *Idiolects Linguist. Uniquenes Appl. Linguist.* **25**(4), 431–447 (2004)
10. Pavelec, D., Justino, E., Oliveira, L.S.: Author Identification using Stylometric Features. *Inteligencia Artif. Rev. Iberoamericana de Inteligencia Artif.* **11**(36), 59–65 (2007)
11. Bozkurt, D., Baglioglu, O., Uyar, E.: Authorship attribution: performance of various features and classification methods computer and information sciences (2007)
12. Taş, T., Görür, A.: Author identification for Turkish texts. *J. Arts Sci.* **7**, 151–161 (2007)
13. Türkoğlu, F., Diri, B., Amasyalı, M.F.: Author attribution of Turkish texts by feature mining. In: *Proceedings of the 3rd International Conference on Intelligent Computing, ICIC 2007 Qingdao, China, LNCS 4681/2007* (2007)
14. Stamatatos, E., Fakotakis, N., Kokkinakis, G.: Automatic text categorization in terms of genre and author. *Comput. Linguist.* **26**(4), 471–495 (2000)
15. Sebastiani, F.: Machine learning in automated text categorization. *ACM Comput. Surv.* **34**(1), 1–47 (2002)
16. Stamatatos, E.: A survey of modern authorship attribution methods. *J. Am. Soc. Inf. Sci. Technol.* **60**(3), 538–556
17. Stamatatos, E.: Author identification using imbalanced and limited training text. In: *Proceedings of the 18th International Conference and Database and Expert Systems Applications, Regensburg*, pp. 237–41. IEEE Computer Society, Germany (2007)
18. Luyckx, K., Daelemans, W.: The effect of author set size and data size in authorship attribution. *Literary Linguist. Comput.* **26**(1), 35–55 (2011)
19. Grieve, J.: Quantitative authorship attribution: an evaluation of techniques. *Literary Linguist. Comput.* **22**, 251–270 (2007)
20. <http://code.google.com/p/zemberek/>
21. Vapnik, V.: *The nature of statistical learning theory*. Springer, New York (1995)
22. Manning, C.D., Raghavan, P., Schütze, H.: *Information retrieval*. Cambridge University Press (2008)

# Noun Phrase Chunking for Turkish Using a Dependency Parser

Mucahit Kutlu and Ilyas Cicekli

**Abstract** Noun phrase chunking is a sub-category of shallow parsing that can be used for many natural language processing tasks. In this paper, we propose a noun phrase chunker system for Turkish texts. We use a weighted constraint dependency parser to represent the relationship between sentence components and to determine noun phrases. The dependency parser uses a set of hand-crafted rules which can combine morphological and semantic information for constraints. The rules are suitable for handling complex noun phrase structures because of their flexibility. The developed dependency parser can be easily used for shallow parsing of all phrase types by changing the employed rule set. The lack of reliable human tagged datasets is a significant problem for natural language studies about Turkish. Therefore, we constructed a noun phrase dataset for Turkish. According to our evaluation results, our noun phrase chunker gives promising results on this dataset.

## 1 Introduction

Noun phrase chunking is a subset of shallow parsing (or text chunking). Shallow parsing consists of dividing sentences into non-overlapping phrases in such a way that syntactically related words are grouped in the same phrase. It can be considered as an intermediate step for full parsing. Each phrase has a name such as noun phrase (NP), verb phrase (VP), etc. Noun phrase chunking is the process of the determination of only noun phrases in a text.

There are many motivations for shallow parsing. By an intuition, when we read a sentence, we read it chunk by chunk [1] and some natural language tasks do not need full parsing [8]. For example, finding noun phrases and verb phrases may be

---

M. Kutlu

Department of Computer Science & Engineering, Ohio State University, OH, USA  
e-mail: kutlu@cse.ohio-state.edu

I. Cicekli (✉)

Department of Computer Engineering, Hacettepe University, Ankara, Turkey  
e-mail: ilyas@cs.hacettepe.edu.tr

enough for information retrieval technologies. For question-answering, information extraction, text mining and automatic summarization, phrases that give us information about time, places, objects, etc. are more significant than the complete analysis of a sentence.

Shallow parsing has been attracting the researchers for decades and one of the earlier works is Church's NP extractor [4] which uses a stochastic model. Ramshaw and Marcus [15] introduce NP chunking as a machine learning problem and they apply transformation based learning by using lexical information. Several groups worked with the same dataset and the same NP definition of Ramshaw and Marcus's study [15]. Argamon et al. [2] use memory based sequence learning in order to determine NPs and VPs without using any lexical information. Cardie and Pierce [3] learn POS tag sequences that form a complete NP to find NPs that are not found in training set. Veenstra [18] uses a cascaded chunking that uses lexical information. Daybelge and Cicekli [5] use memory based system and evaluates the system with a different dataset. Munoz et al. [11] use a network of linear units for recognizing NP and SV phrases.

Most of the studies are performed on English texts. However, there are also some studies for other languages, too. Sobha and Vijay [17] uses transformation based learning (TBL) for Tamil texts. Pattabi et al. [14] applies TBL for three Indian languages: Hindi, Bengali, and Telugu. Sastry et al. [16] uses dynamic programming algorithm for finding best possible chunk sequences for the same three Indian languages. To the best of our knowledge, there is no previous study about shallow parsing or noun phrase chunking of Turkish texts.

In this paper, we present a noun phrase chunker system for Turkish that consists of a dependency parser that uses hand-crafted constraint rules. The system takes a Turkish text as an input. The morphological analyses of words in the text are obtained by a Turkish morphological analyzer [6]. Correct morphological parses of words are found by our own morphological disambiguation tool for Turkish [10] which is a hybrid system that combines statistical information and hand-crafted grammatical rules and transformation based learning rules. After disambiguation of words, we use a dependency parser which uses hand-crafted rules in order to determine noun phrases. The dependency parser creates links between sentence components to represent relationships between them. After creating links, we obtain noun phrases by processing them. The dependency parser we propose can be easily converted to a shallow parser that includes all types of phrases because of its generic structure.

We use a dependency parser for noun phrase chunking. Dependency parsing is a technique that researchers are working on since 1960s for syntactic parsing and a detailed discussion about dependency grammars can be found in [12]. There are also studies for dependency parsing of Turkish sentences. Istek and Cicekli (2006) apply link grammar for parsing Turkish sentences. Oflazer [13] uses extended finite-state approach and he uses violable constraints in order to prevent robustness problem. Eryigit and Oflazer [7] worked on statistical dependency parsing techniques. Our study is in the category of constraint dependency grammars. We manually define rules that are used by the dependency parser. The rules act as finite state machine and when all constraints of a rule are satisfied, we construct a link for representing

relationship between corresponding sentence component. We connect links between sentence components called Token which are same with inflectional groups in [13].

The rest of the paper is organized as follows. Section 2 discusses general structures of noun phrases in Turkish. We give the details of our noun phrase chunking tool in Sect. 3. Our evaluation results are given in Sect. 4. Section 5 contains concluding remarks.

## 2 Noun Phrase Structure in Turkish

A single noun, pronoun or proper noun can be a noun phrase without any modifier in Turkish. A noun which is modified by other words can be a noun phrase (NP) too. The modified noun is the main noun and other words are modifiers of that main noun. In complex NPs, a sub-NP can also be a part of a modifier and the main noun can be modified by the modifier containing that sub-NP.

Modifiers can be words with different part of speeches such as adjective, noun, number, genitive pronoun. Modifying noun can be in nominative case or genitive case. Main noun has to take third personal possessive marker depending on the modifier type. If the modifier is a genitive pronoun, the possessive marker of the main noun has to agree with that genitive pronoun.

A modifier may not modify an actual derived word, and it can modify a part of that derived word. For example, in the noun phrase “yeşil başlıklı kız” (the girl with green cap), the word “başlıklı” (with cap) which an adjective derived from a noun modifies the noun “kız” (girl). On the other hand, the word “yeşil” (green) does not modify the word “başlıklı” and it modifies “başlık” (cap) of the word “başlıklı”. This fact demonstrates that we need the morphological structures of words in noun phrase chunking in Turkish. For this reason, our noun phrase chunker works on tokens (parts of words) not on words. We morphologically parse all words, and divide derived words from their derivation boundaries into tokens (inflectional groups). Thus, a derived word consists of more than one token and underived words consist of single tokens. For example, the word in “kitapçidaki” (the one in the book store) has following morphological parse.

kitap+Noun+A3sg+Pnon+Nom^DB+Noun+Agt+A3sg+Pnon+Loc^DB+Adj+Rel

This morphological parse is divided into three tokens from derivation boundaries.

Token <sub>1</sub>	kitap+Noun+A3sg+Pnon+Nom
Token <sub>2</sub>	^DB+Noun+Agt+A3sg+Pnon+Loc
Token <sub>3</sub>	^DB+Adj+Rel

The last token of a derived word (Token<sub>3</sub>) is known as *head-token* and holds last inflections, and its other tokens (Token<sub>1</sub> and Token<sub>2</sub>) are called as *non-head-tokens*. During the creation of links between tokens, only head-tokens of derived words can be modifiers in links and non-head-tokens cannot be modifiers. On the other hand, all tokens of derived words can be modified by a modifier.

### 3 Dependency Parser for Noun Phrase Chunking in Turkish

Our dependency parser connects links according to a rule set which consists of rules that determine the restrictions for constructing links between tokens. These rules also define scores and types of links to be connected and put extra constraints when needed. Putting restrictions only to tokens that will be connected is not enough for handling complex structures because they need more information in the analysis of a text, such as context or background information. Therefore, rules contain generic functions that can define constraints using morphological and semantic information. The usage of these type constraints in the rules eases the job of rule designers and it allows handling complex structures. As we construct links between tokens, we extract noun phrases according to the links.

We use links to represent relationship between tokens. In order to understand structures of links, specifications for constructing links are listed as below.

- A token can modify only one token however, a token can be modified by more than one tokens.
- A link can be constructed between only two tokens and there can be at most one link between two tokens. No crossing links are allowed
- The modifier token is called *Source* of the link and the modified token is called *Target*. The links have one direction which is from *Source* to *Target*. Since Turkish is head-final, links are normally from left to right.
- Only head tokens of words can modify a token. On the other hand, non-head tokens of a word can be modified by another token.
- Each link has a name representing the relation type between tokens. The non-head tokens are connected to the tokens at their right with a DB link.

#### 3.1 Rule Structure

We use a set of hand-crafted rules for connecting links between two tokens. A rule consisted of 5 parts which are *Source*, *Target*, *Constraints*, *Priority* and *LinkName*. A rule template is as follows.

CONSTRUCT a link between Token<sub>i</sub> and Token<sub>j</sub> with name of *LinkName* with a *Priority* IF Token<sub>i</sub> can be a *Source* and Token<sub>j</sub> can be a *Target* and all *Constraints* are satisfied.

By defining names to links, we can distinguish types of modifications. For example, if a modifier gives quantity information about a modified token, we use a certain link name (Dn) and if the modifier gives information about quality or color of the modified token, we use another link name (A). In our dependency parser, 30 link types are defined for different modifier types.

The priority of a constructed link is a positive number where a bigger number means a higher priority. The priorities of the rules are determined manually by the

rule designer. Priorities of rules can be learned from a big tagged dataset, when it is available. This is left as future work because of the lack of a big dataset.

*Source* and *Target* fields in a rule are specific properties describing which tokens can be the source and the target of a link respectively. In fact, these fields can be seen constraints for tokens that will be considered as a source token or a target token. A *Source* field (or a *Target* field) is a pair consisting of a *type* that specifies which parts of the token will be checked and a *value* that specifies the value of those parts. Types can be categorized into four groups: *TokenBased*, *WordBased*, *ContextBased* and *SpecialTypes*. *TokenBased* types specify which parts of a single token will be checked. On the other hand, *WordBased* types specify the parts of more than one token which belong to a single word. *ContextBased* types specify the parts of the words appearing immediately left or right of the considered token. *ContextBased* types can be used to catch collocations such that the part of speech of that collocation is different than part of speeches of its words. There are also some types that do not belong to first three types and are called as *SpecialTypes*. For example, a list of words can be considered for source/target fields, and that list is treated as a special type. *TokenBased* types specify the properties of the only considered token. A *TokenBased* type can be as follows:

*Part of Speech* The part of speech of a token are checked with the given value in the rule and its inflectional morphemes are ignored.

*Part of Speech and Partial Inflections* The part of speech and some of its inflections of the token are checked.

*Part of Speech and All Inflections* The part of speech and all inflections of the token are checked. Only stems can vary with this type.

*Full Token* An exact match with the token is checked.

A derived word can have more than one token and *WordBased* types can check a token of a derived word together with its other tokens that are on the left of that considered token. In the example in (1), if  $Token_2$  is considered as a Source/Target token by a rule with a *WordBased* type,  $Token_1$  can also be checked together with  $Token_2$ . Similarly, if  $Token_3$  is considered,  $Token_1$  and  $Token_2$  can also be checked together with  $Token_3$ .

kitapçıda (the one in the book store)

$$\underbrace{\text{kitap} + \text{Noun} + \text{A3sg} + \text{Pnon} + \text{Nom}}_{\text{Token}_1 (\text{kitap})} \wedge \underbrace{\text{DB} + \text{Noun} + \text{Agt} + \text{A3sg} + \text{Pnon} + \text{Loc}}_{\text{Token}_2 (\text{çı1-da})} \wedge \text{DB} + \underbrace{\text{Adj} + \text{Rel}}_{\text{Token}_3 (\text{ki})} \tag{1}$$

*WordBased* types are as follows:

*Surface Form* The surface form corresponding to the token is checked with the given value in the rule. The corresponding surface form of a token is the part of the surface form of the word up to that token. The corresponding surface forms for tokens in (1) are *kitap*, *kitapçıda* and *kitapçıda ki*, respectively.

*Starting with Surface Form* The corresponding surface form of the token has to start with the given value. For example, if the value in a rule with this type is *kitapçı*, this rule can accept Token<sub>2</sub> and Token<sub>3</sub> but it does not accept Token<sub>1</sub>.

*Starting with Surface Form and Partial Inflections* The corresponding surface form of the token has to match with the value and that token has to contain inflectional morphemes given in the rule.

*Lexical Form* The corresponding lexical form of a token is the part of the lexical form of the word up to that token. The corresponding lexical form of the token has to match the value in the rule.

*Starting with Lexical Form* The corresponding lexical form of the token has to start with the value in the rule.

*Stem with Partial Inflections* The stem of the corresponding word has to match with the value in the rule, and the token has to contain given partial inflections.

Although Source/Target fields of a rule give constraints for possible source and target tokens, extra constraints has to be satisfied before that rule is considered as an applicable rule. All constraints in *Constraints* field of a rule has to be satisfied before that rule can be treated as an applicable rule for possible source and target tokens. The constraints are connected to each other with AND boolean operator.

A constraint can check whether a token satisfies certain conditions. The position of the token in a constraint is given relative to source or target token. A simple constraint is a boolean function which checks whether that token satisfies certain properties. A constraint also checks a range of tokens whether satisfies certain conditions or not. A constraint can also contain boolean operators (OR, AND, NOT) to represent more complex cases.

### 3.2 *Dependency Parser Algorithm*

Our dependency parser finds noun phrases in a given sentence by creating links between tokens of that sentence. First, all words of the given sentence are morphologically parsed by our Turkish morphological analyzer [6] in order to get lexical forms of words in the sentence. After the morphological analysis phase, all words of the given sentence are morphologically disambiguated by our Turkish morphological disambiguator [10] in order to get single morphological parses for all words. The lexical form of a non-derived word is a single token and the lexical form of a derived word is divided from derivational boundaries in order to obtain its tokens. Thus, the given sentence is converted into a list of tokens where each token is the lexical form of a root word or a derivational morpheme with possible inflections.

In order to create links between tokens of a sentence, the tokens of the sentence are processed from the last token to the first token. If a new link is created in a pass on tokens of the sentence, another pass is performed. Iterations are repeated until no more new link is created in a pass. The creation of a link a previous pass may cause the creation of another link in a new pass. Before iterations start, all tokens of

derived words are linked with special links (called DB link) with highest priorities to indicate that they are parts of a single word.

Each iteration starts from the last token and it is moved to the left token when the process of a token is finished. If the current token is a head-token, it is treated as a possible source token of a link. If the current token is not a head-token, we move to the left token. When the current token is a head-token, an applicable rule which can accept the current token as a source token is selected from the rule list. Every token on the right of the current token is a possible target token for the selected rule. If a token is accepted as the target token by the rule and all constraints of the rule are satisfied, a link from the source token to the target token is created with the link name and the priority in the rule. All tokens and all applicable rules are checked to create links.

When a new link is being created, it is checked whether it crosses an available link. If it crosses, this new link is not created or that available link is deleted in order to create this new link. Here, the priorities of these two links play role in the outcome. If the priority of the available link is higher, it is kept. Otherwise that available link is deleted and the new link with a higher priority is created.

After all links of a sentence are created, its noun phrases are extracted by analyzing its link structure. Some links can be parts of noun phrases, but themselves cannot be noun phrases. All main noun phrases are extracted together with all sub-noun-phrases in those main noun phrases. A noun phrase can be a part of another noun phrase. If a noun phrase is not a part of another noun phrase, it is called as a main noun phrase. Otherwise, it is called as sub-noun-phrase.

A sample noun phrase extraction from a sentence is explained here by giving results of each step explicitly. Let's consider the sentence "*yeşil başlıklı kız büyük kapıyı açtı*" (the girl with green cap opened the big door). First, all words of the sentence are morphologically analyzed and disambiguated. The correct morphological parses of words that are morphological disambiguator are as follows:

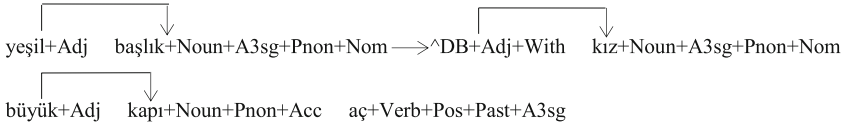
<i>yeşil</i>	yeşil+Adj	(green)
<i>başlıklı</i>	başlık+Noun+A3sg+Pnon+Nom <sup>^</sup> DB+Adj+With	(with cap)
<i>kız</i>	kız+Noun+A3sg+Pnon+Nom	(girl)
<i>büyük</i>	büyük+Adj	(big)
<i>kapıyı</i>	kapı+Noun+Pnon+Acc	(door)
<i>açtı</i>	aç+Verb+Pos+Past+A3sg	(opened)

Then all morphological parses of derived words are separated from derivational boundaries to get tokens of the sentence. Since "*başlıklı*" is the only derived word in the sentence, its morphological parse is separated into two tokens from its derivation boundary. Morphological parses of all other words in the sentence are treated as single tokens. The tokenized sentence is as follows.

yeşil+Adj başlık+Noun+A3sg+Pnon+Nom<sup>^</sup>DB+Adj+With kız+Noun+A3sg+Pnon+  
Nom büyük+Adj kapı+Noun+Pnon+Acc aç+Verb+Pos+Past+A3sg



Our link construction algorithm using rules are applied to the tokenized sentence to construct noun phrase links. The following links are obtained for the sentence.



Following two noun phrases are extracted from the sentence.

- *yeşil başlıklı kız* (girl with green cap). This NP also contains the sub-NP “*yeşil başlık*” (green cap).
- *büyük kapı* (big door).

## 4 Evaluation

In order to evaluate our Turkish noun phrase chunker, three datasets (D1, D2, D3) are created using a simple tagger tool, and noun phrases in those three datasets are tagged by Turkish native speakers. Taggers tagged noun phrases using surface forms of words. In the datasets D1 and D2, words are morphologically disambiguated by our disambiguator and words are approximately disambiguated with %97 correctness. On the other hand, morphological parses of the words in D3 are corrected by humans and it can be assumed that D3 is morphologically disambiguated with %100 correctness.

Statistical information about datasets is given in Table 1. Main NPs are directly components of sentences and they are not sub-NPs. The forth row gives the number of main NPs and the extraction of main NPs may be enough for some applications. The fifth row gives the total number of all NPs including sub-NPs.

Total 98 NP chunker rules are created analyzing Turkish grammar for noun phrases. When rules are prepared, structures of noun phrases in the dataset D1 are used by the rule developer. But the datasets D2 and D3 are not used in the development of the rules. All noun phrases including sub-NPs in all datasets are found. Some NPs can be found correctly, but their sub-NPs may not be found correctly. On

**Table 1** Statistical information about datasets: D1, D2 and D3

Features	D1	D2	D3
Number of words	5695	3854	3152
Average length of main NPs	2.04	1.59	2.05
Total number of main NPs	1743	1208	990
Total number of NPs including sub-NPs	2511	1398	1424

**Table 2** Raw and precision-recall-FMeasure results for datasets

DataSet	# of main NPs	# of total NPs	# of main NPs found	# of total NPs found	# of main NPs found correctly	# of total NPs found correctly
D1	1743	2511	1832	2533	1552	2173
D2	1208	1398	1243	1422	1038	1206
D3	990	1424	999	1403	934	1313
	TotalNPs			MainNPs		
DataSet	Precision	Recall	F-measure	Precision	Recall	F-measure
D1	0.858	0.865	0.862	0.847	0.890	0.868
D2	0.848	0.863	0.855	0.835	0.859	0.847
D3	0.936	0.922	0.929	0.935	0.943	0.939

the other hand, all sub-NPs of a main NP can be found correctly, but that main NP cannot be found correctly. For this reason, we obtained the results for main NPs and total NPs in datasets. The results are given in Table 2.

The first chart in Table 2 gives the raw results for the datasets and the second chart gives precision, recall and f-measure values for the datasets. The results of the dataset D1 is slightly higher than the results of D2 because D1 is used in the rule development. We made sure that chunker rules do not overfit noun phrases in the dataset D1. This can be seen from the results of D1.

There are many reasons that our chunker cannot find all correct NPs. One reason is that the complex sentence structures and the capability of our rules to cover all NP structures in Turkish. But, one of other reasons is the effect of morphological disambiguation. We used a morphological disambiguator that works approximately with %97 accuracy and we used it to tag words in the datasets D1 and D2. In order to measure the performance of our chunker with a perfect morphological disambiguation, we created the dataset D3 where all words corrected by humans. The results for D3 in Table 2 are much higher than other datasets. This indicates that the improvement in the morphological disambiguation also increase the performance of the NP chunker.

## 5 Conclusion

In this paper, we presented an NP Chunker for Turkish and it is implemented using a dependency parser which uses constraint based handcrafted rules for NP chunking. Our dependency parser has many distinctive features. It uses a scoring algorithm in constructing links for overcoming ambiguity problem. In addition, we have defined powerful constraining rules that allow rule designers to use semantic and morpho-

logical information together and handle very complex structures. By changing rules, full sentence parsing can be constructed, too.

Our dependency parser connects links between tokens rather than words. Therefore, morphological analysis and morphological disambiguation are crucial for our NP Chunker in order to determine correct tokens. We used our own morphological analyzer and morphological disambiguator for this process. As expected, the performance of the morphological disambiguation affects directly the performance of NP chunking.

We have also constructed three small datasets for testing our NP Chunker. These datasets consisted of 3941 main noun phrases at total. We implemented a tool for easing tagging process. This tool can be used for increasing size of dataset in future. Our NP Chunker gives promising results and the performance can increase by getting better chunking rules on Turkish language.

## References

1. Abney, S.: Parsing by chunks. In: *Principle-Based Parsing*, pp. 257–278 (1991)
2. Argamon, S., Dagan, I., Krymolowski, Y.: A memory-based approach to learning shallow natural language patterns. In: *Proceedings of ACL'98. Association for Computational Linguistics*, pp. 67–73 (1998)
3. Cardie, C., Pierce, D.: Error-driven pruning of Treebank grammars for base noun phrase identification. In: *Proceedings of COLING'98. Association for Computational Linguistics*, pp. 218–224 (1998)
4. Church, K.W.: A stochastic parts program and noun phrase parser for unrestricted text. In: *Proceedings of the Second Conference on Applied Natural Language Processing*, pp. 136–143. Austin, Texas (1988)
5. Daelemans, W., Bosch, A.V., Zavrel, J.: Forgetting exceptions is harmful in language learning. *Mach. Learn.* **34**, 11–14 (1999)
6. Daybelge, T., Cicekli, I.: A rule-based morphological disambiguator for Turkish. *Proceedings of Recent Advances in Natural Language Processing (RANLP 2007)*, pp. 145–149. Borovets, Bulgaria (2007)
7. Eryiğit, G., Oflazer, K.: Statistical dependency parsing of Turkish. In: *Proceedings of 11th Conference of the European Chapter of the Association for Computational Linguistics (EACL 2006)*, Trento, Italy (2006)
8. Hammerton, J., Osborne, M., Armstrong, S., Daelemans, W.: Introduction to special issue on machine learning approaches to shallow parsing. *J. Mach. Learn. Res.* **2**, 551–558 (2002)
9. Istek, O., Cicekli, I.: A link grammar for an Agglutinative language. In: *Proceedings of Recent Advances in Natural Language Processing (RANLP 2007)*, pp. 285–290. Borovets, Bulgaria (2007)
10. Kutlu, M., Cicekli, I.: A hybrid morphological disambiguation system for Turkish. In: *Proceedings of the 6th International Joint Conference on Natural Language Processing (IJCNLP 2013)*, Nagoya, Japan (2013)
11. Munoz, M., Punyakanok, V., Roth, D., Zimak, D.: A learning approach to shallow parsing. In: *Proceedings of EMNLP-WVLC'99. Association for Computational Linguistics* (1999)
12. Nivre, J.: *Dependency grammar and dependency parsing*. Vaxjö University: School of Mathematics and Systems Engineering: Technical Report MSI report 05133 (2005)
13. Oflazer, K.: Dependency parsing with an extended finite-state approach. *Comput. Linguist.* **29**, 515–544 (2003)

14. Patabhi, R.K., Vijay, S.R., Vijayakrishna, R., Sobha, L.: A text chunker and hybrid POS tagger for Indian languages. In: Proceedings of IJCAI-07 Workshop on “Shallow Parsing for South Asian Languages” (2007)
15. Ramshaw, L.A., Marcus, M.P.: Text chunking using transformation-based learning. In: Proceedings of the Third Workshop on Very Large Corpora, ACL (1995)
16. Sastry, G.R., Chaudhuri, S., Reddy, P.N.: An HMM based part-of-speech tagger and statistical chunker for 3 Indian languages. In: Proceedings of IJCAI-07 Workshop on “Shallow Parsing for South Asian Languages” (2007)
17. Sobha, L., Vijay, S.R. Noun phrase chunking in Tamil. In: Proceedings of the MSPIL-06, Bombay, pp. 194–198 (2006)
18. Veenstra, J.: Fast NP chunking using memory-based learning techniques. In: Proceedings of the Eighth Belgian-Dutch Conference on Machine Learning (1998)

# Enabling Secure and Collaborative Document Sharing in BIM Processes

Carlo Argiolas, Nicoletta Dessì, Maria Grazia Fugini and Barbara Pes

**Abstract** During the lifecycle of construction projects, documents and resources should be shared by partners in a collaborative and secure way. Starting from the Building Information Modelling (BIM) paradigm, we propose the *Smart BIM Folder (SBF)*, a cloud-oriented framework enabling project teams to share construction documents on collaborative nodes, respecting security requirements of documents sharing. The core idea is to consider the *SBF* as a single, collectively-operated virtual resource accessed via a cloud-based catalogue. The paper proposes mechanisms for document sharing respecting both the typicality of *BIM* projects and security requirements of stakeholders in such sharing. We discuss cloud architecture issues related to the implementation of the *SBF* as a cloud service.

## 1 Introduction

Architects, engineers and construction managers have always faced a difficult task in coordinating multiple phases during the design and the construction of Architecture/Engineering/Construction (AEC) products. Building Information Modelling (BIM) has helped in facing this difficulty by providing a paradigm enabling the creation of an accurate model (BIM model), which defines the digital representation of the physical and functional characteristics of an AEC product. Devised to cover the

---

C. Argiolas

Dipartimento di Ingegneria Civile, Ambientale e Architettura,  
Università di Cagliari, Cagliari, Italy  
e-mail: argiocar@unica.it

N. Dessì (✉) · B. Pes

Dipartimento di Matematica e Informatica, Università di Cagliari, Cagliari, Italy  
e-mail: dessi@unica.it

B. Pes

e-mail: pes@unica.it

M.G. Fugini

Dipartimento di Elettronica, Informazione e Bioingegneria,  
Politecnico di Milano, Milano, Italy  
e-mail: mariagrazia.fugini@polimi.it

early stages of construction design, the BIM model has proven useful throughout the entire life cycle of AEC products, including detailed design of construction components, project and cost management, construction development, facility operation, and so on [1].

With the BIM advent, the stages of an AEC project are now progressively being shifted from 2D-based construction drawings and documents to 3D-based BIM models. These represent viewpoints of the AEC products from the various perspectives of project participants, such as design teams, main contractors and subcontractors, asset owners, and so on. In practice, the BIM model integrates (graphically, and in text and table format) a 3D model with *multidisciplinary attributes and parameters* aimed at saving project time, money and resources. Operating via a shared BIM model, interdisciplinary team members can define or modify geometric, position, material, format, dimensional and other design attributes, sharing these attributes with the whole AEC work teams, rather than working in isolation on their own project portion. Accordingly, decisions can be taken with the consensus of various teams and project roles.

It has been observed [2] that *interdisciplinary collaboration* is an important component of BIM-driven processes, which should be considered as dynamic processes able to produce, communicate and analyze product models rather than a modelling technology “per se”. At the same time, critics have been expressed about BIM, seen as a standalone system framework that restricts access by the project stakeholders to a common set of data and resources [3].

Currently, cloud technology is seen as a suitable way to deal with the standalone nature of traditional BIM since the cloud can lead to higher levels of cooperation, providing an effective real-time communication platform for project members [4].

Moreover, recent research works [5] demonstrate that existing commercial BIM tools and platforms [6, 7] are still inadequate to support a fully multidisciplinary collaboration and only offer functionality to integrate drafting tools through a database structure storing information about the construction process (i.e. scheduling data, costs, maintenance information, manufacturers’ details etc.).

In practice, during the design of an AEC product, collaboration and decisional processes refer to the interconnection of different sources of information which can be a part of the BIM model or contained in other sources, such as a central repository containing project catalogues and documents accessible to all co-located or distributed team members. In this perspective, the central repository is a facilitator of BIM collaborative practices; however, designing its architecture brings about several open issues related to construction documents, namely:

(a) *dynamicity*: documents are frequently revised as the project proceeds and occasionally accessed when decisions need to be taken (e.g. the building orientation, building materials choice);

(b) *business assets are owned by various stakeholders* (project participants, external organizations, government agencies etc.). Hence, documents should be secured and their sharing kept under control. However, control is difficult in current practices, also due to security and privacy requirements, which hold in a different way across different organizations and are difficult to be harmonized [8].

Moreover, documents rely on different categories of data, often extracted from the BIM platform, such as engineering and project management data, documents about the construction lifecycle (e.g. scheduling, and work assignment), or financial data (costs and resources). It is expected that the integration of cloud and BIM technologies will result in the second generation of BIM [4] that will enable more effective collaboration platform for project team members.

Indeed, Cloud Computing technology offers an effective support to share documents, regardless of physical location. However, while opportunities for cloud are attracting the attention of BIM communities, there remains the need for critical examination of their implication on successful adoption within the BIM collaborative requirements and critical research questions still exist in terms of the use of cloud computing for secure collaborative exchange within BIM [9].

To face the above challenges, this paper proposes a framework named *Smart BIM Folder (SBF)*, a novel approach for flexible and secure document sharing among BIM partners. Our approach provides support for sharing documents, while it does not cover operation about BIM modelling processes. Document sharing is achieved via a cloud-based catalogue, which provides information about documents stored in distributed repositories. Security is given by defining access control rules, restricting access to resources to selected teams of project members. Moreover, sharing is decentralized, so that project participants can leverage existing documents from other project teams, but can also create their own associations and collaborations on the fly. Authorization is dynamically provided allowing users to set their own sharing policies, without affecting the policies of other users.

The *SBF* framework is inspired by the concept of dataspace [10], a paradigm for data integration which, to the best of our knowledge, has not been applied in the BIM domain. Additionally, we propose an implementation of *SBF* as a cloud application to allow for easy development, flexibility and security.

The paper is organized as follows. Section 2 presents a background about the management of BIM documents. Section 3 illustrates the *SBF* proposed framework and security modelling. Section 4 focuses on cloud architectural aspects of the proposed framework. Section 5 reviews related work and draws our conclusions.

## 2 BIM Documents

As depicted in Fig. 1, BIM technologies originate a new vision of participants involved in the construction process. According to the owner's requirements, Architects design the architectural model of a building, which is then completed by structural Mechanical, Electrical and Plumbing (MEP) engineers. The resulting building design defines all technical details about the building to be constructed. Critical information about a building are stored and/or extracted from the BIM by the facility manager or by contractors. Work scopes can be isolated by subcontractors, in order to evaluate opportunities about pre-assembling some parts off-site. Information about scheduling is extracted by suppliers for delivering products just in time and avoiding to stock on-site.

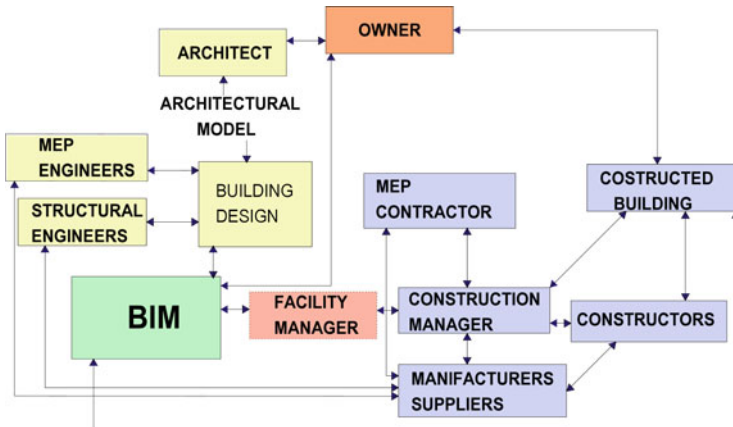


Fig. 1 BIM stakeholders

As the use of BIM grows, so does the amount of documents while the geographic distribution of construction partners favors the dispersion of information in multiple products through the building life cycle. The main feature of software environments supporting secure and dynamic document sharing would turn out to be effectiveness in empowering the activity of teams and professionals. Effectiveness brings about two requirements: (1) devising mechanisms for document distribution and sharing; (2) ensuring dynamic secure access to documents.

The first requirement calls for methods for effective exploitation of web resources. Albeit Web-based technologies have promoted several integration and collaboration technologies, BIM applications are a step behind other sectors (e.g. manufacturing) [5] due to the complexity of the domain.

The second requirement needs documents accessed in a secure way, considering user roles, and physical resources described by documents in a cybersecurity style [11]. Dynamic adaptive access controls should be in place to allow new users/teams to enter/leave/change their privileges on a need-to-know basis, keeping security compliant with the overall security policies stated for document sharing.

Above requirements are not fully satisfied by the current BIM technology which provides capability for data exchange and sharing within a virtual 3D environment which is a centralized and accessible repository for more integrate communication between project actors. Indeed, it has been stressed [9] that significant concerns over security and privacy within BIM processes affects project teams on adopting digital collaborative exchange technologies because of their greater risk as compared to traditional paper-centric communications.

As previous mentioned, the integration of BIM and cloud technologies is considered a critical step and a promising solution to mitigate risks about secure collaboration. It can be regarded as a viable alternative to the state of data exchange and storage of current BIM technologies.



Specifically, Cloud computing assembles large networks of virtualized services (including hardware and software resources) which can be abstracted into three layers [12]: Software as a Service (SaaS), Platform as a Service (PaaS), and Infrastructure as a Service (IaaS). Hardware and software resources are delivered at IaaS and PaaS while SaaS relies on application services.

Our proposal focuses on the PaaS layer which provides a platform (i.e. NoSQL databases, programming languages, libraries, APIs, environments and tools) enabling users to build applications which, in turn, can be hosted and provided as services at SaaS level. PaaS applications are independent of external resources and their internal organization, exploit a close integration with web servers and standard protocols and their rapid development and updates are enabled. Moreover, databases provided as cloud services improve data management in terms of elasticity and scalability.

### 3 The Smart BIM Folder (SBF)

With the aim of satisfying the above requirements, *SBF*:

- introduces a model for sharing documents within a scenario of multidisciplinary cooperating participants involved in a building construction process;
- adopts the Attribute-Based Access Control (ABAC) paradigm [8] for dynamic security rules;
- relies on cloud computing as a viable architecture for document sharing along the lifecycle of a building construction process.

Basically, the *SBF* acts as a librarian who retrieves documents required by a user, as long as he is allowed to read them for privacy reasons. It provides means for crossing the boundaries of local teams to allow documents produced by each team to be extended and reused within other teams, while ensuring that only the allowed user has the privileges to access the data. The idea is that team members should switch between different thematic documents, possibly linked to specific project activities, but they should avoid unnecessary data transfer on the web to find a document about a given topic. Under these aims, the *SBF* is a dynamic environment enabling the project participants to search and explore the set of construction documents in a user-friendly and in a security-aware way. The *SBF* has the main following key design features:

- *SBF* is based on dataspace [10, 13] for modelling web resources relevant to a particular organization and its business/activities, regardless of their format and location. A dataspace consists of components (also called participants) and of a set of relationships among them. The participants are individual data sources, e.g. relational databases, XML schemas, unstructured or partially structured information, and so on. Each participant contains information about the kind of data it contains, the data allocation, the storage format and the querying mechanisms.
- Within the *SBF*, data spaces are populated with classes of Objects that express abstract views on users and documents exposed by the organization users belong

to. Objects are featured by attributes storing data values useful to identify and searching objects. An attribute is a “name-value” pair (e.g., “Type = Report”). Security attributes are the secrecy level of a document, or the authentication method to access it. Contexts group documents/physical resources with homogeneous security requirements (e.g., documents/physical resources related to public areas of a building).

The *SBF* is composed of two dataspace: the Document dataspace (Dd) and the User dataspace (Ud). Within the Dd, documents are organized in layers; each contains all documents related to a particular defined phase of the construction process (e.g., design, construction, maintenance etc.). A Document has a unique identifier, is described by a set of attributes which express its properties (e.g. web address of the document, document type, data of creation, and so on). The dataspace are populated by structural and security relationships:

- (1) Structural relationships provide insights about documents belonging to the same layer (for example, two documents about the design of a building) or inter-relationships between documents belonging to different layers in order to generate new types of documents from the aggregation of existing ones.
- (2) Security relationships link a document to the physical resources it refers. Being one-to-many, security relationships can have different security requirements. For example, a Document D can be linked to office areas (Relationship1) and to the physical layout of pipelines (Relationship2): while Relationship1 is “confidentially”, Relationship2 can be “top secret”.<sup>1</sup>

Stakeholders and project members (users from now on), participate in the Ud. A user has a unique identifier and a set of attributes stored at configuration time or dynamically from an LDAP or a token received from the authentication services used at login time. Member-of relationships between users define project teams or hierarchical organizations in project management. We rely on User Roles to assign privileges. User Roles are stored into a subsidiary archive kept updated by the administration or security staff within BIM processes.

A catalogue stores data about participants in both dataspace and includes security rules. According to ABAC model, these rules, defined as usual as <Subject, Object, Privilege> triples, are based on attributes of dataspace participants and dynamically enrich the user roles. Here, for the sake of simplicity, we consider Subjects with security levels and Objects with sensitivity levels. To show an example about ABAC in the dynamic BIM scenario, let us consider the following security policy:

*Policy 1:* A user with Role “Structural Engineer” and Access Level less than “3” is allowed to Read an Engineering document (Object) related to a single physical resource r, e.g., “jail underground”, whose sensitivity level is higher than “2”.

Using the attributes of dataspace participants, privileges about the role “Structural Engineer” can be dynamically specified by defining the following <Rule> in

---

<sup>1</sup> We use terms of mandatory security although mandatory security is not actually adopted.

XACML style [14], which holds in the Context Jail Documents, so that security policies are respected since the rule holds in a confined domain [15]:

*Rule 1:* DefineRule(read, permit) in Context = “Jail Documents”  $\leftarrow$  (s.Role = “Structural Engineer”)  $\wedge$  (s.AccessLevel < “3”)  $\wedge$  (o.Group = “engineering-document”)  $\wedge$  (o.SensitivityLevel > “4”)  $\wedge$  (r.SensitivityLevel > “2”).

The *SBF* manages this security rule that defines the allowed actions so that controls can apply during the necessary time, while they can be subsequently revoked when the need to access the Object is concluded. Such temporary security rule can for example hold during a construction phase: when the *SBF* is notified that a given construction phase of the BIM is about to start, documents needed by the Construction Manager Role are granted to him. Accordingly an ordered list of access rules on all the needed documents (Objects) are activated by the *SBF*. Access control rules are so adapted according to the phase “initiation”. Upon termination of the phase, the security rule is revoked.

## 4 Architectural Aspects

Given the abstract nature of the dataspace paradigm, the architectural choices about the implementation of *SBF* functionalities rise a number of related issues.

First, relational databases are inadequate to support the dataspace paradigm: when rules are stored in a single relational table, their evaluation becomes potentially slow (due to the amount of self-joins). The less selective the rule, the more problematic these joins. The lack of just-in-time evaluations of rules results in a disconnection between the rule-based policies and the resources (i.e. the documents) that they protect.

Secondly, documents are physically distributed and accessible by a potentially large community of users. However, access to shared documents must be uniform, no matter where the document is. Because the number of accesses grows exponentially with the number of communications among teams, issues of data access latency and service partitioning must be considered.

To face the above issues, the cloud offers a viable architecture and new deployment models providing flexibility and efficiency. In particular, the PaaS layer of the Cloud provides a platform enabling users to build applications hosted and provided as services. PaaS applications are independent of external resources and exploit a close integration with web servers and standard protocols.

Based on these considerations, the *SBF* architecture is a software environment deployed at PaaS level and accessed by users through a web browser or a lightweight desktop or mobile app (Fig. 2). This environment is made up of two components: a database, which stores the dataspace catalogue, and a PaaS application supporting document search and access.

The database contains attributes about users and their security levels, documents and how they will be accessed, allocated and connected. Documents are exposed

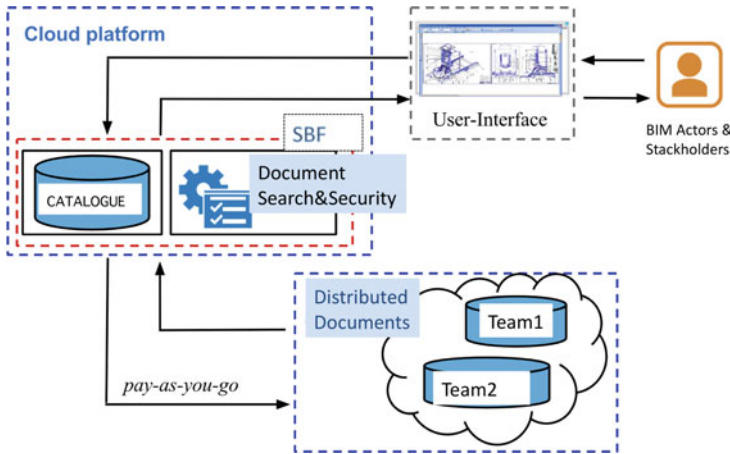


Fig. 2 The SBF as a PaaS application

in web repositories made available by single organizations and accessed via API. This allows project participants to expose only documents they want to share with other project teams. Relationships between documents and users are mapped into a set of items in the form of a subject-predicate-object expression where the terms “subject” and “object” denote two dataspace participants while the term “predicate” expresses the type of relationship between the subject and the object. Also a security rule conforms to such expression, where “predicate” corresponds to “privilege” (read, write, etc.).

The NoSQL nature of cloud databases gives great flexibility in structuring dataspace participants. These are logically modeled as objects having key-value properties and belonging to particular classes (e.g. the “users” and “documents” classes). Moreover, their relations can be directly mapped into a list where each term contains the identifier of two objects and a term or a code which details the type of the relationship in which they are involved. Differently from relational databases, data are not normalized and their duplication favors the easy evaluation of attributes.

The PaaS application is a procedural component for monitoring users in accessing documents they require, and deals with searching and extracting documents from distributed resources and evaluating rules for policies application. Documents are accessed according to a pay-as-you-go approach: searched documents are released after they are accessed by the users or stored into a buffer during a certain period of time (for example 24 h) and then released.

## 5 Related Work and Conclusions

A review of research about system integration and collaboration in architecture, engineering, construction and facility management (AEC/FM) industry is presented in [5] which details the state of the art within some recent ICT paradigms including Web

Services and Semantic Web. However, this review does not investigate cloud computing. Towards the integration of construction data in AEC, a cloud-based approach is presented in [16] that proposes a novel cloud application model for construction data and a series of BIM tools integrated with a special platform for project collaboration. A comprehensive review about BIM and mechanism for their constructions is given in [17]. Reference [9] presents an interesting study about the potential of cloud computing in enabling interoperable processes based on binding several construction applications through a single repository platform. Specifically, answers to questionnaires given by professionals confirm the opinion that cloud computing would increase efficiency and productivity. Moreover, using an integrated BIM process through a cloud service was considered as a key benefit as respect to the integration of component parts of the building modeled in disparate software programs. In our work [18] we have detailed the use of ABAC to dynamically authorize subjects to objects respecting the security policies.

This paper has presented a framework for secure document sharing of construction documents within BIM processes with security. To foster the vision of a single documental resource, we adopted the dataspace paradigm for modelling documents and users as participants indexed by a catalogue. We have outlined the benefits of the adoption of a cloud-based solution. Our work is preliminary to a specification of the proposed framework. This will be the focus of our future work.

**Acknowledgments** Research is supported by RAS (Regione Autonoma della Sardegna—Legge regionale 7 agosto 2007, n. 7), Project “*DENIS: Dataspace Enhancing the Next Internet in Sardinia*”.

## References

1. Cerovsek, T.: A review and outlook for a Building Information Model (BIM): a multi-standpoint framework for technological development. *Adv. Eng. Inform.* **25**(2), 224–244 (2011)
2. Eastman, C., Teicholz, P., Sacks, R., et al.: *BIM handbook: a guide to building information modeling for owners, managers, designers, engineers and contractors*, 2nd edn. Wiley, New York (2011)
3. Chuang, TH., Lee, BC., Wu, IC.: Applying cloud computing technology to BIM visualization and manipulation. In: *Proceedings of 28th International Symposium on Automation and Robotics in Construction*, pp. 144–149 (2011)
4. Wong, J., Wang, X., Li, H., et al.: A review of cloud-based BIM technology in the construction sector. *J. Inf. Technol. Constr.* **19**, 281–291 (2014)
5. Shen, W., Hao, Q., Mak, H., et al.: Systems integration and collaboration in architecture, engineering, construction, and facilities management: a review. *Adv. Eng. Inform.* **24**(2), 196–207 (2010)
6. <http://www.autodesk.com/products/revit-family/overview>. Accessed 03 June 2015
7. <http://www.bentley.com/en-US/Products/MicroStation/>. Accessed 03 June 2015
8. Sandhu, R.: Attribute-based access control models and beyond. In: *Proceedings of the 10th ACM Symposium on Information, Computer and Communications Security*, pp. 677–677 (2015)

9. Abdul-Majeed, M., Mahdjoubi, L., Booth, C.: Challenges to BIM-cloud integration: implication of security issues on secure collaboration. In: IEEE 5th International Conference on Cloud Computing Technology and Science, pp. 209–214 (2013)
10. Franklin, M.J., Halevy, A.Y., Maier, D.: From databases to dataspace: a new abstraction for information management. *SIGMOD Rec.* **34**(4), 27–33 (2005)
11. Valenzano, A.: Industrial cybersecurity: improving security through access control policy models. *IEEE Ind. Electron. Mag.* **8**(2), 6–17 (2014)
12. Hayes, B.: Cloud computing. *Commun. ACM* **51**(7), 9–11 (2008)
13. Atzori, M., Dessì, N.: Dataspace: where structure and schema meet. *Learn. Struct. Schemas Doc. Stud. Comput. Intell.* **375**, 97–119 (2011)
14. Rissanen, E.: eXtensible access control markup language (XACML) version 3.0 (2012)
15. Fugini, M.G., Teimourikia, M.: Access Control Privileges Management for Risk Areas. *ISCRAM-med* **2014**, 98–111 (2014)
16. Jiao, Y., Wang, Y., Zhang, S., et al.: A cloud approach to unified lifecycle data management in architecture, engineering, construction and facilities management: Integrating BIMs and SNS. *Adv. Eng. Inform.* **27**, 173–188 (2013)
17. Redmond, A., West, R.: The use of cloud enabled building information models—an expert analysis. *Australas. J. Constr. Econ. Build.* **12**(4), 53–67 (2012)
18. Dessì, N., Fugini, M., Garau, G., Pes, B.: Architectural and security aspects in innovative decisional supports. In: Proceedings of ITAIS13 Conference, Milano, Italy (2013)

# Tactical Graphics Description Language

İsmail Kiliñç, Hüseyin Ateş, Bülent Özhorasan and Hüseyin Korkmaz

**Abstract** Military tactical graphics are used to visualize and plan various command and control features on a tactical map. MIL-STD-2525 and APP-6 series standards make definition of tactical graphics. Since hundreds of complex tactical graphics are defined in standards, a sophisticated design approach is needed in order to achieve a successful software library implementation. A domain specific language named TGDL (Tactical Graphic Description Language) is proposed in this work and implemented on a desktop and web map software to define tactical graphics. This paper introduces TGDL and design approach to implement tactical graphics as a software library. TGDL effectiveness is also analyzed by using tools and metrics.

**Keywords** DSL · MIL-STD-2525 · APP-6 · GIS · Tactical graphics · TGDL

## 1 Introduction

MIL-STD-2525 series standards define common warfighting symbology for command and control systems and help establish interoperability between systems and units. The latest version of MIL-STD-2525 is “MIL-STD-2525D” [1] and was released on 2014. APP-6 [2] standards are subset of MIL-STD-2525 and used mainly in NATO. The work in this article is based on MIL-STD-2525C [3] which was released in 2008. The latest standard is different from previous by means of naming conventions of graphics. There are also some new added graphics and changes in

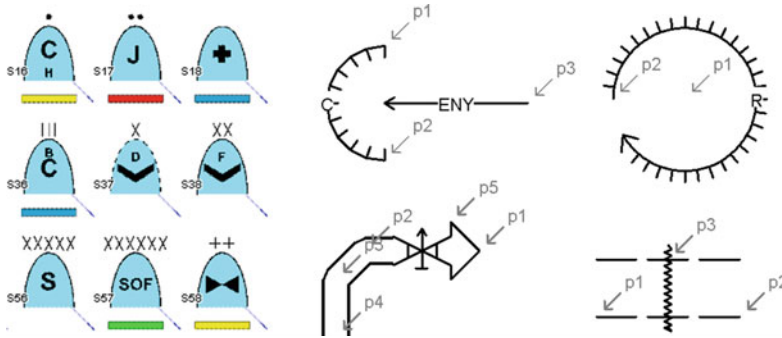
---

İ. Kiliñç (✉) · H. Ateş · B. Özhorasan · H. Korkmaz  
HAVELSAN A.Ş., Ankara, Turkey  
e-mail: ikilinc@havelsan.com.tr

H. Ateş  
e-mail: hates@havelsan.com.tr

B. Özhorasan  
e-mail: bozhorasan@havelsan.com.tr

H. Korkmaz  
e-mail: hkorkmaz@havelsan.com.tr



**Fig. 1** Some examples of tactical icons and tactical graphics

existing graphics also in new standard. Proposed approach to define tactical graphics for implemented version is generic and may be used to implement latest version or any similar standard

Symbology is categorized as tactical symbols and tactical graphics as shown in Fig. 1. Tactical symbols are icon based and have rules to show point features on a map in which the size and shape of a tactical icon remains fixed regardless of the map scale. Tactical graphics are point, line or area based and are differentiated from tactical symbols since line and area geometry presentation may change according to map scale and projection. Tactical icon definitions are simpler than tactical graphics. It is possible to define point graphics by standard vector definition languages such as SVG [4, 5]. However, tactical graphic definition is more complex than tactical icons since those graphics are geographic and has complex shapes.

Before TGD (Tactical Graphics Definition Language) arrived into the scene, C2 (Command and Control) applications implemented tactical graphics by using GIS libraries. As time passed, GIS software has supported tactical graphics directly and C2 applications obtained tactical graphics services from GIS libraries. Currently, well known GIS software products have built-in tactical graphics support.

ArcGIS is one of the GIS software that supports tactical graphics with its Military Overlay Editor (MOLE). It is possible to add custom graphics but it is limited to default renderers. If a custom graphics needs a new renderer, new renderer needs to be written with the support of ArcGIS library. Writing a renderer is complex and needs advance programming knowledge [6].

LuciadMap is another well-known GIS software in C2 domain, mostly used by NATO nations. It is possible to add custom graphics by adding new shape and renderers extended from LuciadMap GIS software library. Code needs to be written by utilizing the knowledge of computer graphics, Java graphics and OpenGL [7].

WorldWind is an open source GIS software library that was initiated by NASA. Later, library was expanded with additional tactical graphics through the efforts of WorldWind open source community. Currently, 75 % of tactical graphics set has been added to WorldWind [8]. Consequently, the need for an easy-to-use tactical graphics implementation/library has not been adequately fulfilled yet.



A domain-specific language (DSL) is a programming language or executable description language that offers, through appropriate notations and abstractions, expressive power focused on, and usually restricted to, a particular problem domain [9]. The tactical graphics can be considered as complex graphics as a composition of simple graphics such as point, label, multiline and polygons with complex relations. The main mission of DSL is to produce or manipulate simple graphics according to descriptors of tactical graphics, control points and feature attributes. TGDL is a script based sequential executable description language.

One motivation that triggered the development of a domain specific language was the need for two general purpose language implementations. Library implementation of MIL-STD-2525C tactical graphics was already developed in Java prior to TGDL development. But it is also required to implement same symbology standard in JavaScript after Java implementation has been completed.

Another motivation was the multiple symbology issue. Existence of different versions of the same standards, appearance of same graphics in different standards, and customer request to customize symbology according to their specific needs required a proposal of a dynamic symbol definition environment to adapt upcoming symbology requirements a necessary. Domain specific language approach solves multiple symbology issue since editing a language file is the only thing to do to support a new symbology.

Tools that implement tactical graphics, definition of TGDL and related examples are covered in the following section. Section 3 contains design and implementation of TGDL in JavaScript language and prior to conclusion, Sect. 4 contains the analysis of TGDL by using theoretical and empirical methods.

## 2 Tactical Map Graphics Description Language (TGDL)

A sample tactical graphic definition in TGDL is shown in Fig. 2 with its respective identification, control points and rules. The instruction set for sample's TGDL is given in Appendix A. The given graphic is defined in MIL-STD-2525-C with G\*TPB symbol identifier and defined in APP-6B as G\*T\*GB symbol identifier. Both symbol identifiers represent the same tactical graphic.

Default control points are defined in a rectangle in which coordinates for lower left corner and upper right corner are (0, 0) and (1, 1) respectively. In addition to its use in graphic overview creations, default control points may also be used on a user interface component and can also be used to initialize tactical graphic on a map. Geographic control points are calculated by extending overview shape of tactical graphic to viewing area using map scale and map center. However, application can also retrieve control points from a database or other data source to construct tactical graphics.

Rules part defines the rendering and editing behavior of tactical graphic. "CON P3 1 2" means Point 3 should be perpendicular to Point 1 and Point 2. "A P1" and "A P2" defines the points for P1 and P2 line. "N" is used to discriminate line and polygons

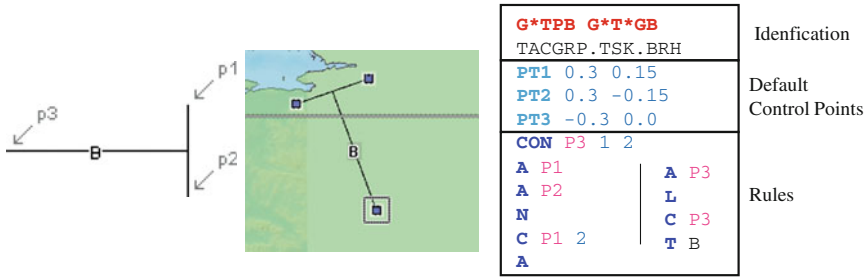


Fig. 2 A tactical graphic definition in TGDL

in multiline or multipolygon. “C P1 2” is used to accumulate center of P1 and P2. Accumulation is used to store calculated position in a variable to use for following rules. “A” adds accumulated point to point list. “A P3” adds the P3 to point list. “L” constructs multiline using point list previously constructed. “C P3” accumulates center of last accumulated point and P3 and “T B” puts text B to accumulated point, which is the center of the line perpendicular to P1 and P2.

First graphic in Fig. 3 is icon based and contains several textual information with different orientations and offsets. Second graphic is polygonal, but, have a special line pattern which is defined by a line starting with PZ. Last graphic is linear and has two labels for first and last point. Each label defined in definition is a parameter and related parameter value is specified either by user or application.

A challenging problem is the definition of line and polygon styles for symbolized boundaries. Some symbolized boundary definitions in TGDL and respective renderings are given in Fig. 4. Scale value is construed to define ratio of symbol size to graphic size first. Then points for symbol pattern are followed by referencing symbol starting point.

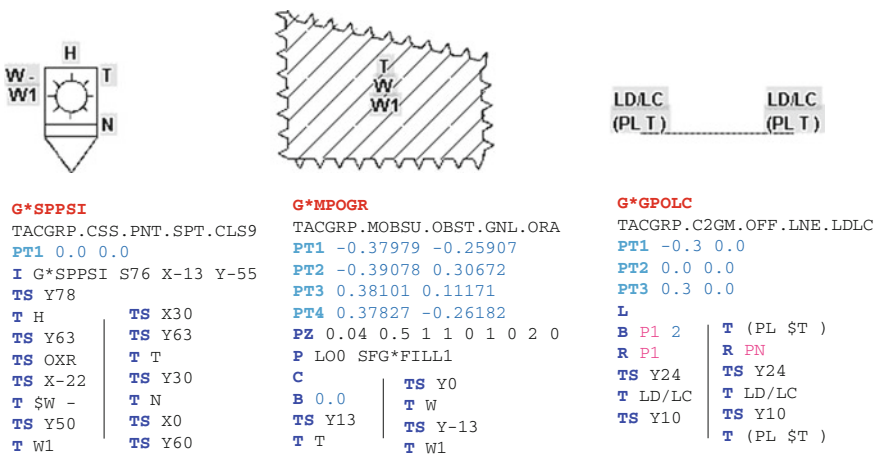


Fig. 3 Example tactical graphic definitions in TGDL

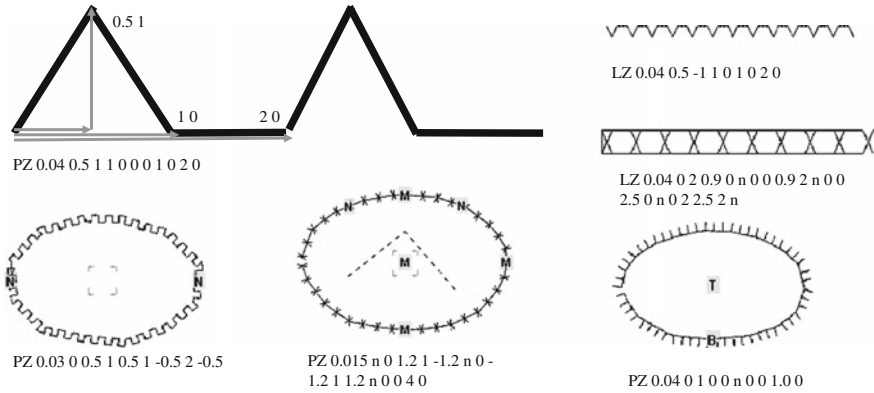


Fig. 4 Symbolized boundary definitions in TGDL

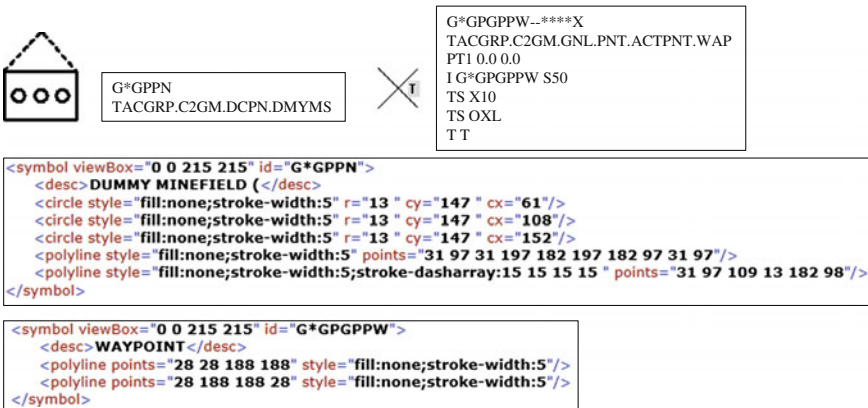


Fig. 5 Icon based tactical graphic definitions

Some tactical graphics are simple icons while others may reference icons but have textual or geometric components as shown in Fig. 5. It is assumed that if there is no other definition, then, identification part for a tactical graphic that referenced icon shares the same name with tactical graphic as in first graphic. In second graphic, “I G\*GPGPPW S50” line indicates that “G\*GPGPPW” icon should be positioned on center point and symbol size should be 50. The second graphic also contains a text definition which is oriented to left and has an offset of 10 pixels Icon definitions are not direct part of TGDL. It is assumed that reference icons are accessible via a service, an image file or an SVG definition.

### 3 Implementation in TMAP-W

In TMAP-W project, TGDL is implemented using pure JavaScript with a simple API shown in Fig. 6. *IGISTacticalGraphicManager* is a factory to create new tactical graphic instances and a manager to load TGDL files and manage loaded templates. An abstract tactical graphic and a listener interface are included in API to handle graphic events such as editing and selection. When a TGDL file is loaded by a manager, file is parsed and *IGISTacticalGraphicTemplate*'s are created for each graphic definition on TGDL implementation file. Each template decodes related section of TGDL.

In many interpreter implementations, the general approach used to execute a speedy script follows fetch-decode-execute model used in classical von Neumann computer architecture [7]. In our interpreter implementation, fetch-decode-execute model was used with some differences. Fetch phase is executed by *TacticalGraphicManager*; then, fetched lines are decoded into *TemplateCommand* by *TacticalGraphicTemplate* and execution phase is executed by *TemplateCommand* implementations. Because of performance issues, fetch and decode phases are executed once and stored in *TacticalGraphicTemplate* as an execution scenario. The class diagram for interpreter's component relationship is given in Fig. 6.

TGDL is a straightforward scripting language without any branching operation where each command is executed line-by-line on a single execution context instance. From this point of view, TGDL execution converges to chain of responsibility pattern where each line is an item in the chain presented by *TemplateCommand* implementations, after fetch and decode phases are completed [10]. Start and end commands automatically added to chain when script is decoded. A *TacticalGraphicExecutionContext* is added to manage state of executions.

TGDL interpreter renders tactical graphics according to TGDL where implementation class diagram is shown in Fig. 7. Created primitive graphics are updated instead

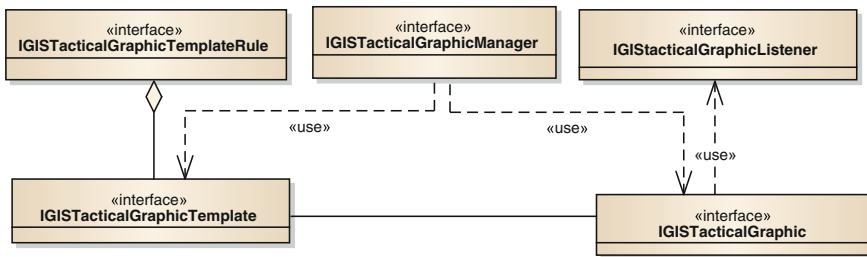


Fig. 6 Tactical API class diagram

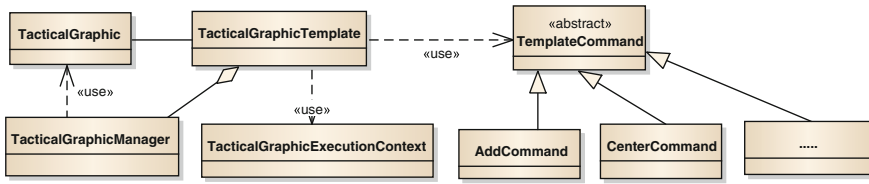


Fig. 7 Tactical implementation class diagram

of recreation to improve performance. For each instruction in TGDL, concrete implementations are done as a derivative of abstract command, *TemplateCommand*. TGDL interpreter implementation is easily extensible by registering any custom implementation of *TemplateCommand*.

### 4 DSL Analysis

For qualitative assessment of the developed DSL, Framework for Qualitative Assessment of Domain Specific Languages (FQAD) was used.

FQAD is used for creating the perspective of the evaluator, comprehending the purpose of the assessment and opting fundamental DSL quality characteristic to guide the evaluator in the evaluation. FQAD adapts and integrates the ISO/IEC 25010:2011 standard, CMMI maturity level evaluation approach [11, 12].

Assessment result of the evaluation is shown in Table 1. Results are obtained via consensus after evaluations of four different development team members. There are three success values: Incomplete, Satisfactory and Effective. The most successful value is Effective which means the DSL satisfies the intended purpose. Overall success rating for Tactical Graphics Description Language is Effective.

Prior to TGDL implementation, a legacy implementation without DSL was present. LOC (Line of Code) comparisons for the legacy implementation, NASA World Wind implementation and TGDL can be seen in Table 2. LOC specifies the size of component and is correlated with development time. World Wind partially implements tactical graphics (341 of 430) but also contains meteorology and ocean graphics (181).

Usage of TGDL in a general purpose language reduces the SLOC and development time approximately 5 times based on metrics for legacy Java implementation and TGDL Java implementation.

TGDL implementation has 8678 lines for 430 tactical graphics and average lines for a single graphic is about 20 lines. It means that a candidate new graphic may be implemented with 20 lines and implementation time may be less than an hour.

**Table 1** Assessment Results for TGDL according to FQAD

Functional suitability	Completeness	Effective
Usability	Comprehensibility	Effective
	Learnability	Effective
	Language helps users achieve their tasks in a minimum number of steps	Effective
	Likeability, user perception	Effective
	Operability	Effective
	Attractiveness	Satisfactory
	Compactness	Effective
Reliability	Model checking ability	Satisfactory
	Correctness	Effective
Maintainability	Modifiability	Satisfactory
	Low coupling	Satisfactory
Productivity	The development time improvement	Effective
	The amount of human resource used for improvement	Effective
Extendibility	Mechanisms for users to add new features	Effective
Compatibility	DSL is compatible to the domain	Effective
	Using DSL to develop models fits in the development process	Satisfactory
Expressiveness	A problem solving strategy can be mapped into a program easily	Satisfactory
	Uniqueness	Satisfactory
	Orthogonality	Effective
	The language constructs correspond to important domain concepts.	Satisfactory
	DSL does not contain conflicting elements	Satisfactory
	DSL is at the right abstraction level	Effective
Reusability	Reusability	Satisfactory
Integrability	Integrability	Satisfactory
Overall success		Effective

**Table 2** Line of Code Comparisons

	Legacy implementation (Java)	TGDL (JavaScript)	TGDL (Java)	Tactical graphics implementation in NASA World Wind (Java)
Physical LOC	13432	1629	2779	10990

Currently there is no metric to measure the time to define tactical graphics using TGDL however it is obvious that code development in a general purpose language is much harder than in a domain specific language.

TGDL is limited to define tactical graphics in military standards. It is not possible to define 3D objects or animations using TGDL. However if military standards evolve to include such graphics TGDL shall be evolved by adding new instructions to define such improvements. Each graphic may have only one color. Colors specify the hostility of the tactical graphic entity. If multiple colored graphics needed than new instructions should also be added to TGDL.

## 5 Conclusion

A solution for portrayal and editing of tactical graphics via domain specific language is proposed in this paper. First, current tools and approaches to solve the problem are summarized. Then, architecture, syntax and instruction set for TGDL are explained via examples. Paper is concluded with the presentation of implementation in web architecture and the analysis of TGDL.

Solution is complete since it covers all of the tactical graphics in MIL-STD-2525C. Some tools, such as NASA World Wind, also implement tactical graphics but do not cover all graphics types.

Approach presented in this paper is extendible and customizable. Hardcoded solution for the symbology problem requires code updates to support new symbology set or version update even for an invalid interpreted tactical graphic. However, it is only needed to edit TGDL file in order to define a new tactical graphic. Since a compiler is not required, only a successful interpreter like our TMAP-W implementation is adequate. Currently, the same language and definition file in two interpreter implementations, JavaScript based web TMAP-W and Java based desktop software TMAP-J, are used. If the declaration of an item in definition file is changed, then both software behaves exactly the same.

TGDL analysis carried out using FQAD shows that an effective language is proposed for tactical graphics. However, language efficiency may be improved by using some looping operators for the definitions of recurring patterns existing on some graphics. Abstractions may also be used to reduce the line of codes for TGDL.

Currently, MIL-STD-2525C and APP6-B standards are supported, but it is planned to extend implementations by including MIL-STD-2525D, APP6-C and national standard MS-76. It is obvious that TGDL will reduce the development and maintenance costs.

## Appendix A: Instruction set for TGDL

Item	Description
First line	Name of tactical graphic
Second Line	Description
PT1 0.0 0.0	First control point position
PT2 -0.25 0.0	Second control point position
D PI 2	Accumulate distance between PI and P2
B PI 2	Accumulate bearing between PI and P2
C PI 2	Accumulate center of PI and P2
T Text	Add text to accumulated point
ARP11.00	Add reckoned position from PI to distance*1.0 and angle 0
N	Add null point, ends a geometry and starts new one in same type
A	Add accumulated position
API	Add PI
L	Add line
SL	Add spline
LA	Add anticipated Line String
CRA	Add anticipated corridor
CR	Add corridor
CW 0.5	Corridor width is 0.5
P	Add polygon
RI 2.0 45.0	Accumulate intersection between last accumulated point and last added line
CON P3 1 2	P3 should be perpendicular to PI and P2
CON P3 1 2	P3 should be perpendicular to PI and P2
LSOI	Line start offset is 1
D PN 12	Distance of last point to PI and P2 line
R PI 1.0 0	Accumulates reckoned position from PI
CL Text	Add text to center of line
LT Text	Sets left text string for line
TS X10	Text style 10 pixel X offset
TS Y10	Text style 10 pixel Y offset
TS OXL/R/C	Text style X orientation to left/right/center
\$Parameter	Defines parameter
TR	Rotated text
LZ 0.04 0.5 1 10 2 0 20	Zigzag line string. 0.04 is zigzag scale according to polygon size. Next parameters are points based on pattern starting coordinates.
LZF	Zigzag filled line string.
PZ	Zigzag polygon.
SEC \$AM \$AM1 \$AN2 \$AN3	Sector, radius 1, radius 2, angle 1, angle 2
//	Comment
INFO	Not displayed on MAP, only for information
1 Reflcon S45	Use Reflcon with size 45 point



## References

1. MIL-STD-2525D, DOD, Joint Symbology Standard, June (2014)
2. APP6-B, NATO, Joint Symbology, June (2008)
3. MIL-STD-2525C, DOD, Common Warfighting Symbology, November 2008
4. <http://github.com/Esri/joint-military-symbology-xml/blob/master/svg/README.md>. Accessed 15 March 2015
5. W3C Recommendation, Scalable Vector Graphics 1.1, August 2011, <http://www.w3.org/TR/SVG11>. Accessed 15 March 2015
6. Matthew C., Clark S., Kyle K., Tom H., Mara, D.: Using Military Overlay Editor 9.1 for ArcGIS®, 2001
7. <http://www.luciad.com>. Accessed 15 March 2015
8. <http://goworldwind.org/developers-guide/symbology/tactical-graphic-status/>. March 2015
9. Van Deursen, A, Klint, P., Visser, J.: Domain-Specific Languages: An Annotated Bibliography. ACM Sigplan Notices (2000)
10. Gamma, E., Helm, R., Johnson R., Vlissides J.: Design Patterns: Elements of Reusable Object-Oriented Software, 1st edn. Addison Wesley, Boston (1994)
11. Kahraman G., Bilgen S.: A framework for qualitative assessment of domain specific languages. Springer Softw. Syst. Model. (SOSYM) J. (2013)
12. Kahraman, G.: A Framework for qualitative assessment of domain specific languages, Ph.D. Dissertation Thesis, Middle East Technical University, Electrical and Electronics Engineering Department, September (2013)

**Part IX**  
**Wireless Networks**

# EASER: Energy Aware Scalable and Reactive Replication Protocol for MANETs

Saeed Nourizadeh Azar, Kaan Karaagacli and Oznur Ozkasap

**Abstract** Mobile ad hoc networks (MANETs) depend on the nodes' collaboration to communicate and transfer data, and scaling the network size up greatly increases the energy needed to transfer data among far away nodes. To preserve nodes' energy and increase the network lifetime, data replication protocols have been proposed, which mainly increase data availability by creating nearby local copies of required data. In this work, first we provide a review of energy aware data replication protocols in MANETs. Then, by considering nodes' energy consumption, we propose EASER: Energy Aware Scalable and rEactive data Replication protocol. Our simulation results and comparison with SCALAR, energy aware ZRP and AODV protocols show that EASER provides improved network lifetime and data accessibility as the network size scales up with considering node energy levels.

**Keywords** MANET mobile ad hoc network · Data replication · Energy aware · Connected dominated set · Virtual backbone

## 1 Introduction

The main characteristic of MANETs is maintaining the network without a fixed infrastructure or controlling station. Nodes can move in any direction independently, and hence the frequent topology changes are not easy to predict [1]. Individual nodes together set up the network in MANET that makes possible to increase the range of communication among nodes, allowing to cover large geographical areas [2].

---

S.N. Azar (✉) · K. Karaagacli · O. Ozkasap  
Department of Computer Engineering, Koc University, Istanbul, Turkey  
e-mail: sazar13@ku.edu.tr

K. Karaagacli  
e-mail: kkaraagacli@ku.edu.tr

O. Ozkasap  
e-mail: oozkasap@ku.edu.tr

Two key challenges in MANETs are finding suitable paths for delivering data and increasing network lifetime. The focus toward preserving the energy of the nodes is increasing drastically and becomes a design requirement to reduce energy consumption in the network. Data replication plays a crucial role in MANETs to improve performance, reliability and availability. However, due to limited memory space of nodes, selecting appropriate data for replicating should be done carefully. Although methods such as SAF, DAFN, DCGT [3, 4] and SCALAR [5] have been proposed for increasing data availability, they do not consider the energy consumption aspect.

In this study, we first provide an overview on energy aware data replication protocols in MANETs based on the key characteristics identified. Addressing the findings and building upon our previous work SCALAR [5], we propose EASER: energy aware scalable and reactive data replication protocol. To transmit data in EASER, a virtual backbone method is used to maintain connection among nodes.

**Contributions** EASER is a distributed approach for data lookup and replication in large scale MANETs. By creating energy aware virtual backbone and taking nodes' energy levels into account, EASER improves network lifetime and reduces average energy consumption. For creating virtual backbone, unlike [6] which has network partitioning and node unreachability problems, we introduce new rules for ensuring that these problems are avoided in the constructed virtual backbone. To the best of our knowledge, using nodes neighbor information and considering energy to replicate data in the virtual backbone are novel aspects. Through modeling and extensive large scale simulations on JiST/SWANS, we analyze the effect of using energy aware virtual path and data lookup on data accessibility, network lifetime and network traffic. Performance results show that EASER protocol increases the network lifetime and outperforms SCALAR, energy aware ZRP (E-ZRP) and AODV (Ad hoc On demand Distance Vector) [7] protocols in terms of data availability and delay.

The rest of this paper is organized as follows. Section 2 provides our overview of energy aware replication protocols in MANETs. Section 3 presents details of the EASER protocol. Section 4 describes our extensive simulation results and performance evaluation of EASER, followed by conclusions in Sect. 5.

## 2 Related Work

Two fundamental characteristics of MANETs are mobility and multi-hop communications [8]. Due to mobility, nodes can easily go out of range of communication and lose access to the network. As a result, they cannot act as routers to deliver data in multi-hop manner. **Data replication** involves copying similar data in multiple nodes which is a fundamental technique used in distributed systems (Table 1).

**Table 1** Comparison of energy aware replication methods

	[11]	[12]	[13]	[14]	[16]	[17]	EASER
Remaining power	+	-	+	+	+	+	+
Read-only	+	-	+	-	-	+	+
Localized	+	-	-	+	+	+	+
Energy balancing	+	+	+	+	-	+	+
Access frequency	+	+	+	+	-	-	+

**Table 2** Simulation parameters

MAC protocol	802.11b
Mobility model	Random waypoint
Radio range	100 m
Node speed	1-3 m/s
Initial energy	5000J
Number of nodes	20, 50, 100, 200, 400
Area (m <sup>2</sup> )	316, 500, 707, 1000, 1414
Memory space per node	5 data items
Number of requests	10
Total simulation time	300 s

**Criteria for comparing different energy aware methods** There exist several techniques for data replication, but not all consider energy awareness when replicating data [9, 10]. In Table 2, we compare different data replication techniques which consider energy consumption, through the following criteria. **Remaining power** criterion represents whether a method uses power of nodes or not. **Read-only** parameter indicates whether the replication technique is performed for read-only data or not. **Localized** this criterion determines that each node according to information from constant hop count of neighbors decide to replicate data. **Energy balancing** criterion indicates whether a replication protocol can manage the energy usage among nodes or not. **Access frequency** refers to the fact whether a method considers access frequency when replicating data or not.

**Energy aware data replication protocols** The energy aware WEA-B protocol [4] takes nodes' energy levels into account when replicating data. WEA-B method uses access frequencies of nearby nodes and protects the nodes with low amount of energy to be accessed by distant nodes, so that it achieves both data availability and energy awareness [11].

A combination of pull-based and push-based data delivery methods were proposed in Expanding Ring Replication (ERR) [12]. In the push based method, the server frequently sends data and the client, by checking the channel, can access data, in contrast in the pull-based method the mobile node queries the server for required data. ERR assumes that the advertisement messages should be broadcast just for the

one hop neighbors in order to prevent a flooding all nodes and their responses. As a result, the energy consumed by each node as well as the network traffic load would be decreased.

EENMDRA is a data replication protocol designed for energy efficiency that uses method of [11] for replicating data and predicts node mobility to replicate data before node may go out of communication range. In EENMDRA, after determination of node mobility, nodes' energy consumption and replicating data, the algorithm aims to make balance among data accessibility, delay and energy consumption [13].

Power-Aware Dynamic Adaptive Replica Allocation Algorithm (PADARA) that makes use of the locality of data access was proposed in [14]. The replica allocation mechanism is balanced periodically to decrease the power consumption, and as a result, increase the network lifetime. According to the number of read and write requests, the total power consumption was calculated and after that by using a heuristic algorithm the suboptimal replica allocation scheme can be found.

ZRP protocol [18] does not consider node energy levels, and its energy aware extension E-ZRP [16] was proposed as a mechanism for decreasing unnecessary transmission range, and it aims at preserving nodes' energy and extending the network lifetime. However, it has the tradeoff between power saving and reducing transmission delay.

A method that frequently checks for the residual energy of the node which holds replica and calculates the node lifetime is provided in [17]. If the node lifetime falls below a predetermined value, replica will be redistributed to the node with the highest energy. Node lifetime is calculated based on the current and previous energy consumption of the node.

Using information about nodes' remaining energy, our approach EASER constructs a virtual backbone, and aims at balancing the energy among the nodes and increasing the network lifetime. We assume that replicated data is read-only and consistency of data in different replicas is not considered. Furthermore, node energy levels, data access frequencies and nodes' neighbors information are utilized when replicating data.

### **3 EASER: Energy Aware Scalable and Reactive Replication**

In this section, we provide details of the EASER protocol, namely CDS construction algorithm, distributed implementation, and reactive and energy aware replication mechanism.

#### ***3.1 CDS Construction Algorithm***

A subset of the graph is a dominating set (DS) if each node in the graph is either in the subset or adjacent to the one node which is in the subset. If all the nodes in the subset

(DS) are connected to each other, then the subset is called connected dominating set (CDS). It has the property that every node in the network has the distance at most one-hop from a backbone. Nodes in the backbone responsible for data routing are called dominator nodes [19].

Our approach to create an energy aware virtual backbone is inspired by [6, 20, 21]. We applied marking method in order to construct the CDS. Each node in a CDS graph,  $G = (V, E)$  by using  $m(p)$  which is a marker of node  $p \in V$ , is marked either as true(marked) or false(unmarked). Initially, all nodes are marked as false. We use approach in [6] for CDS construction, and we also add rules to ensure CDS connectivity in the case of mobile nodes. The algorithm has the following properties:

- Every node requires local single hop information about the network topology.  $N(p)$  is the set of the all neighbor of node  $p$ . The set which involves node  $p$  and its neighbors is represented by  $N[p]$  and defined as  $N[p] = N(p) \cup \{p\}$ .
- The resulting set of nodes forms a DS, and every node in the set is connected to one or more dominator nodes.
- All other non-dominator nodes are directly connected to one or more nodes in the resulting set.

For each node  $p$  there is a unique number,  $un(p)$ , and an initial energy level,  $eng\_lv(p)$ .  $V'$  is the set of graph vertices in  $V$  which are marked True. We suppose that the graph  $G_{sub}$  is the subgraph of  $G$  made by  $V'$ , i.e.,  $G_{sub} = G[V']$  [6]. The heuristic rules for reducing the size of DS generated through the marking process are:

**Rule 1** Assume two marked nodes  $p$  and  $q$  in  $G_{sub}$ . The marker of  $p$  is changed to false if one of the following conditions holds:

1.  $N[p] \subseteq N[q]$  in  $G$  and  $eng\_lv(p) < eng\_lv(q)$
2.  $N[p] \subseteq N[q]$  in  $G$  and  $un(p) < un(q)$  when  $eng\_lv(p) = eng\_lv(q)$

Rule 1 specifies that when all neighbors of  $p$  is covered by  $q$ , and remaining energy of  $p$  is smaller than  $q$ ,  $p$  can be deleted from  $G_{sub}$ . Also if both of them have equal energy levels, if  $un(p)$  is less than  $un(q)$  then  $p$  can be deleted from  $G_{sub}$ .

**Rule 2** Suppose that  $q$  and  $r$  are two nodes which are marked. Also these nodes are adjacent to marked node  $p$  in  $G_{sub}$ . If one of the following conditions are satisfied then the algorithm set the marker of  $p$  to the false.

1.  $N(p) \subseteq N(q) \cup N(r)$ , but  $N(q) \not\subseteq N(p) \cup N(r)$  and  $N(r) \not\subseteq N(q) \cup N(p)$  in  $G$
2.  $N(p) \subseteq N(q) \cup N(r)$ , and  $N(q) \subseteq N(p) \cup N(r)$  but  $N(r) \not\subseteq N(q) \cup N(p)$  in  $G$ ; and satisfying any of following a or b conditions:
  - a.  $eng\_lv(p) < eng\_lv(q)$ , or
  - b.  $eng\_lv(p) = eng\_lv(q)$  and  $un(p) < un(q)$
3.  $N(p) \subseteq N(q) \cup N(r)$ ,  $N(q) \subseteq N(p) \cup N(r)$  and  $N(r) \subseteq N(q) \cup N(p)$  in  $G$ ; and satisfying any of following a, b or c conditions:

- a.  $eng\_lv(p) < eng\_lv(q)$ , and  $eng\_lv(p) < eng\_lv(r)$
- b.  $eng\_lv(p) = eng\_lv(q) < eng\_lv(r)$  and  $un(p) < un(q)$
- c.  $eng\_lv(p) = eng\_lv(q) = eng\_lv(r)$  and  $un(p) = \min\{un(p), un(q), un(r)\}$

Rule 2 shows that if closed neighbors of the  $p$  is supported by the union of  $q$  and  $r$  in rule 2.1, if  $q$  and  $r$  are not covered by the union of two other node, then node  $p$  will be deleted from  $G_{sub}$ . Rule 2.2 indicates that, if  $r$  is not supported by union of  $q$  and  $p$ , but  $p$  is supported by the union of  $q$  and  $r$ , and  $q$  is supported by the union of  $p$  and  $r$ , and also if the energy level of  $p$  is less than that of  $q$  or the  $un(p)$  is less than  $un(q)$  when their energy levels are equal then node  $p$  can be deleted from  $G_{sub}$ . Rule 2.3 specifies that, when all the neighbors of  $q$ ,  $p$  and  $r$  is supported by the union of the two other, and any of the following condition are satisfied in that case  $p$  can be deleted from  $G_{sub}$ . The conditions are: a. if energy level of  $p$  is the smallest among  $q$  and  $r$ ; or b. if the energy level of  $p$  and  $q$  are equal but less than of  $r$ , and the  $un(p)$  is less than  $un(q)$ ; or c. the energy levels of  $q$ ,  $p$ , and  $r$  are equal and unique number of  $p$  is the smallest among  $q$ ,  $p$ , and  $r$  [6, 20].

### 3.2 Distributed Implementation

CDS construction algorithm is executed periodically to maintain the backbone among nodes considering the random mobility of nodes. However, we notice that the CDS construction method of [6, 20] cannot always create a connected DS and dominator nodes may become unconnected to each other. As a result, nodes in the network may become unreachable through the dominating set and the network becomes partitioned that reduces data accessibility drastically. We observed that the problems encountered in [6] arise from the fact that nodes trust each other (for not going unmarked), and go to an unmarked state simultaneously. That is, by concurrent execution of the algorithm at each node, in some cases node  $p$  checks its dominator neighbors  $q$  and  $r$ , and finds out that their open neighbors cover all neighbors of  $p$ . So  $p$  trusts  $q$  and  $r$ , and changes its status to false (pruning phase). At the same time,  $q$  checks the neighbors of  $p$  and  $r$ , and finds that they cover its neighbors, therefore  $q$  also decides to change its status to false. This can also happen for node  $r$ . The update in status of  $q$  and  $r$  to false causes the node or its neighbors to become unreachable. Therefore, to solve this problem we modified CDS construction algorithm as follows:

1. When node  $p$  runs the algorithm, if it returns true for two randomly picked dominator neighbors ( $q$  and  $r$ ), node  $p$  may be pruned.
2. Before pruning, check if the nodes trusted ( $q$  and  $r$ ) will be pruned. Pick another dominator node  $d_1$  and run the algorithm for  $(p, q, d_1)$  and  $(p, r, d_1)$ . In addition, pick another node  $d_2$  and run the algorithm for  $(q, d_1, d_2)$  and  $(p, d_1, d_2)$ . These checks contain all the results about  $q$  or  $r$  that can be calculated from node  $p$ . If any of these checks returns true it means that  $q$  or  $r$  will be pruned.



### 3.3 Reactive and Energy Aware Data Replication Mechanism

We use a reactive replication method, that is a replication decision will be done when a data is received by a node. In the previous work SCALAR [5], the cost function uses request frequency history and distance (number of hops) information to the data owner. In order to replicate data considering the node energy information, we use an enhanced cost function which not only considers access frequency and energy level, but also considers access frequency of the neighbors. We assume that replicated data is read-only and consistency of data in different replicas is not considered. Our energy aware replica allocation method in the virtual backbone is inspired by [4]. The assumptions about nodes and data items are the following:

- Host nodes ( $H_1, H_2, H_3, \dots, H_n$ ) move based on random waypoint mobility.
- Every data item ( $d_1, d_2, d_3, \dots, d_n$ ) stores in a certain node and is read-only.

The following describes the cost function used when host  $H_a$  requests access to data item  $d_{fresh}$  which does not belong to itself.

(1) If  $H_i$  has enough memory space it replicates  $d_{fresh}$ . Otherwise,  $H_i$  sends data request to nodes within  $h(\geq 1)$  hops. The request involves host id of  $H_i$  and all data entry stored by  $H_i$  and  $d_{fresh}$ .

(2) When a node,  $H_c$ , get a request, it sends a response to  $H_a$ . The response message contains the id of  $H_a$  and  $H_c$ , access frequencies of  $H_c$  to data, and flags which shows  $H_c$  stored the data specified in the message.

(3) If  $H_a$  receives response message, it calculates  $\Delta_{a,k \rightarrow fresh}$ , for each data item held by  $H_a$  with

$$\Delta_{a,k \rightarrow fresh} = \delta(\alpha_{a,fresh} - \alpha_{a,k}) + \left(\frac{E_i}{E_{init}}\right)\lambda\left(\frac{A'_{a,fresh}}{S_{a,fresh+1}} - \frac{A'_{a,k}}{S_{a,k}}\right) \quad (1)$$

Here, data in node memory is represented by  $k$ . The total access to  $d_k$  from node  $H_a$  is  $\alpha_{a,k}$ . The value of  $E_i$  and  $E_{init}$  are current and initial amounts of battery power of  $H_a$ . Total access frequencies to  $d_k$  from one hop neighbors of node  $H_a$  is shown by  $A_{a,k}$ . The value for  $S_{a,k}$  is the whole  $d_k$  in one hop neighbors. By using  $\delta$  and  $\lambda$  as weights it is possible to give priority to access frequencies from the node or its neighbors. The value of  $\Delta_{a,k \rightarrow fresh}$  represents the changes in the total data accesses when replacing the  $d_k$  with  $d_{fresh}$  [11].

(4) Data item  $d_k$  with highest value of  $\Delta_{a,k \rightarrow new}$  is chosen by  $H_a$  to replace with  $d_{fresh}$ . By decreasing the node power the value in second term of Eq. 1 will be decreased. Therefore, in this case mobile node give priority to its access frequency.

## 4 Simulations and Performance Results

We implemented EASER protocol on JiST/SWANS Scalable Wireless Ad hoc Network Simulator [22]. Our simulation parameters are presented in Table 2. We analyze the performance of EASER and compare it with SCALAR, E-ZRP and AODV protocols.

ZRP protocol [15] aims to make use of the advantages of two routing methods, namely proactive and reactive routing. However, ZRP does not take the energy level of nodes into consideration and packets are forwarded with full power without considering the nodes position inside the zone. In E-ZRP [16], a method for decreasing unnecessary transmission range was proposed to preserve nodes' energy. According to Friis free-space equation,  $P_r = \frac{P_t}{(4\pi d)^2}$ , where  $P_r$  is the received power and will be significantly reduced by increasing distance  $d$  between nodes. E-ZRP considers that when node power reaches 50% of its initial energy, it halves the transmission power  $P_t$ . By decreasing nodes energy, not only nodes reduce their transmission power to keep batteries capacity but also they do not disturb communication among nodes.

### 4.1 Performance Metrics

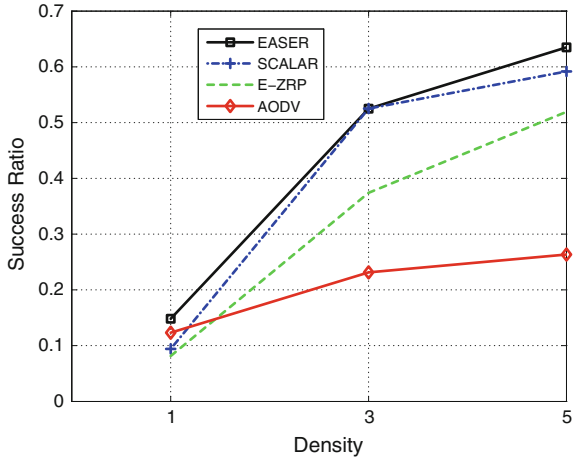
The following metrics are used to analyze system performance. **Success ratio** is the ratio of the number of successful access requests to the number of all access requests. **Average remaining energy** metric shows the ratio of total remaining energy of all nodes to the number of nodes. **Network lifetime** refers to the time at which the first network node runs out of energy.

### 4.2 Analysis Results

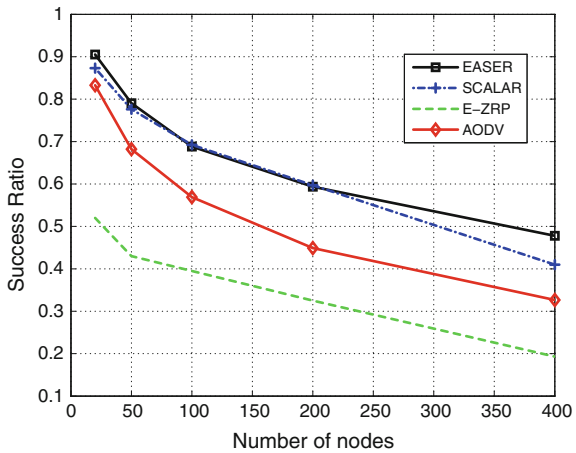
**Network density** is defined as the expected number of a node's neighbors in the network [2]. To compare different methods, we consider different values for density i.e. 1, 3, 5. Figure 1 shows the success ratio of EASER, AODV, SCALAR and E-ZRP for different network densities, where EASER outperforms the other methods due to its efficient replication mechanism. Figure 2 shows that with a fixed density the success ratio of EASER has the largest value, especially when the network size scales up.

**Average energy and network lifetime** Nodes in a virtual path collaborate in several send and receive operations and their power will be depleted more quickly.

**Fig. 1** Effect of network density



**Fig. 2** Data accessibility



If the energy level of a node is smaller than the other nodes' energy level, it will be replaced with a new one. As shown in Fig. 3, EASER achieves the highest average energy compared to other protocols. Furthermore, it also helps improving the network lifetime. As given in Fig. 4, there is no occurrence of node death, up to 100 nodes during the simulation time, but after that by increasing the number of nodes the network lifetime for AODV plunges drastically. In EASER, the network lifetime is improved significantly and no death of nodes occurs during the simulation time.

**Network traffic** Nodes spend energy during send and receive processes. By decreasing network traffic, it would be possible to save nodes' energy. The network

Fig. 3 Average energy

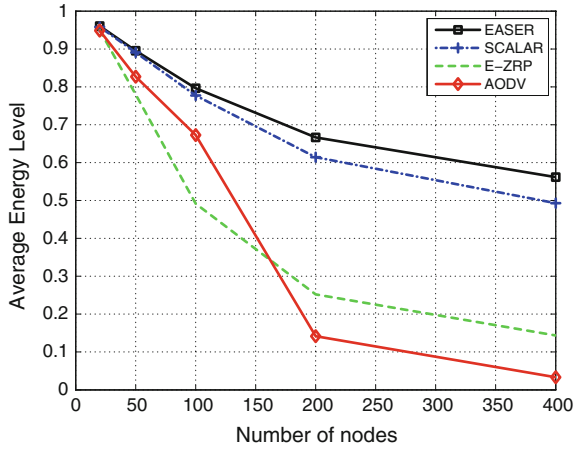
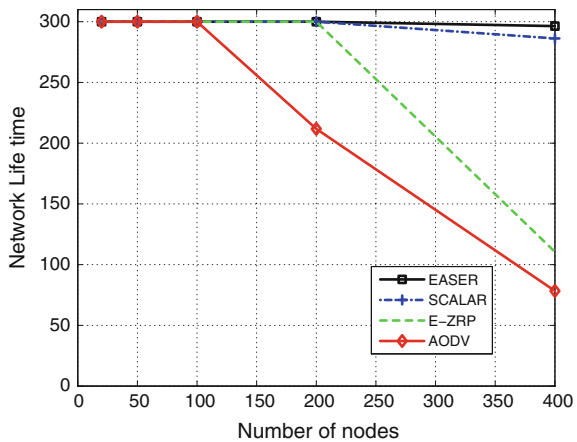


Fig. 4 Network lifetime

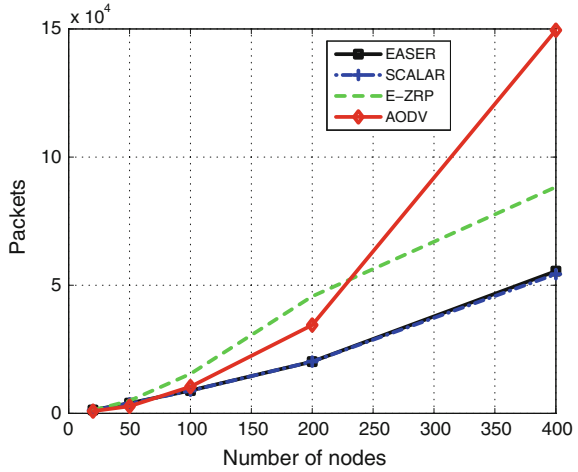


traffic is also an important criterion for the power aware method. According to the number of sent and received packets as shown in Figs. 5 and 6, the highest number of sent and received packets are observed in AODV, and the smallest number of packet transmissions are observed in EASER.

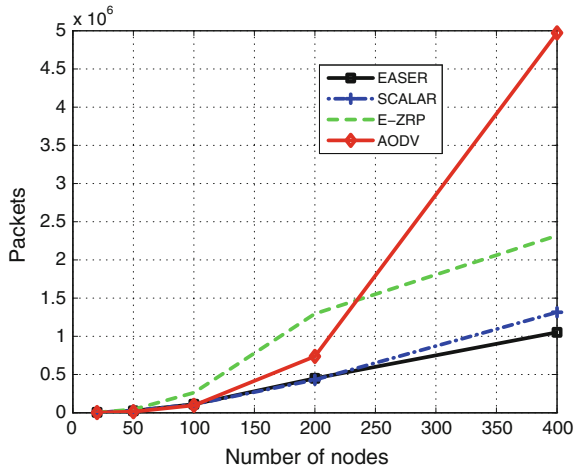
### 5 Conclusions

In this study, we provide a review of energy aware data replication protocols in MANETs. By addressing the findings in existing mechanisms and considering nodes energy levels when constructing virtual backbone, we propose EASER: energy aware scalable and reactive data replication protocol. The aim is to reduce the number of

**Fig. 5** Number of sent packets



**Fig. 6** Number of received packets



nodes participating in lookup and replication in order to preserve network energy. Our simulation results and comparison with SCALAR, E-ZRP and AODV protocols show that EASER provides improved network lifetime and data accessibility in several network scenarios investigated.

**Acknowledgments** This work was partially supported by the COST (European Cooperation in Science and Technology) framework, under Action IC0804, and by TUBITAK (The Scientific and Technical Research Council of Turkey) under Grant 109M761

## References

1. Kaur, N., Kumar, N.: Scarp: scalable congestion aware replication protocol. In: IEEE International Conference on Advanced Communication Control and Computing Technologies (ICAC-CCT), pp. 1554–1558 (2014)
2. Karl, H., Willig, A.: Protocols and architectures for wireless sensor networks. Wiley, New York (2007)
3. Hara, T., Madria, S.K.: Data replication for improving data accessibility in Ad Hoc networks. *IEEE Trans. Mob. Comput.* **5**(11), 1515–1532 (2006)
4. Hara, H.: Energy efficient data access in mobile ad hoc networks. In: Mobile Data Management (MDM), IEEE Eleventh International Conference, pp. 326–330 (2010)
5. Atsan, E., Ozkasap, O.: Scalar: scalable data lookup and replication protocol for mobile ad hoc networks. *Comput. Netw.* **57**(17), 3654–3672 (2013)
6. Zhao, Y., Jie, W., Li, F., Lu, S.: On maximizing the lifetime of wireless sensor networks using virtual backbone scheduling. *IEEE Trans. Parallel Distrib. Syst.* **23**(8), 1528–1535 (2012)
7. Das, S., Perkins, C., Royer, E.: In: Ad hoc on Demand Distance Vector (AODV) Routing, Mobile Ad-hoc Network (MANET) Working Group, IETF, vol. 81 (2002)
8. Rao, A., Kumar, P., Chauhan, N.: Energy efficient dynamic group caching in mobile ad hoc networks for improving data accessibility. In: International Conference on Recent Trends In Information Technology (ICRTIT), pp. 372–376 (2012)
9. Derhab, A., Badache, N.: Data replication protocols for mobile ad-hoc networks: a survey and taxonomy. *IEEE Commun. Surv. Tutorials* **11**(2), 33–51 (2009)
10. Padmanabhan, P., Gruenwald, L., Vallur, A., Atiquzzaman, M.: A survey of data replication techniques for mobile ad hoc network databases. *VLDB J. Int. J. Very Large Data Bases* **17**(5), 1143–1164 (2008)
11. Shinohara, M., Hara, T., Nishio, S.: Data replication considering power consumption in ad hoc networks. In: IEEE International Conference on Mobile Data Management, pp. 118–125 (2007)
12. Thanedar, V., Almeroth, K.C., Belding-Royer, E.M.: A lightweight content replication scheme for mobile ad hoc environments, pp. 125–136. Springer (2004)
13. Mukilan, P., Wahi, D.A.: EENMDRA: efficient energy and node mobility based data replication algorithm for manet. *Int. J. Comput. Sci. Issues (IJCSI)* **9**(3) (2012)
14. Wang, Y., Yang, K.: Research of power-aware dynamic adaptive replica allocation algorithm in mobile ad hoc networks. In: Parallel and Distributed Processing and Applications, pp. 933–944. Springer (2005)
15. Haas, Z.J., Pearlman, M.R., Samar, P.: The zone routing protocol (ZRP) for ad hoc networks (2002). <https://tools.ietf.org/html/draft-ietf-manet-zone-zrp-04>
16. Kim, K.-J., Koo, H.-W.: Optimizing power-aware routing using zone routing protocol in manet. In: IEEE International Conference on Network and Parallel Computing Workshops. IFIP, pp. 670–675 (2007)
17. Pushpalatha, M., Venkataraman, R., Ramarao, T.: Energy aware data sharing in mobile ad hoc networks. In: Recent Trends in Networks and Communications, pp. 378–387. Springer (2010)
18. Perkins, C.E.: Ad hoc networking. Addison-Wesley Professional, Boston (2008)
19. Yuanyuan, Z., Jia, X., Yanxiang, H.: Energy efficient distributed connected dominating sets construction in wireless sensor networks. In: Proceedings of the International Conference on Wireless Communications and Mobile Computing. ACM (2006)
20. Lu, M., Wu, J., Cardei, M., Li, M.: Energy-efficient connected coverage of discrete targets in wireless sensor networks. In: Networking and Mobile Computing, pp. 43–52. Springer (2009)
21. Wu, J., Gao, M., Stojmenovic, I.: On calculating power-aware connected dominating sets for efficient routing in ad hoc wireless networks. In: Parallel Processing, pp. 346–354 (2001)
22. Barr, R.: Swans-scalable wireless ad hoc network simulator (2004). <http://jist.ece.cornell.edu/docs.html>

# Fractional Frequency Reuse Based Adaptive Power Control Scheme for Interference Mitigation in LTE-Advanced Cellular Network with Device-to-Device Communication

Sok Chhorn, Si-o Seo, Ji-eun Song, Suk-ho Yoon, Seung-yeon Kim and Choong-ho Cho

**Abstract** In this paper, we propose a frequency planning and transmission (Tx) power management scheme, called *Fractional Frequency Reuse based Adaptive Power Control* (FFR.APC) Scheme for *orthogonal frequency division multiple access* (OFDMA) and *time division duplex* (TDD) based on LTE-Advanced *device-to-device* (D2D) network. In the proposed scheme, the *macro base station* (mBS) and *D2D senders* (D2DSs) service *macro user equipments* (mUEs) and *D2D receivers* (D2DRs) use different frequency bands and optimal Tx power chosen as D2D links' locations in inner and outer zones in different periods to reduce interference substantially. Simulations show the proposed scheme outperforms D2D networks with *soft frequency reuse* (SFR) and *Fractional Frequency Reuse* (FFR) schemes in terms of the system throughput and outage probability for mUEs and D2DRs.

**Keywords** LTE-advanced · Device-to-device communication · Interference avoidance · Fractional frequency reuse · Resource allocation · Power control

---

S. Chhorn (✉) · S. Seo · J. Song · S. Yoon ·  
S. Kim · C. Cho

Department of Computer and Information Science, Korea University, Seoul, Korea  
e-mail: chhorn168@korea.ac.kr

S. Seo  
e-mail: sioseo@korea.ac.kr

J. Song  
e-mail: rhdbwls123@korea.ac.kr

S. Yoon  
e-mail: bluepig5@korea.ac.kr

S. Kim  
e-mail: kimsy8011@korea.ac.kr

C. Cho  
e-mail: chcho@korea.ac.kr

## 1 Introduction

Due to the fast development of mobile communication technologies, more and more users tend to download content on mobile device. Recently, the amounts of traffic being treated by cellular networks have increased as mobile multimedia services have become popular. In particular, the *macro base station* (mBS) handles more traffic than in the past years because of the fast-growing needs of high data rate services. However, the radio resources in cellular networks are limited and the installation cost of mBS is high. In [1] it has been proposed to handle the local peer-to-peer traffic in a reliable, scalable, and cost-efficient manner by enabling direct *device-to-device* (D2D) communication as an underlay to the IMT Advanced cellular network.

D2D communication is the technology enabling *macro user equipments* (mUEs) to communicate directly with each other without the help of mBS. Before the resource allocation and data transmission (Tx) phase, two devices need to find each other, i.e., the peer discovery phase and D2D connection setup. The peer discovery phase is relatively independent of the D2D communication phase. Existing work can be classified into centralized and distributed approaches [2]. D2D communication can offload the traffic handled by mBS and reduce end-to-end Tx delay since end users are able to directly exchange data without intervention of mBS. However, D2D links may generate high interference to mUEs located in their communication areas if they use the same spectrum with the mUEs for data Tx [3–5].

To solve this problem, research on reducing interference between D2DRs and mUEs in cellular networks supporting D2D communication was conducted. When earlier studies are examined, in [3], they suggested the technique of using D2D links' channels first if frequency bands were not already being utilized in the mBS, and observing the D2D links' channel status if all frequency bands are used and getting assigned the best channel from mBS. Several schemes of inter-cell interference mitigation are being considered in *orthogonal frequency division multiple access* (OFDMA) networks, such as *fractional frequency reuse* (FFR) [6] and *soft frequency reuse* (SFR) [7]. Partial reuse adopts different reuse factors for the cell center and cell edge. Thus, partial reuse schemes can achieve a much higher network capacity compared to traditional frequency reuse schemes and can simultaneously reduce inter-cell interference compared to a *frequency reuse factor* (FRF) of 1. However, since the cell edges use a higher reuse factor, the cell edge spectral efficiency may be significantly degraded compared to the cell center.

In [8], a resource allocation scheme was proposed for D2D networks with a FFR system. The proposed scheme in [8] uses four resource groups to reduce interference of mUEs and D2D links. But the proposed scheme uses only FFR based frequency planning of mUEs and D2D links in inner and outer zones respectively. In addition, the strength of *D2D Tx power* ( $P_{D2DS}$ ) and performance for mUEs and D2D links in inner and outer zones is not analyzed. To solve this problem we propose a *Fractional Frequency Reuse based Adaptive Power Control* (FFR.APC) Scheme for OFDMA and *time division duplex* (TDD) in the LTE-Advanced D2D network. In this scheme,



we control and optimize the  $P_{D2DS}$  based on the user requirement to mitigate the interference.

Thus, the proposed scheme increases the overall system throughput by limiting the cross-tier interference at mBS below a predefined threshold level, i.e., maximum cross-tier interference that mBS can tolerate, by ensuring the optimization of D2DRs power allocation at each power adjustment phase. The simulation results are provided to prove the advantages of FFR.APC scheme compared with SFR and FFR schemes in terms of the system throughput and outage probability for mUEs and D2DRs.

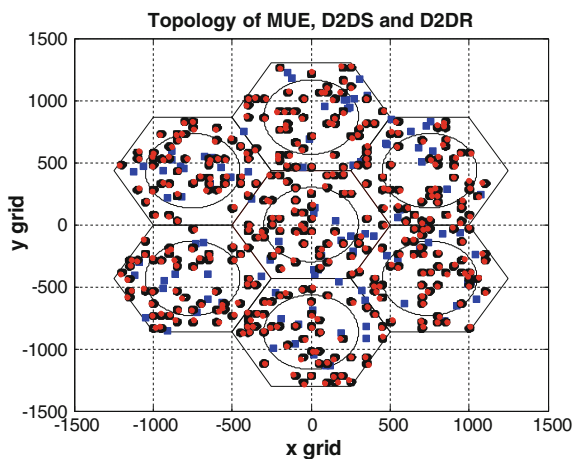
The paper is organized as follows: In Sect. 2, we describe the system model. In Sect. 3, the proposed scheme aiming to efficiently use the bandwidth and control the power of D2DRs and mUEs is elaborated and described. Performance evaluation is discussed in Sect. 4 and finally, Sect. 5 concludes the paper.

## 2 System Model

### 2.1 System Topology

We consider a typical two-tier D2D communication network architecture where D2Ds are overlapped on the macrocell to analyze the system performance. Figure 1 describes the system topology between the macrocell and D2D. We assume that there are  $X$  hexagonal macrocells, here  $X=7$ , and a set of  $X$  mBSs,  $X = |\mathbf{X}|$ ,  $\mathbf{X} = \{1, 2, \dots, X\}$ , is installed at the center of each macrocell as shown in Fig. 1, while a set of  $Y$  D2DSs,  $Y = |\mathbf{Y}|$ ,  $\mathbf{Y} = \{1, 2, \dots, Y\}$  and a set of  $F$  D2DRs,  $F = |\mathbf{F}|$ ,  $\mathbf{F} = \{1, 2, \dots, F\}$ . The inter-site distance and radius of the inner zone are, respectively,  $D_{inter}$  and  $D_{inner}$  in meter (m). We assume that each mBS is located in the center

Fig. 1 System topology



of each macrocell and has a cell *identification* (ID,  $1 \sim 7$ ). For example, mBS with cell ID= $i$  is described as  $mBS_i$ . mUEs and D2DSs are randomly deployed in the macrocell coverage and are stationary. In each cell, D2DRs are separated from their corresponding D2DSs with a distance  $q$ , where  $q$  is a uniform random variable in [1 20] m. The target cell is the center macrocell  $mBS_i$ , and interfering neighbor mBSs to mUEs and D2DSs in each cell site of  $mBS_i$ .

## 2.2 Signal Power Model

The signal power received,  $P_r$ , at mUE and D2DR from mBS and D2DS can be expressed as

$$P_r = P_t \times L \times 10^{-(PL/10)}, \quad (1)$$

where  $P_t$  is the Tx power of mBS and D2DS,  $PL$  is path loss,  $L$  is the shadowing effect, with a log-normal distribution with a zero-mean and a standard deviation of  $\sigma$ .

We consider a path loss model for mUE and D2DR, where  $PL_{mUE_{i,m}}$  is the link between the  $mBS_i$  and the  $m$ -th mUE,  $mUE_{i,m}$ , in the coverage of  $mBS_i$  and  $PL_{D2DR_{i,j,h}}$  is the link between the  $j$ -th D2DS and the  $h$ -th D2DR,  $D2D_{i,j,h}$ , in the  $j$ -th D2DS coverage of  $mBS_i$ , as shown in (2) and (3).

The path-loss is modeled according to the micro-urban models ITU-R report [9]. We apply different path-loss models to D2DRs and mUEs as given in Eqs. (2) and (3) [10]. The path-losses of the micro-urban models for D2DRs ( $PL_{D2DR_{i,j,h}}$ ) and mUEs ( $PL_{mUE_{i,m}}$ ) are expressed as

$$PL_{D2DR_{i,j,h}} = 40 \log_{10} d[km] + 30 \log_{10} f_c[MHz] + 49, \quad (2)$$

$$PL_{mUE_{i,m}} = 36.7 \log_{10} d[m] + 40.9 + 26 \log_{10}(f_c[GHz]/5), \quad (3)$$

where  $d$  represents distance between a sender and a receiver, and  $f_c$  means the carrier frequency of the system.

## 3 The Proposed FFR.APC Scheme

In the FFR scheme, different resources are allocated to mUEs and D2DRs according to their locations. The mBS services mUEs in the inner zone using RBs in  $RB_{inner}$  whereas mUEs in the outer zone use RBs in  $RB_{outer}$ . D2DSs in the inner zone service their D2DRs using RBs in  $RB_{outer}$ , except the same sub-channel group for mUEs. D2DSs in the outer zone service their D2DRs using RBs in  $RB_{inner}$  and  $RB_{outer}$  [8]. In the SFR Scheme, D2DSs allocate random RBs in the *downlink* (DL) subframe ( $RB_{inner}$  and  $RB_{outer}$ ) for D2DRs in conventional D2DS networks, as shown in Fig. 2.

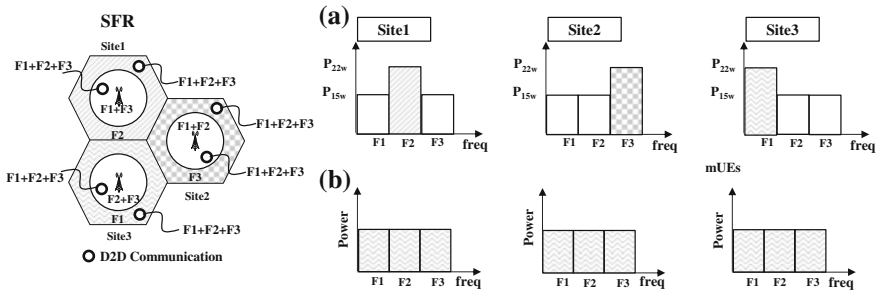


Fig. 2 Conventional SFR scheme for mUEs and D2DRs

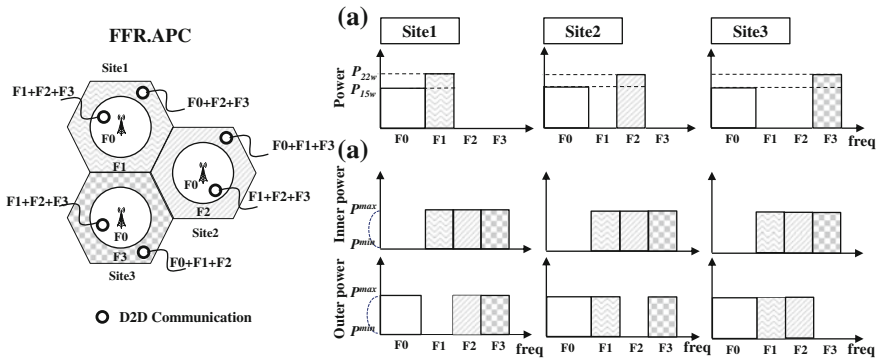


Fig. 3 Proposed Resource Allocation Scheme

Thus, D2DSs cause substantial interference for mUEs in the outer zone while D2DRs in the inner zone are affected by substantial interference from the mBS [7].

In the proposed FFR.APC scheme, mBSs and D2DSs service mUEs and D2DRs using different RBs in  $RB_{outer}$  to reduce strong interference of the outer zone, as shown in Fig. 3. The mBS services mUEs in the inner zone using RBs in  $RB_{inner}$  whereas mUEs in the outer zone use RBs in  $RB_{outer}$ . However, D2DSs in the inner zone service their D2DRs using RBs in  $RB_{outer}$ , except the same sub-channel group for mUEs in the same site, because D2DRs in the inner zone are affected by substantial interference from the mBS.

Also, D2DSs in the outer zone service their D2DRs using RBs in  $RB_{inner}$  and  $RB_{outer}$ , because D2DRs are affected by small interference from the mBS. In addition mBSs and D2DSs's powers are controlled. We increase the powers of mBSs for mUEs in the outer zone. In the FFR.APC scheme, the D2DS uses  $Z$  RBs in the  $a$ -th sub-channel for the D2DR using optimal Tx power according to SINR threshold and gives weak interference to the mUEs and neighbor D2DSs

We explain how the adaptive power control scheme works for D2DS. In our scheme, D2DS first with the full Tx power. This creates interference and the interference is even further accelerated when more than one D2DRs simultaneously share the same RB with another D2DR or mUE for Tx and these D2DRs use full Tx power

for their Tx. In order to mitigate and prevent this further generation of interference, D2DRs sharing the RB are set to transmit with its minimum power,  $P_{D2DS}^{min}$  as shown in (4).

$$P_{D2DS} = P_{D2DS}^{min} \tag{4}$$

If this power satisfies the required SINR of D2DR threshold,  $\gamma_{threshold}$ , RB is then allocated for this D2DR. If the D2DR threshold is not satisfied, the D2DS power is then increased iteratively until its threshold satisfies as shown in (5)

$$P_{D2DS} = \begin{cases} P_{D2DS}^{min} & \gamma_{D2DR} \geq \gamma_{threshold} \\ \sum_{n=1}^N n P_{D2DS}^{min} & \gamma_{D2DR} < \gamma_{threshold} \end{cases}, \tag{5}$$

where  $n=\{1, 2, \dots N\}$  with N being the number of iterations to satisfy the required D2DR SINR,  $\gamma_{D2DR}$  and is limited to the D2DS max power.

The D2DR Tx power adjustment will stop if the condition (6) is satisfied.

$$\gamma_{D2DR} \geq \gamma_{threshold}, \tag{6}$$

where  $\gamma_{D2DR}$  is D2D Received SINR and  $\gamma_{threshold}$  is the required SINR of D2DR.

## 4 Performance Evaluation

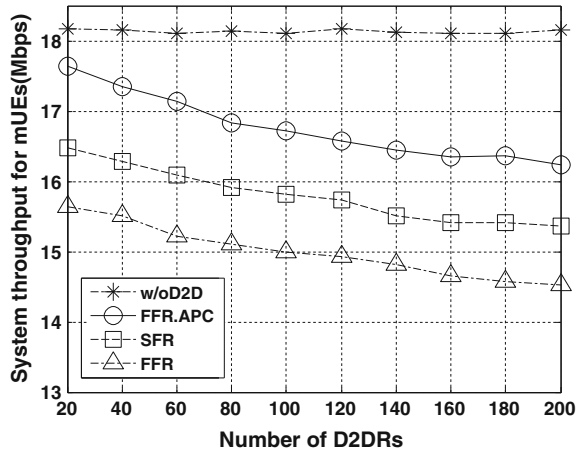
We evaluate the performance measurement in terms of three measurement indicators, i.e., signal to *interference plus noise ratio* (SINR), system throughput and outage probability (Table 1).

The detailed explanation and the description of the performance measurement are given [11]. We compare the performance of the proposed scheme to the network without D2D and the scheme which FFR and SFR to D2D links. Four cases are

**Table 1** System parameters

Parameter	Value
Carrier Frequency	2 GHz
Bandwidth for DL	10 MHz
Bandwidth of sub-channel	180 KHz
mBS/D2D radius	866 m/20 m
mBS Tx power ( $P_{mBS}$ )	Inner zone = 41.7 dBm(15 W), Outer zone=43.42 dBm(22 W)
D2DS Tx power ( $P_{D2DS}$ )	$P^{min} = 8$ dBm(6.3 mW), $P^{max} = 24$ dBm(251 mW)
Noise power density ( $N_o$ )	-174 dBm/Hz

**Fig. 4** System throughput for mUEs



considered: 1) consider having only mUEs (w/oD2D), 2) the FFR scheme 3) the SFR scheme, and 4) FFRAPC presenting our proposed scheme.

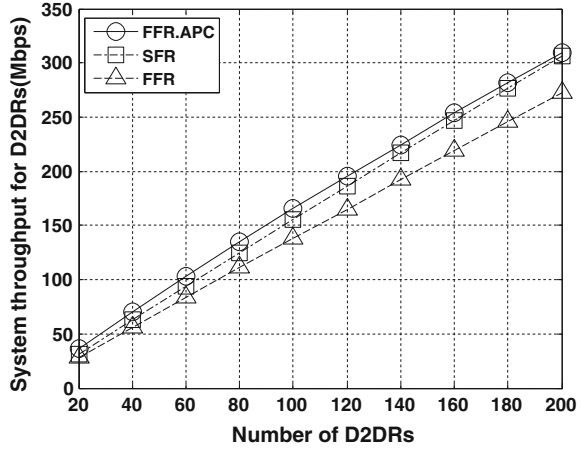
Figure 4 describes the results of system throughput for mUEs, as the number of D2DRs increases. The traditional cellular system that is without D2DSs always has the same performance about 18.5Mbps but the results of the FFR, SFR, and FFR.APC schemes decrease as the number of D2DSs increase because the interference from the D2DSs increase. The proposed scheme has better performance of system throughput than those of the SFR and FFR schemes. The throughput of the FFR.APC with 200 D2DSs is approximately decreased 20 % comparing to the traditional cellular system. Also, the FFR.APC scheme increased by 8 % and 17 % compared to those of the SFR and FFR schemes, respectively. The reason is that FFR.APC has less interference from D2DS even though it uses all RBs.

Figure 5 describes total throughput for D2DRS and the results of throughput increase linearly as the number of D2Ds increases. All the D2DR system throughput for the proposed scheme are higher than those of the SFR and FFR schemes. The results of the proposed scheme with 200 D2DSs are approximately increased by 14 % and 2 % compared to those of the FFR and SFR schemes, respectively.

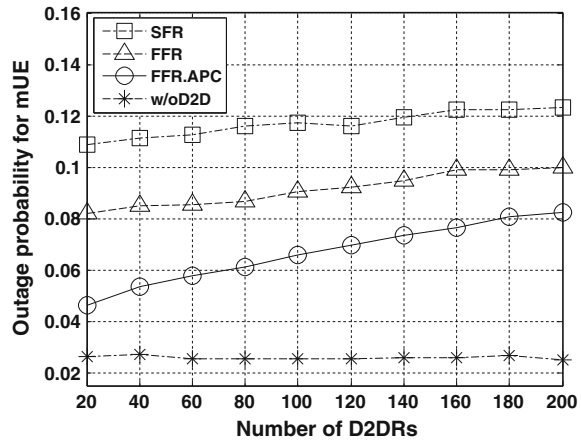
Figure 6 describes the results of total outage probability as the number of D2DSs increases. The results increase as the number of D2DSs increases, as does the increasing interference from the D2DSs. The traditional cellular system that is without D2DSs always has the same outage probability, about 0.023. The results of the proposed scheme are lower than SFR and FFR schemes. The results of the proposed scheme with 200 D2DRs are approximately decreased by 22 and 42 % compared to those of the FFR, SFR system.

Figure 7 describes the system throughputs of center and edge with the number of 200 D2DRs. System throughput of mUEs in the proposed schemes are higher than the other schemes. In particular, since it uses frequency resources that are different from a D2D link at the cell-edge, higher performances were displayed. FFR.APC

**Fig. 5** System throughput for D2DRs



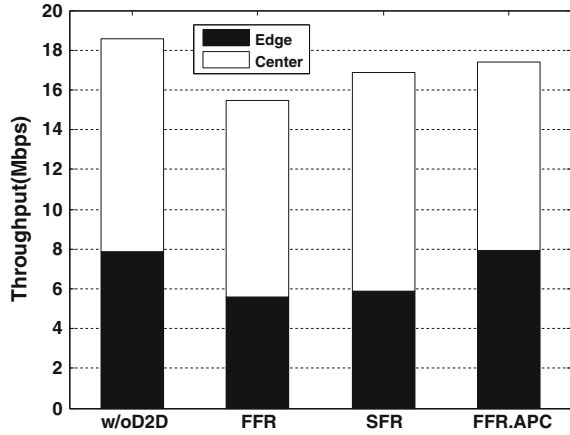
**Fig. 6** The outage probability for mUE



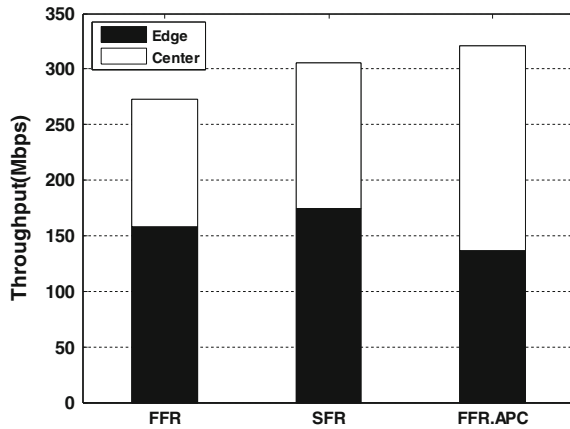
scheme can be improved by optimizing the cell-edge, the cell-center radius, and the allocation of RBs in cell-center and edge mUE such that the overall network throughput is maximized.

Figure 8 describes the system throughputs of center and edge. The System throughput of D2DRs in the proposed scheme is higher than the other schemes. However, when the cell-edge performance of the suggested FFR.APC scheme is compared with the SFR scheme, The SFR scheme's performance was better because of the cell edge uses a higher reuse factor, but it can be seen that the performance of the suggested FFR.APC scheme is higher due to the performance difference at the cell-center. Therefore, the interference is reduced significantly in the comparison with other scheme. Also, D2D can control and optimize the Tx power based on the user requirement to mitigate the interference in the proposed scheme

**Fig. 7** System throughput for mUEs



**Fig. 8** System throughput for D2DRs (center and edge)



## 5 Conclusion

In this paper, we proposed a novel *Fractional Frequency Reuse based Adaptive Power Control* (FFR.APC) scheme to enhance the *downlink* (DL) system performance for D2D networks. In the proposed scheme, the mBS and D2DSs service mUEs and D2DRs in the inner and outer zones with different frequency bands and *transmission* (Tx) power, to reduce substantial interference. Simulation results showed the proposed scheme outperforms D2D networks with SFR and FFR schemes in terms of system throughput and outage probability for mUEs and D2DRs. We discovered that the FFR.APC is the better scheme for improving the performance of the D2D network. Based on this study, we plan to conduct additional future works in the fields of frequency planning schemes and power management technology that have been improved in a sectorized environment which considers a directional antenna.

**Acknowledgments** This research was supported by “Energy Efficiency Network System Development for Heterogeneous Distributed power management and energy balancing Service”.

## References

1. Janis, P., et al.: Device-to-device communication underlying cellular communications systems. *Int. J. Commun. Netw. Syst. Sci.* **2**(3), 169–178 (2009)
2. Lei, L., Zhong, Z., Lin, C., Shen, X.: Operator controlled device-to-device communications in LTE-advanced networks *IEEE wireless communication*, 96–104 (2012)
3. Doppler, K., Rinne, M.P., Wijting, C., Ribeiro, C.B., Hugl, K.: Device-to-device communication as an underlay to LTE-advanced networks. *IEEE Comm. Magazine* **47**(12), 42–49 (2009)
4. Doppler, K., Rinne, M., Jnis, P., Ribeiro, C. B., Hugl, K.: Device-to-device communications; functional prospects for LTE-advanced networks. In: *Proceedings of IEEE International Conference on Communication Workshops*, 1–6 (2009)
5. 3GPP TS 36.211 v10.0.0, Evolved universal terrestrial radio access (E-UTRA); physical channels and modulation (2011)
6. WiMAX Forum, *Mobile WiMAX part 1: A technical overview and performance evaluation* (2006)
7. R1-050841, Huawei, Further Analysis of Soft Frequency Reuse Scheme, 3GPP TSG RAN WG1#42, Aug. 29–Sep. 2 (2005)
8. Chae, H.S., Gu, J., Choi, B.G., Chung, M.Y.: Radio resource allocation scheme for device-to-device communication in cellular networks using fractional frequency reuse, *Asia-Pacific Conference on Communications (APCC)*, 58–62 (2011)
9. Yoon, D., Cho, K., Lee, J.: Bit Error Probability of Mary Quadrature Amplitude Modulation, *IEEE VTC-Fall*, **5**:2422–2427 (2000)
10. Janis, P., Koivunen, V., Ribeiro, C., Korhonen, J., Doppler, K., Hugl, K.: Interference-aware resource allocation for dvice-to-device radio underlying cellular networks. In: *Proceedings of the IEEE 69th Vehicular Technology Conference*, 1–5 (2009)
11. Kim, T.S., Lee, K.H., Ryu, S., Cho, C.H.: Resource allocation and power control scheme for interference avoidance in an lte-advanced cellular networks with device-to-device communication. *Int. J. Control Autom.* **6** (2013)



# Influence of the Management Protocols on the LTE Self-configuration Procedures' Performance

Mariusz Slabicki and Krzysztof Grochla

**Abstract** The performance of SNMP, NETCONF and CWMP in LTE self-configuration scenario is investigated. The time required to perform the autoconfiguration of the LTE eNodeB is evaluated and a part of the LTE SON procedures. The influence of the size of the firmware file and the number of nodes being configured by a single network management server is verified. The results show that the SNMP significantly outperforms NETCONF and CWMP in terms of time required to finish the self-configuration, especially for base stations with small firmware size.

## 1 Introduction

The amount of data transmitted in cellular networks is rapidly growing during last few years [7, 15]. The network operators extend their networks to respond to the growing demand, by adding more base station and increasing the radio frequency reuse. The shrinking cell size and growing number of small, low cost base stations leads to greater complexity of the network. The large amount of managed devices in modern cellular network makes it very hard to manually control and regulate the parameters of the base stations to achieve the optimal performance. To address this issue the concept of Self Organizing Networks (SON) has been proposed by the 3GPP organization as a part of the LTE standard to minimize the OPEX and CAPEX of the network. According to the SON concept the parameter tuning is done automatically based on measurements. The SON function may be executed locally, or in a centralized location: the Operation and Maintenance (OAM) system. The OAM communicates with the network devices using the network management protocols. The protocols allow to monitor the state of the network devices, gather statistics of

---

M. Slabicki · K. Grochla (✉)  
Institute of Theoretical and Applied Informatics, Polish Academy of Sciences,  
Gliwice, Poland  
e-mail: kgrochla@iitis.pl

M. Slabicki  
e-mail: mslabicki@iitis.pl

the network performance and change the configuration of the wireless equipment when needed.

The self-organization functions of a mobile network consist of three elements: the autoconfiguration (also called self-configuration), self-optimization and self-healing [11]. The autoconfiguration is the first element of the SON procedures and provides the initial configuration for the base station, speeding up the initial deployment of the base stations. The self-optimization functions perform ongoing monitoring of the network operation conditions and react to the changes, realizing functions such as mobility robustness optimization, load balancing and maintenance of neighbor relation [16]. The self-healing functions react to failures of network devices and allow to reconfigure the network in case of a temporal unavailability of some of the devices (e.g. a base station).

In this work we concentrate on the performance of the autoconfiguration procedures. We consider the scenario of a newly added base station (eNodeB). During the boot up the eNodeB needs to register in operation and maintenance (OAM) server. We assume that the eNodeB downloads a firmware file and receives its configuration via a network management protocol. Our main goal is to analyse how the selection of network management influences the performance of the LTE autoconfiguration procedures. We have evaluated three popular network management protocols: SNMP [20], NETCONF [6] and CWMP (TR.069) [9]. We also analyse how the size of the firmware file influences the time required to perform the initial configuration.

The rest of the paper is organized as follows: we present the brief review of the literature in Sect. 2, next we describe the model used for performance evaluation in Sect. 3. In the following Sect. 4 we discuss the results of the simulations. we conclude the work in Sect. 5.

## 2 Literature Review

The use cases and procedures related to the self-configuration of the LTE network have been defined in the LTE standard [1]. The SON functions can be classified into three types: centralized, distributed and hybrid. In this paper we concentrate on the centralized self-configuration, as it was defined in [1]. The SON concept is continuously evolving [13] and new research aims in turning the LTE network into an unified self-managed system, which controls the complex network environment as a single entity [5].

The self-configuration procedures for the LTE eNodesB have been defined in the Release 8 of the standard [17]. The performance of the self-configuration procedures was not deeply investigated in the literature, as it is not a problem for large cell deployments. But with wider adoption of femtocells the number of base stations in the LTE network may become very large and the efficient self-configuration of all network nodes is crucial [3]. The performance of the three network management protocols have been deeply investigated, but not in relation to the SON scenario. In the paper [10] authors performed a quantitative analysis of the performance of the SNMP and

NETCONF protocols, proving by simulation that NETCONF provides performance comparable to SNMP for complex networks, but the model neglected the influence of delays related to transmission messages through the network. Authors of [18] consider usage of resources in SNMP and Netconf protocols implemented on embedded devices, but there is no evaluation of the influence of the network conditions on the transmission efficiency. There is quite a few papers analysing a single management protocol performance over wired or wireless channel, such as [14] or [12] for SNMP, but they do not provide the comparison to other network management protocols. In our previous work we have compared the performance of SNMP, NETCONF and TR.069 over a wireless network, but we have not considered firmware download phase, which is crucial for element of the LTE eNodeB's autoconfiguration [19].

### 3 Model Description

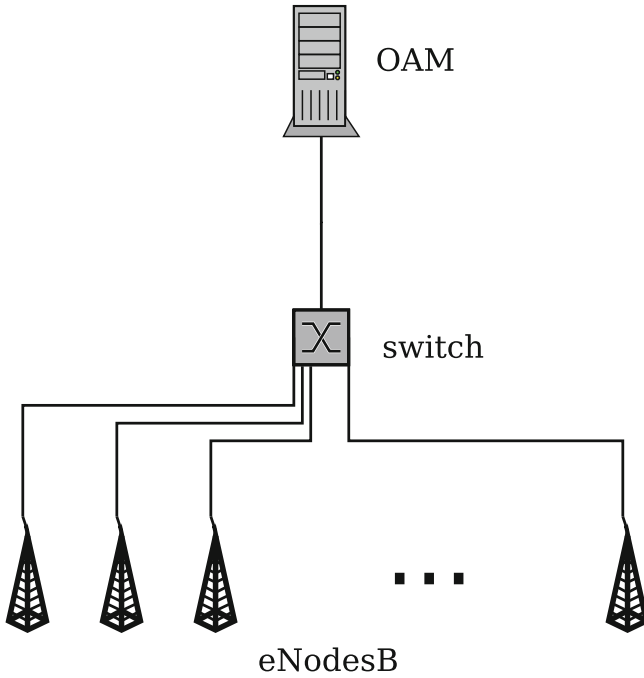
As defined in [1] the self-configuration procedure for LTE eNodeB starts by obtaining the IP address, next the eNodeB provides the eNodeB with basic information about itself. On the next step the OAM provides the addresses and path to download the firmware and the initial configuration. The OAM model was very simple, with no delay for processing firmware request. The eNodeB downloads the file with the firmware, fetches the configuration and after 1 second (simulated reboot time) starts the operation. In our study we omit the IP address assignment phase, as it is an typical use of DHCP protocol and has been deeply analyzed in the literature for both wired [4] and wireless [8] networks. We concentrate on the two most time consuming steps: the firmware download and the configuration phase.

To evaluate the time required for the eNodeB self-configuration we have implemented a model in OMNeT++ [21] with INET Framework, which adds to OMNeT++ support for simulating network based on TCP/IP. We assumed a star topology, where multiple eNodesB communicate with a single OAM server through a high speed (1Gbit/s) wired link. The schema of the simulation scenario is shown in Fig. 1.

Presented simulation scenario looks as follows:

- eNodeB connects via TCP connection to OAM, and sends request to download Firmware,
- OAM sends Firmware,
- eNodeB closes connection,
- after a Restart Timeout, eNodeB opens a connection (by one of the management protocols),
- eNodeB sends a request for configuration,
- OAM sends configuration,
- eNodeB closes connection.

We have evaluated scenarios with different firmware sizes: from 0.5 to 16 MB, and number of eNodesB from 10 to 320. Size of configuration packets were estimated



**Fig. 1** Simulation schema, all connections are Gigabit Ethernet

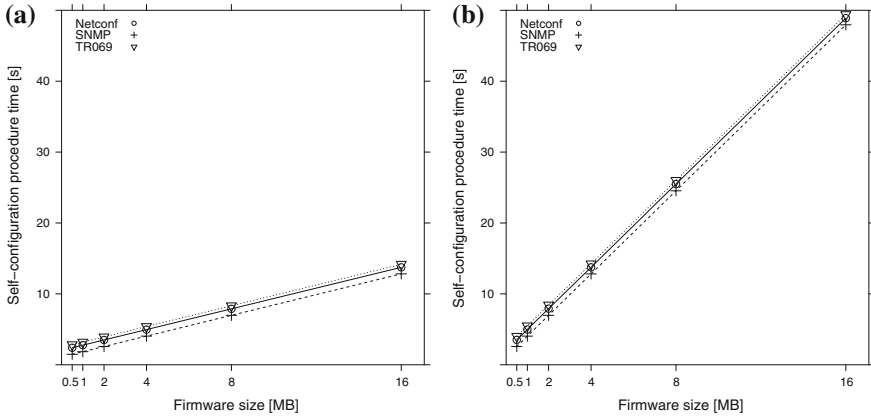
**Table 1** Configuration packets size, in Bytes

	SNMP	Netconf	TR0.69
GET	4386	9152	30,890
RESPONSE	8144	14,438	38,054

on the number of parameters that can be set in eNodeB (it is shown in [2]), and protocol overhead. Overall packets sizes are shown in Table 1. Packets used to control connections used by management protocols had the same sizes as in the paper [19].

## 4 Results

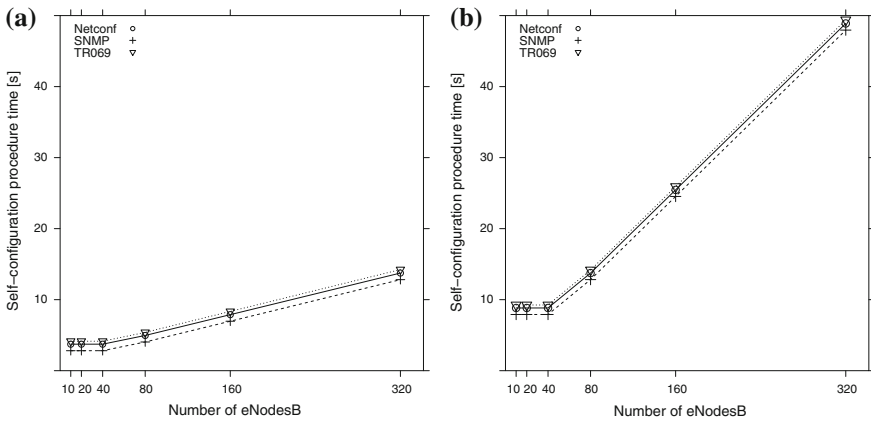
The Fig. 2 shows how much time is required for the the autoconfiguration procedures to finish depending on the management protocol used. The time needed for self-configuration procedure increases linearly as the firmware size increases. The standard deviation for all time intervals was smaller than 1 % and is not presented on the figures, as well as 0.95 confidence intervals which are smaller than 0.1 % of the mean value. The differences between all three protocols considered depends on the size of the firmware and the number of clients. The SNMP is faster for all considered



**Fig. 2** Time required for the self-configuration procedures for different management protocols. **a** 80 nodes, **b** 320 nodes

cases, with TR.069 requiring largest amount of time to finish the self-configuration phase. The selection of the protocol is more important when small firmware size is used, as when the 0.5MB file is considered, the SNMP outperforms TR.069 by almost 50%. The two graphs (for 80 eNodesB and 320 eNodesB) on Fig. 2 show also that the rate of the time increase depends highly on the number of eNodesB that are configured.

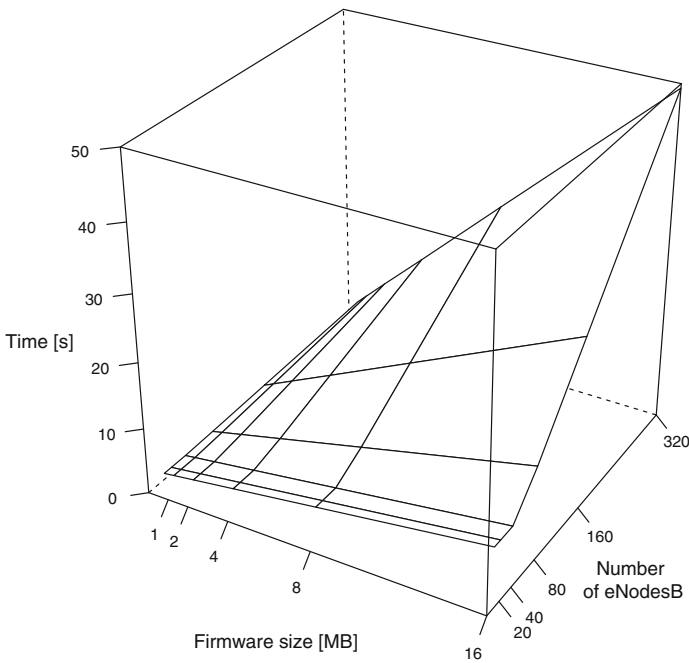
The simulation results show that the time needed for the self-configuration procedure, depends on the number of eNodesB in the network, as it can be seen on Fig. 3. The characteristic of time needed for self-configuration procedure can be divided for



**Fig. 3** Time required for the self-configuration procedures for different number of eNodesB. **a** 4 MB, **b** 16 MB

two parts, separated by a breakthrough point. The results of the simulation suggest that the breakthrough point is approximately 40 for the model considered. If number of eNodesB is below particular level, the time is constant. But if the number is higher than mentioned level, the time increases linearly. This shows that for small number of nodes the time needed to finish the autoconfiguration procedures is almost constant. For larger number of nodes the transmission speed is limited by the network throughput and as more nodes are transmitting similar data the bandwidth available per node gets smaller and the time need to finish the transmission is higher. These graphs show also that the rate of increase differs with the firmware size.

To demonstrate the dependency on the firmware size and number of nodes, the time needed for Self-configuration procedure is presented in these two domains in Fig. 4. This graph shows results for TR069—other protocols have similar characteristic. We can observe that the breakthrough point dividing the constant part and the linear increase in time for number of base stations is more clear when the base station use large firmware, for small files the plot is more linear.



**Fig. 4** Time required for the self-configuration procedures for different number of eNodesB and different firmware sizes

## 5 Conclusions

Our research shows that the management protocol selection is important factor determining the time required for self-configuration procedures when the firmware they need to download is small. In such a case the TR.069 requires approximately twice the time needed for SNMP. The NETCONF showed only slightly better performance comparing to TR0.69. This difference is caused by more efficient binary encoding used in SNMP, comparing to XML used by two other protocols.

The size of the firmware used by the eNodeB more heavily influences the performance of the self-configuration procedures in LTE networks than the selection of the network management protocol. The time required for the eNodeB to start the operation is counted in tens of seconds and needs to be taken into account when designing a self-healing and high availability procedures. However the variance of this time is very low (with standard deviation lower than 1 % of the measured value), what makes it very easily predictable.

**Acknowledgments** This work was funded by the Polish National Centre for Research and Development, under research grant nr LIDER/10/194/L-3/11/: project “Optimization and load balancing in next generation wireless networks”, <http://projekty.iitis.pl/zosb>.

## References

1. 3GPP: Telecommunication management; self-configuration of network elements; concepts and requirements. TS 32.501, 3rd Generation Partnership Project (3GPP) (2013). <http://www.3gpp.org/DynaReport/32501.htm>
2. 3GPP: Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN); Overall description; Stage 2. TS 36.300, 3rd Generation Partnership Project (3GPP) (2014). <http://www.3gpp.org/dynareport/36300.htm>
3. Barbieri, A., Damjanovic, A., Ji, T., Montojo, J., Wei, Y., Malladi, D., Song, O., Horn, G.: Lte femtocells: system design and performance analysis. *IEEE J. Sel. Areas Commun.* **30**(3), 586–594 (2012)
4. Brik, V., Stroik, J., Banerjee, S.: Debugging dhcp performance. In: Proceedings of the 4th ACM SIGCOMM Conference on Internet Measurement, pp. 257–262. ACM (2004)
5. Eisenblatter, A., Gonzalez Rodriguez, B., Gunnarsson, F., Kurner, T., Litjens, R., Sas, B., Sayrac, B., Schmelz, L.C., Willcock, C.: Integrated self-management for future radio access networks: vision and key challenges. In: Future Network and Mobile Summit (Future Network Summit), 2013, pp. 1–10. IEEE (2013)
6. Enns, R., Bjorklund, M., Schoenwaelder, J.: Netconf configuration protocol. *Network* (2011)
7. Foremski, P., Gorawski, M., Grochla, K.: Source model of tcp traffic in lte networks. In: Information Sciences and Systems 2014, pp. 125–135. Springer International Publishing (2014)
8. Grochla, K., Buga, W., Dzierzega, J., Pacyna, P., Seman, A.: Autoconfiguration procedures for multiradio wireless mesh networks based on dhcp protocol. In: IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks & Workshops, 2009, WoWMoM 2009, pp. 1–6. IEEE (2009)
9. Group, D.W., et al.: TR-069 CPE WAN Mangement Protocol Issue: 1 Amendment 5 version 1.4 (2013)

10. Hedstrom, B., Watwe, A., Sakthidharan, S.: Protocol Efficiencies of NETCONF versus SNMP for Configuration Management Functions. Ph.D. thesis, Masters thesis, University of Colorado (2011)
11. Hu, H., Zhang, J., Zheng, X., Yang, Y., Wu, P.: Self-configuration and self-optimization for lte networks. *IEEE Commun. Mag.* **48**(2), 94–100 (2010)
12. Johnson, R.B.: Evaluating the use of SNMP as a wireless network monitoring tool for IEEE 802.11 wireless networks. ProQuest (2009)
13. Jorguseski, L., Pais, A., Gunnarsson, F., Centonza, A., Willcock, C.: Self-organizing networks in 3g pp: standardization and future trends. *IEEE Commun. Mag.* **52**(12), 28–34 (2014)
14. Kantorovitch, J., Mahonen, P.: Case studies and experiments of SNMP in wireless networks. In: 2002 IEEE Workshop on IP Operations and Management, pp. 179–183. IEEE (2002)
15. Paper, C.W.: Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2010–2015 (2011)
16. Połys, K., Grochla, K.: Mobility robustness in lte based on automatic neighbor relation table. In: *Computer Networks*, pp. 232–241. Springer International Publishing (2015)
17. Ramiro, J., Hamied, K.: *Self-Organizing Networks (SON): Self-Planning, Self-Optimization and Self-Healing for GSM, UMTS and LTE*. Wiley, New York (2011)
18. Sehgal, A., Perelman, V., Kuryla, S., Schonwalder, J.: Management of resource constrained devices in the Internet of Things. *IEEE Commun. Mag.* **50**(12), 144–149 (2012)
19. Słabicki, M., Grochla, K.: Performance evaluation of snmp, netconf and cwmp management protocols in wireless network. In: Amir Hussain, M.I. (ed.) *Electronics, Communications and Networks IV: Proceedings of the 4th International Conference on Electronics, Communications and Networks*, 12–15 December 2014, Beijing, China. CRC Press (2014)
20. Stallings, W.: *SNMP, SNMPv2, SNMPv3, and RMON 1 and 2*. Addison-Wesley Longman Publishing Co., Inc. (1998)
21. Varga, A., et al.: The OMNeT++ discrete event simulation system. In: *Proceedings of the European Simulation Multiconference (ESM2001)*, vol. 9, p. 185. sn (2001)



# Subcarrier Allocation for LTE Soft Frequency Reuse Based on Graph Colouring

Krzysztof Grochla and Konrad Polys

**Abstract** The paper discusses the parametrization of Soft Frequency Reuse (SFR) method to minimize the inter-cell interference. The optimization method to select the subcarrier allocation based on graph colouring is proposed. It is evaluated using numerical model for uniform (honeycomb) and non-uniform (based on sample deployed network) network topologies. The results prove that the proposed optimization method can significantly improve (up to 20%) the throughput comparing to the random subcarrier allocation. We also show that selection of scheduling policy determines the optimal sub-band allocation and different SFR configuration should be used for different types of schedulers.

**Keywords** LTE · Soft frequency reuse · SFR · Graph colouring · Inter-cell interferences · Scheduling

## 1 Introduction

The new generation wireless network need capacity increase to support the large amount of traffic generated by diverse Internet services. The efficient utilization of scarce radio resources is crucial to maximize the amount of data the network is able to deliver. The amount of data transmitted in cellular networks is rapidly growing during last few years. Thus the wireless networks need to employ multiple means to increase the frequency reuse factor and use the same radio resources in different location and the same moments. Multiple methods to achieve this goal have been proposed, from spatial multiplexing using MIMO antennas, through the heterogeneous structure of the network with pico- and femto-cells deployed in areas with large load of traffic,

---

K. Grochla (✉) · K. Polys  
Institute of Theoretical and Applied Informatics of PAS, Gliwice, Poland  
e-mail: kgrochla@iitis.pl

K. Polys  
e-mail: kpolys@iitis.pl

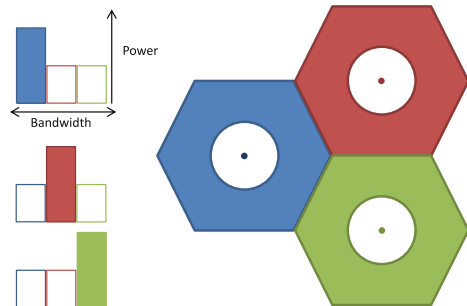
to advanced transmission scheduling methods exploiting the information about the user location to more effectively allocate the subcarriers and time slots to users.

The Long Term Evolution (LTE) has become a leading standard of high speed cellular networks, with almost 500 millions of devices connected worldwide in 2014 [2]. The LTE networks use the OFDMA modulation in the downlink and provides data transfer of up to 300Mbit/s in single cell. LTE is a mobile broadband solution that offers a rich set of features with a lot of flexibility in terms of deployment options and potential service offerings, supporting adaptive modulation and coding and very low delay (less than 10 ms) [1].

In the cellular networks that reuse the same frequency among multiple cells the inter-cell interference become a serious problem. The fractional frequency reuse (FFR) is one of the methods aiming for more effective radio resource allocation and interference mitigation in cellular networks [3]. It has been designed for OFDMA based wireless networks. In FFR scheme the cell bandwidth (all available subcarriers) is divided into two or more parts, of which one part is assigned to the users near to the centre of the cell and the other part(s) to the users near to the border. The eNodeB uses lower transmission power to send the data to the users within the center of the cell (the inner part), and higher power to the rest of the users to maximize the received signal level. There are two types of FFR: strict FFR and soft frequency reuse (SFR) [11]. In SFR the each cell uses one part of the subcarriers (usually not used in the outer parts of neighbouring cells) near the border, and the remaining subcarriers in the centre, as it is depicted on Fig. 1. In strict FFR all cells use in the centre a dedicated part of the bandwidth, which is not used in any of the outer parts by other cells (what would be seen as a dedicated colour for the centres). The number of parts to which the available bandwidth is divided (number of sub-bands) is denoted as the factor  $K$ . Apart from strict FFR and SFR a few different schemes have been proposed, with different methods of grouping of the users into outer/inner part or proposing the sectioning of the cell area into more than two groups [8].

The rest of the paper is organized as follows: we define the subcarrier allocation problem and present the motivation for our work in Sect. 2, next we describe the model used for performance evaluation in Sect. 3. In the following Sect. 4 we discuss how the application of the proposed algorithm for the configuration of SFR influences the throughput offered to the users. We finish with brief conclusion in Sect. 5.

**Fig. 1** Example of SFR schema with reuse factor 3



## 2 Subcarrier Allocation in SFR

In SFR the smaller is the number of sub-bands  $K$ , the larger bandwidth is available in each cell, but the inter-cell interference are larger. When the  $K$  is equal one or two, there will always be a situation when two adjacent cells share the same set of subcarrier in the outer part. The factor  $K = 3$  is the smallest which allows to select the allocation of subcarriers in such a way that there will be no neighbours using the same part of bandwidth on the border. However such ideal allocation is only possible for the network with simple topology, where all cells have only three neighbouring cells. In real life deployment, where cells have different sizes and partially overlapping coverage, there are multiple possible allocations of the parts of the bandwidth which provide different level of inter-cell interference in different parts of the network.

The problem of allocation of the sections of the bandwidth to be used in the outer parts of the cells can be seen as a graph colouring problem. The nodes represent the network cells, the edges is the neighbour relation between the cells and the colour denotes the part of the subcarriers used by a particular cell. The number of available colours is determined by the number of sub-bands  $K$ —if the  $K = 3$  than each of the base stations needs to be assigned with one of the three available colours. The fractional frequency reuse schema is parametrized by three main factors: (1) the method in which the cell is divided into inner and outer part; (2) the proportion between the TX power used for both parts of the cell and (3) the allocation of subcarrier used in the outer part of the cells. The appropriate selection of the threshold by which the cells is divided have been evaluated in e.g. in [14, 17]. The paper [19] provides evaluation of the influence of TX power selection on SFR performance. The joint evaluation of the influence of the factors (1) and (2) on the network performance is given in [9], where also the variability of the cell throughput when user move is evaluated.

The challenges related to the allocation of spectrum to minimize the inter-cell interferences are presented in [16]. The graph-based approach to minimize interference in LTE using coordinated fractional frequency reuse has been initially proposed in [10]. In this approach every cell communicates all data required to draw the interference graph to a central node, called interference coordinator. It is hard to implement, since it requires almost instant communication among base stations. Another graph based approach is presented in paper [4], which use graph colouring to minimize the interference. The algorithm proposed by Chang et al. represents the mobile stations as the graph vertices and generates dynamic FFR scheme. While the static SFR schemes are much easier to implement by the network operator, as require no change of the scheduling algorithm in eNodeB and require only static configuration, we target the problem of finding the optimum selection of subcarrier allocation for static SFR scheme. In [5] the frequency selection is jointly analysed with the selection of optimal selection of number of sub-bands, we extend this approach by considering different schedulers.

### 3 Model Description

To find the optimal allocation of subcarriers per eNodeB we have implemented a numeric model representing the location of the cells, inter-cell interference and allowing to calculate the signal level, modulation and coding scheme used by the users in each place within the modelled network. We have used a PyLTE framework, which is an open source project available on GitHub [13]. It allows to run different LTE network configurations, calculate UE throughput with provision of schedulers and also it is possible to run optimizations of network parameters, such as TX power and SFR configuration.

Two network deployments are considered. The first one is based on the honeycomb topology. The cell size  $R_c$  (Radius of Curvature) is equal 1666 m. Second scenario is based on deployment from [15] where Hannover in Germany was an example. It is more realistic approach than the honeycomb deployment because it is usually impossible to deploy eNBs in such orderly way. The  $R_c$  of cell is set to 400 m. In both scenarios we have placed 1024 UEs as a uniform grid to represent the uniform user distribution. We consider uniform network load and uniform spatial distribution of the users among the network.

Radio resources were allocated with Soft Frequency Reuse (SFR) approach which divides the cell into two regions—depicted in Fig. 1. For the inner area the 2/3 of all subcarriers is available and eNB transmits there with lower power and for the outer area the 1/3 of subcarriers are used with higher power. In both scenarios in the inner part of cell the eNB transmits with power of 37 dBm and in the outer part with power of 40 dBm. The point of the change between both parts is set to 60 % of  $R_c$ , following the optimal parameters found in [7].

### 4 Optimization Results

The goal of our work is to find the best eNBs colouring scheme which gives the highest network throughput. To achieve it, we have implemented a optimization algorithm, using the sum of throughput offered to each user as the optimization goal. Throughput of whole network is calculated as a sum of every UE's throughput. First, the SINR (Signal to Interference and Noise Ratio) from eNB for the every UE is calculated using the SUI radio propagation model [6]. For each user the eNodeB for which the highest signal level is measured is considered as the serving eNB, next basing on the distance it is determined whenever the user is within the inner or outer part of the cell. At the next step the other eNBs which use the same frequency (the same colour for users in the outer part of the cell or all neighbours) are considered as the source of interference. Finally the SINR is used to determine the modulation and coding scheme [12], which with the use of scheduler model is transformed into throughput available for particular user.

We have used the optimization based on Simple Evolutionary Algorithm (SEA). The SEA is a random based solution space searching metaheuristic [18] and was selected because the colour allocation problem is non-linear and the evolutionary algorithms are well suited to such problems. The complexity of the optimization method is determined by the size of the network and the number of generation. In both analysed networks we experimentally determined that there is no improvements when more than 100 generations are used.

The first set of experiments was performed for Round Robin (RR) scheduler, which allocates radio resources consecutively and cyclically to every user as long as there is some traffic to be sent. Every user terminal is equally scheduled without taking the channel quality into account. Depending on the modulation and coding level used by each user, the transmission rate may be very different from user to user. The colour allocations for the honeycomb and Hannover network topology are presented on the Figs. 2 and 3.

The colour allocations for max-min fair scheduler (MMF) are shown on Figs. 4 and 5. The MMF scheduler is assigning each data flow with a transmission time inversely proportional to its data rate, anticipated based on modulation and coding scheme. It guarantees that every user transmits the same amount of data. The allocation for the honeycomb network was practically the same as with the RR scheduler. In less theoretical topology the colour selection was different and the algorithm was less likely to assign two adjacent cells with the same colour; there are 5 such cases for RR scheduler and only one in MMF. This is because the optimization goal function emphasised the minimization of interference for the users near the border of the cell, over the lowered modulation for the users near the centre of the cell.

The Fig. 6 shows quantitative evaluation of the efficiency of the proper subcarrier allocation for the SFR networks. The proposed optimization method outperforms the random allocation of colours for both the honeycomb and Hannover

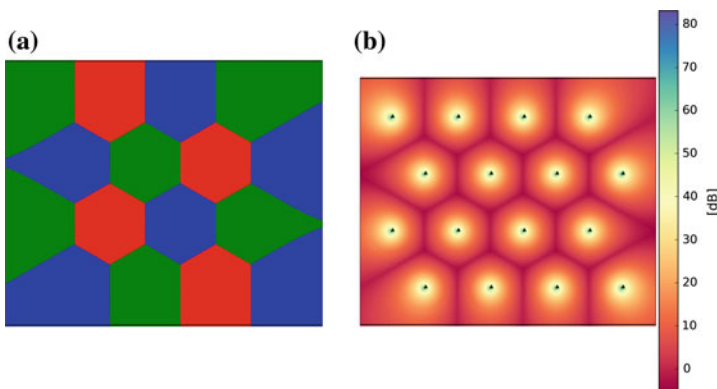


Fig. 2 Honeycomb deployment. Round Robin scheduler. **a** Colour map, **b** SINR map (dB)

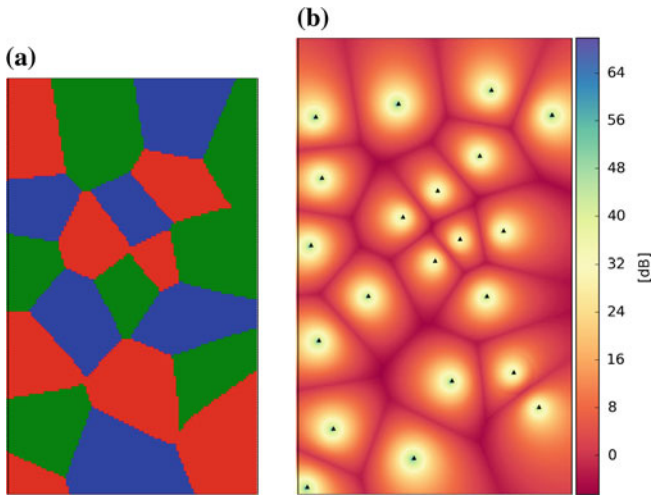


Fig. 3 Hannover deployment. Round Robin scheduler. **a** Colour map, **b** SINR map (dB)

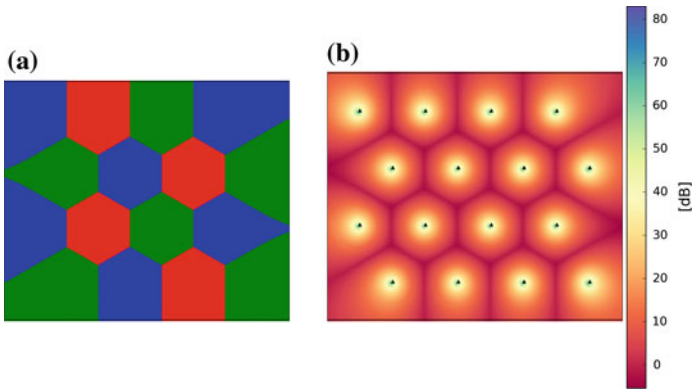
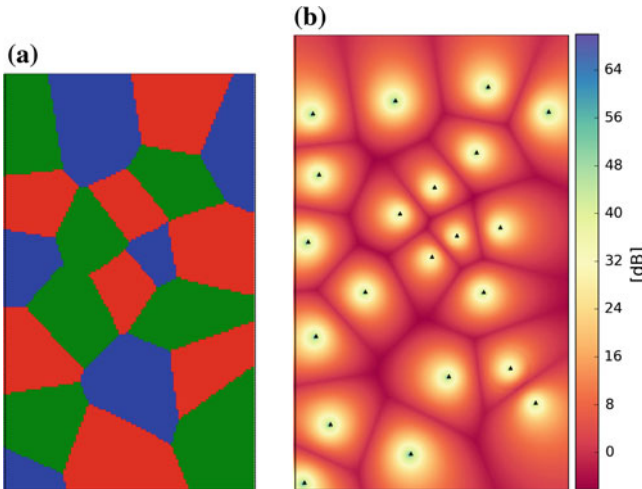
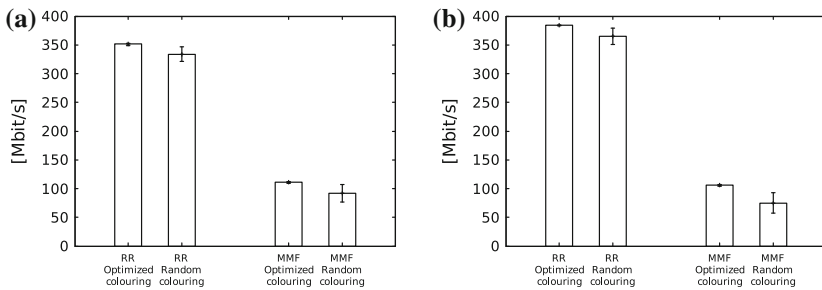


Fig. 4 Honeycomb deployment. Fair scheduler. **a** Colour map, **b** SINR map (dB)

topologies. The difference is significantly higher when max-min fair scheduler is used, the results show that the optimized selection of the SFR configuration provides the total throughput offered to the users higher by more than 20 % than the random selection of colours.



**Fig. 5** Hannover deployment. Fair scheduler. **a** Colour map, **b** SINR map (dB)



**Fig. 6** Average network throughput with Round Robin scheduler (RR), Max-Min Fair scheduler (MMF) for honeycomb (a) and Hannover (b) deployment. Error bars show min and max values

## 5 Conclusions

The SFR allows to increase the network throughput by minimizing the inter-cell interference, however it requires appropriate selection of the sub-band allocation. The optimization method described in the paper allows to select configuration that provides the users significantly higher throughput than the random selection. The results show that the selection of scheduling policy influences the sub-band and different SFR sub-band allocations are optimal for different types of schedulers.

**Acknowledgments** This work was funded by the Polish National Centre for Research and Development, grant no LIDER/10/194/L-3/11, <http://projekty.iitis.pl/zosb>.

## References

1. Ali-Yahiya, T.: *Understanding LTE and Its Performance*. Springer Science & Business Media (2011)
2. Americas, G.: Year-End 2014: Nearly Half a Billion LTE Connections Worldwide. <http://www.4gamericas.org/en/newsroom/press-releases/year-end-2014-nearly-half-billion-lte-connections-worldwide/>. Accessed 15 April 2015
3. Boudreau, G., Panicker, J., Guo, N., Chang, R., Wang, N., Vrzic, S.: Interference coordination and cancellation for 4g networks. *IEEE Commun. Mag.* **47**(4), 74–81 (2009)
4. Chang, R. Y., Tao, Z., Zhang, J., Kuo, C.C.: A graph approach to dynamic fractional frequency reuse (ffr) in multi-cell ofdma networks. In: *IEEE International Conference on Communications, ICC'09*, pp. 1–6. IEEE (2009)
5. Chen, L., Yuan, D.: Soft frequency reuse in large networks with irregular cell pattern: How much gain to expect? In: *IEEE 20th Symposium on PIMRC, 2009*, pp. 1467–1471. IEEE (2009)
6. Erceg, V., Greenstein, L.J., Tjandra, S.Y., Parkoff, S.R., Gupta, A., Kulic, B., Julius, A., Bianchi, R., et al.: An empirically based path loss model for wireless channels in suburban environments. *IEEE J. Sel. Areas Commun.* **17**(7), 1205–1211 (1999)
7. Giambene, G., Yahiya, T.: LTE planning for soft frequency reuse. In: *Wireless Days (WD), 2013 IFIP*. pp. 1–7 (2013)
8. Giambene, G., Le, V.A.: Performance evaluation of different fractional frequency reuse schemes for lte. In: *Euro Med Telco Conference (EMTC), 2014*, pp. 1–6. IEEE (2014)
9. Giambene, G., Yahiya, T.A., Grochla, K., Polys, K., et al.: Resource management and cell planning in lte systems. In: *Wireless Networking for Moving Objects*, pp. 177–197. Springer (2014)
10. Necker, M.C.: Coordinated fractional frequency reuse. In: *Proceedings of the 10th ACM Symposium on Modeling, Analysis, and Simulation of Wireless and Mobile Systems*, pp. 296–305. ACM (2007)
11. Novlan, T., Andrews, J.G., Sohn, I., Ganti, R.K., Ghosh, A.: Comparison of fractional frequency reuse approaches in the ofdma cellular downlink. In: *Global Telecommunications Conference (GLOBECOM 2010), 2010 IEEE*, pp. 1–5. IEEE (2010)
12. Polys, K., Grochla, K.: Mobility robustness in lte based on automatic neighbor relation table. In: *Computer Networks*. Springer International Publishing (2015)
13. PyLTES. <https://github.com/iitis/PyLTES>. Accessed 16 April 2015
14. Qian, M., Hardjawana, W., Li, Y., Vucetic, B., Shi, J., Yang, X.: Inter-cell interference coordination through adaptive soft frequency reuse in lte networks. In: *Wireless Communications and Networking Conference (WCNC), 2012 IEEE*, pp. 1618–1623. IEEE (2012)
15. Rose, D.M., Jansen, T., Werthmann, T., Türke, U., Kürner, T.: The ic 1004 urban hannover scenario-3d pathloss predictions and realistic traffic and mobility patterns. *European Cooperation in the Field of Scientific and Technical Research, COST IC1004 TD (13)*, 8054 (2013)
16. Sun, S., Gao, Q., Peng, Y., Wang, Y., Song, L.: Interference management through comp in 3g pp lte-advanced networks. *IEEE Wireless Commun* **20**(1), 59–66 (2013)
17. Xu, Z., Li, G.Y., Yang, C., Zhu, X.: Throughput and optimal threshold for ffr schemes in ofdma cellular networks. *IEEE Trans. Wireless Commun.* **11**(8), 2776–2785 (2012)
18. Yu, X., Gen, M.: *Introduction to Evolutionary Algorithms*. Springer Science & Business Media (2010)
19. Yu, Y., Dutkiewicz, E., Huang, X., Mueck, M., Fang, G.: Performance analysis of soft frequency reuse for inter-cell interference coordination in lte networks. In: *Communications and Information Technologies (ISCIT), 2010 International Symposium on*. pp. 504–509. IEEE (2010)



# Erratum to: Information Sciences and Systems 2015

Omer H. Abdelrahman, Erol Gelenbe, Gokce Gorbil  
and Ricardo Lent

**Erratum to:**  
**O.H. Abdelrahman et al. (eds.), *Information Sciences  
and Systems 2015*, Lecture Notes in Electrical Engineering  
355, DOI [10.1007/978-3-319-22635-4](https://doi.org/10.1007/978-3-319-22635-4)**

The original version of this book was inadvertently published with incorrect volume number 355. The correct volume number is 363.

---

The online version of the original book can be found under  
DOI [10.1007/978-3-319-22635-4](https://doi.org/10.1007/978-3-319-22635-4)

---

O.H. Abdelrahman (✉) · E. Gelenbe · G. Gorbil  
Department of Electrical and Electronic Engineering, Imperial College, London, UK  
e-mail: o.abd06@imperial.ac.uk

R. Lent  
Department of Engineering Technology, University of Houston, Houston, TX, USA

© Springer International Publishing Switzerland 2016  
O.H. Abdelrahman et al. (eds.), *Information Sciences and Systems 2015*,  
Lecture Notes in Electrical Engineering 363,  
DOI [10.1007/978-3-319-22635-4\\_42](https://doi.org/10.1007/978-3-319-22635-4_42)

E1

# Author Index

## A

Abdelrahman, Omer H., 45, 115  
Aldogan, Deniz, 359  
Ali, Ahsan, 69  
Altop, Duygu Karaođlan, 93  
Argiolas, Carlo, 393  
Ataee, Mehdi-pour Shahin, 191  
Ateş, Hüseyin, 403  
Avcu, Neslihan, 261  
Azar, Saeed Nourizadeh, 417

## B

Barth, D., 319  
Bay, Yasemin, 371  
Bayram, Zeki, 191  
Beyaz, Muhammed, 209  
Birturk, Aysenur Akyuz, 159  
Bougueroua, S., 319

## C

Çađlayan, Mehmet Ufuk, 3  
Çandır, Şemsinur, 339  
Celebi, Erbug, 371  
Cerotti, D., 81  
Cerqueira, Fabio Ribeiro, 307  
Chhorn, Sok, 429  
Cho, Choong-ho, 429  
Çiçekli, Nihan, 349  
Cicekli, Ilyas, 381

Cosar, Ahmet, 209  
Couto, Adriano Donato, 307  
Czachórski, Tadeusz, 251

## D

Danisevskis, J., 147  
Demirel, Hasan, 275  
Dessi, Nicoletta, 393  
Dimitriou, Nikos, 115  
Dokeroglu, Tansel, 209  
Drosou, Anastasios, 115, 199

## E

Ehsani, Razieh, 339

## F

Filis, Konstantinos, 137  
Fourneau, J.M., 223, 241  
Francois, Frederic, 45  
Fugini, Maria Grazia, 393

## G

Gaigeot, M.-P., 319  
Gelenbe, Erol, 35, 45, 105  
Görbil, Gökçe, 105  
Gorbil, Gokce, 115  
Görgün, Onur, 339

Gribaudo, M., 81  
Grochla, Krzysztof, 439, 447  
Günay, Asuman, 295  
Gungor, Tunga, 329  
Gurcan, Fatih, 159

**H**

Hierons, Robert M., 171

**I**

Işıklar, Yunus Emre, 349

**K**

Kadioglu, Yasin Murat, 35  
Kang, Byungseok, 233  
Karaagacli, Kaan, 417  
Kiesling, Elmar, 19  
Kilingç, İsmail, 403  
Kim, Seung-yeon, 429  
Korkmaz, Hüseyin, 403  
Kutlu, Mucahit, 381

**L**

Levi, Albert, 93  
Lyberopoulos, George, 137

**M**

Mautor, T., 223  
McOwan, Peter W., 275  
Mesogiti, Ioanna, 137  
Morozov, Evsey, 57

**N**

Nabiyev, Vasif, 127  
Nabiyev, Vasif V., 295  
Neuhold, Erich, 19  
Nordholz, J., 147  
Nycz, Monika, 251  
Nycz, Tomasz, 251

**O**

Ozgun, Levent, 329  
Özhorasan, Bülent, 403  
Ozkasap, Oznur, 69, 417

**P**

Papadopoulos, Stavros, 115, 199  
Pavloski, Mihajlo, 105  
Pekergin, N., 223  
Pes, Barbara, 393  
Peter, M., 147  
Petschick, M., 147  
Piazzolla, P., 81  
Pincioli, R., 81  
Połys, Konrad, 447  
Pulur, Naim Alperen, 93

**Q**

Quessette, F., 319

**R**

Rumyantsev, Alexander, 57

**S**

Salaht, F. Ait, 223  
Seifert, J.-P., 147  
Seo, Si-o, 429  
Serazzi, G., 81  
Slabicki, Mariusz, 439  
Solak, Ercan, 339  
Song, Ji-eun, 429  
Soyel, Hamit, 275  
Spezia, R., 319

**T**

Taleb, H. Castel, 223  
Theodoropoulou, Helen, 137  
Türker, Uraz Cengiz, 171  
Tzovaras, Dimitrios, 115, 199

**U**

Ulutas, Guzin, [127](#)

Ulutas, Mustafa, [127](#)

Ustubioglu, Arda, [127](#)

Ustubioglu, Beste, [127](#)

**V**

Vetter, J., [147](#)

Vial, S., [319](#)

**Y**

Yaslan, Yusuf, [359](#)

Yildiz, Olcay Taner, [339](#)

Yin, Yonghua, [181](#)

Yoon, Suk-ho, [429](#)

Yurtkan, Kamil, [275](#)

**Z**

Zeugmann, Thomas, [285](#)

Zhu, Yu, [285](#)