Smart Innovation, Systems and Technologies 332

Roumen Kountchev
Rumen Mironov
Kazumi Nakamatsu *Editors*

New Approaches for Multidimensional Signal Processing

Proceedings of International Workshop, NAMSP 2022

KES International

Springer

Smart Innovation, Systems and Technologies

Volume 332

The Smart Innovation, Systems and Technologies book series encompasses the topics of knowledge, intelligence, innovation and sustainability. The aim of the series is to make available a platform for the publication of books on all aspects of single and multi-disciplinary research on these themes in order to make the latest results available in a readily-accessible form. Volumes on interdisciplinary research combining two or more of these areas is particularly sought.

The series covers systems and paradigms that employ knowledge and intelligence in a broad sense. Its scope is systems having embedded knowledge and intelligence, which may be applied to the solution of world problems in industry, the environment and the community. It also focusses on the knowledge-transfer methodologies and innovation strategies employed to make this happen effectively. The combination of intelligent systems tools and a broad range of applications introduces a need for a synergy of disciplines from science, technology, business and the humanities. The series will include conference proceedings, edited collections, monographs, handbooks, reference books, and other relevant types of book in areas of science and technology where smart systems and technologies can offer innovative solutions.

High quality content is an essential feature for all book proposals accepted for the series. It is expected that editors of all accepted volumes will ensure that contributions are subjected to an appropriate level of reviewing process and adhere to KES quality principles.

Indexed by SCOPUS, EI Compendex, INSPEC, WTI Frankfurt eG, zbMATH, Japanese Science and Technology Agency (JST), SCImago, DBLP.

All books published in the series are submitted for consideration in Web of Science.

Roumen Kountchev · Rumen Mironov ·
Kazumi Nakamatsu
Editors

New Approaches for Multidimensional Signal Processing

Proceedings of International Workshop, NAMSP 2022

Springer

Editors
Roumen Kountchev
Technical University of Sofia
Sofia, Bulgaria

Rumen Mironov
Technical University of Sofia
Sofia, Bulgaria

Kazumi Nakamatsu
University of Hyogo
Kobe, Japan

ISSN 2190-3018 ISSN 2190-3026 (electronic)
Smart Innovation, Systems and Technologies
ISBN 978-981-19-7841-8 ISBN 978-981-19-7842-5 (eBook)
https://doi.org/10.1007/978-981-19-7842-5

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Organizing Committee

Honorary Chair

Prof. Lakhmi C. Jain KES International, Selby, UK

General Chair

Prof. Roumen Kountchev Technical University of Sofia, Sofia, Bulgaria

General Co-chair

Prof. Srikanta Patnaik SOA University, Bhubaneswar, India

Co-chairs

Prof. Rumen Mironov Technical University of Sofia, Sofia, Bulgaria
Prof. Parvinder Singh Deenbandhu Chhotu Ram University of Science and Technology, Murthal, Haryana, India

Chair Members

Silai Zhou	Founder of IRnet International Academic Communication Center, China
Bin Hu	Co-Founder of IRnet International Academic Communication Center, China

Publicity Chair

SR Dr. Roumiana Kountcheva	T&K Engineering, Sofia, Bulgaria

International Program Committee

Prof. K. Rao	University of Texas at Arlington, USA
Prof. K. Nakamatsu	University of Hyogo, Japan
Prof. M. Milanova	University of Arkansas at Little Rock, USA
Prof. A. Salem	Ain Shams University, Egypt
Prof. B. Iantovics	University of Medicine, Pharmacy, Sciences and Technology of Targu Mures, Romania
Prof. K. Kpalma	INSA de Rennes, France
Prof. J. Ronsin	INSA de Rennes, France
Prof. I. Kralov	Technical University of Sofia, Bulgaria
Prof. P. Kervalishvili	Georgian Technical University, Georgia
Prof. Yo-Sung Ho	Gwangju Institute of Science and Technology, South Korea
Prof. M. Favorskaya	Siberian State Aerospace University, Russian Federation
Prof. Badrul Khan	Virginia International University, USA
Prof. P. Koprinkova-Hristova	Bulgarian Academy of Sciences, Bulgaria
SR Dr. R. Kountcheva	T&K Engineering, Sofia, Bulgaria
Prof. V. Georgieva	Technical University of Sofia, Bulgaria
Prof. Jair Minoro Abe	University of Sao Paulo, Brazil
Prof. A. Elmaghraby	University of Louisville, USA
Prof. I. Draganov	Technical University of Sofia, Bulgaria
Prof. Omer Jasim	University of Fallujah, Iraq
Prof. A. Bekiarsky	Technical University of Sofia, Bulgaria
Prof. H. Chouiyakh	Université Internationale de Rabat, Morocco
Prof. Pl. Pavlov	Technical University of Sofia, Bulgaria
Prof. S. Bekiarska	Technical University of Sofia, Bulgaria
Prof. A. Manolova	Technical University of Sofia, Bulgaria

Prof. St. Rubin	Space and Naval Warfare Systems Center, San Diego, USA
Prof. Nasreddine Taleb	Djillali Liabes University of Sidi Bel Abbes, Algeria
Prof. N. Hikal	Mansoura University, Egypt
Prof. S. Nagy	Széchenyi István University, Gyor, Hungary
Prof. Bratislav Milovanović	University of Niš, Serbia
Prof. Z. Bojković	University of Belgrade, Serbia
Prof. Mihai Talmaciu	University of Bacau
Prof. Elena Nechita	University of Bacau
Prof. Anna Saro Vijenran	Bharathiar University, Coimbatore, India
Prof. Turki Obaidat	Al-Zaytoonah University of Jordan
Prof. L. Yaroslavsky	Tel Aviv University, Israel
Prof. I. Iliev	Technical University of Sofia, Bulgaria
Prof. V. Balyan	Cape Peninsula University of Technology, Capetown, South Africa
Prof. Atanas Gotchev	Tampere University of Technology, Finland
Prof. Boris Mirkin	Higher School of Economics University, Moscow, Russian Federation

Preface

This book contains papers presented at the International Workshop "New Approaches for Multidimensional Signal Processing-NAMSP 2022" which was carried out during July 07–09, 2022 at the Technical University of Sofia, Bulgaria. The workshop was also part of the "Days of Science" at the Technical University of Sofia, organized in collaboration with the Research and Development Sector at TU-Sofia, Bulgaria. The workshop was supported by the Bulgarian National Science Fund (BNSF) and the Ministry of Education and Science of Bulgaria. Co-organizers of NAMSP 2022 were Interscience Research Network (IRNet) International Academy Communication Center, China; Deenbandhu Chhotu Ram University of Science and Technology, Murthal, Haryana, India, and Interscience Institute of Management and Technology-Bhubaneswar, India. In the workshop, authors participated from China, India, Romania, USA, Egypt, Brazil, Japan, Hungary, and Bulgaria.

The book comprises 23 chapters, divided into the following three groups: Multidimensional Signal Processing (seven chapters), Applications of Multidimensional Signal Processing (eight chapters), and Applications of Blockchain and Network Technologies (eight chapters). The aim of the book is to present the latest

achievements of the authors in the processing and analysis of multidimensional signals and the related applications, to a wide range of readers: IT specialists, engineers, physicians, Ph.D. students, and other specialists.

Sofia, Bulgaria — Roumen Kountchev
Sofia, Bulgaria — Rumen Mironov
Kobe, Japan — Kazumi Nakamatsu
August 2022

Acknowledgments Special thanks to Prof. Lakhmi Jain (Honorary Chair), Prof. Roumen Kountchev (General Chair), Prof. Srikanta Patnaik (General Chair), Prof. Parvinder Singh (Co-chair), Prof. Rumen Mironov (Co-chair), Prof. Mariofanna Moilanova (Plenary speaker), Prof. Laslo Barna Iantovics (Plenary speaker), and Prof. Srikanta Patnaik (Plenary speaker) of the workshop NAMSP 2022. The book editors express their special thanks to book chapter reviewers: Prof. Ivo Draganov, Prof. Mariofanna Milanova, Prof. R. Mironov, Bin Hu, Silai Zhou, and Tian Zhou for their efforts, help, and support for the successful preparation of the book. The editors express their warmest thanks to the excellent Springer team which made this book possible.

Contents

About the Editors

Prof. Roumen Kountchev received his M.Sc. and D.Sc. at the Technical University of Sofia, Bulgaria, and his Ph.D. at LEIS, St. Petersburg, Russia. His scientific areas of interest are digital signal and image processing, image compression, multimedia watermarking, video communications, pattern recognition and neural networks. He has 348 papers published in magazines and conference proceedings, 30 books, 43 book chapters and 19 patents (four international). He had been Principle Investigator of 38 research projects. He is Member of Euro Mediterranean Academy of Arts and Sciences (EMAAS) and President of Bulgarian Association for Pattern Recognition (Member of IAPR); Editorial Board Member of IJBST Journal Group; Member of International Research Institute for Economics and Management (IRIEM), Member of Institute of Data Science and Artificial Intelligence (IDSAI), Member of the Honorable Editorial Board of the nonprofit peer-reviewed open access IJBST Journal Group; Editorial Board Member of *International Journal of Reasoning-based Intelligent Systems*; *International Journal Broad Research in Artificial Intelligence and Neuroscience*; *KES Focus Group on Intelligent Decision Technologies*; *Egyptian Computer Science Journal*; *International Journal of Bio-Medical Informatics and e-Health*; and *International Journal Intelligent Decision Technologies*. He is Editor of books in Springer SIST series.

Prof. Rumen Mironov received his M.Sc. and Ph.D. in Telecommunications from Technical University of Sofia and M.Sc. in Applied Mathematics and Informatics from Faculty of Applied Mathematics and Informatics. He is currently Head of the Department of Radio Communications and Video Technologies at Technical University of Sofia, Bulgaria. His current research focuses on digital signal and image processing, pattern recognition, audio and video communications, information systems, computer graphics and programming languages. He is Member of Bulgarian Association of Pattern Recognition (IAPR) and of Bulgarian Union of Automation and Automation Systems. He is also Editorial Board Member of IJBST Journal Group. He is Author of more than 70 scientific publications and is Co-editor of books published in Springer SIST series.

Dr. Kazumi Nakamatsu received the Ms. Eng. and Dr. Sci. from Shizuoka University and Kyushu University, Japan, respectively. His research interests encompass various kinds of logic and their applications to Computer Science, especially paraconsistent annotated logic programs and their applications. He has developed some paraconsistent annotated logic programs called ALPSN (Annotated Logic Program with Strong Negation), VALPSN (Vector ALPSN), EVALPSN (Extended VALPSN) and bf-EVALPSN (before–after EVALPSN) recently and applied them to various intelligent systems such as a safety verification-based railway interlocking control system and process order control. He is Author of over 150 papers and 20 book chapters and 10 edited books published by prominent publishers. He has served as Editor-in-Chief of the International Journal of Reasoning-based Intelligent Systems (IJRIS) and is now Founding Editor of IJRIS and Editorial Board Member of many international journals. He has contributed numerous invited lectures at international workshops, conferences and academic organizations and is ACM Member.

Multidimensional Signal Processing

Multidimensional Signal Processing and Applications—New Approaches

Roumen Kountchev[1(✉)], Rumen Mironov[1], Ivo Draganov[1], and Roumiana Kountcheva[2]

[1] Technical University of Sofia, 8 Kliment Ohridski Blvd., 1756 Sofia, Bulgaria
{rkountch,rmironov,idraganov}@tu-sofia.bg
[2] TK Engineering, Sofia, Bulgaria

Abstract. This chapter is a short introduction in the contemporary approaches aimed at the multidimensional processing and analysis of various kinds of signals, investigated in related research works, which were presented at the Third International Workshop "New Approaches for Multidimensional Signal Processing", (NAMSP), held at the Technical University of Sofia, Bulgaria in July 2022. Some of the works cover various topics, as: moving objects tracking in video sequences, automatic audio classification, representation of color video чрез 2-level tensor spectrum pyramid, etc., and also introduce multiple applications of the kind: analysis of electromyography signals, diagnostics of COVID based on ECG, etc. Short descriptions are given for the main themes covered by the book, which comprises the following three sections: multidimensional signal processing; applications of multidimensional signal processing, and applications of blockchain and network technologies.

Keywords: Multidimensional signal processing · Tensor decomposition · Deep learning · Blockchain · Medical decision support · Audio classification · Paraconsistent logic · Deep faith network model · Ant colony

1 Introduction

In contemporary life, multidimensional signal processing covers all signal processing performed in complicated systems and networks in manufacturing and development. As far as multidimensional signal processing is a subset of signal processing, it has its own specifics in sense that it deals specifically with data of more than one dimension [1]. Typical examples in this area are the multi-sensor radar information, medical imagery, surveillance systems, big data databases, satellite multi-view and multispectral images, etc., which use multiple sensors to sample signals and form images by manipulating the obtained signals [2, 3]. These specifics imply the use of more complex algorithms, compared to 1D and 2D cases. As a result, one of the big topics of the present day is the development of new efficient methods and algorithms which utilize the natural high correlation and information redundancy existing in multidimensional signals. Together with the application area widening, the number of various new techniques and theories,

R. Kountchev et al. (Eds.): *NAMSP 2022*, SIST 332, vol. 332,, pp. 3–9, 2023.
https://doi.org/10.1007/978-981-19-7842-5_1

aimed at the efficient processing and analysis of such information, is increasing significantly, covering more and more areas and applications. One of the basic mathematical forms used for the efficient representation of multidimensional signals and data, are multidimensional tensors [4, 5]. By using various decomposition methods, each multidimensional tensor could be represented as a weighted sum of first-order coefficients (tensors), arranged in accordance with their weight lessening, after which the small members are neglected. In result, the multidimensional signal processing is transformed to processing of a limited number of first-order tensors, whose weights are higher than a preset threshold value, imposed by the application. Combined with the methods for NN [6] deep learning, tensor decomposition methods became a powerful tool for efficient analysis and processing of multidimensional signals and data in systems for medical decision support [7, 8], object and situation recognition [9, 10], etc. Based on this, and other related approaches, together with original ideas and implementations, significant number of applications and solutions are presented in this book, in such areas as homogeneous and heterogeneous protocols to maximize the lifetime of wireless sensor network for precision agriculture, smart logistic warehouse management with IOT, development of intelligent traction system for bulk cargo terminal and many others.

2 Main Topics Presented at NAMSP'22

In this book, some new aspects are presented in the processing of multidimensional signals of various kind. The book chapters are divided into three sections: (1) Multidimensional signal processing; (2) Applications of multidimensional signal processing; and (3) Applications of blockchain and network technologies.

Section 1 comprises 7 chapters, including this one. In chapter "Video Tracing of Moving Objects by Fusing Three-Term Decompositions", new algorithm is offered for moving objects tracking in video sequences, based on three-term decompositions. The analyzed input video is modelled as a 3-way tensor and over it are applied separately: 3-way-decomposition, Motion-assisted matrix restoration, Robust motion-assisted matrix restoration and Alternating direction method of multipliers. The results obtained for the moving objects detection make the proposed algorithm applicable in fields such as video surveillance, vehicle traffic control, crowd monitoring and others. In chapter "Deep Learning Approaches for Classroom Audio Classification Using Mel Spectrograms", new approach for automatic audio classification of classroom activities, aimed to support the training in science, technology, engineering, and mathematics (STEM), is proposed. For this, various approaches for deep learning are used for classroom audio classification (such as for example, the convolution neural network, and the long short-term memory (LSTM)). In the presented research, the models of three kinds of class activities are evaluated, defined as "single voice", "multiple voice" and "no voice", based on classroom recording. The model is trained on the Mel spectrogram extracted from the recorded audio. Highest accuracy of 97% is acknowledged for the 2D Convolutional Neural Network. In chapter "Tensor Spectral Pyramid for Color Video Sequences Representation, Based on 3D FO-AHKLT", new structure is introduced for representation of video sequence of RGB frames. For this, the RGB components are first transformed into tensors, which after unfolding are represented in the reduced vector space. To realize this,

2-level tensor spectrum pyramid (2LTSP) is used, based on the 3D frequency-ordered adaptive hierarchical KLT (3D FO-AHKLT). The new structure ensures high energy concentration of the transformed tensor into small number of spectrum coefficients, whose calculation needs operations of low computational complexity. The 3D FO-AHKLT is executed in three sequential stages and in each, the input tensor is transformed into a vector with different spatial orientation: horizontal, vertical or lateral (in time). The analysis of the computational cost of 2LTSP and its comparison with the closest decompositions H-Tucker and Tensor Train shows that it decreases together with the vector size growth. Also, there is proved that 2LTSP based on the 3D adaptive fast Walsh–Hadamard transform additionally reduces the needed calculations number. The main advantages of the new structure are its flexibility in respect of the reduced coefficients number in the sequential decomposition levels, the lack of iterative calculations and the low computational complexity. In chapter "Electromyography Signal Acquisition, Processing, Optimization and Its Applications" the contemporary approaches for features extraction and analysis of Electromyography Signals (EMG) are discussed, aimed at their efficient processing, and classification. The approach is based on the development of optimization methods so that to achieve high accuracy in detection of stress and anomalies in biomechanics and movement. In chapter "Research on the Radar Signal Classification Method Based on the Deep Faith Network Model", a method is introduced for classification of radar signals based on the Deep Faith Network Model. Combined with the basic concept analysis of deep learning with multi-layer structure of the hidden layer, this approach is aimed at the detection of the hidden features of data in the construction of multi-layer neural network model. As a result of the investigation, a reliable method for radar signal classification is developed. In chapter "Measuring Machine Intelligence Using Black-Box-Based Universal Intelligence Metrics", the metrics for evaluation of the complicated artificial intelligent (AI) systems intelligence regarding their ability to solve difficult problems are studied, and are presented reliable and accurate intelligence comparisons between such systems. The most important property of the intelligence metrics is its universality imposed by the diversity of the already developed complex intelligent systems. In the research are presented black-box-based intelligence metrics, which can also treat and evaluate such aspects as intelligence changeability and extreme intelligence (manifestation of extremely low or high intelligence in various situations). The offered universal black-box-based machine intelligence metrics is a new tool aimed to support the intelligent systems developers, because it is not influenced by the existing architecture diversity.

Section 2 comprises 8 chapters. At the top of the list (chapter "COVID Detection Using ECG Image Reports: A Survey"), is a survey of a group of research works dedicated to the opportunity for diagnostics of COVID disease by means of electrocardiogram (ECG) image reports. The idea is to detect symptoms and changes in the ECG peaks and intervals. In addition are presented another techniques also used for ECG reports analysis and aimed at the COVID detection. The main limitations and future trends of the investigation are illustrated and outlined very well and is shown that the ECG image report is close to the acceptable level in the COVID disease detection. The conclusion is that ECG reports analysis ensures high efficiency, which in many cases surpasses other similar protocols. In the next chapter "Disease Detection Techniques in Plants: Transition

from Manual to Automation" are analyzed methods for plants disease detection, based on the use of neural networks with deep learning, as AlexNet, DenseNet, ResNet, VGG and GoogLeNet CNN. The results obtained through modeling of these architectures with various datasets show that the last-mentioned CNN ensures accuracy of 99% for plant disease detection. In chapter "On Applying Gradient Based Thresholding on the Canny Edge Detection Results to Improve the Effectiveness of Fuzzy Hough Transform for Colonoscopy Polyp Detection Purposes", the possibilities for efficiency enhancement of the fuzzy Hough transform aimed at the detection of polyps in colonoscopy images, preprocessed with Canny operator, are investigated. The presented algorithm was tested for three open-access databases of colonoscopy images. For this, the authors first define a general threshold for the gradients classification. The modeling results show that it is not possible to define a normalized database-independent threshold which to be used to eliminate the useless edges in the images. Chapter "Development of IoT Indoor Monitoring System for Independent Elderly" is pointed at the need for the development of intelligent monitoring system at home, which to support lonely old people. In this work, an IoT Indoor monitoring system for the independent elderly which allows families to feel at ease even in remote locations, is presented. In chapter "Improving the Process of Evaluating User Stories Using the Paraconsistent Annotated Evidential Logic Eτ", a model for validation of user story (technique used in agile methods to elicit requirements) using the Para-analyzer algorithm, based on the Paraconsistent Annotated Evidential Logic Eτ, is introduced. For this, an investigation based on flexible methods was carried out with a team of researchers. The presented model utilizes the favorable and contrary proofs for each criterion INVEST as input variables. The implemented mathematical model supports decision making and serves as an instrument assisting teams, product owners, managers, etc. In the experimental part, four user stories were analyzed by nine experts, which evaluated the criteria for each of them. The interpretation of the experts evaluations was done on the basis of global analysis in the unit square of the Cartesian plane, where were shown the degrees of favourable/contrary evidence detected. The introduced research demonstrates the ability of the Evidence Noted Paraconsistent Logic Eτ concepts to validate and present perspectives on dealing with situations of uncertainty and inconsistency. In chapter "Neural Network Algorithm Applied in Electrical Engineering Automation", on the basis of the advantages of AI technology applied in the automation of electric equipment diagnostics, is empirically investigates new algorithm based on RBF neural network. The final simulation results show that the algorithm not only permits to identify quantitatively the equipment characteristics, but also—to improve the condition evaluation. The presented application algorithm can accurately identify the feature quantity of the state in the simulation experiment, which proves that it can also complete the learning in judging the new state types. In chapter "Truss Structure Optimization Design Based on FE-PSO-SQP Algorithm", one new Particle Swarm Optimization (PSO) algorithm is offered, in which are overcome the problems related to its low computational accuracy, relatively slow convergence and bad population varieties in the following calculation of the model. To overcome all this, and also—to solve better the problem with the structure optimization, is developed the algorithm FE-PSO-SQP, which combines PSO with the algorithms for Sequential Quadratic Programming (SQP) and Finite Element (FE). A set of programs was developed for the ANSYS software.

For the case when the independent program is used for the simulation calculation of the truss structure optimization, obtained results show that the algorithm FE-PSO-SQP has faster convergence and higher computation accuracy, compared to these of the basic algorithm FE-PSO. In chapter "Practice System of Ant Colony Optimization Algorithm in Business Administration", new approach for business administration optimization is offered, which can also improve the practical work efficiency and reduce the work pressure faced by the industrial and commercial departments. The presented idea is based on the PSO algorithm combined with Ant Colony Optimization (ACO). The new approach was tested experimentally and the obtained results confirm the suggested achievements.

Section 3 comprises 8 chapters. The first one there (chapter "Literature Review of Smart Contracts Using Blockchain Technology") is a survey on the use of Blockchain technologies in various areas, as cryptocurrency, smart contracts, etc., which is a result of its decentralised and digitally distributed peer-to-peer network featuring elevated speed, efficiency and security. Nowadays, Blockchain offers contemporary solutions for contracts management by consensus mechanism, scalability and reliability on 'off-chain' resources. The presented work is focused on rectifying bugs and cybersecurity attacks, such as (for example) re-entrancy attacks with utilization of vaults, GHOST protocol, Bitcoin-NG, botnet C&C command, ERLAY protocol, bug prevention tools like Oyente and SolidiFi and fuzzing tools like ReGuard and Contract-Fuzzer. Chapter "A Comprehensive Study of 5th Generation Scheduling Algorithms" is a survey on the the key points of 5th Generation (5G) communication technology, the basic concepts of the 5G network concepts, its architecture for both standalone and non-stand-alone mode, different existing scheduling algorithms used in 5G, namely Proportional Fair (PF), Modified Least Weighted Delay First (MLWDF), Exponential Proportional Fair (EX-PF), Frame Level Schedular (FLS), Round Robin (RR) and also—a detailed review on various scheduling algorithms anticipated by great number of researchers. This work points out scheduling areas in which improvement is required and the analysis can provide better results in the form of a new strategy. Chapter "A Comparative Analysis of Homogeneous and Heterogeneous Protocols to Maximize the Lifetime of Wireless Sensor Network for Precision Agriculture" is focused on the Wireless Sensor Networks (WSN) which contain big number of battery-powered Sensor Nodes (SN). As a result of the fact that SN are battery-operated and the power source cannot be replaced, is implied the objective to maximize the network lifetime by choosing some low energy path using energy-efficient protocols, so that at the end, the path which needs lowest energy, to be detected. Multiple use of a low-energy part can result in battery discharge for SN in the area and as a sequence—to the creation of energy holes, because all SN in the neighborhood are "dead". The aim is to choose a protocol which to prolong the WSN lifetime through including all SN in the process. The approach is based on the proposed Heterogeneous Energy Efficient and Reliable Routing (HEERR), which is an advanced version of the Distributed Energy Efficient Clustering (DEEC) protocol. The comparison with other hierarchical routing techniques proves that HEERR not only enhances the network lifetime but also increases the throughput. Besides, the authors cleared that the heterogeneous approach is more reliable and energy efficient than the homogeneous approach. Chapter "The Integration Development and Upgrading Path of Industry 4.0 Architecture Industrial Engineering Network Driven by Big Data" is devoted to studies in understanding big background, on

the basis of data driven according to the basic connotation of industrial engineering network integration. The aim is to clear the basic route of industrial architecture engineering network integration, and carry on the empirical research about the mode selection, so that to enhance the competitive advantage of industrial products and upgrade new forms. In chapter "SAAS Application Prospect Analysis in Hrm and Methods to Upgrade the Contemporary System", methods for inventive problem solving to analyze potential new approaches are used, based on preceding scientific works in the field of "Software as a Service". In chapter "All Digital Phase Locked Loop (ADPLL) and Its Blocks—A Comprehensive Knowledge" the systems All Digital Phase Locked Loop (ADPLL) are analyzed, used to synchronize the signals frequency and phase. The investigated ADPLL implements all basic PLL blocks in a digital form. Besides, in this chapter is also presented a detailed search on each block of the ADPLL architecture and the changes in these blocks, introduced together as the research advanced. The best ADPLL solutions are compared in terms of their parameters number, and the implemented techniques. In chapter "On Realization of Smart Logistic Warehouse Management with Internet of Things", an architecture of intelligent warehouse system is offered. The IOT-based system successfully manages the complicated supply chain network by combining cloud servers, Radio Frequency Identification (RFID), devices and sensors, tags, bar codes, mobility Wifi, surveillance and other smart utilities connected into an organic whole, which is equally applicable and feasible by the developers and business stakeholders. Cloud solutions are introduced in the IOT platform, aimed at the intelligent warehouse system mobility improvement, and together with this, to integrate multiple plug-play peripheral devices into one whole. Application Programming Interface (API), varying from Message Queuing Telemetry Transport (MQTT), Constrained Application Protocol (COAP) up to Lightweight Machine to Machine (LWM2M), is applicable and compatible with various plug-play devices, or new-developed contemporary applications. In the IOT and RFID business equipment environment, devices and processes create new efficiency, incomes and various possibilities for warehouse management and control. In chapter "Development and Design of Intelligent Traction System for Bulk Cargo Terminal", new kind of integrated unmanned intelligent traction loading system for bulk cargo terminal is proposed. The system combines the development of key technologies such as sliding contact power supply, vehicle number identification, decoupling control, continuous and accurate loading control and bridging control, and builds a high-quality communication network architecture to solve the defects of the traditional iron ox traction loading process. The presented study effectively improves the bulk cargo shipping capacity, reduces the potential safety hazards in the loading area, and achieves the expected effect of energy saving, cost reduction and efficiency increase in the overall operation process.

3 Conclusions

The research works, described in brief above, introduce the main contents of the studies, presented at the Third International Workshop on New Approaches for Multidimensional Signal Processing (NAMSP'2022), and held at the Technical University of Sofia, Sofia, Bulgaria in July 07–09, 2022. They cover part of the multidimensional signal processing, analysis and applications and outline new trends and ideas in this area.

References

1. Woods, J.: Multidimensional Signal, Image, and Video Processing and Coding, 2nd edn. Academic Press, Amsterdam (2012)
2. Ünsalan, C., Boyer, K.: Multispectral Satellite Image Understanding, From Land Classification to Building and Road Detection. Springer, London (2011)
3. Panagakis, Y., Kossaifi, J., Chrysos, G., Oldfield, J., Nicolaou, M., Anandkumar A., Zafeiriou, S.: Tensor Methods in Computer Vision and Deep Learning. In: Proceedings of the IEEE. arXiv:2107.03436v1 [cs.CV] (2021)
4. Miron, S., et al.: Tensor methods for multisensor signal processing. IET Signal Process., Inst. Eng. Technol. **14**(10), 693–709 (2021)
5. Aidini, A., Tsagkatakis, G., Tsakalides, P.: Tensor decomposition learning for compression of multidimensional signals. IEEE J. Sel. Top. Signal Process. **15**(3), 476–490 (2021)
6. Khan, A., Sohail, A., Zahoora, U., Qureshi, A.: A survey of the recent architectures of deep convolutional neural networks. In: Artificial Intelligence Review, vol. 53, pp. 5455–5516. Springer Nature (2020)
7. Suzuki, K., Chen, Y. (Eds.): Intelligent Systems Reference Library (ISRL). In: Artificial Intelligence in Decision Support Systems for Diagnosis in Medical Imaging, vol. 140. Springer, Cham (2018)
8. Cleophas, T., Zwinderman, A.: Machine Learning in Medicine—A Complete Overview. Springer Science + Business Media (2015)
9. Cyganek, B.: Object Detection and Recognition in Digital Images: Theory and Practice. John Wiley & Sons, Chichester (2013)
10. Aggarwal C.: Neural Networks and Deep Learning. A Textbook. Springer International AG (2018)

Video Tracing of Moving Objects by Fusing Three-Term Decompositions

Ivo Draganov(✉) and Rumen Mironov

Technical University of Sofia, 8 Kliment Ohridski Blvd., Sofia 1756, Bulgaria
{idraganov,rmironov}@tu-sofia.bg

Abstract. In this work, low-rank representation of video with the aim of background modelling and subtraction in order to trace moving objects is investigated based on three-term decompositions. The input video is modelled as a 3-way tensor and over it are applied separately 3-Way-Decomposition (3WD), Motion-Assisted Matrix Restoration (MAMR), Robust Motion-Assisted Matrix Restoration (RMAMR) and the Alternating Direction Method of Multipliers (ADMM). The results from detecting moving objects from the 2 most accurate algorithms (3WD and MAMR) are then combined on a frame basis in order to get more precise results. Two fusing techniques are applied using the logical OR and AND operations. The results are promising and render the proposed algorithms applicable in fields such as video surveillance, vehicle traffic control, crowd monitoring and others.

Keywords: Video tracing · Tensor · TTD · 3WD · MAMR · RMAMR · ADMM · Fusion

1 Introduction

Detecting moving objects in video is one of the key processing techniques in visual scenes analysis. It allows to isolate the objective of interest on a frame by frame basis by its location and further investigate its properties—dimensions, color, texture, mutual disposition to other entities in the video and others. This task is complicated by the presence of a complex background in the general case where its texture and coloration may vary a lot. In the same time the movement of the camera, if present, could cause additional complications in the determination of the gauges of the moving objects. There are numerous algorithms for detecting such objects and one of it is the three-term decomposition of the video with background modelling and subtraction, which is the primary focus in this study.

Zhang et al. [1] propose the use of Three-Term Decomposition (TTD) in combination of alternating directions method with proximal exchange for detecting vehicles in images from Video Synthetic Aperture Radar (VideoSAR). This technique is used within Dynamic Region of Interest (DROI). It allows spotting defocused vehicles. Competitive results to other contemporary algorithms are reported when using single-channel

R. Kountchev et al. (Eds.): *NAMSP 2022*, SIST 332, vol. 332,, pp. 10–22, 2023.
https://doi.org/10.1007/978-981-19-7842-5_2

single-pass system that employs 2D motion features from the polar format algorithm (PFA).

Another approach that incorporates TTD with Low-Rank and Sparse (LRS) representation is used to segment the sea-land coarse from a multisquint spaceborne SAR image [2]. This method relies on both sparse and error terms and the differences on a per couple basis of images are put in a stack and then averaged. A thresholding procedure follows at the end. Avoidance of false alarms for the observed territory is registered.

Detection of moving objects from satellite videos is achieved through low rank representations with improved qualities and together with structured and unstructured sparsity decompositions [3]. The moving objects are detected from a sparse foreground with additive nature using pixel-wise sparsity. The method relies on inexact alternating direction approach for finding the improved decomposition. Precision of the detection is evaluated by the authors as boosted, while there is a reference to extended video model as three-term composition.

VideoSAR analysis of subsequent images is used to spot moving targets with the help of machine learning methods in [4]. It incorporates a defocusing component working in Ku-band covering 360°. Ultra-low speeds are being investigated. In this implementation takes place also the low-rank and sparse decomposition in the form of three-term decomposition. The defocusing element is processed through single-channel single-pass case.

Low-rank and sparse decomposition finds its place in the area of exoplanet detection from directly taken images in a series when Angular Differential Imaging (ADI) is used [5]. Localized subspace projection appears to overcome the difficulties the Principal Component Analysis (PCA) has with non-Gaussian noises in the low subspaces. Detection of objects of interest prove to be in almost real-time mode. It is done by randomized low-rank approximations and thresholding stage which lead to low-rank, sparse and Gaussian noise. This way higher Signal to Noise Ratios (SNR) are achieved and the detection of the objects is more precise.

Object motion detection is much more difficult in the presence of turbulence, which affects the stability of the camera. Oreifej et al. [6] propose a method for turbulence mitigation based on the three-term decomposition which gives three outputs—background, turbulence and object. It is achieved through minimization of three norms—nuclear, Frobenius and 21. The second is discovered to isolate the Gaussian noise and the third—to isolate the object. Additional constraints on the model that represents the Gaussian turbulence could make more discoverable the linearly based motions of the objects.

Another application that makes use of the three-term decomposition is the extraction of moving blood vessels over the complex background of the body in angiography [7]. Hierarchical deep learning networks are used with sensory and control layers incorporating robust PCA to obtain the low-rank and sparse components. Long Short Term Memory (LSTM) plays the role of getting the unstructured areas with a random nature and further on projects them on orthogonal subspaces from where the vessels are being retrieved. Their heterogenous content helps at that stage to separate them from interferences. The Detection Rate (DR) achieved is 0.818 ± 0.078, the Precision is 0.865 ± 0.065 and the F-measure is 0.838 ± 0.055.

A novel three-term decomposition based on low-rank tensor representation for the structure tensor, sparse noise and Gaussian noise is proposed in [8]. It is extension of the existing techniques where 2 types of noises are considered. A dictionary is being built from the denoised tensor. As a last stage, the Alternating Direction Method of Multipliers (ADMM) optimizes the proposed algorithm. Denoising YUV videos by this algorithm yields Peak Signal to Noise Ratio (PSNR) as high as 45.66 dB and in the dominant number of cases it is higher than other decompositions, e.g. R-PCA, R-LRR, SNN, R-TPCA and R-TLRR.

Three-term decomposition finds its place in applications related to atmospheric imaging as well for objects motion detection [9]. Due to the non-uniform deteriorations caused by the atmospheric turbulence, this task is considerably difficult. It is solved by introducing a mask which eliminates the misdetections using a series of thresholding techniques.

Another area of usage of video decomposition is the foreground detection in water monitoring as described by Wu et al. [10]. It is found out that the Robust Principal Component Analysis (RPCA) copes not efficiently enough with the background which often changes in too complex patterns. The requirements for storage space are also challenging. The foreground spatial and temporal continuity are the means into separating it from the background and detecting the moving objects. They are used through the Online Robust Principal Component Analysis (OR-PCA). Three-term decomposition is tested as well in parallel to OR-PCA and the difference in Recall is 0.295, in Precision—0.884, and in F-measure—0.458 in favor to OR-PCA.

The main goal of the current study is to test 4 of the most popular implementations of the three-term decompositions—3-Way-Decomposition (3WD), Motion-Assisted Matrix Restoration (MAMR), Robust Motion-Assisted Matrix Restoration (RMAMR) and the Alternating Direction Method of Multipliers (ADMM). They are applied over test videos with varying content for moving objects detection and the two most accurate ones are embedded in 2 new fusion schemes that produce promising results into the quality of motion segmentation. The rest of the paper is organized as follows—in Sect. 2 the description of the 4 base decompositions is given, followed by experimental results in Sect. 3, and then there is a discussion in Sect. 4 with conclusions in the end in Sect. 5.

2 Algorithms Description

2.1 3-Way Decomposition (3WD)

The 3-Way Decomposition (3WD) could be defined, starting from the equation [6]:

$$\left| \begin{array}{l} \min\limits_{A,O,E} Rank(A), \;\; given \; F = A + O + E \\ \qquad\qquad \|O\|_0 \leq s, \quad \|E\|_F \leq \sigma \end{array} \right. , \tag{1}$$

where F is a matrix of frame, A—matrix of background, O—matrix of object, E—matrix of error, s—cumulative number of pixels, forming the moving object, σ—maximal variance of the noise, $\|.\|_0$—0-norm that counts the number of non-zero entries, $\|.\|_F$—Frobenius norm. Accounting for turbulent environment where the linear motion

of objects is quite distinct from the random movement of the background, the model from Eq. (1) could be modified in the following way [6]:

$$\left| \begin{array}{c} \min\limits_{A,O,E} Rank(A), + given\ F = A + O + E \\ \|\Pi(O)\|_0 \leq s, \quad \|E\|_F \leq \sigma \end{array} \right., \tag{2}$$

where Π is the object confidence map. Then in Lagrange form Eq. (2) could be expressed as [6]:

$$\left| \begin{array}{c} \min\limits_{A,O,E} Rank(A), + \tau\|\Pi(O)\|_0 + \lambda\|E\|_F^2 \\ F = A + O + E \end{array} \right., \tag{3}$$

where τ and λ are weights. After imposing certain relaxation conditions, the optimization problem comes to [6]:

$$\left| \begin{array}{c} \min\limits_{A,O,E} \|A\|_*, + \tau\|\Pi(O)\|_1 + \lambda E_F^2 \\ F = A + O + E \end{array} \right., \tag{4}$$

where $\|.\|_*$ is a nuclear norm, and the Augmented Lagrange Multiplier (ALM) is the approach that leads to the solution of Eq. (4) [6].

2.2 Motion-Assisted Matrix Restoration (MAMR)

In the base of the MAMR method lies the following optimization problem [11]:

$$\min_{\boldsymbol{B},\boldsymbol{F}} \|\boldsymbol{B}\|_* + \lambda\|\boldsymbol{F}\|_1, \ given\ \boldsymbol{W} \circ \boldsymbol{D} = \boldsymbol{W} \circ (\boldsymbol{B} + \boldsymbol{F}), \tag{5}$$

where $\boldsymbol{D}$ is the sequence of frames from the input video, $\boldsymbol{B}$—the background, $\boldsymbol{F}$—the foreground, $\boldsymbol{W}$—matrix of weights, $\|.\|_1$—matrix l_1 norm, $\circ$—multiplication of matrices on element-by-element principle. Following the matrix of weights generation using the information for motion in the video, Eq. (5) could be expressed also as [11]:

$$\min_{\boldsymbol{B},\boldsymbol{F}} \|\boldsymbol{B}\|_* + \lambda\|\boldsymbol{F}\|_1, \ given\ P_\Omega(\boldsymbol{D}) = P_\Omega(\boldsymbol{B} + \boldsymbol{F}), \tag{6}$$

where Ω is linear subspace of entries that definitely belong to the background, $P_\Omega(.)$—projection operator.

The last representation of the MAMR allows the application of the Alternating Direction Method-Augmented Lagrange Multiplier (ADM-ALM) into solving it [11]. It is done by minimizing the Augmented Lagrangian function, according to [11]:

$$\begin{aligned} L(\boldsymbol{B}, \boldsymbol{F}, \boldsymbol{Y}, \mu) = \|\boldsymbol{B}\|_* + \lambda\|\boldsymbol{F}\|_1 + \langle \boldsymbol{Y}, \boldsymbol{W} \circ (\boldsymbol{D} - \boldsymbol{B} - \boldsymbol{F})\rangle \\ + \frac{\mu}{2}\|\boldsymbol{W} \circ (\boldsymbol{D} - \boldsymbol{B} - \boldsymbol{F})\|_F^2, \end{aligned} \tag{7}$$

where $\mu > 0$ is a constant, $\boldsymbol{Y}$—Lagrange multiplier, $<.,.>$—inner product of matrices.

The solutions of $\boldsymbol{B}$, $\boldsymbol{F}$ and $\boldsymbol{Y}$ are obtained in sequential alternating manner [11]:

$$\begin{aligned} \boldsymbol{F}_{j+1} &= \underset{\boldsymbol{F}}{\operatorname{argmin}}\, \lambda\|\boldsymbol{F}\|_1 - \langle \boldsymbol{Y}_j, \boldsymbol{W} \circ \boldsymbol{F}\rangle + \tfrac{\mu_j}{2}\left\|\boldsymbol{W} \circ \left(\boldsymbol{D} - \boldsymbol{B}_j - \boldsymbol{F}\right)\right\|_F^2 \\ \boldsymbol{B}_{j+1} &= \underset{\boldsymbol{B}}{\operatorname{argmin}} \|\boldsymbol{B}\|_* - \langle \boldsymbol{Y}_j, \boldsymbol{W} \circ \boldsymbol{B}\rangle + \tfrac{\mu_j}{2}\left\|\boldsymbol{W} \circ \left(\boldsymbol{D} - \boldsymbol{B} - \boldsymbol{F}_{j+1}\right)\right\|_F^2 \cdot \\ \boldsymbol{Y}_{j+1} &= \boldsymbol{Y}_j + \mu_j \boldsymbol{W} \circ \left(\boldsymbol{D} - \boldsymbol{B}_{j+1} - \boldsymbol{F}_{j+1}\right), \mu_{j+1} = \rho\mu_j \end{aligned} \tag{8}$$

2.3 Robust Motion-Assisted Matrix Restoration (RMAMR)

Ye et al. [11] propose robust variant of the MAMR algorithm, in which the Frobenius norm substitutes the l_1 norm in order to handle better dense noise, when it is present all over the video frame. The optimization problem then could be summarized as [11]:

$$\min_{\boldsymbol{B},\boldsymbol{F},\boldsymbol{G}} \|\boldsymbol{B}\|_* + \lambda\|\boldsymbol{F}\|_1 + \gamma\|\boldsymbol{G}\|_F^2, \ \textit{given}\ \boldsymbol{W} \circ \boldsymbol{D} = \boldsymbol{W} \circ (\boldsymbol{B} + \boldsymbol{F} + \boldsymbol{G}), \tag{9}$$

where $\gamma > 0$ is a constant. The augmented Lagrangian of Eq. (9) is defined as [11]:

$$\begin{aligned} L(\boldsymbol{B}, \boldsymbol{F}, \boldsymbol{G}, \boldsymbol{Y}, \mu) = \|\boldsymbol{B}\|_* + \lambda\|\boldsymbol{F}\|_1 + \gamma\|\boldsymbol{G}\|_F^2 + \langle \boldsymbol{Y}, \boldsymbol{W} \circ (\boldsymbol{D} - \boldsymbol{B} - \boldsymbol{F} - \boldsymbol{G})\rangle \\ + \frac{\mu}{2}\|\boldsymbol{W} \circ (\boldsymbol{D} - \boldsymbol{B} - \boldsymbol{F} - \boldsymbol{G})\|_F^2, \end{aligned} \tag{10}$$

where $\boldsymbol{G}$ is the matrix, containing the error due to the presence of dense noise. Its solution could be obtained from the following equation [11]:

$$\boldsymbol{G}_{j+1} = \underset{\boldsymbol{G}}{\operatorname{argmin}}\, \gamma\|\boldsymbol{G}\|_F^2 - \langle \boldsymbol{Y}_j, \boldsymbol{W} \circ \boldsymbol{G}\rangle + \frac{\mu_j}{2}\left\|\boldsymbol{W} \circ (\boldsymbol{D} - \boldsymbol{B}_j - \boldsymbol{F}_{j+1} - \boldsymbol{G}\right\|_F^2 \tag{11}$$

and it is in closed form [11]:

$$\boldsymbol{G}_{j+1} = \frac{1}{\mu_j + 2\gamma}\left(\boldsymbol{Y}_j + \mu_j \boldsymbol{W} \circ \left(\boldsymbol{D} - \boldsymbol{B}_j - \boldsymbol{F}_{j+1}\right)\right). \tag{12}$$

2.4 Alternating Direction Method of Multipliers (ADMM)

ADMM is an algorithm dealing with the decomposition of dual ascents and having good convergence, which is intended to solve problems of the following kind [12]:

$$\min f(x) + g(z), \ \textit{given}\ Ax + Bz = c, \tag{13}$$

where $x \in \boldsymbol{R}^n$, $z \in \boldsymbol{R}^m$, $A \in \boldsymbol{R}^{pxn}$, $B \in \boldsymbol{R}^{pxm}$, $c \in \boldsymbol{R}^p$. The functions f and g are convex and the optimum of the so posed task in Eq. (13) could be expressed as [12]:

$$p^* = inf\{f(x) + g(z) | Ax + Bz = c\}. \tag{14}$$

Here, similar to some of the previous methods, it is necessary to find the augmented Lagrangian [12]:

$$L_p(x, z, y) = f(x) + g(z) + y^T(Ax + Bz - c) + \left(\frac{\rho}{2}\right)\|Ax + Bz - c\|_2^2. \tag{15}$$

Then, the algorithm needs to go through the following steps [12]:

$$\begin{cases} x^{k+1} = \underset{x}{\operatorname{argmin}} L_\rho\left(x, z^k, y^k\right) \\ z^{k+1} = \underset{z}{\operatorname{argmin}} L_\rho\left(x^{k+1}, z, y^k\right) \\ y^{k+1} = y^k + \rho\left(Ax^{k+1} + Bz^{k+1} - c\right) \end{cases}. \tag{16}$$

In Eq. (16) we have two minimizations—to x and z and one update operation at the end. The parameter ρ defines a step of the update and in the same time $\rho > 0$.

The iterations which are involved in the method of multipliers, applicable to Eq. (13) are the following [12]:

$$\begin{cases} \left(x^{k+1}, z^{k+1}\right) = \underset{x,z}{\operatorname{argmin}} L_p\left(x, z, y^k\right) \\ y^{k+1} = y^k + \rho\left(Ax^{k+1} + Bz^{k+1} - c\right) \end{cases}. \tag{17}$$

2.5 Fusing Three-Term Decompositions

In the current research two new schemes for fusion of the output results from two of the most accurate three-term decompositions are proposed. The first is named FusionOR and it consists of applying the logical OR operation on a pixel level over the two resulting videos from the decompositions. Since the videos are binary, it is enough at least one of the compared pixels to be equal to 1 (white) in order the final output video to contain also a white pixel (object). This algorithm is shown in the listing Algorithm 1 below.

```
// Algorithm 1 - FusionOR
function result.avi = FusionOR (input.avi) {
  uint m, n, k, M, N, K;
  result1.avi = TTD1();
  result2.avi = TTD2();
     m = 1; n = 1; k = 1;
     result.avi = result1.avi || result2.avi;
     do {
        if (result.avi(m,n,k) >= 1)
          result.avi(m,n,k) = 1;
        } while ((m <= M)&&(n <= N)&&(k <= K)); }
```

The expected number of white pixels for FusionOR, that is such that correspond to moving objects in the video, should be greater than any of the number of pixels of the fused videos and it should result in wider areas being covered in the output binary video. In Algorithm 1, the variables m and n are the spatial coordinates of the currently processed pixel and k is the current number of frame. Totally, each frame has M and N pixels and the videos contain K frames.

The second fusion scheme is named FusionAND and in it each pixel from each of the two top accurate decomposition algorithms takes place in a logical AND operation

to generate the final value of the pixel in that location. The algorithm is given in the listing named Algorithm 2 below. It follows the same logic as Algorithm 1 with the exception of the applied operation. Here, both pixels' values from the decompositions must be 1 (white, object) in order to generate final value of 1 (white pixel, belonging to a moving object). It is expected that the total number of pixels forming the areas, presumably belonging to a moving object, should be less than any of the preliminary obtained binary videos.

```
// Algorithm 2 - FusionAND
function result.avi = FusionAND (input.avi) {
  uint m, n, k, M, N, K;
  result1.avi = TTD1();
  result2.avi = TTD2();
     m = 1; n = 1; k = 1;
     result.avi = result1.avi && result2.avi;
     do {
        if (result.avi(m,n,k) >= 1)
          result.avi(m,n,k) = 1;
        } while ((m <= M)&&(n <= N)&&(k <= K)); }
```

3 Experimental Results

The experimental results are gathered over a PC® compatible computer with Intel Core i7-6820HQ processor with 4 cores with base frequency of 2.70 GHz, working in hyperthreading mode. The operation memory size is 64 GB and the hard disk drive has a capacity of 1 TB. The operating system is 64-bit MS® Windows® 10 Professional 21H2 and the test environment is Matlab R2022a. Implementation of the 4 tested decompositions—3WD, MAMR, RMAMR and ADMM is used from the LRS library [13].

The test videos are 6, an excerpt from the LASIESTA database [14]. Namely, these videos are: I_IL_01, consisting of 300 frames, O_CL_01—250 frames, I_OC_2—300 frames, I_SI_01—220 frames, O_RA_02—370 frames, and O_SU_02—400 frames. The size of the frame for all videos is 352 × 288 pixels with 8 bpp bitdepth and a framerate of 10 fps. They are contained in non-compressed AVI files using the RGB colorspace. This database has ground truth in separate videos with all moving objects being marked in color.

Three parameters ease the process of evaluating the accuracy of the decomposition algorithms, first of which is the Detection Rate (*DR*):

$$DR = \frac{TP}{TP + FN}, \tag{18}$$

where *TP* are the True Positives—the number of pixels from moving objects that are detected as such, *FN*—False Negatives—pixels, belonging to moving objects but not detected as such.

The second evaluating parameter is *Precision*:

$$Precision = \frac{TP}{TP + FP}, \tag{19}$$

where FP are the False Positives—the number of pixels, detected as part of moving objects, but are actually part of the background.

The third parameter is F-measure:

$$F - measure = \frac{2.DR.Precision}{DR + Precision}. \tag{20}$$

The average values of *DR*, *Precision* and F-measure, achieved by every of the tested algorithms, are shown in Fig. 1.

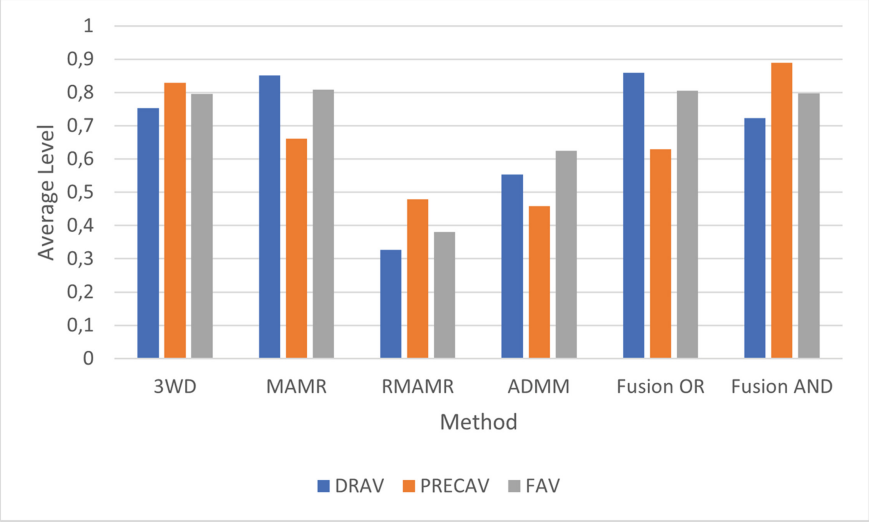


Fig. 1. Average values of *DR*, *Precision* and F-measure

The deviations of *DR*, *Precision* and F-measure are shown in Fig. 2.

During testing two times are registered—Processing Time (PT) and Full Time (FT). The first includes only the time for the particular decomposition while the second includes the input-output operations. Their values are shown in Fig. 3.

The logical OR operation takes 0.0118×10^{-6} s/px and the logical AND—0.0111×10^{-6} s/px.

4 Discussion

From the single decomposition algorithms MAMR has the highest $DR = 0.8505$, followed by the 3WD with a difference of 0.0979, ADMM—with 0.2974 less and

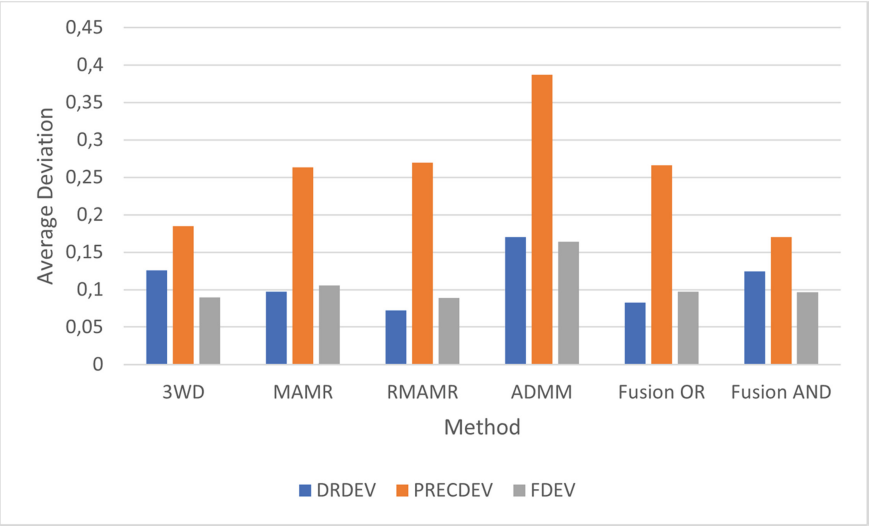


Fig. 2. Deviation values of *DR*, *Precision* and *F*-measure

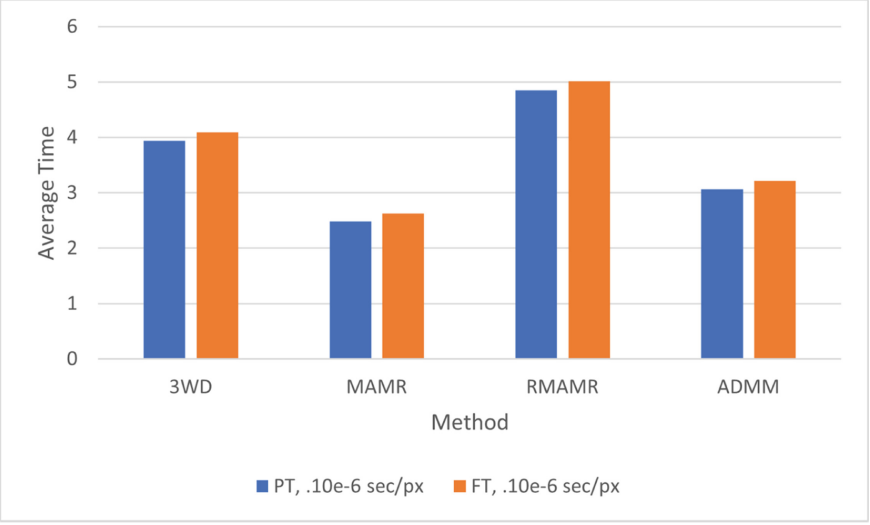


Fig. 3. Measured processing and full times

RMAMR—with 0.5230 less (Fig. 1). The FusionOR scheme between 3WD and MAMR, the two most accurate decompositions, leads to $DR = 0.8598$, while the FusionAND yields $DR = 0.7236$. In the same time, 3WD has the highest *Precision* $= 0.8296$, followed by MAMR with 0.1680 less, RMAMR—with 0.3506 less, and ADMM—with 0.3711 less. FusionAND is the scheme that obtains the highest *Precision* $= 0.889$, while the FusionOR is 0.2591 behind it. According to the *F*-measure, MAMR is first with

0.8082, followed by 3WD—with 0.0128 less, ADMM—with 0.1833 less, and at last from the single algorithms is RMAMR with 0.4274 less. FusionOR has F-measure = 0.8056, being close to FusionAND with F-measure = 0.7971 (Fig. 1).

The detection rate among all 6 tested algorithms deviates with around 0.1 from the complete set of 6 videos (Fig. 2). Given the variety of moving objects and the complexity of the surrounding scenes—people in office environment with variable lighting and occluding objects, moving cars in a parking lot, pedestrians walking in stationary shadows, casting their own moving shadows and others, the average level of deviation does not seem to be intolerably high and it shows the relative robustness of the three-term decompositions. The deviation of the *Precision* is around 2 times higher—changing around 0.2 with the least magnitude of 0.1701 for the FusionAND scheme. FusionOR naturally is higher due to its unifying effect over the groups of pixels around every moving object and it has precision deviation of 0.2662 (Fig. 2). The highest deviation of this parameter is for the ADMM—0.3868. The deviation of the F-measure is very close to 0.1 for all 6 algorithms, being the least for the RMAMR—0.0892.

The fastest decomposition algorithm is MAMR with around 2.5 μs/px, followed by ADMM—1.23 times slower, 3WD—1.59 times slower and RMAMR—1.95 times slower (Fig. 3). Full processing times follow almost the same tendency. Both the logical functions for the fusing schemes are 2 orders of a magnitude faster than any of the decomposition algorithms.

Visually, the results from the parameters for accuracy of the evaluated algorithms are confirmed by analyzing the output videos. A single frame from one of the videos is shown in Fig. 4.a—two persons are crossing a road near a parking lot. One of them is carrying a backpack and walking on broad sunlight, while the other is walking in opposite direction in relatively light shadows carrying a bag under one arm. Both men are followed by their own shadows. The ground truth for this frame is given in Fig. 4.b—it includes just the bodies of the two men, without their shadows or any other object from the scene. The result from decomposition by the 3WD is shown in Fig. 4.c—large portion from the back of the left standing person is missing and part from the right leg of the right standing person is absent. Less are the missing details in the bodies of the two persons in the frame, obtained by the MAMR decomposition (Fig. 4.d). This observation explains the higher *DR* for the MAMR that that for the 3WD (Fig. 1).

In the same time, the false positives in the case of MAMR are less—due to the more shrinked parts of the bodies, which explains the higher *Precision* for the 3WD (Fig. 1). Cumulatively, MAMR is slightly better than 3WD based on the F-measure with a difference of 0.0128. In both single decompositions the shadows of the men are detected as moving objects, almost in their entirety (Fig. 4.c and d). In the resulting frame from the FusionOR scheme (Fig. 4.e) the number of missing pixels from both bodies is even smaller than that from the MAMR (Fig. 4.d), which is expected given the nature of the logical OR function. That is why the *DR* here is the most, but the *Precision* is less than any of the fusing algorithms because of the remaining false positives. The F-measure reaches its maximum for this case.

FusionAND does not generate that much false positives and it suppresses portion of the actual pixels from the bodies (Fig. 4.f) and that is the reason why the *DR* is less than both of the single algorithms, that are taking part in the fusion. The *Precision*, however,

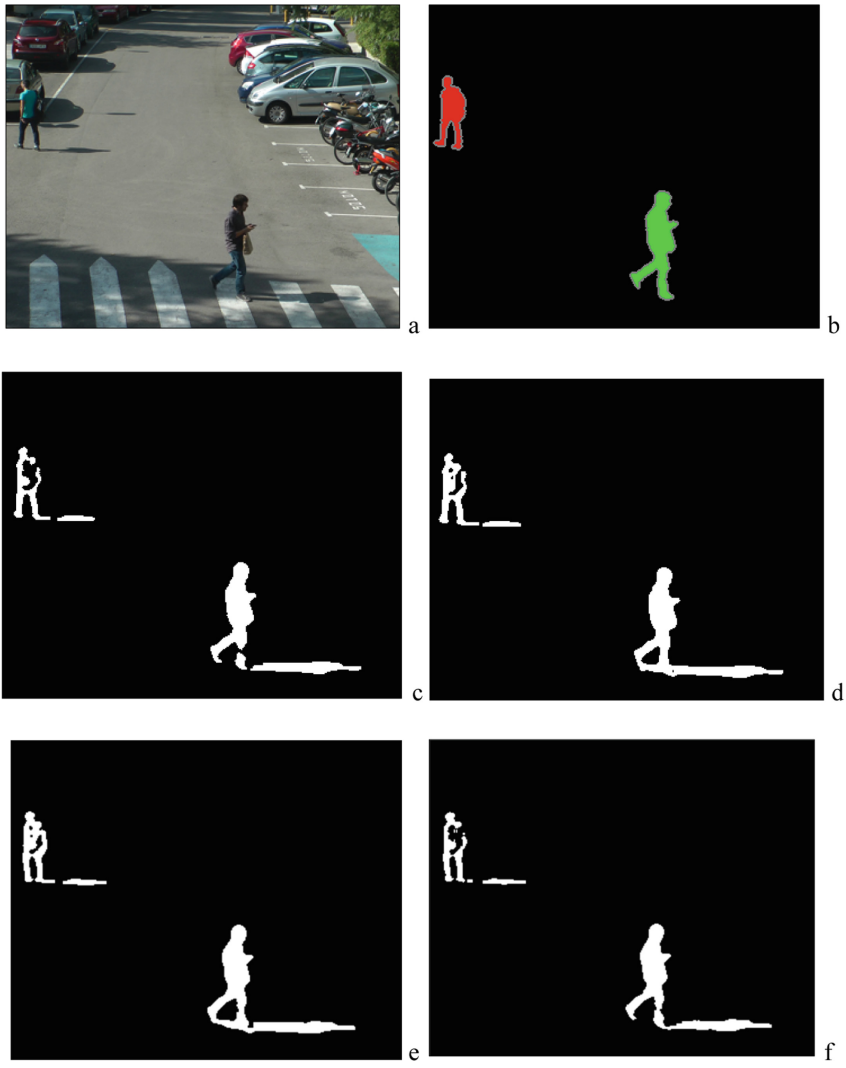


Fig. 4. Detected moving objects: a—original, b—ground truth, c—3WD, d—MAMR, e—FusionOR, f—FusionAND

is reaching its maximum here—0.8891. There are visible discontinuities in the bodies of the pedestrians from the video. The F-measure is very close to that of the FusionOR scheme—only 0.0085 difference (Fig. 1).

In Table 1 is presented a comparison based on the F-measure between the FusionOR and FusionAND schemes from the current study and another test as described by Sobral [15] using the 3WD decomposition. This comparison does not follow strict experimental protocol since these two studies use different video data sets but is at least indicative on the magnitude of achieved efficiency.

Table 1. Detection accuracy comparison

Method	F-measure
Fusion OR	0.8056
Fusion AND	0.7971
3WD Birds, [15]	0.5950

5 Conclusion

In this paper a comparative evaluation is presented among the 3WD, MAMR, RMAMR and ADMM decomposition algorithms used for background modelling and subtraction in order to trace moving objects in video. Two fusing schemes, named FusionOR and FusionAND are evaluated along with the standalone decompositions. The most accurate implementations in terms of detection accuracy is the FusionOR, followed by the MAMR. Highest precision is achieved for FusionAND, followed by the 3WD. Compared based on F-measure the top algorithm is FusionOR, followed by the MAMR. MAMR is also the fastest of all tested implementations, followed by the ADMM, 3WD and RMAMR. Given the average F-measure over the whole video data set the most promising algorithms tend to be the 3WD and MAMR, along with the two fusion schemes when moving objects of general type are considered for detection.

Acknowledgement. This work was supported by the National Science Fund of Bulgaria: KP-06-H27/16 “Development of efficient methods and algorithms for tensor-based processing and analysis of multidimensional images with application in interdisciplinary areas”.

References

1. Zhang, Y., Zhu, D., Wang, P., Zhang, G., Leung, H.: Vision-based vehicle detection for VideoSAR surveillance using low-rank plus sparse three-term decomposition. IEEE Trans. Veh. Technol. **69**(5), 4711–4726 (2020)
2. Li, Y., Yang, W., Jiang, Y., Li, C.: Sea-land coarse segmentation with two and three-term LRS decompositions in multisquint spaceborne SAR imagery. In: 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, pp. 3424–3427. IEEE (July, 2021)
3. Zhang, J., Jia, X.: Improved low rank plus structured sparsity and unstructured sparsity decomposition for moving object detection in satellite videos. In: IGARSS 2019–2019 IEEE International Geoscience and Remote Sensing Symposium, pp. 5421–5424. IEEE (July 2019)
4. Zhang, Y., Ye, Z., Qian, Y., Zhu, D.: Ultra-low velocity MTD based on Ku-band circular VideoSAR system. In: 2021 International Applied Computational Electromagnetics Society (ACES-China) Symposium, pp. 1–2. IEEE (July 2021)
5. Gonzalez, C.G., Absil, O., Absil, P.A., Van Droogenbroeck, M., Mawet, D., Surdej, J.: Low-rank plus sparse decomposition for exoplanet detection in direct-imaging ADI sequences—the LLSG algorithm. Astron. Astrophys. **589**, A54 (2016)
6. Oreifej, O., Li, X., Shah, M.: Simultaneous video stabilization and moving object detection in turbulence. IEEE Trans. Pattern Anal. Mach. Intell. **35**(2), 450–462 (2012)

7. Qin, B., Mao, H., Zhang, R., Zhu, Y., Ding, S., Chen, X.: Working memory inspired hierarchical video decomposition with transformative representations. arXiv preprint arXiv:2204.10105 (2022)
8. Du, S., Liu, B., Shan, G., Shi, Y., Wang, W.: Enhanced tensor low-rank representation for clustering and denoising. Knowl.-Based Syst. **243**, 108468 (2022)
9. Islam, I.B., Elahi, M.T.E., Halder, K.K.: An efficient mask generation method for moving object detection in atmospheric imaging. In: 2019 International Conference on Electrical, Computer and Communication Engineering (ECCE), pp. 1–5. IEEE (February 2019)
10. Wu, C., Chen, H., XiaoLi, Y., Xiong, N.: Total variation regularized OR-PCA for foreground detection in water monitoring. Data Sci. Ind. Internet **2**(1), 56–70 (2019)
11. Ye, X., Yang, J., Sun, X., Li, K., Hou, C., Wang, Y.: Foreground–background separation from video clips via motion-assisted matrix restoration. IEEE Trans. Circuits Syst. Video Technol. **25**(11), 1721–1734 (2015)
12. Boyd, S., Parikh, N., Chu, E., Peleato, B., Eckstein, J.: Distributed optimization and statistical learning via the alternating direction method of multipliers. Found. Trends® Mach. Learn. **3**(1), 1–122 (2011)
13. Sobral, A., Bouwmans, T., Zahzah, E.H.: Lrslibrary: Low-rank and sparse tools for background modeling and subtraction in videos. In: Robust Low-Rank and Sparse Matrix Decomposition: Applications in Image and Video Processing. CRC Press, Boca Raton (2016).
14. Cuevas, C., Yáñez, E.M., García, N.: Labeled dataset for integral evaluation of moving object detection algorithms: LASIESTA. Comput. Vis. Image Underst. **152**, 103–117 (2016)
15. Sobral, A.C.: Robust low-rank and sparse decomposition for moving object detection: from matrices to tensors. Doctoral dissertation, Université de La Rochelle, p. 66 (2017)

Deep Learning Approaches for Classroom Audio Classification Using Mel Spectrograms

Afsana Mou(✉), Mariofanna Milanova, and Mark Baillie

University of Arkansas at Little Rock, Little Rock, AR, USA
{armou,mgmilanova,mtbaillie}@ualr.edu

Abstract. Active learning can be immensely beneficial for science, technology, engineering, and mathematics (STEM) students as well as instructors because they engage themselves with various activities in the classroom so that lectures will be more effective. Automatic audio classification for classroom activities can help us to improve the active learning strategies in classroom. In this paper, we compare the different deep learning approaches (e.g., convolutional neural network, long short-term memory (LSTM)) for classroom audio classification. We evaluate the models on three classes of activity labeled as "single voice", "multiple voice" and "no voice" based on our classroom recording. The model is trained on the Mel spectrogram extracted from the recorded audio. We get the highest accuracy of 97% and F1 score of 0.97 with the 2D Convolutional Neural Network.

Keywords: Active learning · Deep learning · Convolutional neural network · Long short-term memory · Mel spectrograms · Classroom sound classification

1 Introduction

STEM students can maximize their learning and success through active learning pedagogies than the traditional lecturing in the classroom. They can also improve their critical thinking skills and their preservation of knowledge with the help of active learning. So, more students will entice in the STEM disciplines. As this technique is new in this field, STEM faculty also needs to be familiar with it. We need some metrics for the improvement and effectiveness of active learning in classroom so that it is possible to measure the progress [1, 2].

In Wang et al. [3], LENA system (Language Environment Analysis system) [4] is used to classify the three activities ("teacher lecturing", "whole class discussion" and "student group work") from classroom recordings. Random Forest classifier is used for the classification of these three activities. Later, Owens et al. [5] developed a machine learning based algorithm, Decibel Analysis for Research in Teaching (DART) which can analyze the classroom recordings automatically without the need for the human observer. They try to measure the total amount of time spent on "single voice" (e.g., lecture), multiple voice (e.g., pair discussion), and no voice (e.g., clicker question thinking) activities with an ensemble of binary decision trees. They analyze around 1486 recordings of 1720 h of audio and got 90% of accuracy. Later, in Cosbey et al. [6], deep

R. Kountchev et al. (Eds.): *NAMSP 2022*, SIST 332, vol. 332,, pp. 23–30, 2023.
https://doi.org/10.1007/978-981-19-7842-5_3

and recurrent neural network is used for improving the DART system. They procured the 7.1% and 10.1% frame error rate respectively for the previously seen and unseen instructors. It is also important to identify the participation of the teacher and student in the classroom. Li et al. [7] proposed Siamese neural model to identify the teacher and student in the classroom activities. They use log Mel-filter bank for feature extraction from the audio and evaluated the model on LSTM, Gated Recurrent Unit (GRU), and also on the attention mechanism to differentiate the classroom activity. In another approach [8], they used multimodal attention layer which uses attention based neural network for classroom activity detection.

Slyman et al. [9] evaluate the performance of deep fully connected, convolutional, and recurrent neural network models for 9-way classification performance with 5-way and 4-way simplifications of the task. Their best model obtains frame-level error rates of 6.2%, 7.7% and 28.0% for the 4-way, 5-way, and 9-way classification tasks considering mel-filterbank, OpenSmile, and self-supervised acoustic features.

Hersey et al. [10] showed convolutional neural network shows promising results for image classification as well as audio classification. They analysis several models such as fully connected Deep Neural Networks (DNNs), Alex Net, VGG, Inception, and ResNet on 5.24 million of data. Later Palanisamy et al. [11] proposed audio classification using an ensemble of ImageNet pretrained DenseNet model and achieves 92.89% and 87.42% accuracy on ESC-50 and UrbanSound8K datasets. Sequential nature of the audio can be captured in LSTM architecture instead of fixed frame. It uses a series of memory gates for better capturing of long-range feature sequences [12]. Later in Lezhenin et al. [13], LSTM is used for environmental sound classification. They extracted features using Mel Spectrogram from UrbanSound8K dataset and analyzed the model using 5-fold cross-validation. LSTM showed good result compared to the CNN. Audio features can be extracted in various ways such as time, frequency, ceptrum, time frequency etc. Mel Frequency Cepstral Coefficients (MFCC) can be collected from power modulus of the signal spectrum. Short Time Fourier Transform or Spectrogram is useful to collect more features in intermediate stage [14].

In this paper, we compared the 2D Convolutional Neural Network and LSTM model on classification of our recorded classroom activities of two instructors and extracts audio features from the Mel spectrograms.

2 Methodology

Deep learning shows tremendous success in image classification task and achieved state-of-the-art results in numerous fields. Deep learning also showing promising result in audio classification task also such as music genres classification, environmental sound classification. Figure 1 shows the proposed pipeline of the paper.

A detailed description for the proposed work is given as follows.

2.1 Mel Spectrogram

As we mentioned earlier, in a numerous way we can extract features from the labeled audio data in audio classification task such as Mel filter bank, MFCC, Mel Spectrogram.

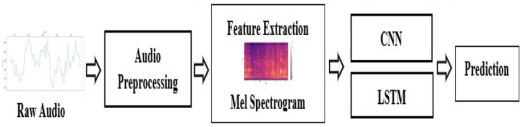


Fig. 1. Face pipeline of the proposed method.

In this paper, we use Mel spectrogram to extract the audio features from classroom recordings. Generally, sound can be represented in either time domain or frequency domain. Fourier transforms help us to convert the sound from time to frequency domain which is known as Spectrum.

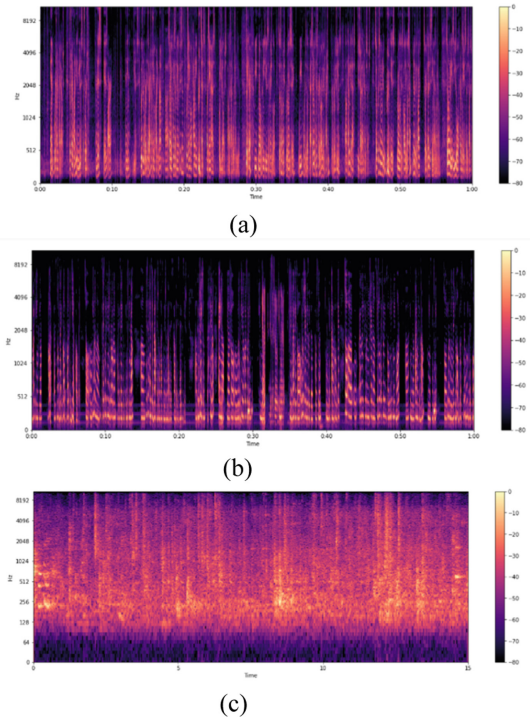


Fig. 2. Mel spectrograms of different classroom activity. (a) Single voice, (b) multiple voice, (c) no voice

Most of the sound we hear is a composite signal. Fourier transform can break down that signal into each separate signal. With general Fast Fourier transform we cannot get the time localized information about the signal. But Short-Time Fourier Transform (STFT) represents varying frequencies over time by computing Fast Fourier transform on the small segment of the signal. Spectrogram can be created by applying STFT in small chunks or windows. Mel spectrogram is helpful to represent the nonlinearity of human pitch. In Fig. 2 we can see the images of Mel spectrograms of classroom activities

[14, 15]. The equation to convert normal frequency, f to Mel scale, $M(f)$ is:

$$M(f) = 1125 \log\left(1 + \frac{f}{700}\right) \tag{1}$$

2.2 Convolutional Neural Network

Convolutional neural network (CNN) is a special branch of deep neural network. Numerous State of the art results for image recognition and classification task is achieved through the CNN. Pixels are used as an array of width, height, and depth for input in CNN. Instead of transferring all the weights from one layer to another, it only transfers only the weights from the previous layer to next layer. As a result, we get better weight management with CNN. The first layer of the CNN architecture is convolutional layer. Different Features are extracted using a different filter or kernel in each layer of the CNN with pixel array such as identity, edge detection, sharpen, box blur, Gaussian blur etc. Stride also helps to reduce the size of parameters by controlling the overlapping in the matrices. Sometimes, some of the information on the outer edge cannot captured during the kernel slide through the image. Padding can help on this process. It also helps to sustain the image size. The following formula is used for zero padding. Pooling is used for down sampling. Generally, max pooling takes the highest value in a smaller region helps to reduce the image resolution. Fully connected layer works as traditional neural network where all the nodes are connected to both previous and upcoming layer [16–18].

The equation for convolutional layer is,

$$f_l^k(p, q) = \sum_c \sum_{x,y} i_c(x, y).e_l^k(u, v) \tag{2}$$

Here, $i_c(x, y)$ represents element of input image and $e_l^k(u, v)$ means index of the element of the kernel.

The equation for zero padding, O is following:

$$O = 1 + \frac{N + 2P - F}{S} \tag{3}$$

Here, N for input size, F for filter size, P for zero padding layer number and S for stride.

2.3 Long Short-Term Memory

Recurrent Neural Network (RNN) is a special type of neural network which works excellent with the sequential data. Main difference of RNN from traditional neural network is that it passes back some information to input layer for better extraction of the feature. But RNN is not good with very long sequence of data. It may face problems such as gradient vanishing or exploding. LSTM can be beneficial to solve these issues. In LSTM, a memory cell resides which can store information for long time and the gates can also control the input and output of the memory cell [19, 20].

3 Experimental Setup

3.1 Dataset

We did not use any available or synthetic data for our experiment. We collected total 9 h of the recordings of the two instructors. Then with the help of an audio editor, Audacity we labeled the recordings into three classes. They are single voice as instructors lecture, multiple voice as group activities and no voice as classroom task. Sample rate of all the recorded audio is 44,100 Hz.

3.2 Dataset Description

First, we clean the audio data to remove the dead space. If there are numerous dead spaces in the audio, it will be difficult to differentiate the audio. All the audio segments may look alike. Any segment less than 20 dB is removed from the audio. For computational simplicity, all the audios have same frame size of 60 s. We take 6 h of data as training dataset and 3 h of data as test dataset. Keras Audio Preprocessors (Kapre) [21] is used to extract audio features from the Mel spectrograms on real time GPU.

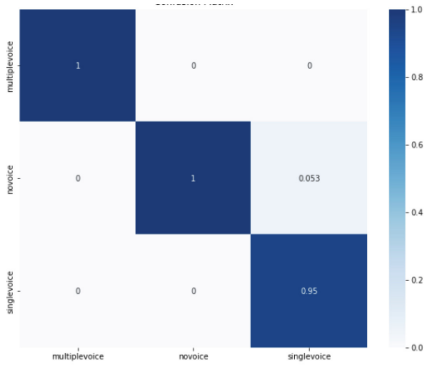
4 Results

To evaluate the classroom activity classification, we performed two tests using 1D CNN, 2D CNN, and LSTM algorithm using TensorFlow in Python. In 2D CNN, we considered five convolutional layer and four Max-pooling layer and at last the dense layer. Similarly, in LSTM we use the bidirectional unit and dense layer at the end of the model. In all the three model, Adam optimizer is used. To extract the features from the Mel spectrograms, we considered 128 Mels, n_fft of 512, window size of 400, 16 kHz sample rate and hop length of 160 as hyperparameters. We use STFT and Hann window to form the Mel spectrograms.

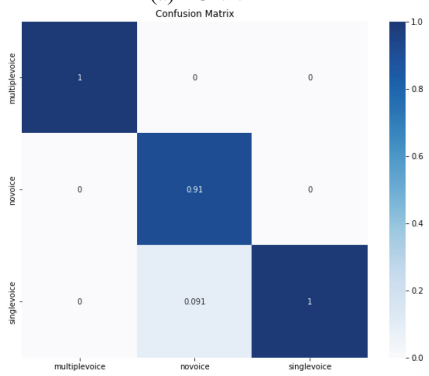
Table 1. Comparison of different model.

Model	Test 1				Test 2			
	Accuracy	Precision	Recall	F1 Score	Accuracy	Precision	Recall	F1 Score
1D CNN	0.81	0.91	0.73	0.81	0.86	0.86	0.81	0.83
2D CNN	0.97	0.98	0.97	0.97	0.95	0.97	0.97	0.97
LSTM	0.92	0.90	0.94	0.92	0.95	0.94	0.94	0.94

Table 1 shows the result of the two test we conducted. For 1D CNN we get the accuracy of 0.86 and F1 score of 0.81. We got the highest the accuracy with the 2D CNN of 0.97 and F1 score of 0.97. LSTM has also showed promising result for the classification task. It got the accuracy of 0.95 and F1 score of 0.94.



(a) CNN



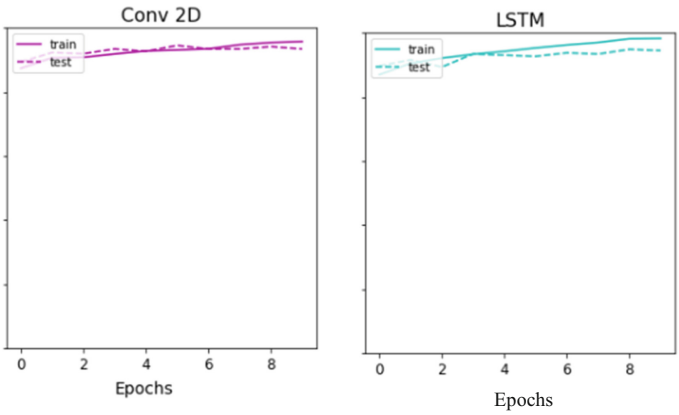
(b) LSTM

Fig. 3. Confusion matrix

In Fig. 3, we show the confusion matrix for 2D CNN and the LSTM model. In 2D CNN, it almost accurately detected the activities. But we observe that sometimes, it detects no voice as single voice. In case of LSTM, we observe the difficulty to detect the no voice activities. In Fig. 4, we illustrate the comparison of training and testing accuracy for 10 epochs.

5 Conclusion

In this paper, we compare the different deep learning approach for classroom activity classification using Mel spectrograms. Raw audio data does not perform good for extracting the features. If sound classification problem is approached as image classification such as converting it to a Mel spectrogram, it can show promising result. Based on two experimentations, we observe that 2D convolutional neural network shows promising result which is 97% accuracy for classroom audio classification. In future research, instead of using only one feature different, concatenation of different feature extraction methods can be applied. Using pre-trained model can also help for better learning of the



(a) CNN (b) LSTM

Fig. 4. Accuracy

features and it can train quickly. Other deep learning architectures or ensemble model can also be used for further improvement of this research.

Acknowledgments. This work was supported by the National Science Foundation under Award No. OIA-1946391 (DART), DCSTEM Seed Grant and National Science Fund of Bulgaria: KP-06-H27/16 Development of efficient methods and algorithms for tensor-based processing and analysis of multidimensional images with application in interdisciplinary areas.

References

1. Freeman, S., Eddy, S.L., McDonough, M., Smith, M.K., Okoroafor, N., Jordt, H., Wenderoth, M.P.: Active learning increases student performance in science, engineering, and mathematics. Proc. Natl. Acad. Sci. **111**(23), 8410–8415 (2014)
2. Olson, S., Riordan, D.G.: Engage to excel: producing one million additional college graduates with degrees in science, technology, engineering, and mathematics. In: Report to the President. Executive Office of the President (2012)
3. Wang, Z., Pan, X., Miller, K.F., Cortina, K.S.: Automatic classification of activities in classroom discourse. Comput. Educ. **78**, 115–123 (2014)
4. Ganek, H., Eriks-Brophy, A.: The Language Environment Analysis (LENA) system: a literature review. In: Proceedings of the joint workshop on NLP for Computer Assisted Language Learning and NLP for Language Acquisition, pp. 24–32 (2016)
5. Owens, M.T., Seidel, S.B., Wong, M., Bejines, T.E., Lietz, S., Perez, J.R., Sit, S., et al.: Classroom sound can be used to classify teaching practices in college science courses. Proc. Natl. Acad. Sci. **114**(12), 3085–3090 (2017)
6. Cosbey, R., Wusterbarth, A., Hutchinson, B.: Deep learning for classroom activity detection from audio. In: ICASSP 2019–2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 3727–3731. IEEE (2019)
7. Li, H., Wang, Z., Tang, J., Ding, W., Liu, Z.: Siamese neural networks for class activity detection. In: International Conference on Artificial Intelligence in Education, pp. 162–167. Springer, Cham (2020)

8. Li, H., Kang, Y., Ding, W., Yang, S., Yang, S., Huang, G.Y., Liu, Z.: Multimodal learning for classroom activity detection. In: ICASSP 2020–2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 9234–9238. IEEE (2020)
9. Slyman, E., Daw, C., Skrabut, M., Usenko, A., Hutchinson, B.: Fine-Grained Classroom Activity Detection from Audio with Neural Networks. arXiv preprint arXiv:2107.14369 (2021)
10. Hershey, S., Chaudhuri, S., Ellis, D.P.W., Gemmeke, J.F., Jansen, A., Moore, R.C., Plakal, M., et al.: CNN architectures for large-scale audio classification. In: 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 131–135. IEEE (2017)
11. Palanisamy, K., Singhania, D., Yao, A.: Rethinking CNN Models for Audio Classification. arXiv preprint arXiv:2007.11154 (2020)
12. Dai, J., Liang, S., Xue, W., Ni, C., Liu, W.: Long short-term memory recurrent neural network based segment features for music genre classification. In: 2016 10th International Symposium on Chinese Spoken Language Processing (ISCSLP), pp. 1–5. IEEE (2016)
13. Lezhenin, I., Bogach, N., Pyshkin, E.: Urban sound classification using long short-term memory neural network. In: 2019 Federated Conference on Computer Science and Information Systems (FedCSIS), pp. 57–60. IEEE (2019)
14. Crocco, M., Cristani, M., Trucco, A., Murino, V.: Audio surveillance: a systematic review. ACM Comput. Surv. (CSUR) **48**(4), 1–46 (2016)
15. Stevens, S.S., Volkmann, J., Newman, E.B.: A scale for the measurement of the psychological magnitude pitch. J. Acoust. Soc. Am. **8**(3), 185–190 (1937)
16. Khan, A., Sohail, A., Zahoora, U., Qureshi, A.S.: A survey of the recent architectures of deep convolutional neural networks. Artif. Intell. Rev. **53**(8), 5455–5516 (2020). https://doi.org/10.1007/s10462-020-09825-6
17. Albawi, S., Mohammed, T.A., Al-Zawi, S.: Understanding of a convolutional neural network. In: 2017 International Conference on Engineering and Technology (ICET). IEEE (2017)
18. Liu, R., Yang, X., Chong, X., Wei, L., Zeng, X.: Comparative study of convolutional neural network and conventional machine learning methods for landslide susceptibility mapping. Remote Sens. **14**(2), 321 (2022)
19. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Comput. **9**(8), 1735–1780 (1997)
20. Graves, A.: Long short-term memory. In: Supervised Sequence Labelling with Recurrent Neural Networks, pp. 37–45 (2012)
21. Choi, K., Joo, D., Kim, J.: Kapre: On-gpu audio preprocessing layers for a quick implementation of deep neural network models with keras. arXiv preprint arXiv:1706.05781 (2017)
22. Dangol, R., Alsadoon, A., Prasad, P.W.C., Seher, I., Alsadoon, O.H.: Speech emotion recognition UsingConvolutional neural network and long-short TermMemory. Multimed. Tools Appl. **79**(43), 32917–32934 (2020)
23. Ma, X., Yang, H., Chen, Q., Huang, D., Wang. Depaudionet, Y.: An efficient deep model for audio based depression classification. In: Proceedings of the 6th International Workshop on Audio/Visual Emotion Challenge, pp. 35–42 (2016)
24. Scarpiniti, M., Comminiello, D., Uncini, A., Lee, Y.-C.: Deep recurrent neural networks for audio classification in construction sites. In: 2020 28th European Signal Processing Conference (EUSIPCO), pp. 810–814. IEEE 2021

Tensor Spectral Pyramid for Color Video Sequences Representation, Based on 3D FO-AHKLT

Roumen Kountchev[1](✉) and Roumiana Kountcheva[2]

[1] Technical University of Sofia, Sofia 1000, Bulgaria
rkountch@tu-sofia.bg
[2] TK Engineering, Sofia 1582, Bulgaria

Abstract. In this work, new structure is proposed aimed at the representation of color video—a sequence of RGB frames. For this, each of the color components is first transformed into a tensor, and then these tensors are represented in the reduced vector space. To achieve this, is used the Two-Layer Tensor Spectrum Pyramid (2LTSP) based on the n-level 3D Frequency-Ordered Adaptive Hierarchical KLT (3D FO-AHKLT). The new approach ensures high energy concentration of the transformed tensor into a small number of spectrum coefficients, together with low computational complexity. The 3D FO-AHKLT execution comprises three consecutive stages. In each stage, the input tensor is transformed into a vector with various spatial orientation: horizontal, vertical or lateral (in time). The computational complexity of 2LTSP is evaluated and compared with that of H-Tucker and Tensor Train decompositions and the analysis shows that it decreases together with the vector size growth. In this work is also proved that the implementation of 2LTSP based on the 3D Adaptive Fast Walsh-Hadamard Transform results into additional significant reduction of the needed calculations. The main advantages of the presented pyramidal structure are its flexibility regarding the number of retained spectrum coefficients in the first and second decomposition levels, the lack of iterative calculations, and the computational complexity. All these qualities open new implementation possibilities in application areas aimed at the information redundancy reduction in color video sequences, the use of relatively small number of features needed for objects recognition, etc.

Keywords: Video tensor representation · Two-layer tensor spectrum pyramid · 3D frequency-ordered adaptive hierarchical KLT · Adaptive directional vectorization

1 Introduction

One of the main tasks when processing of 3-channel video sequences is concerned, is their transform into a new space of reduced dimensionality. For this, the sequence of video frames is packed into three tensors corresponding to color components R, G, B, on which are applied various methods for tensor decomposition. One of the "classic" methods [1–3] is the Canonical-Polyadic Decomposition (CPD), where the tensor $\mathcal{X}$ is represented

R. Kountchev et al. (Eds.): *NAMSP 2022*, SIST 332, vol. 332,, pp. 31–43, 2023.
https://doi.org/10.1007/978-981-19-7842-5_4

as a sum of R rank-one tensors. To this group belongs also the method of Tucker, for which the tensor $\mathcal{X}$ is decomposed into a core tensor and a set of factor matrices. The Tucker model is also known as Higher-Order Singular Value Decomposition (HOSVD) [2]. The Tensor-Train (TT) decomposition [3] expresses $\mathcal{X}$ as the product of third-order tensors (decomposition cores). The Hierarchical Tucker decomposition (H-Tucker) [4] and some of its modifications [5] are based on the calculation of the eigenvalues and eigenvectors of the decomposed tensor. The basic attribute of these methods is that they are optimal with respect to minimization of the mean square approximation error derived from the "truncation" of the low-energy components. For the calculation of the components retained after the "truncation", are used iterative methods which need relatively high number of operations to achieve the requested accuracy. In [6] is offered the TT-based hierarchical decomposition of high-order tensors, for the calculation of which is used the Tensor-Train Hierarchical SVD (TT-HSVD). This approach permits parallel processing, which accelerates the calculations. Unlike the TT-SVD algorithm, the TT-HSVD is based on SVD applied to matrices of smaller dimensions. In result is evaluated that the complexity of TT-SVD needs R^2N^3 operations more than that of TT-HSVD (here, R denotes the rank of the hypercubical tensor of size N).

For tensor decompositions execution are also used various orthogonal 3D transforms, such as: Fast Fourier Transform (FFT), Discrete Cosine Transform (DCT), etc. [7]. This approach is distinguished by its flexibility regarding the transform choice, based on the processed data contents analysis.

As an alternative, in this work is presented the new Two-Layer Tensor Spectrum Pyramid (2LTSP), based on the 3D Frequency-Ordered Adaptive Hierarchical KLT (3D FO-AHKLT) [8, 9]. 3D FO-AHKLT is highly efficient in respect of tensor energy concentration into a relatively small number of spectrum coefficients, and computational complexity [9]. As a result, the offered pyramidal tensor decomposition 2LTSP does not need iterations and has low computational complexity.

In the next Sects. 2, 3 and 4 are explained respectively the 2LTSP structure for color RGB video representation, the basic operations used for the calculation of the spectrum pyramid, and the vectorization specifics for a tensor, comprising a sequence of video frames. In Sect. 5 is evaluated the computational complexity of the 2LTSP decomposition, and Sect. 6 presents the Conclusions.

2 Structure of the Two-Layer Tensor Spectrum Pyramid for RGB Color Video Representation

The new structure is a tensor spectrum pyramid based on the 3D Frequency-Ordered Adaptive Hierarchical Karhunen–Loeve Transform (3D FO-AHKLT). In result, each cubical tensor $\mathcal{X}$ of size $N = 2^n$, consisting of the nonnegative components x(i, j, k), is decomposed in correspondence with the relation below [9]:

$$\mathcal{X} = (1/N^3) \sum_{u=0}^{N-1} \sum_{v=0}^{N-1} \sum_{l=0}^{N-1} s(u, v, l)[\mathbf{k_u}(1) \circ \mathbf{k}_v(2) \circ \mathbf{k}_l(3)] \quad (1)$$

Here, coefficients $s(u, v, l)$ denote the elements of the spectrum tensor $\mathcal{S}$, which is of same size, as $\mathcal{X}$. Each coefficient $s(u, v, l)$ represents the weight of the basic tensor

$\mathcal{K}_{u,v,l}$—the outer product of the three vectors:

$$\mathcal{K}_{\mathrm{u,v,l}} = \mathbf{k}_{\mathrm{u}}(1) \circ \mathbf{k}_{\mathrm{v}}(2) \circ \mathbf{k}_{\mathrm{l}}(3). \tag{2}$$

Here "o" denotes the outer product of the two column-vectors ($\mathbf{x} \circ \mathbf{y} = \mathbf{x}.\ \mathbf{y}^{\mathrm{T}}$), and $\mathbf{k}_{\mathrm{u}}(1) \circ \mathbf{k}_{\mathrm{v}}(2) \circ \mathbf{k}_{\mathrm{l}}(3)$ are the basic vectors, obtained after the execution of the three stages of 3D FO-AHKLT. In each stage, 1D FO-AHKLT is applied on the vectors calculated after directional vectorization of the corresponding input tensor. The stages of 3D FO-AHKLT are executed sequentially (i.e., S_1, S_2, S_3), as shown on Fig. 1.

2LTSP is suitable for processing of color videos represented as sequences of RGB frames, packed correspondingly as cubical tensors $\mathcal{X}_{\mathrm{R}}$, $\mathcal{X}_G$, $\mathcal{X}_B$, each of size $\mathrm{N} = 2^{\mathrm{n}}$. The block diagram of 2LTSP which corresponds to the direct transform of RGB tensors of size N = 8, is shown on Fig. 2. The elements of the input tensors $\mathcal{X}_R$, $\mathcal{X}_G$ and $\mathcal{X}_R$ are $\mathrm{x_R(i, j, k)}$, $\mathrm{x_G(i, j, k)}$, and $\mathrm{x_B(i, j, k)}$, respectively. These tensors correspond to groups, which comprise the sequences of N color video frames in the discrete time moments k = 1, 2,…, N. The value of N is defined by the length of the group of sequential frames of relatively high inter-frame correlation. Each video frame is divided into small blocks of size N × N which are after that represented by the corresponding three matrices **R**, **G**, and **B**. After applying 2LTSP on tensors $\mathcal{X}_R$, $\mathcal{X}_G$, $\mathcal{X}_R$, they are transformed into vectors, defined in the reduced multidimensional spectrum space. This transform is reversible—from the so obtained vectors, through Inverse 2LTSP (2LITSP), are restored the input tensors $\mathcal{X}_R$, $\mathcal{X}_G$, $\mathcal{X}_R$.

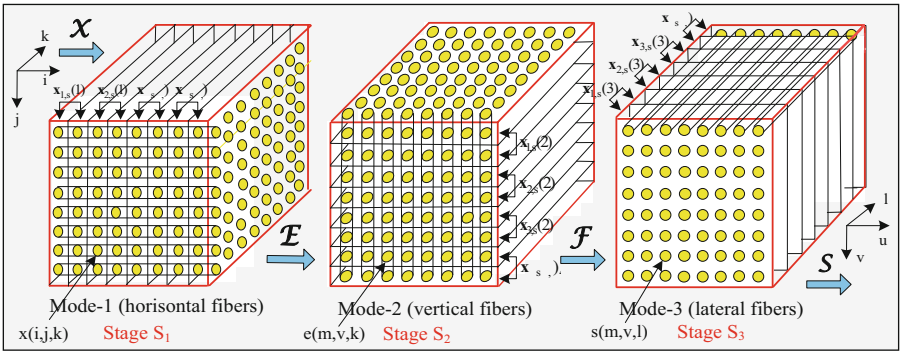


Fig. 1. Three stages of 3D FO-AHKLT for the tensor $\mathcal{X}$, of size 8 × 8 × 8.

In each stage $\mathrm{S_i}$ (for i = 1, 2, 3), is executed 1D FO-AHKLT whose computational graph is shown in Fig. 3, for the case N = 8 [8, 9]. The following notations are introduced in the figure: $\cos\theta_{p,q} = c_{p,q}$ and $\sin\theta_{p,q} = s_{p,q}$, for p = 1, 2, 3 and q = 1, 2, 3, 4. The basic operation in this graph is the Karhunen-Loeve Transform ($\mathrm{KLT_{2\times2}}$) with transformation matrix of size 2 × 2. Here $\mathrm{KLT_{2\times2}}$ is determined by the expression:

$$\mathbf{y}_s(p, q) = \begin{bmatrix} \mathrm{c_{p,q}} & \mathrm{s_{p,q}} \\ \mathrm{s_{p,q}} & -\mathrm{c_{p,q}} \end{bmatrix} \mathbf{x}_s(p, q) \text{ for } \mathrm{s} = 1, 2, .., \mathrm{N}^2, \tag{3}$$

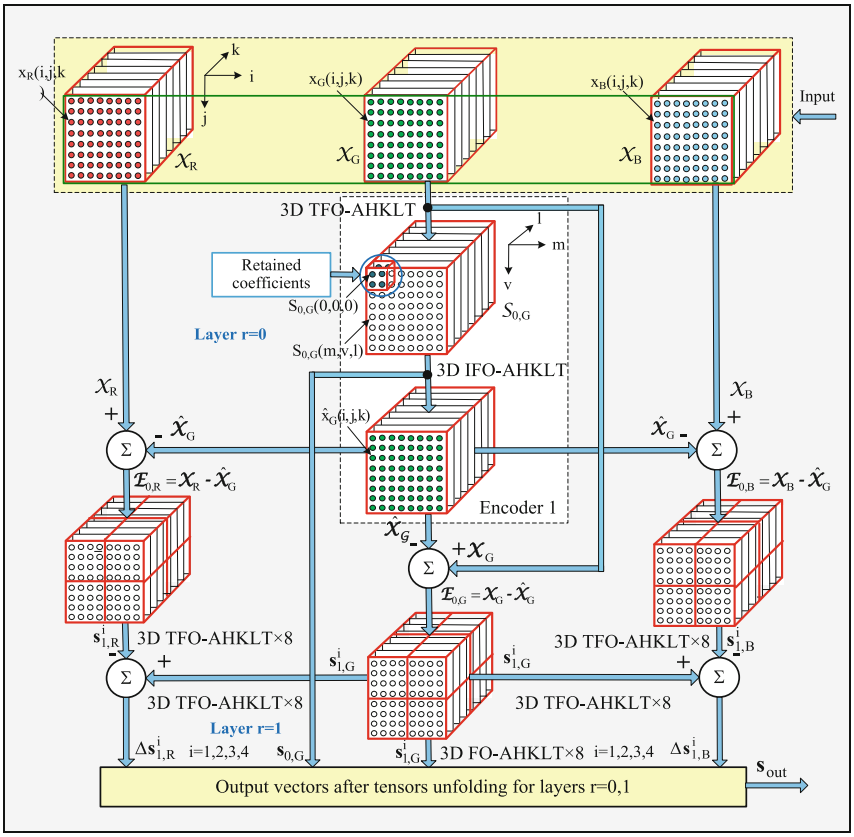


Fig. 2. Block diagram of the 2LTSP for the direct transform of tensors $\mathcal{X}_R$,$\mathcal{X}_G$,$\mathcal{X}_B$, each of size 8 × 8 × 8.

where $\mathbf{y}_s(p,q) = [y_{1,s}(p,q), y_{2,s}(p,q)]^T$, $\mathbf{x}_s(p,q) = [x_{1,s}(p,q), x_{2,s}(p,q)]^T$, and $E\{x_{1,s}(p,q)\} = E\{x_{2,s}(p,q)\} = 0$ (here $E(x_s) = (1/N^2)\sum_{s=1}^{N^2} x_s$ is an averaging operator); N^2 is the number of vectors $\mathbf{x}_s(p,q)$ and $\mathbf{y}_s(p,q)$ for p = 1, 2, 3, q = 1, 2, 3, 4, and N = 8; $\theta_{p,q} = \arctan\{2k_3(p,q)/[k_1(p,q) - k_2(p,q) + \sqrt{(k_1(p,q) - k_2(p,q))^2 + 4k_3^2(p,q)}]\}$; $k_1(p,q) = E\{x_{1,s}^2(p,q)\}$; $k_2(p,q) = E\{x_{2,s}^2(p,q)\}$; $k_3(p,q) = E\{x_{1,s}(p,q)x_{2,s}(p,q)\}$.

The transform is executed for the tensor $\mathcal{X}$ of size 8 × 8 × 8. For this, the tensor is unpacked (vectorized) adaptively in one of the three possible directions: horizontal, vertical, or lateral for u = 1, 2, 3, respectively. The choice of the vectorization direction u is based on the correlation analysis results [9] for the vectors calculated for the input tensor, $\mathcal{X}$. The vectorization direction should be chosen before the execution of each stage S_i of 3D FO-AHKLT.

In correspondence with the graph, shown on Fig. 3, for the hierarchical level p = 1,2 is checked if the decorrelation condition $\Lambda_p = \sum_{i=1}^{N}\sum_{j=1}^{N}\left[k_{y_p,i,j}(u)\right]^2_{(i\neq j)} \leq \delta$ of

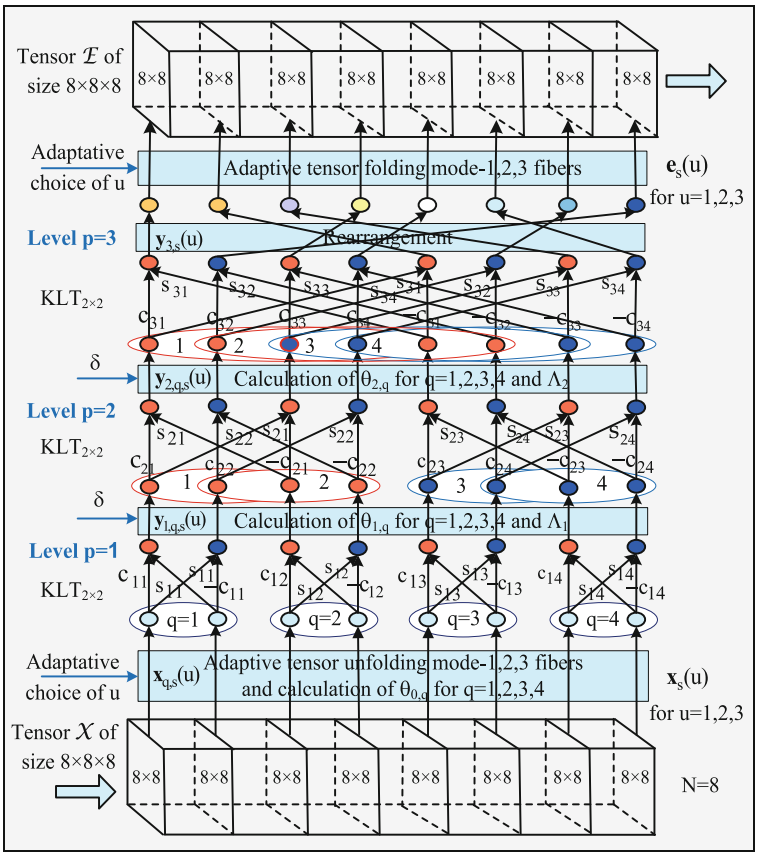


Fig. 3. Three-level direct 1D FO-AHKLT with adaptive vectorization of tensor $\mathcal{X}$ of size 8 × 8 × 8

the transformed vectors $\mathbf{y}_{p,,s}(u)$ is satisfied [9]. Here, $k_{y_p,i,j}(u)$ is the (i, j)th element of the correlation matrix $\mathcal{K}_{\mathbf{y}_p}(u)$, and δ is a threshold. This check is done at the end of the calculations for the corresponding level 1, or 2. When the operations for all levels of 1D FO-AHKLT are finished, through unfolding is restored the output tensor $\mathcal{E}$, of size 8 × 8 × 8. Then, it becomes the input tensor for the next (second) stage of 3D FO-AHKLT. The tensor $\mathcal{E}$ is vectorized adaptively with new vector orientation u = 1, 2, 3 and on it is applied again 1D FO-AHKLT. In the last, third stage, the output tensor $\mathcal{F}$ from the second stage, is transformed similarly into the spectrum tensor $\mathcal{S}$, of size 8 × 8 × 8.

3 Calculation of the Two-Layer Tensor Spectrum Pyramid

In correspondence with the 2LTSP block diagram from Fig. 2, on the tensor $\mathcal{X}_G$ is applied the direct 3D Truncated FO-AHKLT (3D TFO-AHKLT). Out of the total of 256 elements of the spectrum tensor $\mathcal{S}_G$, are retained only $n_0 = 8$ elements $s_G(u, v, l)$ for

u, v, l = 0,1, which correspond to the lowest frequencies of the 3D spectrum. In these coefficients is concentrated the main part of the energy of the input tensor, $\mathcal{X}_G$.

In the *first layer* (r = 0) of 2LTSP, the tensor $\mathcal{X}_G$ is transformed through Coder-1 into the spectrum tensor $\mathcal{S}_G$, which performs the direct 3D TFO-AHKLT in the structure of Encoder-1. At the output of Decoder-1, which executes the Inverse 3D FO-AHKLT (3D IFO-AHKLT), is obtained the approximated tensor $\hat{\mathcal{X}}_G$. The approximation accuracy depends on the number of retained spectrum coefficients (n_0) of the tensor $\mathcal{S}_0$, G. These coefficients are the components of the corresponding n_0-dimensional vector, $\mathbf{S}_{0,G}$. At the end of the execution of layer r = 0 are calculated the difference tensors $\mathcal{E}_{0,\mathrm{R}} = \mathcal{X}_\mathrm{R} - \hat{\mathcal{X}}_\mathrm{G}$, $\mathcal{E}_{0,G} = \mathcal{X}_G - \hat{\mathcal{X}}_G$, $\mathcal{E}_{0,\mathrm{B}} = \mathcal{X}_\mathrm{B} - \hat{\mathcal{X}}_\mathrm{G}$.

In the *second layer* (r = 1) of 2LTSP, each difference tensor $\mathcal{E}_{0,G}$ is divided into 4 sub-tensors of size (M/2) × (N/2) × 3 and for each is executed the corresponding coding/decoding. As a result are obtained the output vectors, each of length n_1 (the number of coefficients at the outputs of each encoder in level r = 1), and are calculated the spectral vectors $\Delta\,\mathbf{s}^i_{1,R} = \mathbf{s}^i_{1,G} - \mathbf{s}^i_{1,R}$, $\mathbf{s}^i_{1,G}$, and $\Delta\,\mathbf{s}^i_{1,B} = \mathbf{s}^i_{1,G} - \mathbf{s}^i_{1,B}$, for i = 1, 2, 3, 4. As a result of the high inter-frame correlation of tensors $\mathcal{X}_R$,$\mathcal{X}_G$,$\mathcal{X}_B$, the values of significant part of the coefficients of the difference vectors $\Delta\mathbf{s}^i_{1,R}$ and $\Delta\mathbf{s}^i_{1,B}$ are close, or equal to zero. In this way, after the hierarchical 2LTSP decomposition, the input tensors are transformed into an output vector of small length which has many zero values, i.e., the features' space is reduced at minimum calculation cost. To reduce the calculations, in accordance with Fig. 2, is used the inter-color correlation which ensures the relation $\hat{\mathcal{X}}_R \approx \hat{\mathcal{X}}_G \approx \hat{\mathcal{X}}_B$.

The main relations which represent the operations from Fig. 2, are given below:

1. For the tensor $\mathcal{X}_G$:

$$\mathcal{X}_G = \hat{\mathcal{X}}_G + \mathcal{E}_{0,G} = \hat{\mathcal{X}}_G + \hat{\mathcal{E}}_{0,G} + \mathcal{E}_{1,G}, \tag{4}$$

$$\mathcal{E}_{0,G} = \mathcal{X}_G - \hat{\mathcal{X}}_G,\ \ \mathcal{E}_{1,G} = \mathcal{E}_{0,G} - \hat{\mathcal{E}}_{0,G}. \tag{5}$$

$$\mathbf{s}_{0,G} = \ Coder\,(\mathcal{X}_G),\ \ \hat{\mathcal{X}}_G = \ Decoder(\mathbf{s}_{0,G}) = Decoder\,[Coder(\mathcal{X}_G)], \tag{6}$$

$$\mathbf{s}_{1,G} = Coder\,(\mathcal{E}_{0,G}),\ \ \hat{\mathcal{E}}_{0,G} = Decoder\,(\mathbf{s}_{1,G}) = Decoder[Coder(\mathcal{E}_{0,G})]. \tag{7}$$

In the above relations, $\mathcal{E}_{1,G}$ is the residual tensor which is calculated only in case that the tensor $\mathcal{X}_G$ should be restored on the basis of vectors $\mathbf{s}_{0,G}$ and $\mathbf{s}_{1,G}$ in layers r = 0, 1.

2. For the tensor $\mathcal{X}_R$:

$$\mathcal{X}_R = \hat{\mathcal{X}}_G + \mathcal{E}_{0,R} = \hat{\mathcal{X}}_G + \hat{\mathcal{E}}_{0,R} + \mathcal{E}_{1,R}, \tag{8}$$

$$\mathcal{E}_{0,R} = \mathcal{X}_R - \hat{\mathcal{X}}_G \quad \mathcal{E}_{1,R} = \mathcal{E}_{0,R} - \mathcal{E}_{0,R}, \tag{9}$$

$$\mathbf{s}_{1,R} = Coder(\mathcal{E}_{0,R}),\ \ \hat{\mathcal{E}}_{0,R} = Decoder\ \ (\mathbf{s}_{1,R}) = Decoder\ \ [Coder\ (\mathcal{E}_{0,R})], \tag{10}$$

$$\Delta\mathbf{s}^i_{1,R} = \mathbf{s}^i_{1,G} - \mathbf{s}^i_{1,R} \text{ for i} = 1, 2, 3, 4. \tag{11}$$

In the relations above, $\mathcal{E}_{1,R}$ is the residual tensor, which is calculated only in case that the tensor $\mathcal{X}_R$ is calculated using the vectors $\mathbf{s}_{0,R}$ and $\mathbf{s}_{1,R}$ in layers r = 0, 1.

3. For the tensor $\mathcal{X}_B$:

$$\mathcal{X}_B = \hat{\mathcal{X}}_G + \mathcal{E}_{0,B} = \hat{\mathcal{X}}_G + \hat{\mathcal{E}}_{0,B} + \mathcal{E}_{1,B}, \tag{12}$$

$$\mathcal{E}_{0,B} = \mathcal{X}_B - \hat{\mathcal{X}}_G,\ \mathcal{E}_{1,B} = \mathcal{E}_{0,B} - \hat{\mathcal{E}}_{0,B}, \tag{13}$$

$$\mathbf{s}_{1,B} = Coder(\mathcal{E}_{0,B}),\ \hat{\mathcal{E}}_{0,B} = Decoder\,(\mathbf{s}_{1,B}) = Decoder\,[Coder(\mathcal{E}_{0,B})], \tag{14}$$

$$\Delta\mathbf{s}^i_{1,B} = \mathbf{s}^i_{1,G} - \mathbf{s}^i_{1,B} \text{ for i} = 1, 2, 3, 4. \tag{15}$$

In the relations above, $\mathcal{E}_{1,B}$ is the residual tensor, which is calculated only in case that the tensor $\mathcal{X}_B$ is calculated using the vectors $\mathbf{s}_{0,R}$ and $\mathbf{s}_{1,B}$ in layers r = 0,1.

The output vectors $\mathbf{s}_{0,G}$,$\Delta\mathbf{s}^i_R$,$\Delta\mathbf{s}^i_B$ for i = 1, 2, 3, 4, are of length n_0 (for the first), and n_1 (for the remaining two). The number of decomposition layers is set depending on the size of matrices **R**, **G**, **B** which correspond to the color components. Together with this size, grows the number of 2LTSP decomposition levels.

The restoration of input tensors $\mathcal{X}_R$,$\mathcal{X}_G$,$\mathcal{X}_B$ from the output vectors, shown in Fig. 2, is executed in correspondence with the block diagram from Fig. 4. Below are given the basic relations from which are obtained the restored tensors, $\mathcal{X}'_R$,$\mathcal{X}'_G$,$\mathcal{X}'_B$:

$$\mathcal{X}'_G = \hat{\mathcal{X}}_G + \hat{\mathcal{E}}_{0,G}, \tag{16}$$

where: $\hat{\mathcal{X}}_G = Decoder(\mathbf{s}_{0,G}) = Decoder\,[Coder(\mathcal{X}_G)]$,

$$\hat{\mathcal{E}}_{0,G} = Decoder\ (\mathbf{s}_{1,G}) = Decoder\ [Coder\,(\mathcal{E}_{0,G})],$$

$$\mathcal{X}'_B = \hat{\mathcal{X}}_G + \hat{\mathcal{E}}_{0,B}, \tag{17}$$

$$\hat{\mathcal{E}}_{0,B} = Decoder\ (\mathbf{s}_{1,B}) = Decoder\ [Coder\ (\mathcal{E}_{0,B})] \tag{18}$$

$$\mathcal{X}'_B = \hat{\mathcal{X}}_G + \hat{\mathcal{E}}_{0,B}, \tag{19}$$

$$\hat{\mathcal{E}}_{0,B} = Decoder\quad (\mathbf{s}_{1,B}) = Decoder\ [Coder\ \ (\mathcal{E}_{0,B})]. \tag{20}$$

The differences between the restored and original tensors are defined by the relations:

$$\mathcal{X} - \mathcal{X}'_R = \mathcal{E}_{1,R},\ \mathcal{X} - \mathcal{X}'_G = \mathcal{E}_{1,G},\ \mathcal{X} - \mathcal{X}'_B = \mathcal{E}_{1,B} \tag{21}$$

In this case, tensors $\mathcal{E}_{1,R}$, $\mathcal{E}_{1,G}$, $\mathcal{E}_{1,B}$ represent the errors of the restored tensors $\mathcal{X}'_R$,$\mathcal{X}'_G$,$\mathcal{X}'_B$. In result of the implementation of 3D FO-AHKLT in the 2LTSP structure, the mean square error of the restored tensors is minimized.

4 Lateral Tensor Vectorization in Direction "TIME"

In accordance with Fig. 1, stages S_1 and S_2 start with the vectorization of tensors $\mathcal{X}$ and $\mathcal{E}$, with horizontal and vertical vectors orientation (i.e., directions i and j respectively). In the third stage S_3, the vectorization is lateral and the vectors are oriented at the direction k (discrete time). In order to get higher correlation between vectors components in the direction k, they must be defined so that to minimize the inter-frame differences. In case that the video camera is at a fixed position, the statistical analysis shows that the inter-frame differences are small, while for a moving camera almost all pixels are changed. In the last case, the video information in the time moment k is displaced at a relatively small distance, compared to the preceding moment, k–1. The difference between two neighbor frames could be defined by the space-time displacement vector $\mathbf{d}_{k,k-1} = [d_i, d_j]^T$, in accordance with the relation:

$$\mathbf{d}_{k,k-1}(d_i = u_0, d_j = v_0) = \min\left\{\sum_{i=0}^{M-1}\sum_{j=0}^{N-1}|x(i,j,k) - x(i-u, j-v, k-1)|\right\} \tag{22}$$

for $|u| \leq \Delta_u$ and $|v| \leq \Delta_v$, where u = 0, ± 1,.., ± a, v = 0, ± 1,.., ± b, $\Delta_u = \{-a, +a\}$, $\Delta_v = \{-b, +b\}$.

Here a and b denote the maximum displacements of the current frame k, compared to the preceding frame, k-1. In the relation above, x(i,j,k) is the value of one voxel of the tensor $\mathcal{X}$ which comprises K matrices (frames), and u_0 and v_0 are the components of the space-time displacement vector, $\mathbf{d}_{k,k-1}$.

As an example, on Fig. 5 is shown a sequence of four frames (K = 4). The first three vectors $\mathcal{X}_i = [x_{1,i}, x_{2,i}, x_{3,i}, x_{4,i}]^T$ for i = 1,2,..,M × N are obtained after lateral vectorization of the tensor $\mathcal{X}$, without videocamera translation. These three vectors are colored in black. For the case when interframe displacement exists, the same three vectors are colored in blue and the corresponding displacement vectors $\mathbf{d}_{2,1}$, $\mathbf{d}_{3,2}$, $\mathbf{d}_{4,3}$, are colored in red. The elements of the blue vectors $\mathcal{X}_i$ (for i = 1, 2, 3) represent the pixels from the sequence of four frames which have very close spatial position, and are highly correlated. In order to accelerate the search of the minimum value for the double sum from Eq. (22), could be defined a limited number of possible small displacements of the frame k contents, compared to these of the frame k-1. For example, in [10, 11] are defined 13 spatial displacements between the two sequential frames, which correspond to the values of u_0 and v_0, represented as a number of pixels: 0, ±3, ±5, ±7 (i.e. 13 values in total). These specific features of the tensor $\mathcal{F}$ vectorization in the third stage of 3D FO-AHKLT, included in the 2LTSP, ensure higher efficiency of the transform aimed at the tensor $\mathcal{X}$ decorrelation enhancement.

5 Evaluation of 2LTSP Computational Complexity

For the evaluation of the computational complexity of 2LTSP should be taken into account this of the algorithm 3D FO-AHKLT. In accordance with the results given in [9],

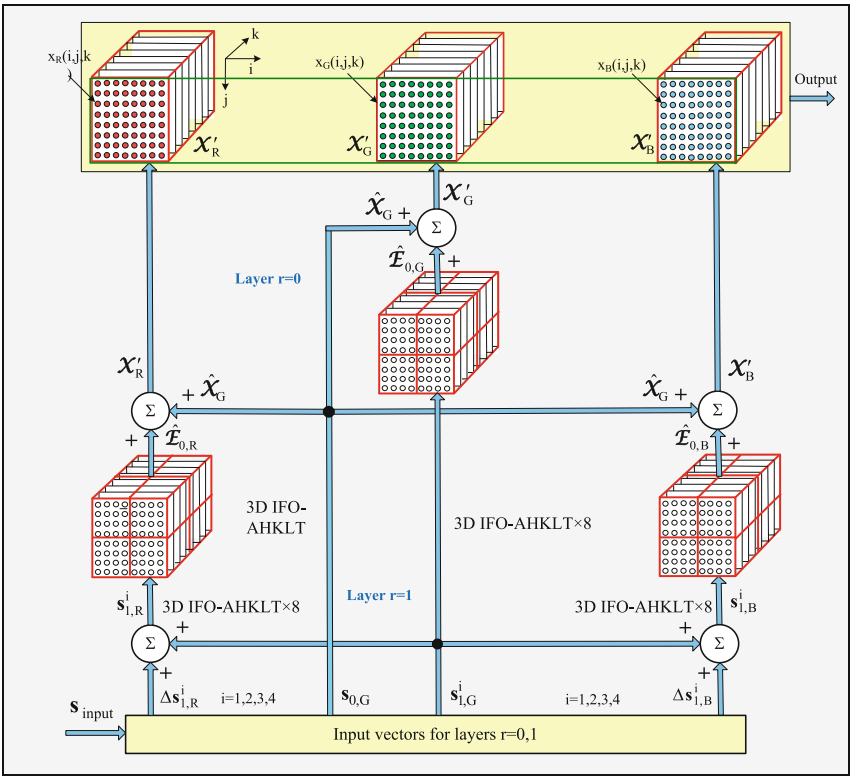


Fig. 4. Block diagram of the two-layer tensor spectrum pyramid for the restoration of tensors $\mathcal{X}_R$,$\mathcal{X}_G$,$\mathcal{X}_B$, of size $8 \times 8 \times 8$

the total number of operations needed for the execution of the direct/inverse transform of the cubical tensor $\mathcal{X}$ of size $N = 2^n$, without coefficients truncation, is:

$$O^{3D\ FO}_{AHKLT}(n) = 2^{2n+1}\left[2n\left(9 \times 2^{n-1} + 1\right) + 3\right] \tag{23}$$

The computational complexity of the 3D Adaptive Fast Walsh-Hadamard Transform (3D-AFWHT) with a kernel $WHT_{2\times2}$, for the particular case of 3D FO-AHKLT, is:

$$O_{3D\text{ - }AFWHT}(n) = 2^{2n}\left[3 \times 2^n n + 2(2n + 3)\right] \tag{24}$$

For the H-Tucker and Tensor Train (TT) decompositions of a cubical tensor $\mathcal{X}$ of size $N = 2^n$, in accordance with [9], the number of needed operations is:

$$O_{HT}(n) = 2^{3n}\left(2^{n+1} + 3\right), \quad O_{TT}(n) = 3 \times 2^{4n} \tag{25}$$

In correspondence with the block diagram of 2LTSP from Fig. 2, in Encoder-1 are executed once the direct and inverse 3D FO-AHKLT for a cubical tensor of size $N = 2^n$, 24 times—the direct 3D FO-AHKLT for cubical tensors of size $N = 2^{n-1}$, and 5

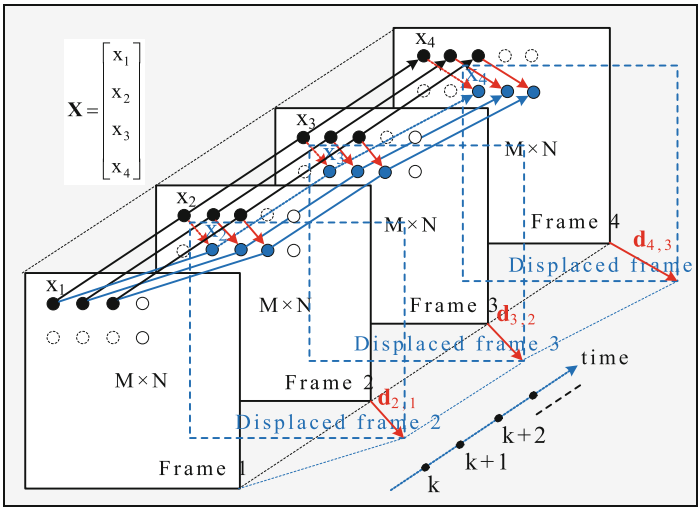


Fig. 5. Vectorization of tensor $\mathcal{X}$ in direction k (discrete time), with inter-frame displacement

times—element-by-element subtraction for the couples of cubical tensors of size $N = 2^n$. Hence, the total number of operations for the direct transform through 2LTSP, is:

$$\begin{aligned} O^{\text{2LTSP direct}}_{\text{3D FO - AHKLT}}(n) &= 2 \times 2^{2n+1}\left[2n\left(9 \times 2^{n-1} + 1\right) + 3\right] \\ &+ 24 \times 2^{2n-1}\left[2(n-1)\left(9 \times 2^{n-2} + 1\right) + 3\right] + 5 \times 2^{3n} \\ &= 2^{3n}(90n - 50) + 2^{2n+1}(5n + 24) \end{aligned} \tag{26}$$

In correspondence with the block diagram of 2LTSP for the inverse transform, shown on Fig. 4, are executed: one inverse 3D FO-AHKLT for a cubical tensor of size $N = 2^n$, 24 times—the inverse 3D FO-AHKLT for cubical tensors of size $N = 2^{n-1}$, and four times—element-by-element subtraction for the couples of cubical tensors of size $N = 2^n$. Hence, the total number of operations for the inverse transform through 2LTSP, is:

$$\begin{aligned} O^{\text{2LTSP inverse}}_{\text{3D FO - AHKLT}}(n) &= 2^{2n+1}\left[2n\left(9 \times 2^{n-1} + 1\right) + 3\right] \\ &+ 24 \times 2^{2n-1}\left[2(n-1)\left(9 \times 2^{n-2} + 1\right) + 3\right] + 4 \times 2^{3n} \\ &= 2^{3n}(72n - 50) + 2^{2n+1}(14n + 9) \end{aligned} \tag{27}$$

From the above relations it is easy to notice that the computational complexities of 2LTSP for the direct and inverse transforms are close, and the complexity of 2LTSP for the inverse transform is a little lower. In case that in the block diagram of 2LTSP for the direct transform the 3D FO-AHKLT is replaced by the 3D-AFWHT, the relations (26) and (27) are modified as follows:

$$O^{\text{2LTSP direct}}_{\text{3D - AFWHT}}(n) = 2 \times 2^{2n}\left[3 \times 2^n n + 2(2n + 3)\right]$$

$$+ 24 \times 2^{2n-2}\left[3 \times 2^{n-1}(n-1) + 2(2n+1)\right] + 5 \times 2^{3n}$$
$$= 2^{3n}(15n-4) + 2^{2n+1}(74n+9) \tag{28}$$

$$O_{3D\text{ - AFWHT}}^{2LTSP\ inverse}(n) = 2^{2n}\left[3 \times 2^{n}n + 2(2n+3)\right]$$
$$+ 24 \times 2^{2n-2}\left[3 \times 2^{n-1}(n-1) + 2(2n+1)\right] + 4 \times 2^{3n}$$
$$= 2^{3n}(11n-4) + 2^{2n+1}(14n+14) \tag{29}$$

For the decomposition of same cubical tensors $\mathcal{X}_R$,$\mathcal{X}_G$,$\mathcal{X}_B$, of size $N = 2^n$ through H-Tucker and Tensor Train (TT), the needed number of operations is correspondingly:

$$O_{HT}^{RGB}(n) = 3 \times 2^{3n}\left(2^{n+1}+3\right), \quad O_{TT}^{RGB}(n) = 9 \times 2^{4n} \tag{30}$$

Then, from Eqs. (28) and (30) follow the relations:

$$y_1(n) = \frac{O_{HT}^{RGB}(n)}{O_{3D\text{ FO - AHKLT}}^{2LTSP\ direct}(n)} = \left[3 \times 2^{n-1}\left(2^{n+1}+3\right)\right] / \left[2^n(45n-25) + 5n + 24\right] \tag{31}$$

$$y_2(n) = \frac{O_{TT}^{RGB}(n)}{O_{3D\text{ FO - AHKLT}}^{2LTSP\ direct}(n)} = \left[9 \times 2^{2n-1}\right] / \left[2^n(45n-25) + 5n + 24\right] \tag{32}$$

Correspondingly, from (29) and (30) are obtained the relations:

$$y_3(n) = \frac{O_{HT}^{RGB}(n)}{O_{3D\text{ - AFWHT}}^{2LTSP\ direct}(n)} = \left[3 \times 2^{n}\left(2^{n+1}+3\right)\right] / \left[2^n(15n-4) + 2(74n+9)\right] \tag{33}$$

$$y_4(n) = \frac{O_{TT}^{RGB}(n)}{O_{3D\text{ - AFWHT}}^{2LTSP\ direct}(n)} = \left[9 \times 2^{2n}\right] / \left[2^n(15n-4) + 2(74n+9)\right] \tag{34}$$

Following the relations (31), (32), (33) and (34), was prepared Table 1. From the results obtained for $n = 3, 4, \ldots, 10$ it follows that for $n \geq 4$, the values of relations $y_3(n)$ and $y_4(n)$ are higher than 1, while the values of $y_1(n)$ and $y_2(n)$ get higher than 1 for $n \geq 7$.

Table 1. Evaluation of the computational complexity 2LTSP direct/3D FO-AHKLT and 2LTSP direct/3D-AFWHT compared to HT/RGB and TT/RGB decompositions

n	3	4	5	6	7	8	9	10
$y_1(n)$	0.23	0.33	0.50	0.79	1.33	2.30	4.05	7.24
$y_2(n)$	0.29	0.45	0.71	1.17	1.98	3.44	6.06	10.84
$y_3(n)$	0.39	1.11	2.12	3.92	7.11	11.38	23.05	41.72
$y_4(n)$	0.48	1.53	3.04	5.75	10.54	16.98	34.48	62.49

For the case, when part of the high-frequency coefficients of 3D FO-AHKLT (respectively, 3D-AFWHT) are "truncated", the values of the relations given above, increase. Together with the growth of the levels n of the transforms 3D FO-AHKLT (respectively, 3D-AFWHT), the computational complexity of 2LTSP decreases faster than these of H-Tucker and TT decompositions. From the analysis of Eqs. (30), (31), (32) and (33) it follows that the efficiency of 2LTSP increases for tensors of big size (Fig. 6).

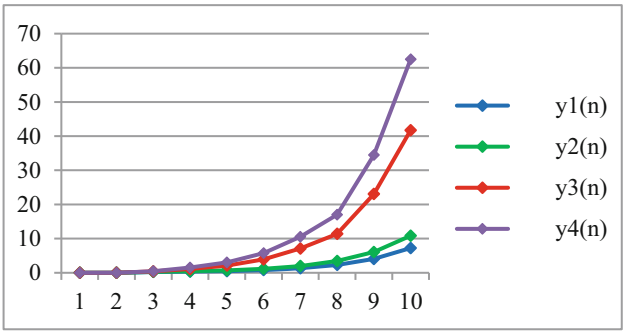


Fig. 6. Relative complexity of 2LTSP direct/3D FO-AHKLT and 2LTSP direct/3D-AFWHT compared to HT/RGB and TT/RGB decompositions

6 Conclusions and Future Work

The main advantages of the 2LTSP decomposition are generalized below as follows:

1. The objects features' space is reduced by using the inter-frame correlation for the processed sequence of tensors (video frames). The error, which defines the approximation accuracy of the restored tensors, depends on the number of retained coefficients n, calculated by the Encoder in each pyramid level;
2. The new decomposition has relatively low computational complexity, which permits real-time performance. The implementation of 2LTSP based on 3D-AFWHT as a particular case of 3D FO-AHKLT, results in significant reduction of the computational complexity, compared to H-Tucker and TT decompositions;
3. The ability to handle a longer input sequence, exploration of temporal dependencies and better performance in case of partially observable scenes. The related statistical analysis estimates that the inter-frame correlation at 25 frames per second usually covers a time interval of up to 12 consecutive frames.

The future work on the offered pyramidal structure will be aimed at investigations on the abilities for its application in various areas, such as the information redundancy reduction in color video, the extraction of a minimum number of features needed for objects recognition, etc.

Acknowledgement. This work was funded by the Bulgarian National Science Fund: Project No. KP-06-H27/16: "Development of efficient methods and algorithms for tensor-based processing and analysis of multidimensional images with application in interdisciplinary areas".

References

1. Panagakis, Y., Kossaifi, J., Chrysos, G., Oldfield, J., Nicolaou, M., Anandkumar, A., Zafeiriou, S.: Tensor methods in computer vision and deep learning. In: Proceedings of the IEEE. arXiv: 2107.03436v1 [cs.CV] 7 Jul (2021)
2. De Lathauwer, L., De Moor, B., Vandewalle, J.: A multilinear singular value decomposition. SIAM J. Matrix Anal. Appl. **21**(4), 1253–1278 (2000)
3. Oseledets, I.: Tensor-train decomposition. SIAM J. Sci. Comput. **33**(5), 2295–2317 (2011)
4. Grasedyck, L.: Hierarchical singular value decomposition of tensors. SIAM J. Matrix Anal. Appl. **31**, 2029–2054 (2010)
5. Ozdemir, A., Zare, A., Iwen, M., Aviyente, S.: Extension of PCA to higher order data structures: an introduction to tensors, tensor decompositions, and tensor PCA. Proc. IEEE **2018**(106), 1341–1358 (2018)
6. Zniyed, Y., Boyer, R., Almeida, A., Favier, G.: A TT-based hierarchical framework for decomposing high-order tensors. SIAM J. Sci. Comput. **42**, A822–A848 (2020)
7. Rao, K., Kim, D., Hwang, J.: Fast Fourier Transform: Algorithms and Applications. Springer, Dordrecht, The Netherlands (2010)
8. Kountchev, R., Mironov, R., Kountcheva, R.: Complexity estimation of cubical tensor represented through 3D frequency-ordered hierarchical KLT. MDPI, Symmetry (Open Access) **12**, 1605 (2020)
9. Kountchev, R., Mironov, R., Kountcheva, R.: Third-order tensor decorrelation based on 3D FO-HKLT with adaptive directional vectorization. MDPI, Symmetry (Open Access) **14**, 854 (2022)
10. Lee, D., Ko, H., Kim, J., Bovik, A.: On the space-time statistics of motion pictures. JOSA A **38**(7), 908–923 (2021)
11. Chen, M., Goodall, T., Patney, A. Bovik, A.: Learning to Compress Videos without Computing Motion. arXiv:2009.14110v3 [eess.IV] 27 March (2022)

Electromyography Signal Acquisition, Processing, Optimization and Its Applications

Amol Kumar[1(✉)], Manoj Duhan[1], and Poonam Sheoran[2]

[1] Department of Electronics and Communication Engineering, Deenbandhu Chhotu Ram University of Science and Technology, Murthal, Sonepat, Haryana, India
kumaramol25@gmail.com

[2] Department of Biomedical Engineering, Deenbandhu Chhotu Ram University of Science and Technology, Murthal, Sonepat, Haryana, India

Abstract. With the ever-increasing role of biomedical signals in the field of Science and technology, electromyogram approach is considered as important technique using EMG signals to monitor muscular activities for stress detection abnormalities and activation level and to study the biomechanics of various movements. EMG signal acquisition and the processing part are being updated day by day in terms of accuracy and artifact removal which makes the analyses part more reliable. This paper discusses the efficient EMG signal acquisition, processing, feature extraction, classification and optimization methods to attain high recognition accuracy using EMG signals.

Keywords: EMG · Signals · Acquisition

1 Introduction

EMG is one of the physiological signal acquisition methods to record electrical activity of skeletal muscle. It reveal complex forms of anatomical features [1]. The signal indicates muscle activity under different stimuli [2]. Abnormal muscle activity, which gives rise to large number of diseases due to poor condition, can be easily detected using EMG signals. It is also used to measure stress, i.e., physical stress while doing muscular exercises [3]. EMG is a reliable measure for physical stress test [4], as it measures muscle activity (contraction), amplitude and frequency. When we acquire an EMG signal it consists of noise, so different techniques such as anti-aliasing filters etc. have to be used to remove the noise artifacts. A novel approach known as deep learning is used in medical imaging which comes from a big data in an end-to-end learning manner to medical tasks.

2 Electromyogram Signal Initiation

EMG is an electrical activity of muscles. EMG signal is generated by contracting the muscles. EMG can be measured with the help of electrodes. At the time of contraction of muscles, charges particles are generated in muscle membrane. The movement of charge

R. Kountchev et al. (Eds.): *NAMSP 2022*, SIST 332, vol. 332,, pp. 44–55, 2023.
https://doi.org/10.1007/978-981-19-7842-5_5

particle allows electrical current generation. The electrical current generated inside the muscles converted into electric energy. The difference between two point of the electrode can be calculated as voltage. The voltage we get on surface of muscles is affected by resistance of surrounding muscles. Figure 1 shows contraction of muscles information (A) gives information about the 10% maximum voluntary contraction that shows low power level of sEMG signal (B) give information about different motor unit firing rates (C) give information about isolated sEMG in first vertebral muscle (D) increase in duration activity by decrease in muscle fiber transfer speed EMG signal acquisition with the help of bipolar electrodes [4].

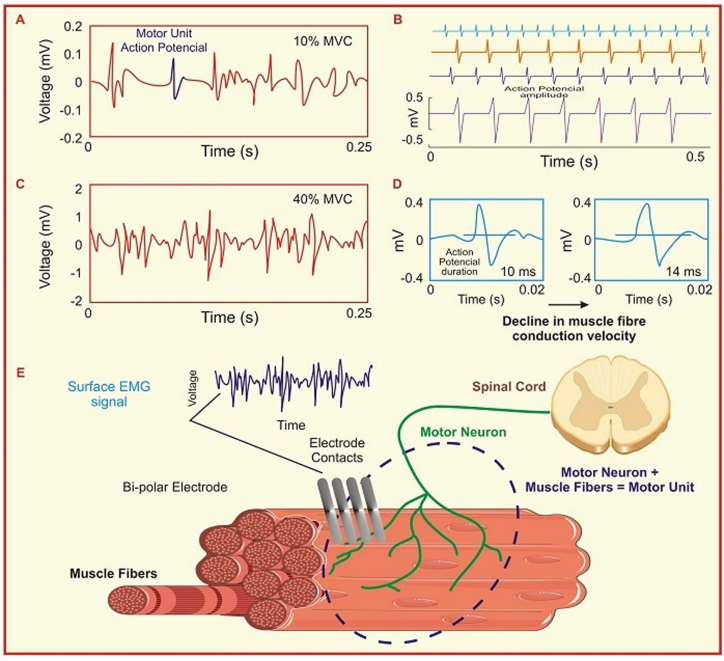


Fig. 1. Isolated surface EMG signal at low power level (10% of maximum voluntary contraction, MVC) (A) high power level, 40% MVC (B) different motor unit firing rates (C), isolated surface EMG signal in the first vertebral muscle and (D) increased duration activity with a decrease in muscle fiber transfer speed (E) EMG signal from a muscle using a bipolar electrode.

Transmission of signals to muscles occurs by active motor-neurons (discharged) as shown in Fig. 1A [5–7]. Every motor unit has a different size of action potential (MUAP) wave, Fig. 1B, showing the potential of the recorded power over time. Action potential size depends on characteristics of motor unit and position of muscle fibers depends on the recording electrode [8]. Action potential (functional capacity) of all active motor units generates approximate EMG signals on the skin surface. The motor units (individual) firing time can be deduced from recorded sEMG signals using special sEMG decay mechanisms [9].

3 EMG Signal Acquisition

Electrodes are used for acquiring EMG signal, i.e., surface (non-invasive) and needle (invasive) electrodes. In Non-invasive/surface method, electrodes and sensors are placed on the muscle surface while in needle method; the muscle is penetrated by electrodes (Fig. 2).

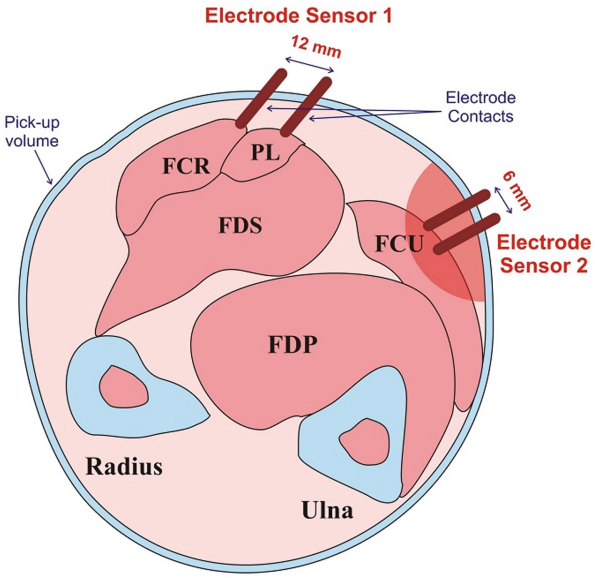


Fig. 2. Schematic view of forehead of different muscles.

The size of the needle electrodes is 1 sq. mm wide approx. The abnormalities in muscle contraction, nerve damage, muscle or brain damage, physical stress cause motor disorders that can be identified by an EMG signal diagnosis [10]. The principle behind the use of surface electrodes is the detection of chemical balance between body surface and skin by electrolytic communication consists of two types—Gelled and Dry EMG electrode. In surface EMG, we put surface electrodes on the cleaned part of skin and are 0.5–2.5 cm wide. These are based on non-invasive techniques to measure and capture the EMG signals. When the muscle is in a relaxed state, the EMG plot shows either more/less noise or less EMG baseline of muscle although it should not exceed 3–5 microvolt's although 1–2 is achievable. Interference noise or trouble detection devices can result in increased base activity or muscle (Figs. 3 and 4).

Motor unit action potentials (MUAPs) also known as raw EMG signals are the basis for any investigation of neurological and neuromuscular problems. MUAP is total contraction of muscle fiber in motor units and are easily detected near electrodes, resulting in lower MU distances and more noise. A group of MUAPs resulting from MUU is

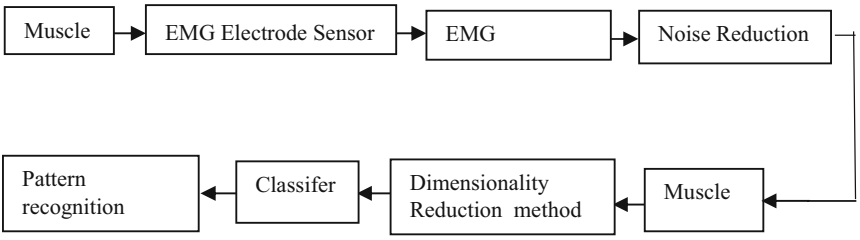


Fig. 3. Diagrammatic representation of acquiring EMG signals.

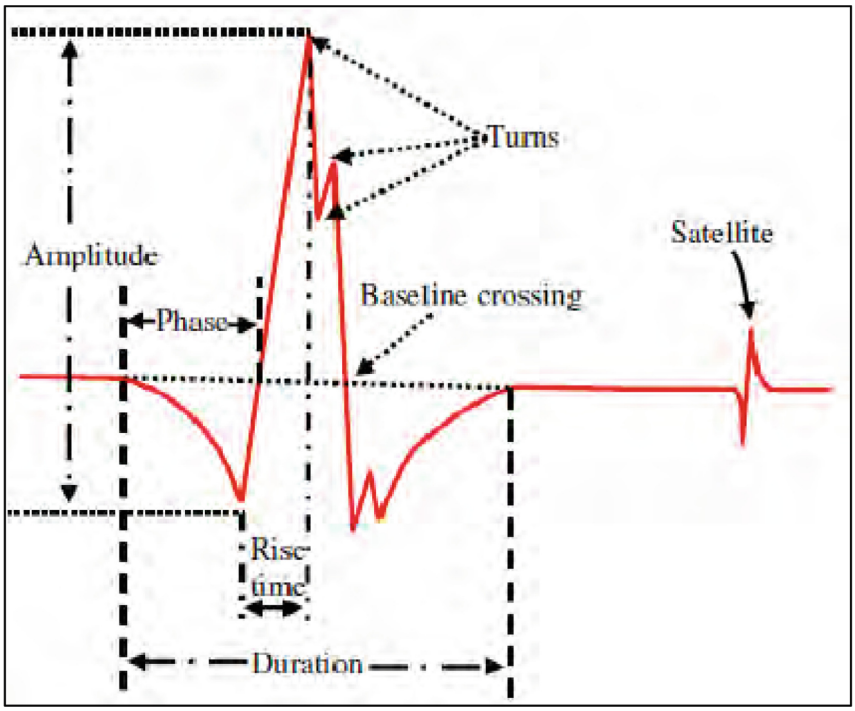


Fig. 4. The individual wave form of EMG signal.

called MUAPT, received signals from single electrode can be specified as follows [11].

$$\text{EMGt} = \sum_{m=1}^{N} \text{MUAPT}m(t) + n(t) \tag{1}$$

In above equation, MUAPT $m(t)$ refers to transient function of MUAPT, $n(t)$ function indicates parts of signal that are neither connected to motion capacity. The characteristics of this sign depend upon size, contraction duration, muscles steady state, lethargy and sweating of the skin (Table 1).

Table 1. MUAP abnormalities and anatomical changes [12].

MUAP abnormalities	Physical changes
Increase in amplitude	Muscle fibers loss and increase in connective tissues
Decrease in amplitude	Muscles fibers group
Decrease duration	Muscle fibers loss
Increased duration	Muscle fibers increases
Increased spike duration	Muscle diameter varies and increase in thickness at ends
Increase in number of turns and phases	Slow conduction of terminal axons/and increased diameter of muscle fiber
Increase in firing rate	Loss of motor units
Increase in the jiggle	Neuromuscular transmission occurs
Decrease in amplitude	Muscles fibers group

4 Factors Affecting EMG Signals

The EMG signals amplitude range lies from 0–10 mV. While acquiring the EMG signals noise affects the electrical signals, so, it is very essential to use high quality equipments to reduce noise. A lot of factors affect the EMG signals as [13] mentioned below:

(1) Inherent noise in electronic equipment: Till date no equipment is available which does not generate noise. Only we can reduce noise by using good-quality tools.
(2) Ambient noise: Source is electromagnetic radiation. It is impossible to avoid this noise as being living on the surface of the earth.
(3) Motion Artifact: Motion artifacts skewed the information; therefore, it leads to irregularity in the data. It is caused either due to the electrode interface or cable. By designing good-quality electronic devices this artifact can be reduced.
(4) Signal Instability: EMG signal is within 0–20 Hz frequency considered as unstable, so this range should be avoided.
(5) Intrinsic: It includes active motor unit, fiber diameter, fiber composition, depth of electrodes between muscle and tissue surface.
(6) Extrinsic: It includes factors such as surface area, electrode shape, distance of electrodes, muscle location etc.

4.1 Essential Requirements for Acquisition of EMG Signals

For acquiring high EMG signal, signal distortion should be as low as possible to avoid unwanted filters. For analysis of EMG signal positive value is analyzed so that half-wave rectifier can be used.

4.2 EMG Signal Detection

The methods available for detection of EMG signal includes-

1. Single-threshold/limit method: It is used for comparing EMG signals having fixed threshold and is observed by computer-based technique. Now days this method is not satisfactory since measured results depends upon fixed threshold value, i.e., the relationship between the probability of detection P_{dk} and the probability P_{γ} that a noise sample is above the threshold γ, is expressed as equation given below [13].

$$P_{dk} = \exp\left(\frac{In\left(P_{\gamma}\right)}{1 + 10^{SNR/10}}\right) \tag{2}$$

   The above equation was also supported by Reema and Vijay [14].
2. Double-threshold/limit method: It was proposed by Bornato in 1998, and detects higher detection probability as compared to a single threshold. Lanyi and Adler [15] suggested Bornato method as complex, expensive, therefore Lanyi and Adler investigated new algorithm (improved method) which provides high sensitivity, consistent and efficient muscle-to-off detection with low computational costs.

4.3 EMG Signal Decomposition

EMG signal is decomposed to study the underlying mechanisms of nerve control which is done with wavelet spectrum and PCA of wavelet coefficients. Wavelet transform include other decomposition methods like EMD/ICA. EMD (empirical mode decomposition) is an effective way to decompose signals. Zvokelj et al. [16] combined different methods that can distort the signal into different time scales. EMG decomposition study was carried out by various scientists and researchers. Among them most of the methods are restricted to laboratories research, specialized equipments, complex algorithms and large-scale data to train the standard models [17].

4.4 EMG Signal Processing

The raw signals offer useful information, can be quantified [18]. This information can be converted into useful and accurate form of signals [19] with the help of various signal processing methods such as (Fig. 5)-

1. Wavelet Analysis/Transform: It is represented by Greek letter psi (ψ). Wavelets are used for different modes of operation and to decompose and reconstruct primary signal. The generated wavelet components are used to remove noise at each scale. Luo and Shen et al. [20] investigated reconstruction algorithm to reduce/eliminate the noise of surface EMG. Gaofeng et al. [21] proposed MODWT algorithm for noise filtering.
2. The main advantages of this wavelet transform are that the windowed Fourier transform can be overcome by scaling the bandwidth of the filter, which is inversely related to the frequency. Secondly in a signal, small wavelets isolate very fine details and very large wavelets isolate coarse details.

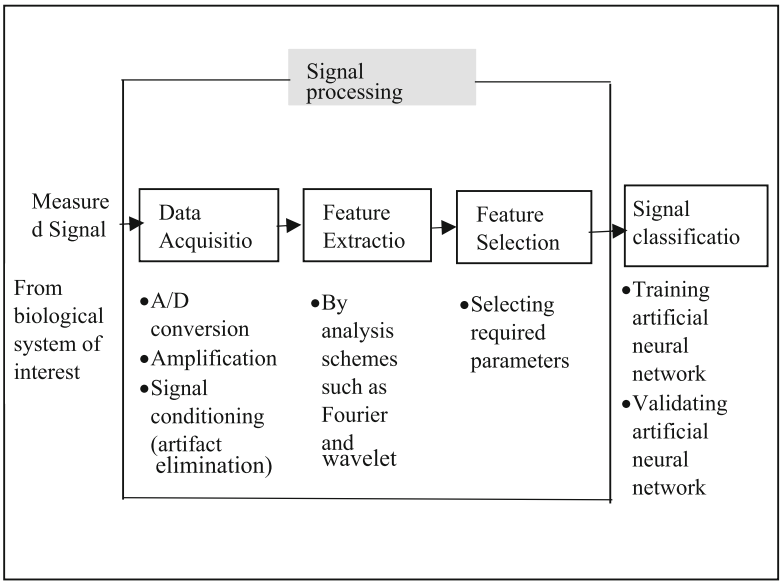


Fig. 5. Signal processing stages.

3. Time-frequency analysis: This method used for analyzing the sEMG signals with FT and wavelet analysis theory. Chan et al. [22] applied theory of wavelet analysis for the extraction of time-frequency eigen values (Table 2).
4. Autoregressive model: This algorithm recognizes the factor of the time sequence which is recorded according to an AR model, different models are mentioned in Table 3.
5. Artificial intelligence: It includes an artificial neurological network (ANN) which is used to classify a variety of signals created by medical images and biomedical acoustic signals. Furthermore used in identifying myopathy with an accuracy level above 87%, it also supports Markov Model.

Linear discriminant analysis and support vector machines. Ganesh [33] performed brachial biceps EMG with EMD. Venkatraman [34] propounded ANN model for more accuracy in classifying dependent variables as compared to logistic regression model. Ram et al. [35] proposed SVM classifiers to be more accurate than ANN in term of speed and well-being (Fig. 6).

4.5 Feature Extraction and Classification

Image processing, machine learning are recently used for detecting any normal or abnormal pattern of any disorder. Machine learning is comprised of supervised and unsupervised learning. In supervisory learning, good outputs are obtained from input and output data using predictive models. Some popular algorithms comprise of decision trees, SVM, ANN, and LDA. Xiuwu et al. [36] suggested a PSO-SVM for identification of

Table 2. List of signal processing methods with features and accuracy.

Authors	Features	Feature selection method	Classifier	Movement	Accuracy (%)
Hichem et al. [23]	Discrete wavelet transform, Wavelet Cepstral Coefficient (WCC) coefficients		WCC-ANN	Lower limb muscles	100
Sara et al. [24]	TD, FD and TFD	FS/MLE	KNN, MLP	Upper limb movements	97
Annachiara et al. [25]	Localization of maximum signal energy	Continuous Wavelet Transform (CWT)		Muscle recruitment during walking	98
She et al. [26]	The complex Morlet wavelet	TLLD	LDA	Wrist and hand flexion and extension	98
Satish et al. [27]	Amplitude mean and maxima of wavelet coefficients	Discrete wavelet transform	ANN	Closed and open palm, wrist extension	93.25
Ercan et al. [28]	Mean, standard deviation, average power	DWT	Random forest decision	Use of DWT and MSPCA for identifying neuromuscular disorders	96.67
Yipeng et al. [29]	Maximum value, multi-scale value and singular values of wavelet coefficients	DWT, multiple mother wavelets	BPNN	(1) Higher and lower step (2) Ascending and descending	98.7
Firas et al. [30]	Wavelet coefficients energies	DWT	GRNN	Eight hand motions	95
Zhang et al. [31]	Energy content ratio of every wave multiplier	WPT	BPNN	Forearm and elbow rotation	95
Muthusamy et al. [32]	AR model coefficient and wavelength	DWT	PNN and GRNN	Wrist motions	99

upper limb movements using EMG signal and Wavelet Packet Transform, they obtained 90.66% recognition rate. Though, PSO-SVM technique give accuracy of 96% [37].

Table 3. List of different autoregressive model.

Feature	Extraction	Application	Strength	Weakness	Suitable	Classification	Accuracy (%)
Fourier transform	Time	Frequentness	Time shift	Variable	Small changes	Neural network	92.47
Wavelet transform	Time	Frequency	Good frequency	Localization	Complicated correlation	ANN (Artificial Neural Network)	92.47
Autoregressive	Time	Frequency	Classify the surface and intra muscle	Not being dependent on training data size		QDF (Quadratic Discriminate Function)	92.36
PSE frequency			Optimized for the calibration	No unique feature	Extraction	ANN (Artificial Neural Network)	94.00
SLEX	Time evolved	Spectral analysis	Overcome the shortcoming of conventional Fourier based spectral analysis	Library components		ANN (MLP)	98.00

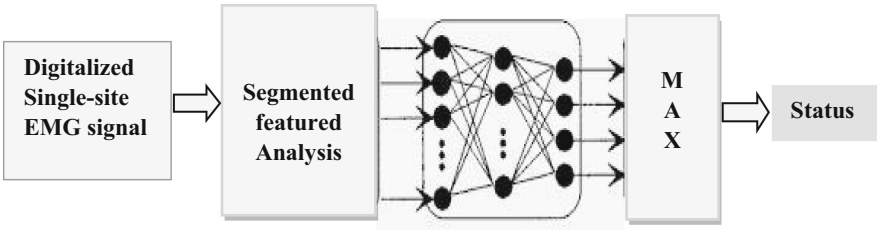


Fig. 6. Electromyogram classification by ANN approach.

4.6 Optimization of the EMG Signals

It can be done with various methods such as Particle Swarm Optimisation (PSO), K-Nearest Neighbor (KNN) and Support Vector Machine (SVM). Among them, PSO is the best method for feature extraction, though KNN method and SVM technique shows 99.7 and 99.6% accuracy [38].

5 Applications of EMG

There are many applications for using EMG such as neuro-muscular physiology, medical diagnosis, and muscle stress management. When muscle activity begins under any force, the accuracy of EMG signals is affected by some physical factors including, age of the person, temperature, muscle examined and technical factors such as type of needle electrode, characteristics of recording surface, preamplifier and amplifier, etc. The above factor reduces robustness as well as the accuracy of recognition. Hence, surface electromyography (EMG), which records changes in muscle stimulation [39], has become an attractive method for assessing muscle metabolic reactions during exercise and muscle fatigue, in speech recognition, robotics and in diagnosis.

6 Conclusions and Future Scope

The present article covered efficient methods which are widely employed in sEMG data analysis. For EMG signal acquisition, best method is “improved method”, and for processing is “DWT and WCC”, while for feature extraction is “DWT” and for optimisation is “PSO technique”. These method is/are considered as best to attain high recognition accuracy that can efficiently capture EMG signals. Furthermore, a deep-learning approach is useful for bringing signal analysis for medical applications as futuristic research work.

References

1. Rajat, E.S., Iqbal, K., White, G., Holtz, J.K.: A review of EMG techniques for detection of gait disorders. In: Artificial Intelligence—Applications in Medicine and Biology, pp. 1–23 (2019)
2. Peter, K.: The ABC of EMG: A Practical Introduction to Kinesiological Electromyography. NoeraxonInc USA (2005)
3. Ayten, O.A., Niraj, K.J.: Keep the stress away with SoDA: stress detection and alleviation system. IEEE Trans. Multi-Scale Comput. Syst. **3**, 269–282 (2017)
4. Lara, M.M., Giuseppe, D.V., Madeleine, M.L.: Analysis and biophysics of surface EMG for physiotherapists and kinesiologists: toward a common language with rehabilitation engineers. Front. Neurol. (2020). https://doi.org/10.3389/fneur.2020.576729
5. Kamen, G., Caldwell, G.E.: Physiology and interpretation of the electromyogram. J. Clin. Neurophysiol. **13**, 366–384 (1996)
6. Toshio, M., Jan, D.S., Roberto, M.: Basic physiology and biophysics of EMG signal generation. In: Electromyography Physiology Engineering and Noninvasive Applications, pp. 1–25. IEEE Press Piscataway, NJ (2004). https://doi.org/10.1002/0471678384.ch1
7. Dario, F., D., Roberto, M., Dick, F.S.: Biophysics of the generation of EMG signals. In: Electromyography Physiology Engineering and Noninvasive Applications, pp. 81–105. IEEE Press Piscataway, NJ (2004). https://doi.org/10.1002/0471678384.ch4
8. Javier, R.F.: Understanding the electrical behavior of the action potential in terms of elementary electrical sources. Adv. Physiol. Educ. **39**, 15–26 (2015)
9. Bert, U.K., Johannes, P.V.D., Machiel, J.Z., Dick, F.S.: Inter-operator agreement in decomposition of motor unit firings from high-density surface EMG. J. Electromyogr. Kinesiol. **18**, 652–661 (2008)
10. Aldo, O.P.: Anatomical Guide for the Electromyographer: The Limbs and Trunk, 5th edn. Charles C Thomas Pub Ltd (2011)
11. Jamileh, Y., Andrew-Hamilton, W.: Characterizing EMG data using machine-learning tools. Comput. Biol. Med. **51**, 1–13 (2014)
12. Yong, N., Xiangjun, Z., Shanan, Z., Yingchun, Z.: Surface EMG decomposition based on K-means clustering and convolution kernel compensation. IEEE J. Biomed. Health Inform. **19**(2), 471–477 (2015)
13. Reaz, M.B.I., Hussain, M.S., Mohd-Yasin, F.: Techniques of EMG signal analysis: detection, processing, classification and applications. Biol. Proced. Online **8**, 11–35 (2006)
14. Reema, J., Vijay, K.G.: Review of electromyography signal with detection, decomposition, features and classifier. Theor. Int. J. Comput. Sci. Eng. **7**(5), 487–500 (2019)
15. Lanyi, X., Andy, A.: An improved method for muscle activation detection during gait. In: Canadian Conference of Electrical and Computer Engineering, vol. 1, pp. 357–360 (2004)
16. Zvokelj, M., Zupan, S., Prebil, I.: EEMD-based multiscale ICA method for slewing bearing fault detection and diagnosis. J. Sound Vib. **370**, 394–423 (2016)
17. Ales, H., Dario, F., Marco, G., Roberto, M., Damjun, Z.: Estimating motor unit discharge patterns from high-density surface electromyogram. Clin. Neurophysiol.: Off. J. Int. Fed. Clin. Neurophysiol. **120**(3), 551–562 (2009)
18. Manoj, D., Chanderpal, S., Dinesh, B.: Study of signal processing techniques for EMG analysis. Int. J. Biomechatronics Biomed. Robot. **1**(3), 141–148 (2011)
19. Jack, K.: Wavelet Analysis and Classification Surface Electromyography Signals. Auckland University of Technology, Auckland, New Zealand (2005)
20. Luo, Z.Z., Shen, H.X.: Electronic signal denoising method based on Hermite interpolation for wavelet modulus maximum reconstruction. J. Electron. Inf. Technol. **31**(4), 857–860 (2009)

21. Gaofeng, W., Feng, T., Gang, T., Chengtao, W.: A wavelet-based method to predict muscle forces from surface electromyography signals in weight lifting. J. Bionic Eng. **9**(1), 48–58 (2012)
22. Chan, A.D.C., Englehart, K., Hudgins, B., Lovely, D.F.: Myoelectric signals to augment speech recognition. Med. Biol. Eng. Comput. **39**, 500–504 (2001)
23. Hichem, B., Abdenour, H.G., Philippe, R., Karim, A.M., Olivier, B.: Surface EMG signal classification for Parkinson's disease using WCC descriptor and ANN classifier. In: Proceedings of the 10th International Conference on Pattern Recognition Applications and Methods, pp. 287–294. Science and Technology Publications, Lda (2021)
24. Abbaspour, S., Lindén, M., Gholamhosseini, H., Naber, A., Ortiz-Catalan, M.: Evaluation of surface EMG-based recognition algorithms for decoding hand movements. Med. Biol. Eng. Comput. **58**(1), 83–100 (2019). https://doi.org/10.1007/s11517-019-02073-z
25. Annachiara, S., Federica, V., Laura, B., Sandro, F., Francesco, D.N.: Time-frequency analysis of surface EMG signals for maximum energy localization during walking. In: EMBEC and NBC 2017, pp. 978–981. Springer, Singapore (2017)
26. She, Q.S., Ma, P.G., Ma, Y.L., et al.: EMG feature extraction method based on tensor linear Laplacian discriminant. J. Southeast Univ. **47**(6), 1117–1121 (2017)
27. Satish, M., Ruta, K., Faruk, K., Nongmaithem, M.S.: Hand motion recognition from single channel surface EMG using wavelet and artificial neural network. Procedia Comput. Sci. **49**, 58–65 (2015)
28. Ercan, G., Abdulhamit, S.: Comparison of decision tree algorithms for EMG signals classification using DWT. Biomed. Signal Process. Control **18**, 138–144 (2015)
29. Yipeng, Y., Lin, S., Feida, Z., et al.: Multi-feature fusion sEMG pattern recognition based on wavelet transform. J. Transduct. Technol. **29**(4), 512–518 (2016)
30. Al Omari, F., Hui, J., Mei, C., Liu, G.: Pattern recognition of eight hand motions using feature extraction of forearm EMG signal. Proc. Natl. Acad. Sci., India, Sect. A **84**(3), 473–480 (2014). https://doi.org/10.1007/s40010-014-0148-2
31. Zhang, H.K., Dawei, W., Yang, H., et al.: Surface EMG signal acquisition and motion recognition system. Mech. Des. Manuf. **3**(8), 38–40 (2013)
32. Muthusamy, H., Chong, Y.F., Sindhu, R., Bukhari, I., Sazali, Y.: A comparative study of wavelet families for classification of wrist motions. Comput. Electr. Eng. **38**(6), 1798–1807 (2012)
33. Ganesh, N., Suviseshamuthu, E.S., Hung, T.N.: Single-channel EMG classification with ensemble-empirical-mode-decomposition-based ICA for diagnosing neuromuscular disorders. IEEE Trans. Neural Syst. Rehabil. Eng. **24**(7), 734–743 (2016)
34. Venkatraman, R.: Overview of artificial neural network models in the biomedical domain. Bratisl. Med. J. **120**(7), 536–540 (2019)
35. Ram, M.S., Vivek, A., Sangeeta, C., Amod, K.: Comparative analysis of SVM and ANN classifier based on surface EMG signals for elbow movement classification. J. Interdiscip. Math. 153–161 (2020)
36. Xiuwu, S., Kelvin, W., Yun, Z.: Pattern recognition of SEMG based on wavelet packet transform and improved SVM. Optik **176**, 228–235 (2019)
37. Reema, J., Vijay, K.G.: Review of EMG signal classification approaches based on various feature domain. Matter: Int. J. Sci. Technol. **6**(3), 123–143 (2021)
38. Virendra, P.M., Prashant, K., Suman, H.: Optimisation and classification of EMG signal using PSO-ANN. In: Devices for Integrated Circuit (DevIC), Kalyani, India, pp. 191–195 (2019)
39. Andrew, V., Israel, H., Gregory, L., Gabriel, S.T., Taian, M.V.: Interpreting signal amplitudes in surface electromyography studies in sport and rehabilitation sciences. Front. Physiol. **8**(985), 1–15 (2018)

Research on the Radar Signal Classification Method Based on the Deep Faith Network Model

Tianren Sun and Wenhao Dong(✉)

National University of Defense Technology, Beijing, China
dongwenhao19a@nudt.edu.cn

Abstract. In the development of machine learning, deep learning, as a research content that researchers focus on, has a relatively strong ability to represent data. Its core basis is to build a neural network with multi-hidden layer structure and a large number of data, so as to obtain the features contained in the data and improve the accuracy of classification prediction. Radar signal classification is an important technology for radar signal processing. After the emergence of a new radar system, the application advantages of traditional signal classification methods are becoming less and less. However, combined with deep learning, signal classification can be optimized on the basis of automatic learning of data characteristics. Therefore, on the basis of understanding the current research status of signal classification in radar countermeasures, this paper conducts in-depth research on the technical concept of deep learning, and thus proposes a radar signal classification method based on deep belief and network model.

Keywords: Deep learning · Deep training network · Radar · Signal classification

1 Introduction

As an important part of information warfare, the core task of electronic reconnaissance is to obtain the enemy radar signal and obtain the corresponding parameter characteristics to complete the signal risk. However, in the current military industry and electronic technology innovation and development, more and more new radar systems have been widely used in the battlefield electromagnetic environment, which not only aggravates the complexity of practical development, but also affects the application performance of signal classification and identification, and has a negative impact on the strategic decision of self defense and ground attack. Therefore, how to accurately classify and recognize different types of radar signals under complex conditions is the main subject of signal classification research in radar countermeasures at present. Deep learning, as the main content of current technology research, has been widely used in image recognition, computer vision and other fields. Generally speaking, deep learning models are considered to be based on the design of biological brain, which in some ways is a further extension of shallow artificial neural networks. According to the exploration of radar signal classification and identification process as shown in Fig. 1, Zhang Gexiang, Rong Haina et al. proposed a method of classification and identification of radar emitter

R. Kountchev et al. (Eds.): *NAMSP 2022*, SIST 332, vol. 332,, pp. 56–64, 2023.
https://doi.org/10.1007/978-981-19-7842-5_6

signals using similarity coefficient features and support vector in practical exploration. Yang Liming et al. used support vector machine (SVM) as a classifier to classify and identify radar radiation source signals [1, 2].

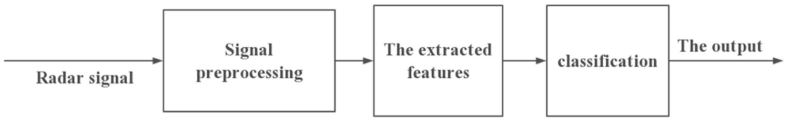


Fig. 1. Classification and recognition flow chart of radar signal

At the same time, with the comprehensive promotion of science and technology and artificial intelligence concept, Guo Xiaobin, Wang Zhuang et al. used Bayesian classifier to complete the classification and identification of radar radiation sources. Compared with the application method with probability approximation criterion as the core, the actual system classification effect is stronger. However, Liu Kai, Wang Jiegui and others applied the concept and algorithm of neural network in the classification and identification of unknown radar radiation sources, and the empirical results proved that the application performance of neural network is very strong. Deep learning, as a deep machine learning model, conventional model contains multiple layers of neural networks, and each layer of neural network is mapped to abstract features in layer upon layer accumulation to complete classification. In the comprehensive development of deep learning technology concept, researchers have put forward many optimization methods of deep learning in practical exploration, such as initialization of network parameters and application of multi-tired activation function, etc. [3, 4]. These methods can not only effectively solve the fitting problems caused by the excessively complex structure of traditional neural network. It can also improve the classification and identification accuracy of radar radiation source signal in practice. For example, the training inquiry conducted by Krizhevsky et al. using the convolutional neural network model improved the accuracy of the actual performance results by 10%, which was far better than the learning and training model of the shallow model. In this paper, combined with the deep belief network model, the radar signal classification method is studied, and the relevant algorithms are verified by simulation, so as to confirm the effectiveness of radar signal classification method based on the deep belief network model.

2 Method

2.1 Deep Belief Network Model

To put it simply, as one of the earliest non-convolution models with deep architecture training as the core, both Boltzmann machines and restricted Boltzmann machines are in the category of self-coding networks in essence. According to the research of scholars to explore the structure analysis of experience and shown in Fig. 2, the Boltzmann machine belongs to use random neural network in the whole connection way constitute a feedback type neural network, all structures are symmetric connections, and no word feedback

between each layer, the overall structure consists of two layers, connected to each other. However, limited Boltzmann machine has a certain difference, the overall structure is no connection within the layer, and between the layers are all connected [5, 6].

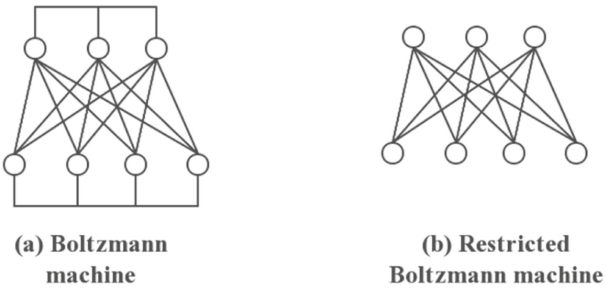


Fig. 2. Structure diagram.

In this paper, the topology and training skills of the deep confidence network are summarized by combining the restricted Boltzmann set. Since this model refers to the type of deep feedforward neural network, the analysis of the training model shown in Fig. 3 shows that the initialization of parameters between layers should be obtained by using the learning mode of the restricted Boltzmann machine. Secondly, we should analyze and understand these four operations according to data, model, optimization objective function and solution. The specific operation is as follows [7, 8].

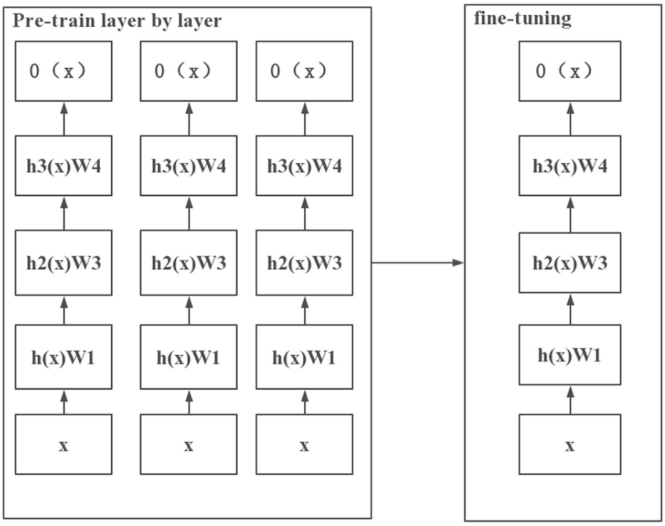


Fig. 3. Training mode of deep neural network.

Data

$$\left\{x^{(n)} \in R^u\right\}_{n=1}^{N} \tag{1}$$

The above formula represents the input data and also the expected output. The data set $\left\{x^{(n)} \in R^u\right\}_{n=1}^{N}$ is divided into two parts. On the one hand, it refers to the data set with class object, also known as the training set; on the other hand, it refers to the data set $\left[x^t\right]_{t=1}^{T}$ without class object, also known as the training set, as shown below:

$$\begin{cases} \left\{x^{(n)}, y^{(n)}\right\}_{n=1}^{N} \rightarrow TrianData \\ \left\{\tilde{x}^{(1)}\right\}_{t=1}^{T} \rightarrow TestData \end{cases} \tag{2}$$

The total number of data sets is $N + T$.

Model. The hyperparameter design of the model $\sigma(.)$ is that L hidden layers exist during feature learning, and the number of nodes on all hidden layers is nl ($L = 1, 2\ldots, L$), the activation function contained in each hidden layer means that the classifier part of the model should use Softmax classifier, or support vector machine or matrix machine can also be selected. The network structure of the model belongs to the deep feedforward neural network. The overall module design is divided into two parts: on the one hand, it refers to feature learning, and the specific formula is as follows:

$$\begin{cases} X_l = \sigma_l(W_l \cdot X_{l-1} + b_l) \in R^{nl} \\ X_0 = x \end{cases} \tag{3}$$

In the above formula, $l = 1,2\ldots, L$.

On the other hand, it refers to classifier design. If the number of categories is K, it can be obtained:

$$\begin{cases} y(k) = \dfrac{e^{(X_L \cdot \theta_k)}}{\sum_{j=1}^{k} e^{(X_L \cdot \theta_j)}} \\ y = \left[y(1), y(2), ..., y(K)\right]^T \end{cases} \tag{4}$$

The waiting to learn parameter is $\theta_k (k = 1, 2, ..., K)$.

Compared with the deep stack network formed by the traditional self-coding network, the only difference lies in feature learning. The parameter initialization network model between layers is as follows:

$$\begin{cases} h\ P(h|v, W, a, b) = \dfrac{1}{P(v) \cdot Z} \cdot e^{\left(b^T \cdot h + v^T \cdot W \cdot h\right)} \\ \hat{v}\ P(v|h, W, a, b) = \dfrac{1}{P(v) \cdot Z} \cdot e^{\left(a^T \cdot h + v^T \cdot W \cdot h\right)} \end{cases} \tag{5}$$

In the above formula, v represents the visual layer, which belongs to the data after normalized input processing; h represents the hidden layer, W represents the weight

connection matrix between the visual layer v and the hidden layer. The symbol a represents multiplicative bias in the visible layer and the symbol b represents multiplicative bias in the hidden layer. Assuming that the parameters (W, a, b) have been specified, the following calculation formula can be obtained by combining the input content v with the hidden layer h:

$$\begin{cases} P(h(i) = 1|v) = \sigma\left(v^T \cdot W_{:,i} + b_i\right) \\ P(h(i) = 0|v) = 1 - \sigma\left(v^T \cdot W_{:,i} + b_i\right) \end{cases} \tag{6}$$

In the above formula, represents the Sigmoid function $\sigma(.)$, $h(i)$ represents the output value of the ith node in the hidden layer h, W_{ij} represents the ith column of the weight matrix W, and b_i represents the ith component of the hidden multiplicative bias.

Similarly, the visible layer can be estimated and analyzed in combination with the hidden layer h, and the specific formula is as follows:

$$\begin{cases} P(v(j) = 1|h) = \sigma\left(W_{:,i} \cdot h + a_j\right) \\ P(v(j) = 0|h) = 1 - \sigma\left(W_{:,i} \cdot h + a_j\right) \end{cases} \tag{7}$$

Optimize the Objective Function. This operation is divided into two processes, the first refers to the initialization of parameters, and the second refers to the feature learning and classification of its design. The optimization objective function of hierarchical parameter initialization is:

$$\min_{\theta} J(\theta) = -\sum_{t=1}^{T} \log P\left(\hat{v}^{(t)}\right) = -\sum_{t=1}^{T} \log \sum_{h} P\left(\hat{v}^{(t)}, h\right) \tag{8}$$

Fourth, solve. This stage is consistent with the traditional deep feedforward neural network, which mainly uses the back propagation algorithm to adjust the whole end-to-end network. According to the optimization objective and function obtained in the above research, the comparative divergence algorithm is used to solve the problem.

2.2 DBN Model Based on Support Vector Machine

Because deep learning model has strong nonlinear function approximation ability, it can effectively show the essential characteristics of data. However, it is difficult for traditional depth models to obtain ideal results when dealing with classification and regression problems. However, support vector machines can map linear and indivisible contents into high-dimensional space and classify data according to solving the optimal partition hyperplane mapped into high-dimensional space. Therefore, this paper studies the effective fusion of deep learning model and support vector machine to build a generation model for classification. According to the following Fig. 4 DBN signal based on Support Vector Machine (SVM) classification flowchart analysis shows that the radar signals formed by the normalized data set, and so as the input of the model, in the end after several network training for characterization of the input data of the characteristic matrix, and then the input into the Support Vector Machine (SVM) implementation of the training, Thus the classification results are obtained [8–10].

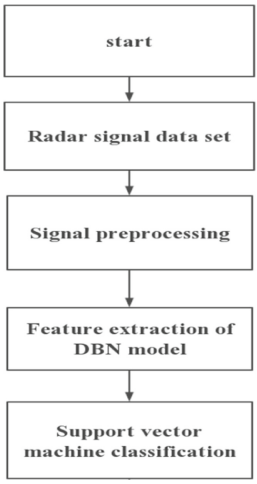


Fig. 4. FLOW chart of DBN signal classification based on SVM

3 Result Analysis

Combined with the classification model and method constructed by the above research, 700 sample data are selected for simulation training, among which 600 samples belong to the training set and 100 samples belong to the test set. In the DBN model with support vector machine as the core, the following results can be obtained by selecting appropriate support vector machine parameters (Table 1).

Table 1. Parameter simulation results of support vector machine.

Penalty parameter C	RBF kernel function parameter g	Classification accuracy (%)
$9.7656*10^{-4}$	$9.7656*10^{-4}$	92.31
0.1	0.1	92.15
0.5	0.1	92.31
1	0.05	92.31
1	0.1	92.38
1	0.2	91.62
2	0.1	92.08

According to the above data set and parameter design, the DBN model and the DBN model with support vector machine as the core are simulated and trained respectively. After 20 times of simulation, the results are shown in the Fig. 5.

Combined with the model comparison results analyzed in the figure above, it can be seen that most of the classification error probability of DBN model is lower than that of DBN model with SUPPORT vector machine as the core, but the actual classification

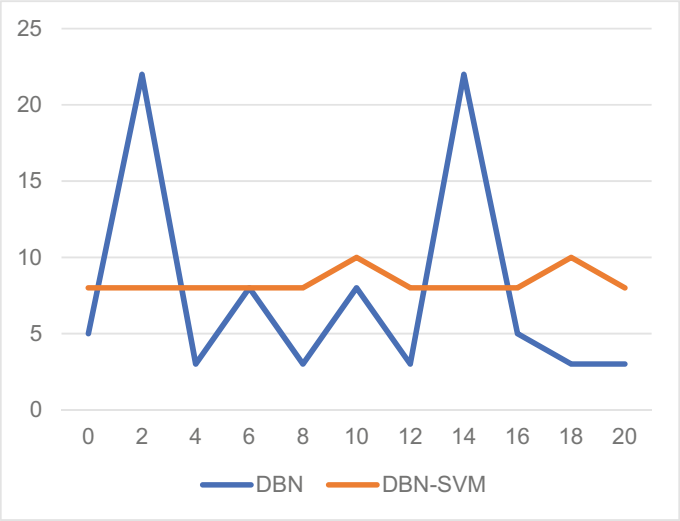


Fig. 5. Comparison results of model training.

probability fluctuates greatly, while that of DBN model with support vector machine as the core has little fluctuation of classification accuracy. The average classification error probability of DBN model can reach 5.85%, and the actual variance is 37.56. The average classification error probability of DBN model based on support vector machine can reach 7.98%, and the actual variance is 0.45. Therefore, after the DBN model optimization by integrating support vector machine, although the actual error rate decreases compared with the original DBN model, the overall model is more robust and has higher stability in practical application [11–13].

At the same time, a comparison between the DBN model with Stacking algorithm and the DBN model with multi-classification gas is conducted. The running time results of the two models are shown in the following Table 2.

Table 2. Time comparison of integrated DBN model.

Model type	Running time per second
DBN model based on Stacking algorithm	89.67
Joint multi-classifier DBN model	85.68

Combined with the comparison results in the table above, it can be seen that the DBN model with combined multi-classifiers has a shorter running time (3.99 s) than the DBN model with Stacking algorithm, which proves that the structure with combined multi-classifiers is more efficient than that with Stacking algorithm in terms of time cost.

In addition, by comparing the classification performance of the two integration models and support vector machines, the results can be obtained as shown in Table 3.

Table 3. Comparison of classification performance of different models.

Model	Classification error rate/%
DBN model based on Stacking algorithm	2.34
Joint multi-classifier DBN model	2.54
Support vector machine SVM	3.23

According to the analysis in the above table, the classification effect of the two DBN models with ensemble learning as the core is more effective than that of support vector machine, which further verifies the advancement of ensemble learning method.

4 Conclusion

To sum up, radar signal classification and recognition is an important part of modern electronic intelligence, reconnaissance and support system. This paper mainly studies the application of current radar signal classification technology from the deep belief network model. Combined with the basic concept analysis of deep learning, it can be seen that it is mainly to learn the hidden features of data in the construction of multi-hidden layer neural network model. Ensemble learning, as an effective learning method of combinatorial optimization, can construct and combine multiple learners to accomplish specific tasks. The final practice results show that the DBN model with ensemble learning as the core can get better prediction results, and the design of the overall system model has a strong generalization level and robustness. Therefore, in the era of big data, researchers should strengthen the application of deep learning models.

References

1. Sun, Y., Tian, R., Liu, C., Guo, Y.: J. Ordnance Equip. Eng. **42**(05), 191–195+226 (2010)
2. Wang, B.: Application of lifting tree in radar signal classification and recognition. Electron. Technol. Softw. Eng. **09**, 78–79 (2021)
3. Ma, Z., Huang, Z., Lin, A.: J. Nav. Univ. Eng. 33(02), 30–36 (2010)
4. Zang, Q., Hong, D., Qian, O., Liu, J., Shang, R.: A signal classification method based on the characteristics of depth. J. Radar Confront. **9**(01), 43–45+53 (2021). https://doi.org/10.19341/j.carolcarrollnki. ISSN: 1009-0401.2021.01.011
5. Sun, Y., Tian, R., Wang, X., Tian, W.: Radar signal recognition under class unbalanced condition. Telecommun. Technol. **61**(03), 321–327 (2010)
6. Hou, C.: Research on Identification Method of In-pulse Modulation Mode of Multi-Component Radar Signal. Harbin Engineering University (2020). https://doi.org/10.27060/dc.nki.Ghbcu.000328
7. Gao, L., Li, Y.: Classification method based on an integrated study of DBN model. J. Harbin Commer. Univ. (Nat. Sci. Ed.) **35**(5), 585–589 (2019). https://doi.org/10.19492/j.carolcarrollnki. 1672-0946.2019.05.016
8. Liu, X., Liu, J., Liu, B., Qin, L., Chen, T.: Research on a new algorithm for radar signal pre-sorting by random forest classification. Sci. Technol. Rev. **37**(13), 93–97 (2019)

9. Li, Y.: Research on Radar Signal Classification Based on Deep Belief Network Model. Harbin Engineering University (2019)
10. Wen, Q.: Research on Classification Method of Complex Radar Signal Based on Neural Network. Xidian University (2018)
11. Jiang, B.: Research on Radar Signal Intra-Pulse Modulation Recognition Based on Time-Frequency Analysis. Nanjing University of Information Science and Technology (2018)
12. Liu, Y.: Research on Feature Information Extraction and Recognition Technology of Radar Signal. Xidian University (2018)
13. Chen, C., Jin, W.: An improved algorithm in the application of radar signal sorting within DNN. Comput. Appl. Res. **4**(4), 1178–1180 + 1185 (2019). https://doi.org/10.19734/j.issn.1001-3695.2017.09.1005

Measuring Machine Intelligence Using Black-Box-Based Universal Intelligence Metrics

Laszlo Barna Iantovics(✉)

George Emil Palade University of Medicine, Pharmacy, Sciences, and Technology of Targu Mures, Gheorghe Marinescu 38, 540142 Targu Mures, Romania
barna.iantovics@umfst.ro

Abstract. Measuring the machine intelligence quotient (MIQ) of intelligent agent-based systems (IABSs) is very important based on the increasing number of intelligent systems applied to real-life problem solving. The most important property of an intelligence metric must be its universality. Developing universal intelligence metrics is difficult based on the very large diversity of intelligent systems. A feasible approach for ensuring the universality of measuring machine intelligence consists in using black-box-based methods able to measure the central intelligence tendency in problem solving. This paper represents a guide for choosing the most appropriate black-box-based intelligence metric for measuring the intelligence of developed IABSs, classification of IABSs in intelligence classes and detection of the IABSs with statistical low and high outlier intelligence. In research where the performance of heuristic and metaheuristic algorithms is studied, the performance indicator is frequently calculated as the mean or the median of experimental evaluation results. There is no consensus agreement regarding which of them is more appropriate. In some cases, both of them are reported. The manner in which it should be decided which of them to be used is scientifically grounded in this paper.

Keywords: Intelligent agent-based system · Machine intelligence · Machine intelligence quotient · Industry 4.0 · Smart factory · Computational hard problem · Central intelligence tendency · Central performance tendency · Heuristic algorithm · Metaheuristic algorithm

1 Introduction

Most intelligent systems are intelligent agent-based systems (IABSs) that could be intelligent agents (IAs) [1] which operate individually or intelligent cooperative multi-agent systems (ICMSs) [2]. The number and diversity of IABSs applied for real-life problem solving in all fields is increasing very fast. For example, the IABSs applied in Industry 4.0 can be mentioned. [3] presents an IABS specialized in solving the problem of sustainable supplier selection. [4] presents an IABS specialized in the decision support used in the chemical process industry.

In this context, measuring the machine/systems intelligence becomes of utmost importance. There are very few studies related to the subject of measuring machine

R. Kountchev et al. (Eds.): *NAMSP 2022*, SIST 332, vol. 332,, pp. 65–78, 2023.
https://doi.org/10.1007/978-981-19-7842-5_7

intelligence [5]. Machine intelligence metrics presented in the scientific literature rely on different philosophies, which hinders their effective comparison. There is no standardization on what machine intelligence is and what should be measured to quantify it [5].

This study investigates the measurement of the artificial complex system intelligence from the viewpoint of difficult problem-solving abilities and highlights the importance of being able to make accurate and robust comparisons in intelligence between multiple intelligent complex systems. The most important property of an intelligence metric is the universality. This is necessary based on the very large diversity of intelligent complex systems. In this sense an important general approach consists in the black-box-based intelligence metrics that should be able to treat aspects like the variability in intelligence, and outlier/extreme intelligence (statistically very low and high intelligence behavior in different situations).

Universal black-box-based machine intelligence metrics represents a useful tool for intelligent systems developers in measuring the intelligence of their systems, with the purpose of comparing it with the intelligence of other systems despite the diversity of their architecture. Research evaluation results must frequently be analyzed using methods based on statistics [6]. The methods for intelligence measuring that were elaborated by the author of this paper, called MetrIntPair [7], MetrIntPairII [8], and ExtrIntDetect [9], which are treated in this paper, are statistically grounded.

The upcoming part of the paper is structured as follows. Section 2 presents a brief survey on state-of-the-art intelligence metrics; Sect. 3 treats the subject of measuring the machine intelligence using the universal black-box-based methods called MetrIntPair, MetrIntPairII and ExtrIntDetect; Sect. 4 treats the machine intelligence quotient measure as central intelligence tendency indicator, and presents the central performance indicator calculus in case of heuristic and metaheuristic algorithms. The conclusions are presented in the final part of the paper.

2 State-of-the-Art Machine Intelligence Metrics

A brief survey on the state-of-the-art metrics for measuring the intelligence of the intelligent agent-based systems is presented in this section.

One of the earliest measurement criteria for machine intelligence was proposed by Alan Turing [10] in 1950. A system able to make computations was considered intelligent if a human evaluator could not decide the nature of the system as being artificial or human, based on questions asked by the evaluator without him seeing who answers.

Schreiner [11] analyzed how to measure and compare the intelligence of the systems. The competition between Watson, developed by IBM, against humans in the popular Jeopardy game was analyzed by Sterret [12]. Newborn [13] presents the well-known competition between the chess grandmaster Garry Kasparov and the chess playing machine named Deep Blue, built by IBM. Besold et al. [14] analyzed diverse problems whose solving by humans is complex, that prospectively could be used as benchmark problems for measuring IABSs intelligence. Detterman [15] suggested a challenge for programs by applying human psychometric intelligence tests. Sanghi and Dowe [16] successfully developed a computer program that surpassed the average human intelligence (by 100) on some psychometric tests [16]. Park et al. [17] proposed the concept

of intelligence task graph for measuring the intelligence of some specific cooperative systems. Anthon and Jannett [18] studied the systems intelligence based on abilities for solving tasks of different complexity. Hernández-Orallo and Dowe [19] developed a so-called universal anytime intelligence test. Legg and Hutter developed an intelligence measure [20] based on the performance of the systems in difficult environments. Hibbard [21] developed an intelligence metric based on the problem-solving ability in complex environments. Liu et al. [22] presented a comprehensive study regarding the analysis of the calculus of the machine intelligence quotient.

The main purpose of endowing agent-based systems with intelligence is to obtain improvements in solving problems of a very high complexity. The human intelligence cannot be completely understood; however no unanimous definition of human intelligence exists. A general definition [23] is given in Britannica: "Human intelligence is a mental quality that consists of the abilities to learn from experience, adapt to new situations, understand and handle abstract concepts, and use knowledge to manipulate one's environment." Brown and Wai [24] outlined that the intelligence level is generally correlated with increased success in life. It is interesting to note that even if human intelligence is not easily understandable, it can be measured. The human intelligence can be measured by intelligence tests able to calculate a so-called intelligence quotient (IQ) [25]. The systems' intelligence can be considered similarly: it cannot be unanimously defined but it can be measured.

Machine intelligence measure is frequently considered based on different kinds of complex-problem-solving abilities. The final goal of measuring machine intelligence is to enable the comparison of the systems and the selection of systems based on their intelligence. It could be also useful if a metric would be able to classify the systems based on their intelligence. Another useful feature could be the identification of the systems with outlier (statistical extreme) intelligence from a set of studied IABSs.

Methods for measuring machine intelligence presented in the scientific literature are based on completely different approaches. However, most of them cannot be compared between them. One of the main drawbacks of many actual metrics consists in their limitation in universality that makes them impossible to be applied in real-life scenarios when one must deal with intelligent systems with a largely diverse architecture. The standardization and universality of the intelligence metrics is an important subject that must be treated by scientific research.

3 Guide for Choosing Between Black-Box-Based Metrics for Measuring the Machine Intelligence

Certain black-box-based approaches for treating the subject of intelligence measure are synthesized in this section. The main purpose was to elaborate some guidelines in choosing the most appropriate method for measuring developed IABSs intelligence.

Concretely, there are presented two metrics/methods called MetrIntPair [7] and MetrIntPairII [8] for intelligence measuring and a method called ExtrIntDetect [9] for detecting IABSs with outlier intelligence (statistical extremes). The presented methods are statistically grounded, universal, and can be applied to agent-based systems generally.

The agent-based systems can be intelligent agents which individually solve problems or multiagent systems where the agents cooperatively solve problems.

In the following there are presented use cases of these methods. The detailed descriptions can be found in the papers in that are published.

3.1 The Specificity of the Black-Box Methods for Intelligence Measuring

The specificity of all the approaches mentioned in the previous paragraph consists in the fact that they operate as black-box-based methods which gives them the advantage of universality. They should not treat aspects such as: studied systems are individual agents (solve the problems individually) or cooperative multiagent systems (solve the problems cooperatively), and what type of architecture they have. We consider that intelligence measuring should be based on their difficult problem-solving ability. It should be noticed that the consideration of a system's intelligence alone is useless. This has relevance in the context of the existence of other IABSs. This fact requires the measuring of the intelligence and comparability of intelligence measurements.

The general idea of the black-box-based methods that we treat consists in the operation in three general steps presented by the generic *Black-box-based intelligence analysis skeleton* (BBSkel). In the following there are used the following notations, n represents the set of studied IABSs. $Iab = \{Iab_1, Iab_2 \ldots, Iab_n\}$, denote the studied IABSs. $MQ = \{Mq_1, Mq_2 \ldots, Mq_n\}$, denote the MIQs of *Iab.*

Skeleton Black-box-based Intelligence Analysis (BBSkel)

Step1: *collecting experimentally obtained difficult problem-solving intelligence evaluation data*

@Making experimental evaluation of difficult problem-solving intelligence of *Iab* (all of the studied IABSs). As result it is obtained in case of each IABS the so-called intelligence indicator data set.

Step 2: *calculus of the MIQs of the studied IABSs*

@The intelligence indicators data are analyzed using diverse statistically grounded methods.

@Based on the analysis results is established the best type of indicator of the central intelligence tendency that will measure the MIQ.

@It is calculated for each studied intelligent system $Iab_1, Iab_2 \ldots, Iab_n$ the machine intelligence quotient $Mq_1, Mq_2 \ldots, Mq_n$.

Step 3: *final step of the machine intelligence measuring and analysis*

@In case of MetrIntPair and MetrIntPairII based on the intelligence measurements results, the studied IABSs intelligence is compared statistically. Finally, the IABSs with the statistically equal intelligence are classified in the same class.

@In case of the ExtrIntDetect method, the IABSs with outlier intelligence (statistical extremes) are identified.

EndBlackBoxMethods

The studied IABSs could be even of a different type, for instance some can be individual agents, while others could be cooperative multiagent systems. In the framework of the

MetrIntPair and MetrIntPairII metrics in the obtaining of the experimental intelligence evaluation data, Step 1 of BBSkel, the so-called pairwise problem-solving intelligence evaluation should be applied. It has the specificity that the problem-solving intelligence is evaluated in pairs. The intelligence of all the studied systems is evaluated for each problem, and the further statistical analysis takes into consideration this aspect of the evaluation.

Table 1 presents comparatively the properties of the studied black-box methods.

Table 1. Properties of the treated black-box methods

Property	MetrIntPair	MetrIntPairII	ExtrIntDetect
N	2	$\geq$2 (any number)	$\geq$3 (any number)
Type	Pairwise	Pairwise	Not-pairwise
Calculate MIQ	Yes	Yes	Yes
Compare the measured MIQs	Yes	Yes	Yes
Classify the IABSs	Yes	Yes	No
Detect outlier intelligence	No	No	Yes
Universal	Yes	Yes	Yes
Other additional properties	Accuracy	Accuracy; robustness	Accuracy; robustness

n denotes the number of studied IABSs. *Type* denotes the type of experimental evaluation of the studied IABSs, that could be pairwise or not-pairwise. *Calculate MIQ* indicate the ability to calculate the machine intelligence quotient. *Compare the measured MIQs* indicates the ability to compare the *MIQs* of the studied systems. *Detect outlier intelligence* indicates the ability to identify outlier intelligence. *Other additional properties* like the accuracy and/or robustness are assured by the specific statistical analyses incorporated in the methods.

3.2 The Universal MetrIntPair Intelligence Metric

Iantovics et al. in 2018 [7] proposed a novel intelligence metric/method called Metric for Pairwise Intelligence Comparison of Agent-Based Systems (MetrIntPair). When used, MetrIntPair can be applied to two studied IABSs specialized in difficult problem solving, denoted in the following as Iab_{a} and Iab_{b}. It can make a measurement by obtaining the machine intelligence quotient Mq_{a} and Mq_{b} of the two studied systems. After this it can compare the intelligence quotient of the two considered systems, by taking into consideration the variability in intelligence, and points out extreme intelligence measurements that should not be taken into account. If the studied systems have a statistically identical intelligence, then they are classified in the same class of intelligence.

In [7] the experimental testing, evaluation and validation of the MetrIntPair metric are presented, using two intelligent cooperative multiagent systems specialized in solving an NP-hard problem.

The universal MetrIntPair intelligence metric can be used for measuring and comparing two intelligent agent-based systems intelligence which can be applied in any field specialized in the solving of the same type of problem.

3.3 The Universal MetrIntPairII Intelligence Metric

MetrIntPair metric presented in the previous section can be successfully applied for the simultaneous intelligence measure analysis of two IABSs. If there are more than two IABSs, a so-called Familywise error rate (FEWER) [26] appears a statistical error that increases with the number of studied IABSs. The paper [8] presents the novel intelligence metric/method called Metric for Pairwise Intelligence Comparison of Agent-Based Systems (MetrIntPairII). MetrIntPairII metric proposed by Iantovics in 2021 [8] eliminates the previously mentioned limitation of the MetrIntPair metric, making it applicable even for very large numbers of studied IABSs.

When used, MetrIntPairII can be applied to n (two or more than two) studied IABSs denoted in the following as $Iab_1, Iab_2 \ldots, Iab_n$. It can make a measurement by obtaining the machine intelligence quotients $Mq_1, Mq_2 \ldots, Mq_n$ of the studied systems. Afterwards it can compare the intelligence quotient of the systems, by taking into consideration the intelligence variability and extreme intelligence measurements. The studied systems with the statistically same intelligence are classified in the same class of intelligence.

In [8] it is presented an experimental testing, evaluation, and validation of the MetrIntPairII metric, using three developed intelligent cooperative multiagent systems specialized in solving a certain type of NP-hard problem.

MetrIntPairII universal intelligence metric can be applied for measuring and comparison of intelligence even of very large numbers of IABSs, no matter in what type of problem solving they are specialized. Two noticeable properties of the MetrIntPairII metric are the accuracy and robustness. These properties are assured by the applied statistical modeling that makes specific processing based on the intelligence indicator data statistical properties.

3.4 The ExtrIntDetect Method for Detecting Systems with Outlier Intelligence

The previously presented MetrIntPair, and MetrIntPair II intelligence metrics can calculate the MIQs of the studied IABSs by obtaining comparable intelligence measures. Finally, they make a classification of the IABSs by including in the same class all the systems which statistically have the same intelligence.

In real-life application it is very important to identify the IABSs from a set of studied IABSs which have outlier intelligence. Outlier intelligence could be statistically extremely high or statistically extremely low intelligence compared with the intelligence of a set of studied IABSs. An IABS has outlier (statistical extreme) intelligence if its MIQ is statistically significantly higher or statistically significantly lower than those others studied IABSs intelligence. Maximal and minimal intelligence does not necessarily mean outlier.

MetrIntPair and MetrIntPair II metrics are not able to detect the systems with outlier intelligence from a set of studied IABSs. Based on this limitation, ExtrIntDetect [9]

method was developed by Iantovics et al. [9]. ExtrIntDetect can identify the systems with low and high outlier intelligence in the framework of a set of studied IABSs.

When used, ExtrIntDetect method can be applied to any number of IABSs (three or more), $Iab = \{Iab_1, Iab_2 \ldots, Iab_n\}$, where $|Iab| = n$ $(n \geq 3)$ denotes the number of studied IABSs. The specific type of calculus of the MIQs is mathematically grounded based on different statistical properties of the experimental intelligence evaluation results data. Initially the machine intelligence quotients $Mq_1, Mq_2 \ldots, Mq_n$ of the studied systems are computed. Afterwards identifies the IABSs with statistical outlier intelligence using a statistically grounded method.

In [9] it is presented an experimental testing, evaluation and validation study performed on six intelligent cooperative multiagent systems specialized in solving an NP-hard problem.

ExtrIntDetect universal black-box-based method can be applied even for a very large number of IABSs specialized in solving the same type of difficult problem no matter with what architecture they are endowed.

ExtrIntDetect can be applied even combined with the MetrIntPairII metric. In case of a set of studied IABSs, MetrIntPairII is applied initially, which calculates the MIQs of the studied systems, compares the MIQs and classifies them in intelligence classes. At the second step ExtrIntDetect method can be applied, which will detect the systems with statistical low and high outlier intelligence.

4 Machine Intelligence Quotient as Indicator of Central Intelligence Tendency

4.1 Calculus of the Central Intelligence Tendency of an IABS

The concept of General Intelligence (G Factor), originally proposed by Charles Spearman in the early years of the 20th century [27], is well-known. In case of all the previously treated methods MetrIntPair, MetrIntPairII and ExtrIntDetect, the calculated machine intelligence quotient of a studied IABS is some type of G factor that summarizes the intelligence level of the systems as a single value.

This section details the proposed calculus of the MIQ of a single IABS specialized in a certain type of difficult problem solving denoted in the following *Ty*.

Initially, as a first step, experimental problem-solving intelligence evaluations are performed, using a set of problems by the *Ty* type, of the studied IABS denoted in the following Iab_e. As result the problem-solving intelligence measures, called intelligence indicator values, are obtained. In the following n denotes the number of experimental problem-solving evaluations; $Iv = \{Iv_1, Iv_2 \ldots, Iv_n\}$ denotes the problem-solving intelligence evaluation results, the so-called intelligence indicator values, which correspond to the intelligence in solving of a set of problems $Pr = \{Pr_1, Pr_2 \ldots, Pr_n\}$.

$Iv_1, Iv_2 \ldots, Iv_n$ data is analyzed and based on the analysis result, the most appropriate indicator of the central intelligence tendency is established, which will represent the type of machine intelligence measure. More concretely if $Iv = \{Iv_1, Iv_2 \ldots, Iv_n\}$ meets the normality assumption (parametric case) then the mean is considered the most appropriate indicator of the central intelligence tendency, $Mq_e = \text{mean}(Iv_1, Iv_2 \ldots, Iv_n)$. Elsewhere,

the median is considered as the most appropriate indicator of the central intelligence tendency (non-parametric case), Mq_e = median (Iv_1, Iv_2…, Iv_n).

Among the most frequently applied goodness-of-fit normality tests, the following can be mentioned: the One-Sample Kolmogorov-Smirnov test (KS test), Shapiro-Wilk test (SW test), Lilliefors test (Lil test), and the Anderson Darling test (Ad test). The Lill test is an adaptation of the KS test. The SW test was proved to have the highest statistical power compared with the previously mentioned tests [28, 29]. The SW test is recommended to be applied in case of samples by a smaller size ($n \leq 30$). It must be noticed that the SW test does not work well with many identical values [28, 29].

The application of a goodness-of-fit normality supposes the elaboration of two hypotheses H_0 and H_1. H_0 denotes the null hypothesis. H_1 denotes the alternative hypothesis. H_0 claims the meeting of the normality assumption. H_1 claims the failing of the normality assumption. The chosen normality test must be applied at a certain significance level that is denoted α, which in most cases is recommended to be set to 0.05. The applied normality test will return a calculated so-called P-value denoted Pv.

The decision rule is the following:

DecisionRule:
If ($Pv > \alpha$) **Then**

H_0 can be accepted, however the normality assumption was passed at the significance level α.

ElseIf ($Pv \leq \alpha$) **Then**

H_0 must be rejected, and accepted H_1, however the normality assumption failed at the significance level α.

EndDecisionRule

Jointly with the numerical evaluation of the data normality the visual interpretation is recommended using the Quantile-Quantile plot (QQ plot) scatterplot. In the case of normally distributed data, the points should fall approximately along this reference line. The visual interpretation can be made as follows: the greater the departure from the reference line, the greater the evidence about the fact that the data do not meet the normality assumption.

4.2 Experimental establishment of the central performance tendency

In the previous section it was presented the specific calculus of the central intelligence tendency that was adopted in [7–9].

Frequently in research that involves elaboration of heuristic and metaheuristic algorithms in the final stage, the mean or the median is calculated as a performance indicator. This paper proposes the performance indicator calculus using the method presented in the previous section. With illustrative purposes in the following, we present an explanatory example.

Travelling salesman problem (TSP) is one of the most well-known NP-hard combinatorial optimization problems whose solving was approached using a large diversity of algorithms [30–34] including genetic algorithms [35, 36]. Frequently, the algorithmic evaluations of the TSP are performed on the Euclidean TSP. In the following we present

the experimental evaluation results of a designed and implemented genetic algorithm applied for solving the Euclidean TSP, which is an NP-hard problem [37].

For the experimental evaluation it was used a desktop computer with AMD Ryzen 9 3900X 12-Core processor, and 80 GB RAM. A map composed of 220 points distributed randomly in the bidimensional Euclidean space was taken into account. For this problem solving, a developed genetic algorithm was applied. The parameters of the algorithms were set to usual values presented in the literature and performing additional experimental evaluations. The population size was set to 100, and the generation number was set to 3000. Table 2 presents the obtained experimental evaluation results denoted $Iv = \{Iv_1, Iv_2 \ldots, Iv_n\}$. $Nr_1, Nr_2 \ldots, Nr_n$ represents the identifiers of the problems.

Table 2. Experimental evaluation results

Problem number	Performance indicator	Shortest tour length
Nr_1	Iv_1	20314.54
Nr_2	Iv_2	20570.31
Nr_3	Iv_3	21364.4
Nr_4	Iv_4	21502.61
Nr_5	Iv_5	20564.18
Nr_6	Iv_6	21484.33
Nr_7	Iv_7	20859.49
Nr_8	Iv_8	21729.4
Nr_9	Iv_9	22796.67
Nr_{10}	Iv_{10}	23807.89

Based on the small sample size by $n = 10$ the SW-test was chosen for the normality verification test. Table 3 presents the results of the SW test and additionally, even if it is not necessary in this case, the Lil test results. In case of both tests, the significance level $\alpha = 0.05$ was considered as the most appropriate. In case of both tests, $Pv \geq \alpha$ and according to the decision rule this leads to the acceptance of H_0, which means the normality assumption is met.

Table 3. Numerical testing of the *Iv* normality

Lil test		SW test	
Statistics	*Pv*	Statistics	*Pv*
0.216	0.2	0.887	0.156

For visual evaluation of the *Iv* normality, the QQ-plot was represented (Fig. 1). The visual representation of the QQ-plot indicates meeting the normality assumption, but

the slight departure of some points from the line indicates the possibility of existence of values that are far from others, the fact that outliers may even exist.

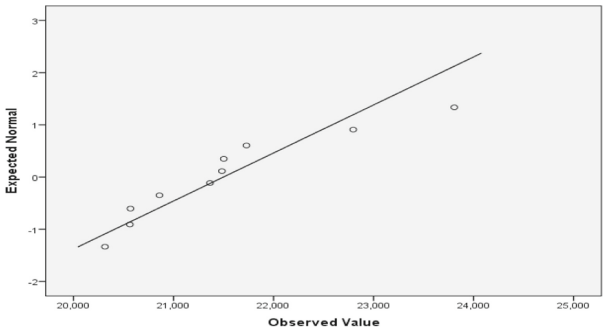


Fig. 1. The QQ plots of Iv for normality visual interpretation

Grubbs test [38–40] is applied for the statistical outliers (low and high) detection when the data is normal or the normality is expectable. Lowest or highest value do not necessarily mean statistical outlier. The outlier's detection is mostly necessary in case of calculus of mean, based on the fact that the mean is more sensitive to outliers then the median. An outlier value influences the mean more significantly than the median. The median is more robust to outliers. When applied, the Grubbs test is able to detect an outlier. After the removing of the outlier, it can be verified if another outlier still exists. The outlier's detection test could be applied consecutively more times until no other outliers are detected.

For the verification of statistical outliers, based on the fact that Iv met the normality assumption and $n \geq 3$ the decision was to apply the Grubbs outliers detection test. At the first application of the two-sided Grubbs outlier's detection test, no outlier was detected, just a value, $Iv_{10} = 23807.89$, that was statically furthest from the rest.

Based on the previous analysis, it can be concluded that the best performance central tendency indicator is the mean $Mq_{\mathrm{e}} = \mathrm{mean}(Iv_1, Iv_2 \ldots, Iv_{\mathrm{n}})$ whose value is 21499.382.

An additional characterization, a descriptive statistic, is presented in Table 4. In some cases, this could be useful for the formulation of some additional remarks.

CL denotes the confidence level. It is recommended that $CL = 95\%$ in most cases. For *CL* other values can be used in some cases, like 90 or 99%. *CI* denotes the confidence interval. It was considered the calculus *CI* of the mean at the 95% *CL*. *LB* denotes the lower bound of the 95% *CI* of the mean. *UB* denotes the upper bound of the 95% *CI* of the mean. *TM* denotes the 5% trimmed mean. *SD* denotes the standard deviation. *SD* is an important measure of the quantity of variation of a set of numerical values. Variance $= SD^2$. Minimum denotes the lowest value. Maximum denotes the highest value. Mode denotes the most frequent value. In the case of our data, the mode cannot be calculated due to the fact that each value appears a single time. Range = Maximum-minimum. The interquartile range presents the spread of the middle half of the distribution. The skewness is a measure of the lack of symmetry. The kurtosis is a measure of the fact that the studied data are light-tailed or heavy-tailed relative to the normal distribution.

Table 4. Descriptive statistical characterization of *Iv*

Indicator		Statistic	Standard error
Mean		**21499.38**	**343.74766**
95% CI of the mean	**LB**	**20721.77**	
	UB	**22276.99**	
TM		21436.9561	
Median		21424.365	
SD		1087.02554	
Variance		1181624.516	
Minimum		20314.54	
Maximum		23807.89	
Mode		N/A	
Range		3493.35	
Interquartile range		1427.44	
Skewness		1.184	0.687
Kurtosis		1.114	4.334

5 Conclusions

Intelligent agent-based systems are applied in all the domains. The number and diversity of intelligent systems is increasing very much. It must be noticed that even if the measuring of the systems intelligence is very important, there are very few methods proposed for measuring machine intelligence. One of the motivations of our research consists in the difficulty in assuring universality in the context of an extremely large diversity of intelligent systems, which could be individual intelligent agents or intelligent cooperative multiagent systems, and that the intelligence is considered at the system's level.

In this paper we took into consideration the measurement of the systems intelligence using universal black-box-based metrics. We discussed how to decide between different metrics, which is most appropriate for specific situations, and how they can be combined. We set forward two metrics/methods called MetrIntPair and MetrIntPairII for intelligence measuring and a method called ExtrIntDetect for detecting IABSs with outlier (statistically extreme) intelligence.

Based on the large number of intelligent systems applied in all the fields and the necessity of measuring the systems intelligence we consider that the methods presented in this paper will represent a great interest to intelligent systems developers who would like to measure their systems intelligence with the purpose of comparing it with the intelligence of other systems.

Acknowledgment. This work was developed in the framework of the CHIST-ERA programme supported by the Future and Emerging Technologies (FET) programme of the European Union through the ERA-NET Cofund funding scheme under the grant agreements, title: Social Network

of Machines (SOON). This research was supported by a grant of the Romanian National Authority for Scientific Research and Innovation, CCCDI-UEFISCDI, project number 101/2019, COFUND-CHIST-ERA-SOON, within PNCDI III.

References

1. Iantovics, L.B.: A novel mobile agent architecture. Acta Universitatis Apulensis **11**, 295–306 (2006)
2. Iantovics, L.B., Enăchescu, C.: Intelligent complex evolutionary agent-based systems. In: Enachescu, C., et al. (eds.) Proc. of the 1st Int. Conf. on Bio-Inspired Computational Methods Used for Solving Difficult Problems-Development of Intelligent and Complex Systems (BICS 2008), American Institute of Physics Proceedings, AIP 1117, pp. 116–124 (2009)
3. Ghadimi, P., Wang, C., Lim, M.K., Heavey, C.: Intelligent sustainable supplier selection using multi-agent technology: theory and application for industry 4.0 supply chains. Comput. Ind. Eng. **127**, 588–600 (2019)
4. Gao, Y., Shang, Z.G., Kokossis, A.: Agent-based intelligent system development for decision support in chemical process industry. Expert Syst. Appl. **36**(8), 11099–11107 (2009)
5. Iantovics, L.B., Gligor, A., Niazi, M.A., Biro, A.I., Szilagyi, S.M., Tokody, D.: Review of recent trends in measuring the computing systems intelligence. BRAIN—Broad Res. Artif. Intell. Neurosci. **9**(2), 77–94 (2018)
6. Iantovics, L.B., Rotar, C., Morar, F.: Survey on establishing the optimal number of factors in exploratory factor analysis applied to data mining. Wiley Interdiscip. Rev.: Data Min. Knowl. Discov. **9**(2), e1294 (2019)
7. Iantovics, L.B., Rotar, C., Niazi, M.A.: MetrIntPair-A novel accurate metric for the comparison of two cooperative multiagent systems intelligence based on paired intelligence measurements. Int. J. Intell. Syst. **33**(3), 463–486 (2018). https://doi.org/10.1002/int.21903
8. Iantovics, L.B.: Black-box-based mathematical modelling of machine intelligence measuring. Mathematics **9**(6), 681 (2021). https://doi.org/10.3390/math9060681
9. Iantovics, L.B., Kountchev, R., Crișan, G.C.: ExtrIntDetect-A new universal method for the identification of intelligent cooperative multiagent systems with extreme intelligence. Symmetry **11**(9), 1123 (2019)
10. Turing, A.M.: Computing machinery and intelligence. Mind **59**, 433–460 (1950)
11. Schreiner, K.: Measuring IS: toward a US standard. IEEE Intell. Syst. Their Appl. **15**, 19–21 (2000)
12. Sterret, S.G.: Turing on the integration of human and machine intelligence. In: Floyd, J., Bokulich, A. (eds.) Philosophical Explorations of the Legacy of Alan Turing. Boston Studies in the Philosophy and History of Science, vol. 324, pp. 323–338, Springer, Cham, Switzerland (2017)
13. Newborn, M.: Kasparov Vs. Deep Blue: Computer Chess Comes of Age. Springer, New York, NY, USA (1997)
14. Besold, T., Hernández-Orallo, J., Schmid, U.: Can machine intelligence be measured in the same way as human intelligence? KI—Künstliche Intelligenz **29**(3), 291–297 (2015). https://doi.org/10.1007/s13218-015-0361-4
15. Detterman, D.K.: A challenge to Watson. Intelligence **39**, 77–78 (2011)
16. Sanghi, P., Dowe, D.L.: A computer program capable of passing I.Q. tests. In: Proceedings of the Joint International Conference on Cognitive Science, 4th ICCS International Conference on Cognitive Science and 7th ASCS Australasian Society for Cognitive Science, Sydney, Australia, 13–17 July 2003, pp. 570–575 (2003)

17. Park, H.J., Kim, B.K., Lim, K.Y.: Measuring the machine intelligence quotient (MIQ) of human-machine cooperative systems. IEEE Trans. Syst. Man Cybern. Part A Syst. Hum. **31**, 89–96 (2001)
18. Anthon, A., Jannett, T.C.: Measuring machine intelligence of an agent-based distributed sensor network system. In: Elleithy, K. (ed.) Advances and Innovations in Systems, Computing Sciences and Software Engineering, pp. 531–535. Springer, Dordrecht, The Netherlands (2007)
19. Hernández-Orallo, J., Dowe, D.L.: Measuring universal intelligence: Towards an anytime intelligence test. Artif. Intell. **174**, 1508–1539 (2010)
20. Legg, S., Hutter, M.: A formal measure of machine intelligence. In: Proceedings of the 15th Annual Machine Learning Conference of Belgium and The Netherlands, Ghent, Belgium, 11–12 May 2006, pp. 73–80 (2006)
21. Hibbard, B.: Measuring agent intelligence via hierarchies of environments. In: Schmidhuber, J., Thórisson, K.R., Looks, M. (eds.) Artificial General Intelligence, pp. 303–308. Springer, Berlin/Heidelberg, Germany (2011)
22. Liu, F., Shi, Y., Liu, Y.: Intelligence quotient and intelligence grade of artificial intelligence. Ann. Data Sci. **4**, 179–191 (2017)
23. https://www.britannica.com/science/human-intelligence-psychology. Accessed 07 June 2022
24. Brown, M.I., Wai, J.: Can you ever be too smart for your own good? Comparing linear and nonlinear effects of cognitive ability on life outcomes. Perspect. Psychol. Sci. **16**(6), 1337–1359 (2021)
25. Neisser, U., et al.: Intelligence: knowns and unknowns. Am. Psychol. **51**(2), 77–101 (1996)
26. Bartroff, J., Song, J.: Sequential tests of multiple hypotheses controlling type I and II familywise error rates. J. Stat. Plan. Inference **153**, 100–114 (2014)
27. Warne, R.T., Burningham, C.: Spearman's g found in 31 non-Western nations: strong evidence that g is a universal phenomenon. Psychol. Bull. **145**(3), 237–272 (2019)
28. Razali, N., Wah, Y.B.: Power comparisons of Shapiro-Wilk, Kolmogorov-Smirnov, Lilliefors and Anderson-Darling tests. J. Stat. Model. Anal. **2**, 21–33 (2011)
29. Stephens, M.A.: EDF statistics for goodness of fit and some comparisons. J. Am. Stat. Assoc. **69**, 730–737 (1974)
30. Kovács, L., Iantovics, L.B., Iakovidis, D.K.: IntraClusTSP—an incremental intra-cluster refinement heuristic algorithm for symmetric travelling salesman problem. Symmetry **10**(12), 663 (2018)
31. Dantzig, G.B., Fulkerson, D.R., Johnson, S.M.: Solution of a large-scale traveling-salesman problem. Oper. Res. **2**, 393–410 (1954)
32. Miller, C.E., Tucker, A.W., Zemlin, R.A.: Integer programming formulations and traveling salesman problems. J. Assoc. Comput. Mach. **7**, 326–329 (1960)
33. Baldacci, R., Hadjiconstantinou, E., Mingozzi, A.: An exact algorithm for the traveling salesman problem with deliveries and collections. Netw. Int. J. **42**, 26–41 (2003)
34. Held, M., Karp, R.M.: A dynamic programming approach to sequencing problems. J. Soc. Ind. Appl. Math. **10**, 196–210 (1962)
35. El-Samak, A.F., Ashour, W.: Optimization of traveling salesman problem using affinity propagation clustering and genetic algorithm. J. Artif. Intell. Soft Comput. Res. **5**, 239–245 (2015)
36. Phienthrakul, T.: Clustering evolutionary computation for solving travelling salesman problems. Int. J. Adv. Comput. Sci. Inf. Technol. **3**, 243–262 (2014)
37. Papadimitriou, C.H.: The Euclidean traveling salesman problem is NP-complete. Theoret. Comput. Sci. **4**(3), 237–244 (1977)
38. Grubbs, F.E.: Sample criteria for testing outlying observations. Ann. Math. Stat. **21**, 27–58 (1950)

39. Barnett, V., Lewis, T.: Outliers in Statistical Data, 3rd edn. Wiley, Hoboken, NJ, USA (1994)
40. Grubbs, F.E.: Procedures for detecting outlying observations in samples. Technometrics **11**, 1–21 (1969)

Applications of Multidimensional Signal Processing

COVID Detection Using ECG Image Reports: A Survey

Mahmoud M. Bassiouni[1(✉)], Islam Hegazy[4], Nouhad Rizk[3], El-Sayed A. El-Dahshan[1,2], and Abdelbadeeh M. Salem[4]

[1] Egyptian E-Learning University (EELU), 33 El-messah Street, Eldoki, El-Giza 11261, Egypt
mbassioun@eelu.edu.eg, seldahshan@sci.asu.edu.eg
[2] Department of Physics, Faculty of Science, Ain Shams University, Cairo 11566, Egypt
[3] Director of Undergraduate Studies, Computer Science Department, Houston University, Houston, USA
njrizk@central.uh.edu
[4] Faculty of Computer and Information Science, Ain Shams University, Abbassia, Cairo 11566, Egypt
{islheg,absalam}@cis.asu.edu.eg

Abstract. COVID-19 is one of the greatest pandemics that threaten individuals, especially the elders. It was first reported in Wuhan, China in 2019. It was discovered recently that COVID-19 disease can be detected using three main protocols. The first protocol is based on Polymerase Chain reaction (PCR), while the second protocol is based on lung chest (ultrasound, X-ray, and CT-Scan), and the final protocol is based on the ECG image reports. This review aims to present a survey on the methodologies and algorithms applied for the detection of COVID disease using electrocardiogram (ECG). In this study, various papers were presented for determining how the COVID can be diagnosed using ECG image reports relying on symptoms and changes in the ECG peaks and intervals. In addition to this, other studies are presented on techniques applied to the ECG reports for the detection of COVID. Also, the main limitations and future works are illustrated. It can be concluded that COVID can be detected with high accuracy using ECG reports and it is even more efficient than other protocols. Finally, based on the performance of the studies it can be shown that the ECG image report is close to an acceptable level in the detection of COVID disease.

Keywords: COVID · Electrocardiogram (ECG) · Deep learning · Pre-trained models · Diagnosis

1 Introduction

Coronavirus is a virus that infects the respiratory tract and can be transferred from one individual to another. It has become a pandemic because it has continued to spread all over the world and caused a significant number of deaths. It is confirmed by the World Health Organization (WHO) [1] that by now more than 6 million people died because of this virus. Therefore, an accurate and efficient diagnosis is required for the detection

R. Kountchev et al. (Eds.): *NAMSP 2022*, SIST 332, vol. 332,, pp. 81–92, 2023.
https://doi.org/10.1007/978-981-19-7842-5_8

of COVID-19 disease. Two main protocols exist for the diagnosis of COVID-19. The first protocol is the Real-Time Reverse Transcriptase-Polymerase Chain Reaction (rRT-PCR). This protocol provides high accuracy in terms of COVID detection, but it requires a long waiting time to achieve results. It is a personal test for each individual and it may take from 4 to 6 h to gain results. The second protocol is based on radio graphical images of the lung chest. The images provide useful information for the detection of the COVID disease, therefore, it is recommended to be done in the early phases of the diagnosis of COVID. Even though these images have shown great success in the detection of COVID, this protocol is high in cost, large in the amount of radiation exposure, and requires some technical skills for capturing images during examination [2].

Based on the former disadvantages, several researchers attempted to find a protocol that is low in cost, accessible, real-time detection of COVID, and less harmful. Therefore, it was discovered that COVID does not only affect the respiratory system, but it affects various origins in the human body, especially the cardiovascular system. The main changes that affect the cardiovascular system have appeared clearly in the Electrocardiogram (ECG). It was found that the COVID patient suffers from Shorting in the RR interval, prolongation in the QT, and changes in the ST. As a result, a third protocol is achieved using ECG image reports for the diagnosis of COVD [3].

Several approaches exist in the fight against COVID and these approaches are based on the Internet of Things (IoT), artificial intelligence, and Industry 4.0 and 5.0 applications. Deep learning is one of the sub-fields of artificial intelligence that has shown high performance compared to other hand-crafted and traditional feature extraction methods. Various deep learning approaches are proposed for the detection of various abnormalities using ECG signals. Some deep learning models are based on the training of the 1D ECG signals [4], while other models provide a 2D representation of the 1D ECG signals using several representations such as higher-order spectral representations, Short Time-Fourier Transform (STFT), Continuous Wavelet Transform (CWT), and scalograms representations [5]. Considering a large number of studies presented for the detection of abnormalities using ECG signals [6], there is a lack in the number of studies presented for the detection of abnormal heartbeats using ECG image reports. As it will be seen, a few studies exist for the diagnosis of COVID using ECG reports. Even though the details of each study will be discussed in terms of acquiring data, filtration of ECG images, feature extraction, and classification. Therefore, this study illustrates three main contributions:

1. Investigation of the studies that show how the COVID can be detected using changes in ECG leads.
2. Presentation of main phases required to detect COVID using ECG image reports.
3. Demonstration of the studies that applied different methodologies for the detection of COVID using ECG reports.

The following sections are structured as follows: the second section will show the related works, while the third section discusses the main phases required to detect COVID from the ECG images, whereas, the fourth section will present the discussion and limitations, and finally the last section contains the conclusion, main findings, and future work.

2 Related Works

Various studies have been revealed to address the main changes that occur in the ECG leads in the COVID patients. The ECG signal consists of five main peaks which are the P, Q, R, S, and T. Wang et al. [7] observed that out of 310 COVID patients, 201 suffer from changes in the distance between S and T peaks and the T wave of the ECG signals. It was also noticed that these patients suffer different types of abnormalities such as sinus tachycardia, atrial arrhythmia, atrial fibrillation (AF), abnormal Q-wave, right bundle branch block (RBBB), and the progression of the R-wave is weak. Pavri et al. [8] performed a study on 75 COVID patients. It was realized that 50.7% of these patients suffer from a shortening in the distance between P and R peaks, while no changes were observed in the remaining patients. In addition to this, Angeli et al. [9] applied an experiment on 50 COVID patients and it was observed that there exist an abnormal ST-T in 30% of the patients, while left ventricular hypertrophy in another 30% of COVID patients, whereas, the remaining patients suffer from several abnormalities such as AF, tachy-brady syndrome and also a rare myocardial infraction and RBBB were also realized.

Another study was performed by Li et al. [10] on 113 COVID patients, 63 of the COVID patients died and 50 survived. It was observed a ventricular arrhythmia with a large influence on the patients who died more than on those who survived. Santoro et al. [11] applied an experiment on 110 patients and he observed a prolongation in the distance between Q and T peaks of 14% of the patients, and also another study was proposed by Jain et al. [12] on 2006 patients and it was observed a clear prolongation in distance between Q and T peaks of 19.7% of the patients. Moreover, a study was presented by McCullough et al. [13] on the 756 COVID patients and it was determined and detected several abnormalities such as RBBB, interventricular block, and contraction of the atrial premature. Lam et al. [14] performed a study on 18 COVID patients and it was observed that 63% of them suffer from atrial trigeminal, while the remaining patients have several abnormalities detected such as depression and prolongation in the distance between the P and the R peaks, RBBB, atrial flutter, and elevation in the segment between S and T peaks. Finally, Bertini et al. [15] experimented on 431 COVID patients and it was found several abnormalities in 93% of the patients. AF, acute right ventricular pressure overload and prolongation between S and T peaks were observed in 22%, 30%, and 4 patients respectively. Finally, based on the presented related work it is obvious that there exist main common changes that occur on the ECG lead to the COVID patients and how the COVID disease can be detected using ECG reports.

3 Main Phases for COVID Detection Using ECG Reports

In this section, the main studies that worked on the detection of the COVID using ECG reports are illustrated in phases. The main methodology of the former studies consists of five phases which are data capturing, de-noising, feature extraction, and classification. Each phase will be illustrated and manifested in detail in the following subsections.

3.1 Data Acquisition

It is required to capture ECG image reports for COVID patients to prove the efficiency of the ECG protocol. Until now, there only one dataset exists that holds COVID ECG image reports that are available online. This dataset is known as the "ECG Images dataset of Cardiac and COVID-19 Patients" [16]. This dataset holds various ECG image reports for five main ECG classes. These classes are COVID (250), Normal (859), Myocardial Infarction (77), Patients with a previous history of MI (PMI) (203), and Abnormal cases (548). This dataset is used in all the studies that used ECG image reports for the detection of COVID disease. Figure 1 shows a COVID ECG image report from the dataset defined in [16]. As shown in the figure it can be seen that the reports consist of three main parts. The first part is some basic information such as id, year, gender, weight, height, and the exam room. The second is the 12 ECG leads of the patient and one prolonged lead. Finally, the third part is also more information related to the heart rate, sampling frequency, date, and time of the ECG report.

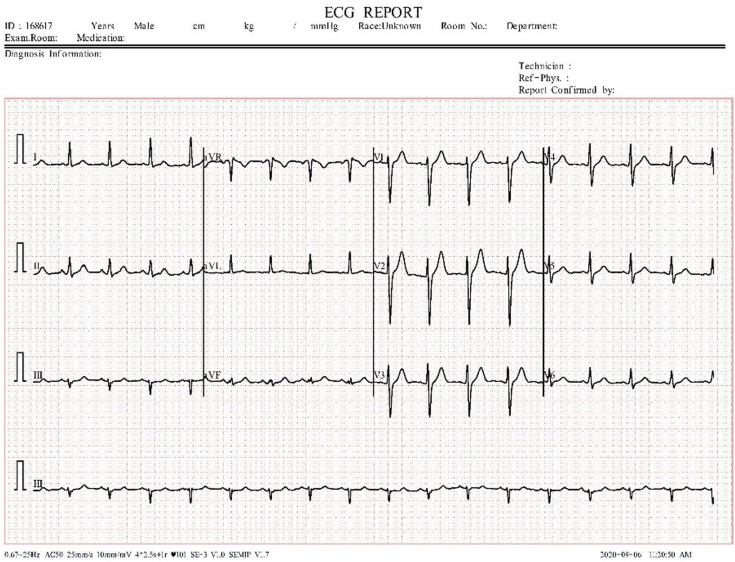


Fig. 1. An ECG image report for a COVID patient

3.2 Data Augmentation

As mentioned in the data acquisition phase, the ECG categories in the dataset defined in [16] are unbalanced in the number of the ECG image reports. Therefore, various studies are applied various geometric transformation methods to balance the number of ECG reports in each class. Attallah [17, 18] has applied several augmentation techniques to increase the number of images in each category and this is performed by applying flipping to the ECG report in the x and y directions. In addition to this, translation, scaling,

and shearing operations are performed on the images with values (−30, 30), (0.9, 1.1), and (0, 45) in x and y respectively. Anwar et al. [19] performed some augmentation operations on the ECG reports balancing the number of images in each class category. These operations are based on flipping, cropping, and perspective distorting. The flipping operation is performed horizontally and vertically with a probability of 0.7, while the cropping is applied central and random, whereas, the last operation is based on distorting the images with a scale and probability of 0.1 and 0.5 respectively.

Irmak [20] applied some geometric transformations to the ECG image report for solving the unbalance problems. These transformations are based on translation, scaling, and rotation to reduce overfitting. Rahman et al. [21] performed the same transformations performed in [20], but they specify the main parameters of these transformations. The rotation was done clockwise and counter-clockwise with an angle between 5 and 10, while the scaling operation was based on reduction and magnification of the ECG report with a percentage of 2.5 and 10% respectively, and the last operation which is the translation is performed horizontally and vertically by 5 to 20%. Finally, Bassiouni et al. [22] applied some augmentation operations based on random rotation, shearing vertically and horizontally based on parameters with different values for each ECG class category.

3.3 Pre-processing

Pre-processing is the process that prepares the ECG image report with enhanced quality and forms for further processing and feature extraction. Several studies applied a filtration methodology to the ECG image reports to improve the ECG reports for further processing. Ozdemir et al. [23] performed a filtration process on the ECG image report based on two main steps. The first step is the removal of the background lines by filtering the input densities with a density map function. After this step is applied the ECG image report still contains traces of the background, therefore, another step is performed by adding a filtering function to remove the interconnected ECG curve pixels. Shahin et al. [24] performed two main steps for the diagnosis of the ECG image reports. The first step is based on segmenting the electrocardiogram image and the process of segmentation has been developed using a rectangular frame to remove undesirable distortions. The second step is applying a density map function to filter the input densities to remove the paper lines of the ECG image reports.

Sobahi et al. [25] applied a filtration methodology for the removal of the background gridlines and the binary noises on the ECG image reports. The methodology is based on applying thresholding operations relying on the green channel of the ECG reports and some morphological operations. Rahman et al. [21] performed a pre-processing step for the diagnosis of the ECG reports. This step is based on applying a gamma correction function to the reports to enhance the quality of these images. Finally, a study was presented by Bassiouni et al. [22] on the filtration of the ECG reports. This process consists of four main steps which are cropping, masking, median filter, and sharpening filter. The cropping removes the header and the footer of the ECG report, while the masking filter is used to remove the red rectangles which are the background of the ECG images, whereas, the median filter is used to remove the black dots found on the reports, and finally, the sharpening filter is applied to sharpen the ECG leads before further processing.

3.4 Feature Extraction

After the augmentation and pre-processing phases are completed, the phase of extracting features from the ECG report is performed. Attallah [17] extracted features from the ECG image report by obtaining the fully connected and max-pooling features of five main pre-trained models (ResNet, Inception, Inception-ResNet, Xception, and DenseNet). Then, the max-pooling features are passed to discrete wavelet transform (DWT), and the features obtained from DWT are integrated with fully connected features using bi-level feature integration. The integrated features are passed to a feature selection step before classification. Another study presented by Attallah [18] for extracting the fully connected features from ten pre-trained models (ShuffleNet, DenseNet-201, ResNet-50, ResNet-18, DarkNet-19, DarkNet-53, InceptionV3, InceptionResnet-V2, Xception, and MobileNet). Then, these are integrated and passed to a hybrid feature selection before classification. In addition to this, Anwar et al. [19] used two pre-trained models for feature extraction. The two models are based on Vgg19 and EfficeintB0, and the EfficientB0 model features outperformed the Vgg19 features in classification. Shahin et al. [24] extracted features from the ECG image reports using six pre-trained models which are Vgg16, Vgg19, InceptionResnetV2, InceptionV3, Resnet50, and Densenet201. The Vgg16 model has surpassed the remaining models in terms of performance. Moreover, Rahman et al. [21] extracted features from the ECG reports based on six main pre-trained models which are ResNet-18, ResNet-50, ResNet-101, InceptionV3, DenseNet-201, and MobileNet-V2. The features obtained from the DenseNet-201 and InceptionV3 produced the highest diagnosis performance compared to other models. Bassiouni et al. [22] extracted features from four pre-trained models which are Vgg16, Vgg19, Resnet-101, and Xception. The combination between the Xception and Temporary Convolution Networks (TCN) showed the most robust feature compared to other pre-trained models.

Irmak [20] has developed a deep learning model based on 18 CNN layers architecture for feature extraction. The layers of the model are 1 image input layer, 1 convolutional layer 1 cross normalization layer, 3 grouped convolution layers, 4 max-pooling layers, 5 ReLU layers, 2 fully connected layers, and 1 dropout layer. Sobahi et al. [25] have developed a deep learning model for extracting features from the ECG image reports. This model consists of 17 layers based on 3D convolutional layers. The layers of the proposed model are 1 (3D) image input layer, 3 (3D) convolutional layers, 3 batch normalization layers, 3 ReLU layers, 2 dropout layers, 2 addition layers, and 1 sigmoid layer, and 1 element-wise multiplication layer, and 1 fully connected layer. Finally, Ozdemir et al. [23] extracted features from the ECG report by passing the ECG reports to the Grey-level Co-Occurrence matrix (GLCM). Then, the features of the GLCM are passed for hexaxial feature mapping. The output of the hexaxial mapping is a 2D image that is passed to a proposed deep learning model. The proposed model consists of 2 convolutional layers, 3 pooling layers, 2 activation layers, 1 convolution block, 1 dense block, 1 flatten layer, and 1 fully connected layer.

3.5 Classification

This phase is based on passing the extracted features to various classifiers for classification. The studies [20, 23, 25] that proposed DL models used Softmax as a classifier at

the end of their architectures. In addition to this, the studies that used a direct pre-trained model [19, 21, 24] on the ECG image reports also applied the Softmax classifier for diagnosis. Moreover, the studies that combined features from different deep learning models applied different classifiers to obtain the highest performance classifier. Attallah [17], applied four different classifiers to the features extracted from the proposed combined approach. These classifiers are Linear Discernment Analysis (LDA), Random Forest (RF), Support Vector Machine (SVM), and voting. The performance of the voting classifier showed the highest performance over other classifiers in different experiments. Also, Attallah [18] used six classifiers for the features obtained from the combined deep learning models. These classifiers are the Decision Trees (DT), RF, Quadratic Discernment Analysis (QDA), SVM, k-Nearest Neighbors (KNN), and LDA. The performance of the DT showed the highest performance on binary class experiments, while RF showed the highest diagnosis result in the multi-class experiments. Finally, Bassiouni et al. [22] applied five main classifiers to the features obtained from the four pre-trained models and the combined model. These classifiers are Softmax, Random Trees (RT), RF, Multilayer Perception (MLP), and SVM. The performance of the SVM classifier was the highest in diagnosis over other classifiers on different experiments using all deep learning models.

4 Discussion and Limitations

This study presents the main phases applied by various studies for the diagnosis of COVID using ECG image reports. The phases start by capturing data, augmenting the data, filtering the ECG reports, extracting features, and classification of the features for an accurate diagnosis. In the data acquisition phase, the dataset available consists of five main classes (COVID, Normal, Abnormal, MI, and PMI). All of the presented studies worked on this dataset because it is the only dataset available online with COVID ECG image reports. In the data augmentation phase, most of the studies applied geometric transformations as an augmentation step before any pre-processing or feature extraction. The geometric transformation operations are based on flipping, rotation, scaling, scaling, translating, and shearing. These operations are applied to balance the number of images in each dataset. In the pre-processing phase, the images are filtered using various techniques in most of the presented studies. Some studies applied density map functions to remove the background lines and the remaining traces of noise, while other studies applied a processing chain that consists of cropping, masking, and filtering to remove the red rectangles and the black dots, and finally, some studies applied thresholding function and some morphological operation intending to filter the ECG image reports. Moreover, the feature extraction phase can be classified into three main types based on the presented studies. The first type is applying a pre-trained model directly on the augmented and the pre-processed ECG data to extract features. The second type is extracting features based on the combination of the features obtained from different pre-trained models or different deep learning architectures. The third type is obtaining features based on a proposed deep learning model. Finally, in the classification phase, most of the studies used Softmax as a classifier and only a few studies used different classifiers for diagnosis. Table 1 shows the performance of the presented studies in the diagnosis of COVID and other ECG diseases in terms of size of ECG reports, number of classes, methodology, and accuracy performance. All the studies used the dataset presented in [16].

Table 1. Performance of various studies for COVID diagnosis based on ECG image reports

Authors	Methodology	Experiments	Performance
Anwar et al. (2021)	EfficientB0 + Softmax	**Multi-class experiment** COVID (250) versus MI (77) versus PMI (203) versus abnormal (548) versus normal (859)	Five-fold cross-validation A = 81.8 without augmentation A = 77.6 with augmentation
Shahin et al. (2021)	Vgg16 + Softmax	**Multi-class experiment** **5-class diagnosis** COVID (250) versus MI (77) versus PMI (203) versus abnormal (548) versus normal (859)	K-fold cross-validation A = 81.39
Ozdemir et al. (2021)	GLCM + Hexaxial Mapping + a proposed deep learning model + Softmax	**Binary-class experiments** 1st: COVID (250) versus normal (250) 2nd: positive (COVID (250)) versus negative (normal (250))	5-fold cross validation 1st: A = 96.20% 2st: A = 93.20%
Attallah et al. (2021)	Fully connected features of 5 pre-trained models + DWT of the max-pooling features + 3 classifiers	**Multi-class experiment** 1st: COVID (250) versus normal (250) versus abnormalities (250) **Binary-class experiment** 2nd: COVID (250) versus normal (250)	10-fold cross validation 1st: A = 91.73% 2st: A = 98.80% Using voting classifier
Attallah et al. (2022)	Fully connected features of 10 pre-trained models + feature selection + classification using six classifiers	**Multi-class experiment** 1st: COVID (250) versus normal (250) versus abnormalities (250) **Binary-class experiment** 2nd: COVID (250) versus normal (250)	10-fold cross validation 1st: A = 91.60% using RF classifier 2st: A = 98.20% using DT classifier

(*continued*)

Table 1. (*continued*)

Authors	Methodology	Experiments	Performance
Sobahi et al. (2022)	A proposed 17 DL layers model based on a 3D convolution network + Softmax	**Multi-class experiments** 1st: COVID (250) versus normal (250) versus abnormal (250) **Binary-class experiment** 2nd: COVID (250) versus normal (250)	10-fold cross validation 1st: A = 92.0 2nd: A = 99.0 TP: 246, FP: 249 FP: 1, FN: 4
Rahman et al. (2022)	DenseNet201 + Softmax	**Multi-class experiments** 1st: COVID (250) versus MI (77) versus PMI (203) versus abnormal (548) versus normal (859) 2nd: COVID (250) versus normal (250) versus abnormalities (250) **Binary-class experiment** 3rd: COVID (250) versus normal (250)	10-fold cross validation 1st: A = 97.83 2nd: A = 97.36 3rd: A = 99.10
Irmak et al. (2022)	A proposed 18 DL model + Softmax	**Multi-class experiments** 1st: COVID (218) versus MI (77) versus abnormal (500) versus normal (713) 2st: COVID (218) versus MI (77) versus abnormal (500) **Binary-class experiments** 3rd: COVID (218) versus MI (77) 4th: COVID (218) versus abnormal (500) 5th: COVID (218) versus normal (713)	1st: A = 82.00 2nd: A = 86.60 3rd: A = 96.70 TP: 46, TN: 43 FP: 3, FN: 0 4th: A = 93.20 TP: 111, TN: 122 FP: 8, FN: 9 5th: A = 98.60 TP: 129, TN: 147 FP: 3, FN: 1

(*continued*)

Table 1. (*continued*)

Authors	Methodology	Experiments	Performance
Bassiouni et al. (2022)	Xception + TCN + SVM	**Multi-class experiments** 1st: COVID (750) versus MI (747) versus abnormal (750) versus normal (856) versus PMI (812) 2st: COVID (250) versus normal (250) versus abnormalities (250) **Binary-class experiments** 3rd: COVID (250) versus normal (250) 4th: positive (COVID (250)) versus negative (normal (250))	1st: A = 99.74 2nd: A = 99.10 3rd: A = 100.0 TP: 250 TN: 250 FP: 0 FN: 0 4th: A = 99.80 TP: 250 TN: 249 FP: 0 FN: 1

A: Accuracy (%) **TP:** True positive **FP:** False Positive **TN:** True Negative **FN:** False Negative

It can be observed from the presented studies that there exist three main limitations. The first limitation is there exists only one dataset that holds COVID ECG image reports and there is not even a digitized COVID ECG form signals. The second limitation is that a few studies targeted the diagnosis of COVID based on ECG lead compared to the studies or the work done using lung chest. Even though it is clear the diagnosis performance obtained from the ECG lead is accurate and can be compared to other protocols. The third limitation is that most of the studies worked on the ECG image report as a single image and no study performed digitization on the ECG leads or worked on each ECG lead as a single image.

5 Conclusion and Future Works

Several researchers around the world address the problem of the COVID diagnosis based on different protocols. Even though there exist two main protocols for COVID diagnosis which are based on PCR and lung chest scan, it would be better to provide a third protocol that can operate in real-time with low cost and achieve a high accuracy performance depending on the ECG image report. In this paper, a comparative study is provided determining different studies with details of their phases in the detection of COVID. It can be concluded from the presented studies that COVID can be diagnosed using ECG reports with high performance. Several suggestions and aspects are concluded for improving the work in this research. It is recommended by the research community to add more effort to this protocol and this can be done by supplying the researchers with more datasets on the ECG image reports related to COVID patients. It is also recommended to perform digitization on the ECG image reports for two main reasons. The first reason is to see if the performance of the digitized ECG is higher or lower than the ECG image report as a single image. The second reason is to obtain a digitized ECG dataset for COVID patients. Finally, it would be a challenge if a system is developed to detect COVID patients using only a single ECG lead.

References

1. World Health Organization: COVID disease. https://www.who.int/emergencies/diseases/novel-coronavirus-2019. Accessed 19 June 2022
2. Haleem, A., Javaid, M., Vaishya, R., Deshmukh, S.G.: Areas of academic research with the impact of COVID-19. Am. J. Emerg. Med. **38**(7), 1524–1526 (2020)
3. Talwar, V., Goel, V., Raina, S., Patnaik, N., Doval, D.C.: Altretamine in advanced pretreated epithelial ovarian carcinoma patients: Experience from a center in north India. Curr. Med. Res. Pract. **6**(3), 109–112 (2016)
4. Bassiouni, M.M., El-Dahshan, E.S.A., Khalefa, W., Salem, A.M.: Intelligent hybrid approaches for human ECG signals identification. Signal, Image Video Process. **12**(5), 941–949 (2018)
5. Bassiouni, M., Khaleefa, W., El-Dahshan, E.A., Salem, A.B.M.: A machine learning technique for person identification using ECG signals. Int. J. Appl. Phys. **1**, 37–41 (2016)
6. Bassiouni, M., Khalefa, W., El-Dahshan, E.S.A., Salem, A.B.M.: A study on the intelligent techniques of the ECG-based biometric systems. In: Recent Advances in Electrical Engineering, pp. 26–31 (2015)

7. Wang, Y., et al.: Electrocardiogram analysis of patients with different types of COVID-19. Ann. Noninvasive Electrocardiol. **25**(6), e12806 (2020)
8. Pavri, B.B., Kloo, J., Farzad, D., Riley, J.M.: Behavior of the PR interval with increasing heart rate in patients with COVID-19. Hear. Rhythm **17**(9), 1434–1438 (2020)
9. Angeli, F., et al.: Electrocardiographic features of patients with COVID-19: one year of unexpected manifestations. Eur. J. Intern. Med. **95**, 7–12 (2020)
10. Li, L., Zhang, S., He, B., Chen, X., Wang, S., Zhao, Q.: Risk factors and electrocardiogram characteristics for mortality in critical inpatients with COVID-19. Clin. Cardiol. **43**(12), 1624–1630 (2020)
11. Santoro, F., et al.: QTc interval prolongation and life-threatening arrhythmias during hospitalization in patients with coronavirus disease 2019 (COVID-19): results from a multicenter prospective registry. Clin. Infect. Dis. **73**(11), e4031–e4038 (2021)
12. Jain, S., et al.: Enhanced electrocardiographic monitoring of patients with coronavirus disease 2019. Hear. Rhythm **17**(9), 1417–1422 (2020)
13. McCullough, S.A., Goyal, P., Krishnan, U., Choi, J.J., Safford, M.M., Okin, P.M.: Electrocardiographic findings in coronavirus disease-19: insights on mortality and underlying myocardial processes. J. Cardiac Fail. **26**(7), 626–632 (2020)
14. Lam, K., McClelland, S., Dallo, M.J.: ECG: essential in care of patients with COVID-19. Med. J. Aust. **213**(10), 476 (2020)
15. Bertini, M., et al.: Electrocardiographic features of 431 consecutive, critically ill COVID-19 patients: an insight into the mechanisms of cardiac involvement. EP Europace **22**(12), 1848–1854 (2020)
16. Khan, A.H., Hussain, M., Malik, M.K.: ECG images dataset of cardiac and COVID-19 patients. Data Brief **34**, 106762 (2021)
17. Attallah, O.: ECG-BiCoNet: an ECG-based pipeline for COVID-19 diagnosis using bi-layers of deep features integration. Comput. Biol. Med. **142**, 105210 (2022)
18. Attallah, O.: An intelligent ECG-based tool for diagnosing COVID-19 via ensemble deep learning techniques. Biosensors **12**(5), 299 (2022)
19. Anwar, T., Zakir, S.: Effect of image augmentation on ECG image classification using deep learning. In: The International Conference on Artificial Intelligence (ICAI), pp. 182–186. IEEE, Pakistan (2021)
20. Irmak, E.: COVID-19 disease diagnosis from paper-based ECG trace image data using a novel convolutional neural network model. Phys. Eng. Sci. Med. **45**(1), 167–179 (2022)
21. Rahman, T., Akinbi, A., Chowdhury, M.E., Rashid, T.A., Şengür, A., Khandakar, A., Ismael, A.M.: COV-ECGNET: COVID-19 detection using ECG trace images with deep convolutional neural network. Health Inf. Sci. Syst. **10**(1), 1–16 (2022)
22. Bassiouni, M.M., Hegazy, I., Rizk, N., El-Dahshan, E.S.A., Salem, A.M.: Automated detection of COVID-19 using deep learning approaches with paper-based ECG reports. In: Circuits, Systems, and Signal Processing, pp. 1–43 (2022)
23. Ozdemir, M.A., Ozdemir, G.D., Guren, O.: Classification of COVID-19 electrocardiograms by using hexaxial feature mapping and deep learning. BMC Med. Inform. Decis. Mak. **21**(1), 1–20 (2021)
24. Shahin, I., Nassif, A.B., Alsabek, M.B.: COVID-19 Electrocardiograms Classification using CNN Models. arXiv preprint arXiv:2112.08931 (2021)
25. Sobahi, N., Sengur, A., Tan, R.S., Acharya, U.R.: Attention-based 3D CNN with residual connections for efficient ECG-based COVID-19 detection. Comput. Biol. Med. **143**, 105335 (2022)

Disease Detection Techniques in Plants: Transition from Manual to Automation

Preeti Yadav(✉) and Parvinder Singh

Dcrust, Murthal, Sonipat, India
Preeti.schcse@dcrustm.org

Abstract. India is an agricultural land with substantial agricultural poverty. The plant diseases impose severe threat to the crop yield, productivity and sustainability each year. In early days, observational methods were adopted by experts and with the evolution of techniques, cultivators started sharing images of plants over distance and the distribution of knowledge was speeded with limited access. The rapid and accurate estimation of plant diseases is still an unmet need of agriculture and led researchers to train and test computer-aided deep learning techniques like ANN, CNN. These techniques are fast, precise and consistent with quantitative information. Plant epidemiologists are adopting and preferring automated disease detection techniques over previous techniques to save cultivators from stressful, time-consuming, and laborious disease detection methods. This is the very first review where we have shown how and why the transition of plant disease detection techniques from naked-eye to automatic detection techniques happened? In this review, we have tried to compile the most recent literature available with special emphasis on deep-learning techniques w.r.t. plant diseases.

Keywords: CNN · SVM · HOG · ANN · AlexNet · ResNet

1 Introduction

Plants have a vital role in our life. Plants provide us food, oxygen and helps in maintaining ecosystem balance. But the plant diseases impose a threat to agricultural productivity and sustainability each year. India is an agricultural land where more than 50% people do farming and are dependent on agriculture for their livelihood [1]. Although, agriculture is an important pillar of Indian economy still more than 10,000 farmers commit suicide each year due to impoverished cultivation [2]. There are several adversities like agro-climatic change, water shortage, monetary problems and higher rate of disease attack that reduces yield by almost 50% and results in sharp decline in percentage of farmers from 71.9% in 1951 to 54.4% in 2001. The reduction in number of cultivators is threatening for the survival of any community in long term [3].

Among these adversities, few are inevitable while few are being taken care of via Indian Govt. but disease-attack rate can be controlled to certain extent only with adequate knowledge and proper action plan for treatment. A good understanding of emerging and rare plant diseases that is accessible by cultivators is indispensable to combat yield losses.

R. Kountchev et al. (Eds.): *NAMSP 2022*, SIST 332, vol. 332,, pp. 93–109, 2023.
https://doi.org/10.1007/978-981-19-7842-5_9

The disease level is measured as disease severity [4]. The rapid, involuntary and precise assessment of disease seriousness is required for security of food, disease management and crop loss prediction. The early diagnosis of diseased state is one of the effective strategies that might ease disease severity but detection of diseases in plants is a tedious job [5]. The plants either exhibit symptoms of disease at community-spread stage or do not exhibit any visible symptoms throughout their lifespan. Hence, these need to monitor via naked eye for longer duration by experts that is time-consuming and produce inconsistent information. Therefore, the tools for automatic disease detection are becoming the valuable resource of information nowadays, especially in large fields and areas lacking specialised technical support [6]. To avoid anomaly, plant disease epidemiologists are nowadays adopting automated deep-learned disease detection techniques. The quick, non-destructive, automated and systematically observed with quantitative data are adding benefits to these techniques [7].

Botanists prefer leaf over other plant parts to compare the variation in characteristics in healthy and diseased plants since leaves are available and collected to study throughout the year. However in seasonal crops, there is an unmet need of non-invasive, quick and accurate estimation of disease severity that causes heavy losses to the annual yield and led to agricultural poverty and reduction in number of cultivators each year. Deep-learning, the present-day technique in computer vision has shown quite favourable results in the literature but its practical utility is still warranted [8]. The successful implementation of these user-friendly and inexpensive techniques at field level might alleviate heavy loss of time and energy and would uplift farmers socio-economically.

Although automation of plant diseases has its own challenges but Convolutional Neural Networks (CNN) are becoming the hopeful methods to overcome these problems. In recent years, several studies have been published for detection of different diseases in various plants as well as detection of new plant species [9] or soil complexity and texture [10] by the use of CNN. Sooner the disease detection is going to be an easy-going task which will be just a click away with the help of CNN! CNN extracts the features from almost any type of image captured with any camera, mobile phones with high or low light intensity and background noise [11]. In order to provide a quick transition of plant disease detection techniques to researchers in agriculture fields, we have compiled various techniques of plant disease detection in this review. Although we have discussed qualitative detection techniques as well but we emphasized on deep learning techniques with a focus on latest techniques based on neural network based detection techniques (artificial neural network and convolutional neural network) published in last two to three years in subsequent sections.

This article has been divided into six sections as described—(i) introduction, (ii) types of detection techniques, (iii) deep-learning techniques, (iv) Artificial Neural Networks (ANN) and (v) Convolutional Neural Networks (CNN) with a compilation of latest studies in more detail and (vi) summary with limitations and future directions.

2 Types of Detection Techniques

2.1 Naked Eye Observation

This technique has been used for long and still widely used in many parts of the world. In this technique an expert farmer who is cultivating crop/s for years and have experienced disease attacks on the fields used to observe the field for a certain period of time and then pass on the information about the disease attack [12]. However, the qualitative data or information generated by this technique has no scientific validation and reports but most of the farmers relied on this technique. The major drawbacks of this technique are inconsistency, laborious, inaccuracy, time-consuming and expert availability to the site. Unfortunately, most of the time there are no experts in the area to give a data based analysis and advise to the farmers. To overcome this problem, Abu-Naser tried to develop an expert system for diseases in papaya plants [13] and safflower plants [14]. This expert system works just like physician do for humans. This expert system utilizes the knowledge database collected from experts and assign symptoms to the already identified diseases. It is a quick assessment of diseases based on symptoms and growing conditions. Even a non-expert cultivator can extract information of disease, causative agent and possible treatment from this expert system just by entering the symptoms of disease [13, 14]. These expert systems are easy to handle and information retrieval but these can be specialised in already well-defined diseases. These systems cannot be applied to new symptom and/or disease.

2.2 Image Capture Method

To overcome the availability of an expert, people adopted to take images of crop fields from different sites, angles and different parts of plants to compare with images of healthy plants or images imprinted within brain of an expert. In this method, the comparison of diseased and healthy plant used to be made based on colour transformed images [6]. The variation in spots between plants used to be identified and categorised for different diseases. By the help of computer perception vision field, handmade manually feature-based methods like Histogram of Oriented Gradients (HOG) and Scale-Invariant Feature Transform (SIFT), which are usually merge with classifiers such as Adaptive Boosting (AdaBoost) or Support Vector Machine (SVM) were adopted for image recognition. A handmade method is known as an algorithm in which human knowledge is implied in its development and the intricate parameters that are included in the process [15].

The major drawback of this method is the high computational cost and time consuming due to the intricate pre-processing, feature extracting, and classifying. Moreover, this method was accessed by limited cultivators and erroneous data based on the quality of images like pixel size, resolution, focal length.

2.3 Paper Grid Method

In this method, the images captured at various angles for larger fields then image processing used to be done for sharpness, magnification and resolution. To collect some informative data, the area of the colour spots used to be calculated via grid method by

superimposing the images on special paper [16]. This was the very first method where some quantitative data was produced. This method was more accurate and precise compared to the afore-mentioned methods but standard protocols need to be followed for image capture and it has its own limitations [17].

2.4 Invasive Methods

Plants are continuously facing threats from multiple pathogens (bacteria, viruses and fungi) like animals. These micro-organisms infect plants and destroy the crop once it spreads over the whole field. Plants, being sedentary and sessile are more prone to community-spread since wind acts as microorganisms' carrier. Researchers started to study the characteristics of diseased plants compared to healthy ones in laboratory set-up [18]. In this method, they used to prepare the temporary and/or permanent slides of sections from different parts of plants and made a repository to characterize each disease. They classified diseases based on causative agent, duration p. but these methods are invasive and time-consuming. Also, we need to sample a large population to check if only few plants are affected or the disease has been spread to large scale. Therefore, we need non-invasive, cheap and easy-to implement techniques for plant disease detection.

2.5 Automated Detection Techniques

Plants lack mobile immune cells like animals but they have developed innate immune system to sense harmful signals and produce a series of protection responses. These protection responses are recognized as pathogen signatures and perceived by recognition of patterns receptors in plants [19]. Based on this principle of pattern recognition, researchers have developed neural network based early detection of pathogen attack. Artificial neural network contains thousands of layers of interconnected nodes. The neurons are processing units. The input layer feeds the signal produced by a multiple linear regression into an activation function that may be nonlinear. The input layer collects input patterns. The output layer classifies the signals to which input patterns may map. Hidden layers fine-tune the input weightings until the neural network's margin of error is minimal. It is hypothesized that hidden layers extrapolate salient features in the input data that have predictive power regarding the outputs. The output of each layer is determined by the current input which is output of previous layer. The final output can be calculated by the use of different formulation for the algorithm as given below (adopted from Zhu et al. [9]). This express deep feature extraction, which accomplishes a utility similar to statistical techniques such as principal component analysis (Fig. 1).

Input Layer to Hidden Layer input:

$$net_i^h = \sum_{j=1}^{r} v_{ij} p_j + b_{hj} \tag{1}$$

Hidden Layer output:

$$a_i = f_h\left(net_i^h\right) = f_h\left(\sum_{j=1}^{r} v_{ij} p_j + b_{hj}\right) \tag{2}$$

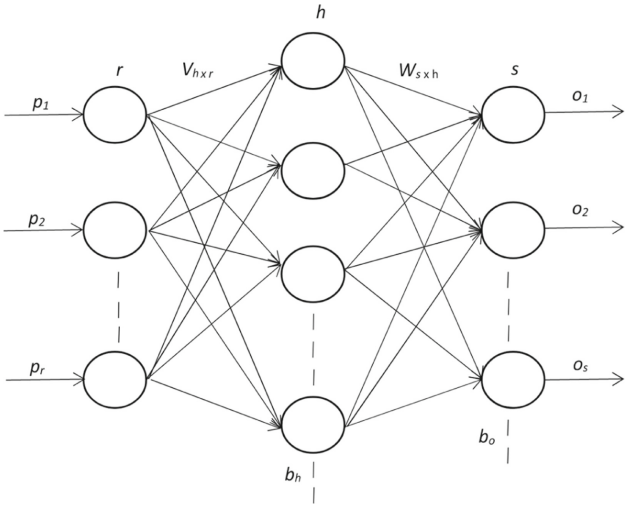


Fig. 1. Back propagation neural network

Output Layer Input:

$$net_k^o = \sum_{i=1}^{h} w_{ki} a_i + b_{ok} \tag{3}$$

Output Layer Output:

$$O_k = f_o\left(net_k^o\right) = f_o\left(\sum_{i=1}^{h} w_{ki} a_i + b_{ok}\right) \tag{4}$$

In these four equations, the terms are explained below:
Training set = $\{(p^1,d^1), (p^2,d^2), (p^3,d^3) \ldots (p^r,d^s)$
Weights between neuron code r and $h = [v_{ij}]_{h*r}$
Threshold value in neuron code $h = b_h$
Input of the first hidden layer = net_i^h
Output of the first hidden layer = a_i
Activation function of the first hidden layer = f_h
Input of the output layer = $net_k{}^o$
activation function of the output layer = f_o
output of output layer = O_k.

Leaves being the most sensitive and highly active part of the plant have the tendency to show the symptoms of the disease at its earliest and also possess neural network in the form of reticulation so leaves are the most favourable part of plant for early detection of plant disease. A number of hyperspectral images of leaves acquired in-order by camera gadget with various resolutions and then processed for deep feature extraction like VGG, ResNet, etc. [20]. We have discussed these automated plant disease detection techniques in detail in following sections.

3 Deep Learning Techniques

Deep learning such as algorithms of CNN, RNN and GAN has been used widely nowadays. In this technique, the principle remains the same to that of conventional naked-eye technique. The only difference between the two is the observer. In conventional naked eye technique, an expert used to compare healthy vs diseased state while in automated techniques, we train a computer to identify the difference between healthy and diseased state by giving several hyperspectral images from plants as input and prepare different test sets as healthy and diseased. This is known as machine learning [21]. Machine learning is a subset of an Artificial Intelligence (AI) technique is a self-adaptive algorithm that processes big data and identified patterns. Deep learning is a subset of machine learning, use a hierarchical level of artificial neural networks to carry out the process of machine learning. In deep learning algorithms, cost function predicts the difference between predicted output and the actual output for training purpose. During training, cost function reduces moderately via gradient descent and ultimately the delta between predicted and actual [22, 23]. The gradient descent is combination of two words i.e. Gradient and descent. The word "Gradient" means the sharpness of a slope and the other word "descent" means decreasing the point of steep. This function is used for minimizing the cost and loss function. The very first attempt of deep learning-dependent disease diagnosis of acquired images was performed in 2016 in which the trained set classified 14 crops and 26 diseases with an accuracy of 99.35% compared to manual diagnosis [24].

3.1 Artificial Neural Networks (ANN)

An artificial neural network (ANN) is the piece of a computing system designed to simulate the way the human brain analyses and processes information. It is the foundation of Artificial Intelligence (AI) and solves problems that would prove impossible or difficult by human or statistical standards. ANNs have self-learning capabilities that enable them to produce better results as more data becomes available. The artificial neural networks are built like the human brain, with neuron nodes connected like a web. While traditional programs build analysis with data in a linear way, the hierarchical function of deep learning systems enables machines to process data with a nonlinear approach [25].

3.2 Convolutional Neural Network (CNN)

Convolutional neural network is a deep learning tool that is composed of multiple convolutional layers and these layers are pooled for feature extraction and connected together for classification [8] as shown in Fig. 2.

The concept of neural networks, which has its roots in artificial intelligence, is swiftly gaining popularity in the development of plant disease detection. A neural network works similarly to the human brain's neural network. A "neuron" in a neural network is a mathematical function that collects and classifies information according to a specific architecture. The network bears a strong resemblance to statistical methods such as curve fitting and regression analysis [26]. Neocognitron and LeNet, the very first CNNs came into existence in 1980s but development in hardware technology evolved deep CNN. The

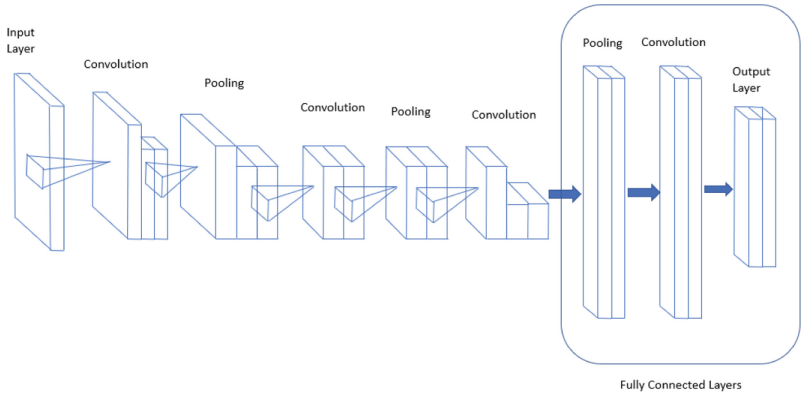


Fig. 2. CaffeNet, an example of feature extraction in CNN architecture

success of AlexNet in ImageNet Large Scale Visual Recognition Challenge (ILSVRC) was the turning point in 2012 that helped in achieving state-of-the-art performance on ImageNet and other benchmark datasets [27]. CNN architecture become complex and more accurate with time as VGG-19 consisted 19 hidden layers, GoogLeNet comprised of 22 layers and ResNet outpaced all of these architectures with 152-layered network [28].

CNN allowed researchers to design systems that can be trained and tested end-to-end (all included in the same process) unlike handcrafted-based methods like SIFT, HoG, SURF, etc. where separate learning algorithms are applied. CNN has fewer parameters than deep networks because of its local perception mechanism and parameter sharing mechanism that reduces parameters [29]. Its principle remains similar to ANN but its feature extractor can adapt to changing input because of image pre-processing in which kernels of each convolutional layer changes and used as input for next layer; so the network generates the best possible result without needing to redesign the output criteria. Due to feature extraction in image recognition, it is now widely used approach in agriculture [30]. Several researchers are using different combinations with different methodologies at each step of CNN and validating the identification of variety of plant species and disease identification [31–33]. The step-by-step detection and identification of plant diseases using this technique in various disease-crop pair has been discussed by Khirade in 2015 [34]. This approach can use input images captured from different camera devices, at variable resolutions, illumination conditions, background etc. as performed by Algergawy on banana leaves via LeNet architecture [11]. The acquired images then pre-processed and segmented from RGB model to lab colour model. After segmentation, images are clustered using Otsu classifier or K-mean clustering or grabcut segmentation for fruits as done by Sharath et al. in pomegranate in 2019 [35]. The following step is feature extraction that is performed to predict shape-oriented parameters like area, colour, degree of rectification and density or solidity; texture-oriented factors like contrast, energy, and homogeneity; and colour aspects like hue, saturation and value in the infected region. The statistical parameters like smoothness, third-order moment, consistency and entropy are also extracted in these multiple-regression based

CNN methods. These extracted parameters are stable irrespective of image changes and thus these methods are comprehensive and descriptive. After the extraction of multiple features then regression analysis is performed and plant diseases are characterized and scored in four categories based on the severity of disease: normal situation (0–25), minor disasters (25–50), medium disasters (50–75) and serious disasters (75–100) [36]. We have discussed several CNN-based plant disease detection strategies in the following sections and compiled these studies in Table 1.

Fuentes et al. in 2017 described three families of convolutional neural network-based detectors for detection of diseases in tomato leaves namely, Faster Region-based Convolutional Neural Network (Faster R-CNN), Region-based Fully Convolutional Network (R-FCN), and Single Shot Multibox Detector (SSD). By using these different detectors on tomato diseases and pests database, they identified nine types of diseases efficiently [37]. Although a handful research papers are available to detect diseases in tomato plants but still there are many research gaps in-between like identification and extraction of several diseases is tough due to variation in symptoms, feature weighing needs to be done at each layer, and also a large number of images are required [28, 37, 38]. Karthik et al. has highlighted these study gaps and tried to combat these issues by attention embedded residual CNN for the first time. They trained the architecture with 95,999 images and validated with 24,001 images, a large number of images were used in this study [39]. They utilised residual extracted features by the initial layers and pooled them in final layers for the fine granular details. Their architecture consisted three blocks of Residual Progressive Feature Extraction (RPFE) with increasing channel at each block and then attention model worked on the top of these RPFEs. To increase the precision, significant features were weighed high compared to other features based on the attention coefficient calculated for each pixel and ~98% accuracy was achieved, a significant increase compared to other studies [39].

The automatic plant disease detection techniques have been extended in fruits as well since it costs considerable yield reduction and subsequently heavy losses. Wang et al. annotated healthy black rot images of apple in the public Plant Village dataset with severity labels. They trained and compared shallow networks of different depth from scratch and fine-tuned top layers of pre-trained deep networks for the precision in prediction of disease severity stage. The performance of fine-tuned models was better than models trained from scratch. The VGG16 model exhibited best result with an accuracy of 90.4%. These results clearly indicated that fine-tuning pre-trained networks might alleviate the problem of insufficient training data or ill-trained sets [36]. Several such studies have been performed with different CNN architecture and the accuracy rate has been compiled in Table 2 (adopted and modified from 40).

Bin Liu trained a dataset of 13,689 synthetic images from apple leaves and run it on AlexNet-based neural network along with two Inception layers to detect four diseases via NAG algorithm. This architecture achieved 97.62% accuracy with reduced parameters by 51,206,928 and improvement by 10.83% compared to standard AlexNet leading to faster convergence [31]. The synthetic images here referred to the expanded images from 1053 natural images of apple leaves provided by two apple experiment workstations installed in China by rotation transformation and mirror symmetry. Another group tried to solve this issue of limited training data via data augmentation. They exploited

Table 1. Parameters used at each step of CNN in different studies

Author's name, year	Crop-disease pair	Image source/capture device	Experimental CNN	Training type	Hardware/software	Results/accuracy (%)
Mohanty et al. (2016)	54,306, 14 plants with 38 classes of diseases	PlantVillage dataset	AlexNet and GoogLeNet	Transfer learning from scratch	Own fork of Caffe	99.35%
Lu et al. (2017)	500, rice	camera	CNN, BP, SVM, PSO	–	–	–
Ferentinos (2018)	87,848, 25 plant species in 58 distinct classes of disease	PlantVillage dataset	AlexNetOWTBn, Overfeat, AlexNet, VGG, and GoogLeNet	Torch7	–	–
Karthik et al. (2020)	95,999, late blight, early blight, and leaf mold on tomato leaves	PlantVillage dataset	–	Residual learning, integrated residual learning, attention-embedded residual learning	–	–
Fuentes et al. (2017)	500, Tomato	Collected from farms	VGG-16, ResNet-50, 101 and 152	Faster-R-CNN, R-FCN, SSD	Intel Core I7 3.5 GHz Processor on two NVidia GeForce Titan X GPUs	R-FCN with ResNet-50 achieves 85.98%

(*continued*)

Table 1. (*continued*)

Author's name, year	Crop-disease pair	Image source/capture device	Experimental CNN	Training type	Hardware/software	Results/accuracy (%)
Wang et al. (2017)	2086, Apple	PlantVillage dataset	VGGNet, Inception-v3, and ResNet50	Building a shallow network fromscratch and transfer learning by finetuning the top layers of a pretrained deep network	Ubuntu workstation equipped with one Intel Core i5 6500 CPU (16 GB RAM), accelerated by one GeForce GTX TITAN X GPU (12 GB memory)	Fine-tuned VGG16 model performs best, achieving an accuracy of 90.4%
Amara et al. (2017)	3700 Banana	PlantVillage project [HS15, MHS16]	LeNet [Le89] architecture, SGD model	–	–	–
Garcia and Barbedo (2019)	1575, 79 diseases affecting 14 plant species	Several different sensors (smartphones, compact cameras, DSLR cameras)	GoogLeNet CNN using Matlab	Pretrained CNN	NVIDIA Titan XP Graphics Processing Unit (GPU)	–

(*continued*)

Table 1. (*continued*)

Author's name, year	Crop-disease pair	Image source/capture device	Experimental CNN	Training type	Hardware/software	Results/accuracy (%)
Mishra et al. (2020)	679, Common rust, Northern Leaf Blight	Smartphone camera (live images) and PlantVillage dataset	NCS and Raspberry pi 3b+ module	–	1.4 GHz 64-bit quad-core processor, 1 GB LPDDR2 SDRAM, full-size HDMI, 4 USB 2.0 ports, BCM 43,438 highly integrated single chip which includes 2.4 GHz WLAN, Bluetooth	Designed deep CNN achieves 98.40% while analysis on captured live images achieves 88.66%
Liu et al. (2017)	1053 inc. to 13,689, Apple with four diseases	Nithish Kannan et al. (2020)	12,206, inc. to 9801, Tomato with four diseases	PlantVillage	ResNet-50	97.62%

Table 2. Accuracy rate (%) when different CNN architecture were used from various types of datasets

Training type model	Shallow	Scratch	Deep
AlexNet	94.15	95.78	99.24
DenseNet	96.53	98.86	99.72
Inception_v3	91.53	97.43	99.76
ResNet34	94.75	98.48	99.67
Squeezenet1_1	96.26	92.49	99.2
VGG13	92.23	97.95	99.49

ResNet50 as a trained model for transfer learning and augmented the data four-fold to the actual data with an accuracy rate of 97%. Some of the frequently used augmentation strategies are scaling, flipping, cropping, rotating, padding, and translation. The library (*torchvision.transforms*) provided by Pytorch is available with several methods of data augmentation that helps in elaborating the dataset size [38].

Though CNN has a long list of aforementioned benefits still there is a limited use of CNN on realistic and varied practical diversity of image datasets. The image datasets available on Plant Village database has a limited number of images and representative in qualitative terms but to obtain quantitative data and proper training of CNN, a large number of images are required. Wallelign presented plant disease detection in FLAIRS-31 in 2018 in which dataset of 12,673 images was downloaded from PlantVillage database and resized to 128*128 pixels. The dataset was then augmented to run on LeNet architecture. Their model consisted 3 convolutional layers followed by pooling layers that ultimately provides 128 kernels with fully-connected 512 neurons. This model was trained using Adam with batch size of 100 for 1000 epochs and the testing accuracy was highest in colored set with a max of 78.74%, suggestive of importance of color feature during feature extraction. However, after augmentation the accuracy rate was increased to 99.32%, indicating towards small dataset hampers the testing accuracy. For training purpose, one should have a good number of images or a large dataset [40].

Even though Mohanty et al. in 2016 demonstrated the technical feasibility of smartphone-assisted disease diagnosis employing HD cameras and high performance processors in mobile devices but they also reported the similar limitation in practical application of these techniques. They used 54,306 images from 38 classes of 14 crops with 26 diseases available on open-database PlantVillage and resized them into 256*256 pixels. They performed all the experiments on three versions of same datasets i.e. coloured, grey-scaled and segmented (shown in Fig. 2) and evaluated the applicability via GoogLeNet and AlexNet on trained dataset or dataset from scratch. They used approx. 60 experimental configurations and found GoogLeNet with trained dataset provides more accurate results up to 99.35% [24] (Fig. 3).

When trained images are different from images captured for testing then accuracy rate reduces substantially. A more diverse training datasets are required to improve the efficacy and prediction precision of these techniques. However, this drawback was

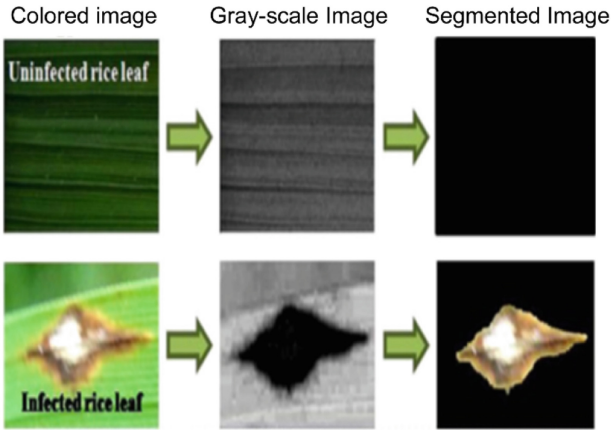


Fig. 3. Coloured, gray scaled and segmented image

undertaken into consideration by Garcia and Barbedo in 2019. They subdivided leaf images from 14 crops with 56 classes into individual spots and lesions that significantly increased number of images from ~2-fold to 20-fold and 79 classes that might identify multiple diseases affecting the same leaf [41]. This group collected images from different databases captured via various sensors like smartphones, compact cameras, DSLR cameras, etc. with variable image resolution ranging from 1 to 24 MP. This group converted PDDB images (~60% images captured in controlled conditions while 40% in practical conditions) into XDB images with some set rules like images from leaves were only considered with 20% healthy tissue area and background was blackened. However, the process of subdividing was done manually to ensure the criteria applied correctly and this might increase the challenge and biasness further. After subdividing, GoogLeNet CNN with MATLAB toolbox was applied using a NVIDIA Titan XP Graphics Processing Unit and accuracy was increased vastly as from 60 to 75% in corn, from 92 to 99% in wheat or 100% in cassava [41].

Another drawback of these techniques is the availability of representative image datasets with the images on homogeneous background that limits its use in practical fields. The number of images (diseased and healthy state) of different crops available on Plant Village Dataset is given in Table 3.

The capturing and labelling of images from the field is a difficult, expensive and time-consuming. This burdensome work can be ameliorated if different research groups make their work available in public domain and enhanced versions of representative databases could be prepared for research purposes and ultimately their practical utility. However, some researchers have developed the concept of social networking to overcome this limitation and to increase its availability in public domain.

Convolution Architecture For Feature Extraction (Caffe) is the first deep learning framework that has been opened for end-user without expertise. The high speed to process massive data with easy-to-use text methodology instead of codes that defines the definition of the model, optimal settings, and pre-training weights supports preference of Caffe model and the corresponding optimizing methods over TensorFlow. Users can

Table 3. No. of images of each plant/crop species available on PlantVillage, an open database

Sr. No.	Name of the crop	Diseased images	Healthy images
1	Apple	1526	1645
2	Blueberry	–	1502
3	Cherry	1052	854
4	Corn	2690	1162
5	Grape	3639	423
6	Orange	5507	–
7	Peach	2297	360

define their own models using the types of neuron layers provided by Caffe. Caffe provides modifiable framework to users that is a BSD-licensed C++ library with Python and MATLAB-based databases for training and testing CNN and other deep learning methods on various datasets [39].

The next hindrance in spreading the use of CNN is the availability of internet in rural areas. To overcome this limitation, Sumita Mishra had recently presented a module that works on cloud platform. This module is consisted of two hardware blocks named NCS and Raspberry pi 3b+ with 1.4 GHz 64-bit quad-core processor, 1 GB LPDDR2 SDRAM, HDMI, and 2.4 GHz WLAN, Bluetooth. The deep convolution network is comprised of convolutional layer, max-pooling layer, drop-out layer, flatten layer and dense layer. This model achieved 99% accuracy after 24 epochs that was stabilised after 18 iterations but this module needs to be checked on more datasets, crop species, and pest-crop pairs [33].

CNN has the potential to diagnose plant diseases efficiently in practical conditions, large field areas still its use lays in initial phase. The more rigorous and active research is required to develop or train more and more datasets from different crop species, crop-disease pairs. The trained datasets should be made available to research groups and hence, composing integrated databases is an unmet need. The use of CNN in plant disease diagnosis can be spread if it would be available to cultivators directly and to achieve this, farmers should be aware of these techniques. The techniques should be user-friendly with no complexity that would escalate the number of images for the database as well as use of CNN in plant disease detection.

4 Summary and Future Directions

Plant diseases cause severe reduction in yield and sustained agriculture. The early and accurate disease detection could ameliorate the losses by proper treatment of crop in time. This would help in mitigating agricultural poverty and might improve the number of cultivators. The manual detection of plant diseases is time-consuming, inconsistent and non-reliable. In lieu of overcoming these limitations, automation of techniques outpaces manual diagnosis by an expert with time. CNN has proved its efficacy in plant disease

detection in different crops like tomato, maize, banana, apple, corn in practical fields. CNN can extract the feature from images captured in different conditions irrespective of illumination, resolution, image size, pixel size. However, CNN still stands in its initial phase of development. It has few challenges like each pixel of the space borne SAR (synthetic aperture radar) imagery is characterized by backscatter phase and intensity in multiple polarizations. Both data sources have multi-temporal nature and different spatial resolutions. Information fusion is thus important in the future to make DL more applicable in this area.

There are multiple DL Architectures, implemented on different plants datasets and get results. Our aim to make a better result with defining the multiple classes of diseases and also get the solution for protecting the diseases at early stage. For this, trying to make a novel approach so that the result in terms of loss and efficiency would be better.

CNNs has their practical utility in not only plant disease detection but it can detect weather conditions, land types, soil texture and quality, fruit counting that signifies yield production. All these data types can be composed in a grid form model and genetic algorithm could be produced to provide optimal solutions to farmers, consumers and the Government to plan efficient strategies for selling, purchasing and food shortage relief. This architecture might help in precision agriculture and benefit agroindustry.

References

1. Sandhu, G.K., Kaur, R.: Plant disease detection techniques: a review. In: International Conference on Automation, Computational and Technology Management, ICACTM 2019, pp. 34–38 (2019). https://doi.org/10.1109/ICACTM.2019.8776827
2. NCRB: NCRB data shows 42,480 farmers and daily wagers committed suicide in 2019—the economic times. The Economic Times (2019)
3. Amutha, D.D.: Present status of Indian agriculture. SSRN Electron. J. (2016). https://doi.org/10.2139/ssrn.2739231
4. Liang, Q., Xiang, S., Hu, Y., Coppola, G., Zhang, D., Sun, W.: PD 2 SE-Net: computer-assisted plant disease diagnosis and severity estimation network. Comput. Electron. Agric. **157**, 518–529 (2019). https://doi.org/10.1016/j.compag.2019.01.034
5. Ferentinos, K.P.: Deep learning models for plant disease detection and diagnosis. Comput. Electron. Agric. **145**, 311–318, (2018). https://doi.org/10.1016/j.compag.2018.01.009
6. Chaudhary, P., Chaudhari, A.K., Cheeran, A.N., Godara, S.: Color transform based approach for disease spot detection on plant leaf. www.ijcst.org (2012). Accessed 21 October 2020
7. Rath, A.K., Meher, J.K.: Disease detection in infected plant leaf by computational method. Arch. Phytopathol. Plant Prot. **52**(19–20), 1348–1358 (2019). https://doi.org/10.1080/03235408.2019.1708546
8. Kamilaris, A., Prenafeta-Boldú, F.X.: Deep learning in agriculture: a survey. Comput. Electron. Agric. (Elsevier B.V.) **147**, 70–90 (2018). https://doi.org/10.1016/j.compag.2018.02.016
9. Zhu, N., et al.: Deep learning for smart agriculture: concepts, tools, applications, and opportunities. Int. J. Agric. Biol. Eng. **11**(4), 32–44 (2018). https://doi.org/10.25165/ijabe.v11i4.4475
10. Raghuveera, E., Kumar, N.P., Yeswanth, A.S., Pavan, L.S.M.: Soil Moisture Monitoring System Using Iot (2019)
11. Amara, J., Bouaziz, B., Algergawy, A.: A Deep Learning-based Approach for Banana Leaf Diseases Classification (2017)

12. Riley, M., Williamson, M., Maloy, O.: Plant disease diagnosis. Plant Heal. Instr. (2002). https://doi.org/10.1094/phi-i-2002-1021-01
13. Abu-Saqer, M.M., Abu-Naser, S.S.: Developing an expert system for papaya plant disease diagnosis. Int. J. Acad. Eng. Res. **3**(4), 14–21 (2019). http://dspace.alazhar.edu.ps/xmlui/handle/123456789/93. Accessed 17 October 2020
14. Salman, F.M., Abu-Naser, S.S.: Rule based system for safflower disease diagnosis. Int. J. Acad. Eng. Res. **3**(8), 1–10 (2019). www.ijeais.org/ijaer
15. Mahlein, A.K.: Plant disease detection by imaging sensors—parallels and specific demands for precision agriculture and plant phenotyping. Plant Dis. **100**(2). Am. Phytopathol. Soc. 241–254. (2016). https://doi.org/10.1094/PDIS-03-15-0340-FE
16. Saradhambal, G., Dhivya, R., Latha, S., Rajesh, R.: Plant Disease Detection and its Solution Using Image Classification. http://www.ijpam.eu (2018). Accessed 17 October 2020
17. Bhagat, M., Kumar, D., Haque, I., Munda, H.S., Bhagat, R.: Plant leaf disease classification using grid search based SVM. In: 2nd International Conference on Data, Engineering and Applications, IDEA (2020). https://doi.org/10.1109/IDEA49133.2020.9170725
18. Fang, Y., Ramasamy, R.P.: Current and prospective methods for plant disease detection. Biosensors (MDPI AG) **5**(3), 537–561 (2015). https://doi.org/10.3390/bios5030537
19. Chen, Y., Fan, K., Hung, S., Chen, Y.: The role of peptides cleaved from protein precursors in eliciting plant stress reactions. New Phytol. **225**(6), 2267–2282 (2020). https://doi.org/10.1111/nph.16241
20. Moghadam, P., Ward, D., Goan, E., Jayawardena, S., Sikka, P., Hernandez, E.: Plant disease detection using hyperspectral imaging. In: DICTA 2017—2017 International Conference on Digital Image Computing: Techniques and Applications, vol. 2017-December, pp. 1–8 (2017). https://doi.org/10.1109/DICTA.2017.8227476
21. Bishop, C.: Pattern Recognition and Machine Learning—CERN Document Server, p. 738. Springer, New York, NY (2006)
22. Brahimi, M., Arsenovic, M., Laraba, S., Sladojevic, S., Boukhalfa, K., Moussaoui, A.: Deep Learning for Plant Diseases: Detection and Saliency Map Visualisation, pp. 93–117. Springer, Cham (2018)
23. Saleem, M.H., Potgieter, J., Arif, K.M.: Plant disease detection and classification by deep learning. Plants (MDPI AG) **8**(11), 468 (2019). https://doi.org/10.3390/plants8110468
24. Mohanty, S.P., Hughes, D.P., Salathé, M.: Using deep learning for image-based plant disease detection. Front. Plant Sci. **7**, (2016). https://doi.org/10.3389/fpls.2016.01419
25. Golhani, K., Balasundram, S.K., Vadamalai, G., Pradhan, B.: A review of neural networks in plant disease detection using hyperspectral data. Inf. Process. Agric. (China Agricultural University) **5**(3), 354–371 (2018). https://doi.org/10.1016/j.inpa.2018.05.002
26. Toda, Y., Okura, F.: How convolutional neural networks diagnose plant disease. Plant Phenomics 1–14 (2019). https://doi.org/10.34133/2019/9237136
27. Sainath, T.N., Mohamed, A.R., Kingsbury, B., Ramabhadran, B.: Deep convolutional neural networks for LVCSR. In: ICASSP, IEEE International Conference on Acoustics, Speech and Signal Processing—Proceedings, pp. 8614–8618 (2013). https://doi.org/10.1109/ICASSP.2013.6639347
28. Francis, M., Deisy, C.: Disease detection and classification in agricultural plants using convolutional neural networks—a visual understanding. In: 6th International Conference on Signal Processing and Integrated Networks, SPIN 2019, pp. 1063–1068 (2019). https://doi.org/10.1109/SPIN.2019.8711701
29. Ponti, M.A., Ribeiro, L.S.F., Nazare, T.S., Bui, T., Collomosse, J.: Everything you wanted to know about deep learning for computer vision but were afraid to ask. In: Proceedings—2017 30th SIBGRAPI Conference on Graphics, Patterns and Images Tutorials SIBGRAPI-T 2017, vol. 2018-January, pp. 17–41 (2018). https://doi.org/10.1109/SIBGRAPI-T.2017.12

30. Too, E.C., Yujian, L., Njuki, S., Yingchun, L.: A comparative study of fine-tuning deep learning models for plant disease identification. Comput. Electron. Agric. **161**, 272–279 (2019). https://doi.org/10.1016/j.compag.2018.03.032
31. Liu, B., Zhang, Y., He, D.J., Li, Y.: Identification of apple leaf diseases based on deep convolutional neural networks. Symmetry (Basel) **10**(1), (2018). https://doi.org/10.3390/sym10010011
32. Zhang, S., Zhang, S., Zhang, C., Wang, X., Shi, Y.: Cucumber leaf disease identification with global pooling dilated convolutional neural network. Comput. Electron. **162**, 422–430 (2019). https://www.sciencedirect.com/science/article/pii/S0168169918317976. Accessed 26 October 2020
33. Mishra, S., Sachan, R., Rajpal, D.: Deep convolutional neural network based detection system for real-time corn plant disease recognition. Procedia Comput. Sci. **167**, 2003–2010 (2020). https://doi.org/10.1016/j.procs.2020.03.236
34. Khirade, S.D., Patil, A.B.: Plant disease detection using image processing. In: Proceedings—1st International Conference on Computing, Communication, Control and Automation, ICCUBEA 2015, pp. 768–771 (2015). https://doi.org/10.1109/ICCUBEA.2015.153
35. Sharath, D.M., Akhilesh, S., Kumar, A., Rohan, M.G., Prathap, C.: Image based plant disease detection in pomegranate plant for bacterial blight. In: Proceedings of the 2019 IEEE International Conference on Communication and Signal Processing, ICCSP 2019, pp. 645–649 (2019). https://doi.org/10.1109/ICCSP.2019.8698007
36. Wang, G., Sun, Y., Wang, J.: Automatic image-based plant disease severity estimation using deep learning. Comput. Intell. Neurosci. **2017**, (2017). https://doi.org/10.1155/2017/2917536
37. Fuentes, A., Yoon, S., Kim, S., Park, D.: A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. Sensors **17**(9), 2022 (2017). https://doi.org/10.3390/s17092022
38. Kaushik, M., Prakash, P., Ajay, R., Veni, S.: Tomato leaf disease detection using convolutional neural network with data augmentation. 1125–1132 (2020). https://doi.org/10.1109/icces48766.2020.9138030
39. Karthik, R., Hariharan, M., Anand, S., Mathikshara, P., Johnson, A., Menaka, R.: Attention embedded residual CNN for disease detection in tomato leaves, Appl. Soft Comput. J. **86**, (2020). https://doi.org/10.1016/j.asoc.2019.105933
40. Wallelign, S., Polceanu, M., Buche, C.: Soybean plant disease identification using convolutional neural network. www.aaai.org (2018). Accessed 26 October 2020
41. Arnal Barbedo, J.G.: Plant disease identification from individual lesions and spots using deep learning. Biosyst. Eng. **180**, 96–107 (2019). https://doi.org/10.1016/j.biosystemseng.2019.02.002

On Applying Gradient Based Thresholding on the Canny Edge Detection Results to Improve the Effectiveness of Fuzzy Hough Transform for Colonoscopy Polyp Detection Purposes

Raneem Ismail, Péter Prukner, and Szilvia Nagy(✉)

Széchenyi István University, Egyetem tér 1, 9026 Győr, Hungary
peter.prukner@ddc.sze.hu, nagysz@sze.hu

Abstract. The possibilities of improving the effectiveness of fuzzy Hough transform calculations in the detection of colonoscopy image processing and polyp detection by gradient based preprocessing is studied. For the Hough transforms a black and white image consisting of line segments is necessary, thus most of the times an edge detected image is used as the basis of the transform. Here, the gradient magnitude values corresponding to the Canny edge pixels of the image are used for determining, which are the typical magnitude values for the polyp contours, in order to remove part of the non-contour edges from the image. Three publicly available databases with images and ground truth masks are used for testing, whether a general threshold for the gradients is applicable, but based on the histograms it seems to be not possible to generate a database independent normalized gradient based domain that can be used for sorting out the unnecessary edges.

Keywords: Colorectal polyp detection · Computer aided diagnosis · Fuzzy Hough transform · Canny edge detection · Image gradient

1 Introduction

Colorectal cancer (CRC) is a rather common lethal cancer type, and its survival rate varies depending on the stage at which it is detected, ranging from more than 95% in early stages to less than 35% in later stages [1]. There are several techniques for screening the gastrointestinal tract and detecting colorectal polyps, but the most effective is still the colonoscopy, which is a special type of endoscopy, developed to be applied at the lower part of the bowel. In its movable head, the coloscope has a camera and a light source, as well as a loop for pulling off the affected tissue for biopsy or removing the entire polyps [2].

Polyp detection and localization in colonoscopy images is the first and most important step that leads to precise final classification and which prevents the development of early-stage precancerous polyps into malicious cancerous polyps if they are detected and excised in time. Currently, automatic polyp detection and localization has a high

R. Kountchev et al. (Eds.): *NAMSP 2022*, SIST 332, vol. 332,, pp. 110–121, 2023.
https://doi.org/10.1007/978-981-19-7842-5_10

importance and one of the world's public health priorities [3], as well as a field of interest for many computer aided diagnostics (CAD) researches.

Recently, Deep Learning (DL) is getting popular in addressing this challenge. Several research studies applied different types of Deep Neural Networks in order to solve this task and help colonoscopists enhance their diagnosis performance by automatically detecting and localizing colorectal polyps [4–6]. Besides various neural networks, many polyp characterization methods based on the computation of some shape-based [1, 7] and texture-based [8, 9] feature descriptors over the entire image or a tile of it are also reported in the literature.

In [10] the classical Hough transform is proven to be rather effective in detecting possible polyp outline segments, and an, intelligent algorithm for determining whether the found curve really belong to a polyp was proposed, too. In [11], we suggested to replace the classical Hough transform to fuzzy Hough transform, as it is more tolerant to imprecisions of the points of the curve, and in real life polyp images, the polyps are not completely circular.

However, the Hough transform, especially the fuzzy version requires a lot of calculations, if the edge detection algorithm, which is the first step of the transform, detects too many edges. For this reason, as many non-polyp edges as possible should be suppressed. The colonoscopy images have edges not only at the borders of the polyps, but also at bowel folds, reflections and at the edges of the area that can be sensed by the camera. In many cases impurities have also visible edges, veins can show rather large contrast to the bowel walls and sometimes even the structure of the bowel tissue is visible, although these edges are very often blurry and have intensity steps that are mostly less steep and smaller in magnitude. As these smaller magnitude intensity change steps are producing unnecessary edges with most of the edge detecting algorithms, it would be advisable to remove them from the picture before the Hough transform. Also, the intensity changes around the sensed area as well as around the reflections are mostly larger than those around the polyps, except if the polyp is in lateral view in front of the dark continuation of the bowel. These observations gave us the idea to propose a method to remove those edges that have too small or too high gradient magnitude but if possible, only those edges that are not around the polyp contour.

For this purpose, the distribution of the gradient magnitudes at the edge pixels is studied in this contribution in 2 regions. First in the complete image, and second in the surroundings of the borderline of the ground truth mask. The ground truth mask contours could not be used because the manually drawn masks often did not meet exactly the borders of the polyps, and also because the polyp borders were often not fully visible because of the lighting conditions, bowel folds or impurities.

The purpose is to find out, whether it is possible to determine a general gradient domain that is interesting for polyp detection, and thus making it possible to decrease the calculation load, and increase the speed of the calculations in a fuzzy Hough transform-based colonoscopy polyp detection algorithm.

The paper is organized the following way. In Sect. 2, the necessary mathematical and scientific background will be outlined, from the basic idea of Hough transform and its fuzzy version through the gradient filtering and Canny edge detecting. Section 3 contains the setup of the calculations, the used databases, and the experiments as well as

the metrics used for determining the effectiveness of the preprocessing method. Section 4 gives the results, and a short conclusion is written in Sect. 5.

2 Background Considerations

2.1 Classical and Fuzzy Hough Transform

Hough transform (HT) is one of the most efficient algorithms used in the computer vision and objects detection area since the first appearance of its classical version for machine analysis of bubble chamber pictures by P.V.C Hough in 1959 [12]. Later, a number of interested researchers adjusted Hough transform to create new versions.

In 1994, Han, Kóczy and Poston introduced the fuzzy Hough transform [13] to find fuzzy lines and circles in the noisy environments by approximately fitting the data points to given parametric shapes.

Hough transforms have been widespread for lines and circles detection applications especially in the fields of lanes and roads detection [14], in industrial automation domain [15], and for medical images diagnosis as well [16].

The efficiency of Hough transform application has been studied in the field of polyp detection. In [10] authors used classical Hough transform to define the possible regions of interest (ROIs), where the polyps might exist in 300 videoendoscopy images. The classical Hough transform based on Canny–Deriche filter edge detection method allowed a good detection of the ROIs containing a polyp. On the other hand, in some samples, the polyp was even visually difficult to detect due to the surrounding noise that resulted in multiple alternative weaker circles which increased the False Positive Rate (FPR) of the classification system. To improve the overall method performance, they followed the Hough transform step by computing the textural features from co-occurrence matrices which were used later within a boosting-based approach in the final classification step.

To improve the results of the classical Hough transform, as a first step, we studied, how useful the fuzzy Hough transform in finding not precisely circular-shaped colorectal polyps in [11]. We found that the circular fuzzy Hough transform results in circles that fit the polyp at least as well as the classical Hough transform output, but may result in one circle instead of multiple ones. These resulting circles can be used as initial masks to more refined methods, like an active contour segmentation methods [17, 18].

2.2 Gradient Filtering, Gradient Magnitude

Mathematically, partial derivatives of a 2D continuous function represent the amount of change along each dimension. An edge in an image is a rapid change in the image intensity in a small region. This leads us to use gradient filtering in image processing field for detecting edges. The gradient of the image intensity function at each image point is a 2D vector with two elements identified by the derivatives in the horizontal and vertical directions. From these two numbers at each pixel, we can find both the strength of the magnitude of the edge as well as the orientation of the edge.

The previously mentioned are all done in the continuous domain. When it comes to the digital image and the intensity function has been sampled at image discrete points

we can actually apply the gradient operator by convolving the image with a kernel which creates the difference of the neighboring pixels for each pixel.

A variety of gradient operators have been proposed such as Roberts, Prewitt, and Sobel. They use convolutional filters to detect the variation in both the x and y direction of the picture. For our purposes we used Sobel type gradient, which has kernel of

$$Sobel_x = \begin{bmatrix} 1 & 0 & -1 \\ 2 & 0 & -2 \\ 1 & 0 & -1 \end{bmatrix}, Sobel_y = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{bmatrix}. \quad (1)$$

2.3 Canny Edge Detection

Canny edge detection was firstly introduced by John Canny in 1986 [19]. It is the most widely used edge detection technique in many computer vision and image processing applications, as it focuses not only on high gradient image points, but also on the connectedness of the edge points, thus it results in very nice, edge-like images, that is close to the human concept of edges.

Canny is a multistep operator which consists of the following four main steps. First, the original image is smoothed using a Gaussian filter in order to get rid of the undesired noise which makes the edge detection process more reliable and accurate. Then, the gradient magnitude and direction of the smoothed image is computed. As a third step, the Non-Maximum Suppression (NMS) technique is used. For the final step, two hysteresis thresholds (high and low thresholds) are calculated. In this step, edge points that have gradient values bigger than the higher threshold are defined as strong edges. The edge points with gradient values lower than the lower threshold are discarded. For the residual edge points, they are classified as suspicious, weak edge points, and their connectivity is investigated. Finally, a weak edge point is considered an edge pixel only if an adjacent pixel is definitely a strong edge pixel; otherwise, it is not an edge pixel [20].

3 Applied Methods and Evaluation Metrics

Because Hough transform involves mapping each pixel in the edge detected image, it is a time-consuming procedure. Furthermore, the amount of computation necessary to achieve Hough transform grows exponentially as the number of curve parameters rises; as a result, this makes the transform impractical in terms of the time-consuming process and the requirements of great storage and computational capabilities. The research community has investigated the different aspects of the Hough transform limitations and suggested different approaches as solutions.

Some researchers attempted to reduce the computational time [21, 22], while others worked on improving the voting procedure which is an advantageous modification for both processing cost and object detection accuracy [23]. In the present paper we focus on the increasing of the computational effectiveness by decreasing the number of unnecessary points to be fuzzy Hough transformed.

As the basic idea, we recall that the gradient magnitudes of the colorectal polyp borders are expected to be neither too large, nor too small, at least in a significantly large

part of the polyp contour. To test this assumption, we used three colonoscopy image databases, the CVC Clinic [1], the CVC Colon [3], and the ETIS Larib [10], all of them publicly available; they contain different number of polyps with different resolution and different image quality. All images have polyps on them, and the pixels belonging to the polyps, i.e., the ground truth masks are given in separate images. An example can be seen in the top row of Fig. 1. It is visible, that the polyp contour is not always visible, and the ground truth mask is often a couple of pixels off from the real edges; this is the reason, why an extension (to a finite width ring) of the manually drawn contour was necessary, when determining whether an edge belongs to the contour or not. The preprocessing of the image as well as the evaluation of the data are listed in steps in the followings.

1. Reflection removal: As a first step in decreasing the number of unnecessary pixels, the reflections (and thus their contours) were removed from the images (see Fig. 1.a). The reflection removal was based on the histogram of the image pixel intensities.

   i. The highest (and also the lowest) intensity peak in the histogram was cut off and then the histogram was stretched back to its original 0–255 domain.
   ii. The pixels belonging to the highest peak of the histogram were collected into a "white mask".
   iii. The "white mask" was extended and smoothened into the environment, similarly to the method detailed in [24].

2. Polyp contour: The contour (Fig. 1.c) of the ground truth mask (Fig. 1.b) was determined. The *No. of pixels in mask contour* was calculated.
3. Preparation of the contour "ring mask": The contour is extended into a ring (Fig. 1.d) by selecting the first x nearest neighbors of all the contour. The width of the ring mask is based on the image size (it is $x = 3$ for database [1], $x = 5$ for [3], and $x = 10$ for [10]).
4. Edge detection: 3 by 3 median filtering and Canny edge filtering, see Fig. 1.e. Calculating the *total No. of edge pixels*.
5. Collecting the edge pixels that belong to the polyp contour: The edge detected image (Fig. 1.e) was multiplied with the extended polyp contour "ring mask" (Fig. 1.d). Counting the number of the arising white pixels, i.e., the *No. of edge pixels in ring mask*.
6. Calculation of the gradient magnitudes: By using Sobel gradients (1), like in Fig. 1.f.
7. Normalization: For every picture in all the 3 databases the edge-masked gradients were normalized to the interval [0, 1]. This step makes a universally applicable method, even if the pixel intensity dynamic range varies for all the images.
8. Gradient weighted edges: The edge filtered image (Fig. 1.e) was multiplied by the gradient filter output (Fig. 1.f), thus receiving Fig. 1.g type images.
9. Histograms of gradient value distribution: 10-bin histograms of all the images were generated, and studied to determine the dynamic distribution of the gradients for the edges of both the ring mask (i.e., the surroundings of the ground truth mask contour) and the full image. Sample histograms can be found in Fig. 2. It is visible, that there are cases, when the distribution of the ring mask edges (yellow) and the full image edges (cyan) are not similar, their support is different, like the one on the

left hand side, while in other cases, both distributions follow the same tendency, just the magnitude is smaller.

10. Thresholding: The pixels belonging to the upper and lower parts of the histograms (below and above the thresholds) were set to value 0 (unmasked). The total number of edge pixels, the total number of edge pixels in the ring masks, as well as the total number of pixels in the edge contour were determined for the thus arising images, too.

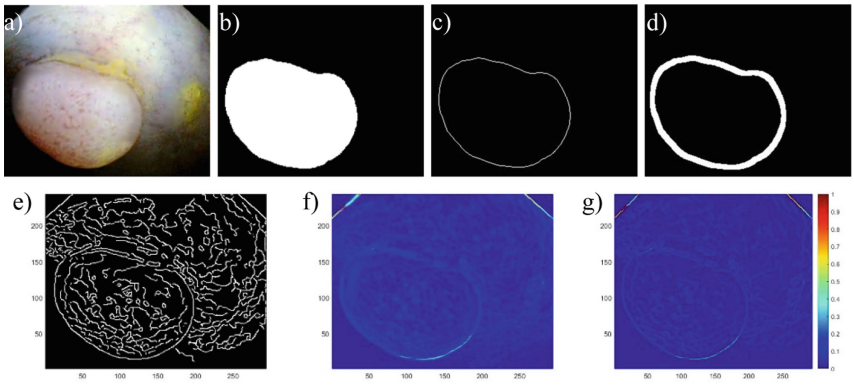


Fig. 1. Upper row: (a) a sample preprocessed image from database CVC Clinic [1], (b) its ground truth polyp mask, (c) mask contour, and (d) extended contour ring mask, i.e., the 3 nearest neighbors in all directions for all the contour pixels. Lower row: (e) the Canny edge detected version of the image, i.e., the edge mask, (f) the gradient magnitude filtered image, and (g) the edge masked gradient magnitude image. It is visible, that subplot (g) is the multiplication of subplots (e) and (f), i.e., it contains the gradient values only where the edge mask value is 1 (white pixels in (e)). Subplot (a) has a 3×8 bit colour depth, the black and white picture contain only 0 (black) and 1 (white) values, while the last 2 subplots share the same colourmap.

Two metrics were used to qualify the edge finding efficiency. First, the ratio between the number of edge pixels in the ring mask around the polyp contour and the total number of edge pixels, i.e.,

$$R\_calc = \frac{No.\ of\ \ edge\ pixels\ in\ the\ ring\ mask}{total\ No.\ of\ \ edge\ pixels}. \tag{2}$$

This ratio characterizes the goodness of the edge detecting with respect to the Hough transforms and polyp detection, the larger this number is, the less non-mask contour edges are found, thus the less unnecessary calculation is needed in the classical or fuzzy Hough transform.

Second, the relation to the ideal mask contour was given by the ratio of the number of edge pixels in the ring mask and the number of pixels in the actual ground truth contour, i.e.,

$$R\_edge = \frac{No.\ of\ \ edge\ pixels\ in\ the\ ring\ mask}{No.\ of\ \ pixels\ in\ mask\ contour}. \tag{3}$$

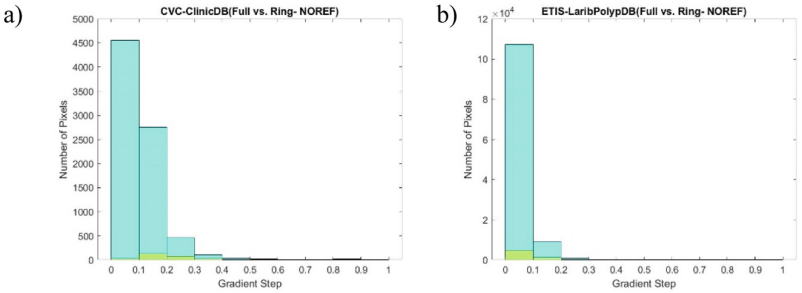


Fig. 2. Individual histogram samples of the edge pixels regarding their normalized gradient values. The teal columns represent the whole image edge pixels, whereas the yellow ones only the ring mask edge pixels in the given normalized gradient intervals.

This determines the goodness of the polyp contour finding; this number should be as close to 1 as possible.

4 Results

The histograms are not as visibly different as in the first case in Fig. 2: often, the distribution of the gradients in the contour edges was rather similar to the distribution of all the edges. The total histograms for the 3 databases are summarized in Figs. 3 and 4. Both linear and logarithmic scale plots are given for the same histograms beside each other to make it possible to see the small details in the higher normalized gradient domains as well as the dominance of the lower gradient valued pixels among all the pixels. Not only the histograms of the total image edge pixels are shown, but also the histograms of the edge pixels within the ring masks.

A systematic test was carried out to determine the most applicable normalized gradient magnitude range that can be used for filtering out some of the unnecessary edge pixels and keeping as many necessary ones as possible. There were no really optimal domains, as even though *R_calc*, became more advantageous, the *R_edge* got less good for many images. (I.e., even though the number of unnecessary pixels decreased, the number of edge pixels around the polyp contour decreased too much). However, for the databases CVC Clinic and CVC Colon, the removal of the lowest 10 percent of the gradient magnitude dynamic range resulted in a still recognizable polyp contour with much lower total number of remaining edge points to be Hough transformed, for most of the cases. In the case of the very good quality, detailed, large sized images of database ETIS Larib [10], the edges behave in a rather similar way both for the full image and for the ring mask around the ground truth polyp contour; they are concentrated in the lowest 20% of the dynamic range of the gradient magnitude, thus the removal of the edges corresponding to the lowest 10% of the gradient magnitude dynamic range is not applicable.

As an example, for database CVC Clinic, a comparison between three different threshold ranges is given in Fig. 5. In the first subplot the total number of edge pixels is shown. The 2nd and 3rd subplots present the metrics *R_calc* and *R_edge*, respectively.

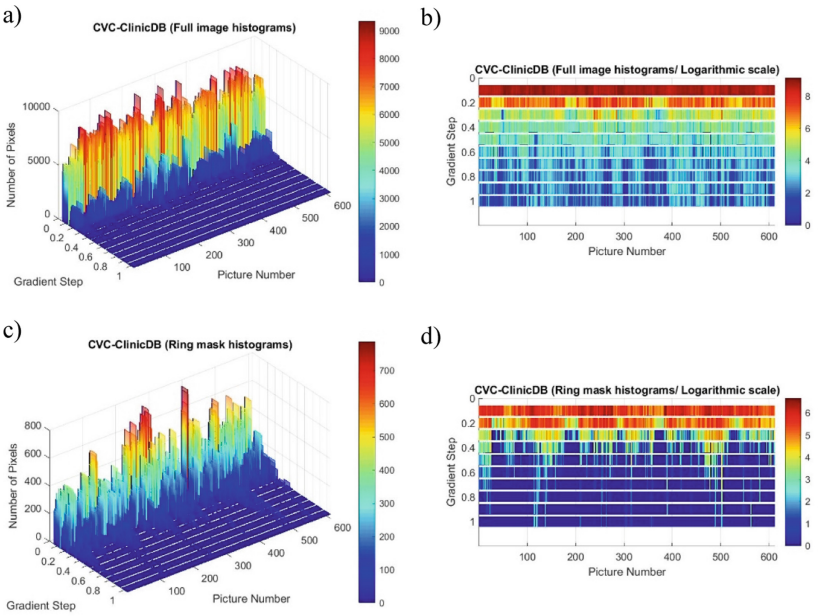


Fig. 3. 10-bin histograms of the normalized gradient values (a)–(b) for the edge points in the full image, and (c)–(d) for the edge points in the ring masks, for database CVC Clinic. The horizontal axes are the normalized gradient magnitudes and the picture number in the given database. Linear (1st column) and logarithmic (2nd column) scale plots of the same histogram are both given, because the smaller valued histogram parts at the higher normalized gradient values cannot be seen well in the linear scale first column plots. The logarithmic scale plots are shown from top view to make the complete set of data visible without columns covering the ones behind them.

Removing the edges with lower gradient magnitude decreases the number of edge pixels significantly for most of the images (as it can be seen in Fig. 5.a, however, based on Fig. 5.b, there are image groups, where the number of the edge pixels around the polyp contour decreases almost proportionally to the total edge pixel number (i.e., the green line in Fig. 5.b does not increase compared to the previous ones), which is of course not beneficial. Also, the ratio that represents the goodness of finding the mask in the subplot c) drops much too close to 0 for many pictures, endangering the effectiveness of the Hough transforms.

5 Conclusion

In this paper, the distribution of the gradient magnitudes corresponding to edge pixels of colonoscopy image databases were studied in order to determine, whether it is possible to remove those edges that do not belong to polyp contours based on gradient magnitude thresholding. The goal of the study was to determine, if the effectiveness of fuzzy Hough transform can be improved by thus decreasing the number of pixels to be transformed, and still keeping that many edge points, that the polyp contour would be detectable.

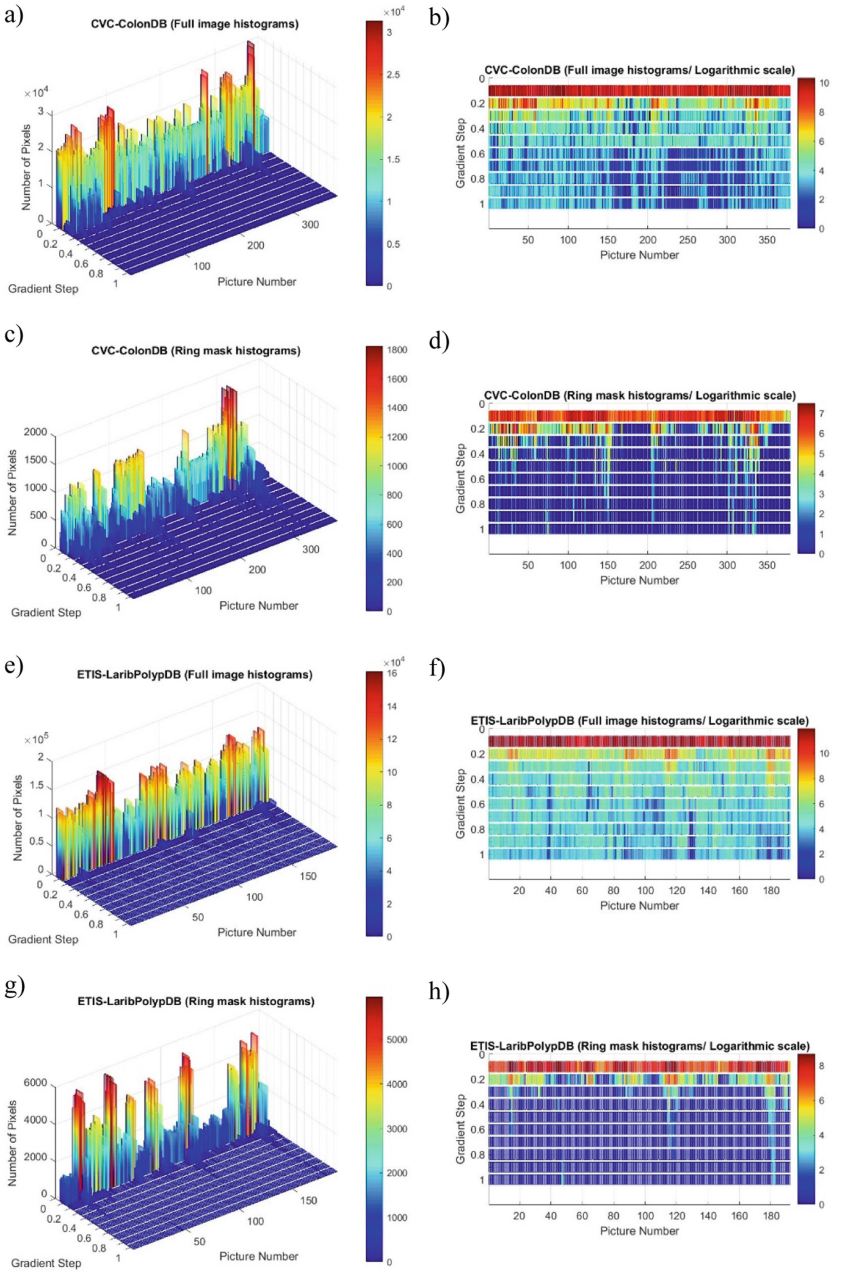


Fig. 4. 10-bin histograms similar to Fig. 3, for databases CVC Colon (from (a) to (d)) and ETIS LaribPolypDB (from (e) to (h)).

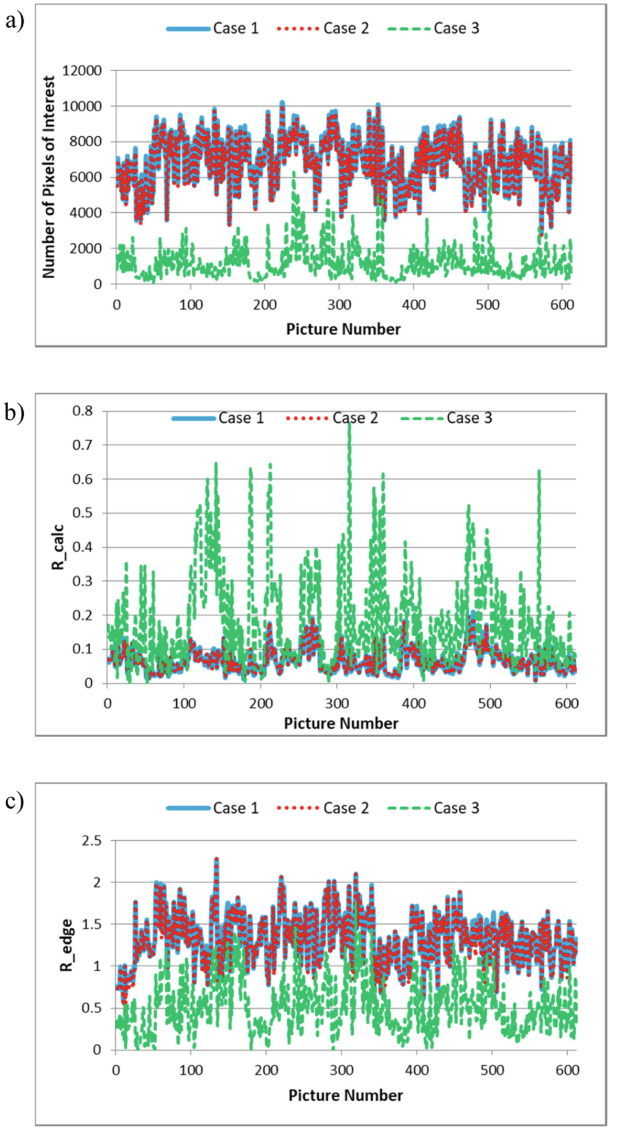


Fig. 5. Plot (a) The number of pixels of interest, i.e., the total number of edge pixels, (b), the ratio between the number of edge pixels in the ring mask and the total number of edge pixels *R_calc* and (c) the ratio between the number of edge pixels in the ring mask vs. the number of pixels in the ground truth mask's contour *R_edge*, for the database CVC Clinic. The horizontal axis is always the picture number in the given database. The normalized gradient thresholds in the 3 cases are the following, Case 1: [0, 1], Case 2: [0, 0.4], Case 3: [0.1, 0.4].

Individual histograms were created for all the images of 3 different public databases about the distribution of the gradient values in the edge pixels in the whole image and in the near surroundings of the manually drawn ground truth mask contours.

Also quantities for determining the effectiveness of the pixel number reduction and the matching with the mask contour were introduced to evaluate the goodness of the gradient magnitude histogram based thresholding.

Generally, despite some promising individual results, there were no uniform gradient ranges between the three studied databases. Thus we concluded, that it is not possible to give a universal gradient magnitude threshold domain that would assist to achieve both the low total number of edge pixels and still allow to find the polyp contour by Hough transform. This indicates that the calculations load of the Hough transform using the current (without thresholding) version of Canny detected images remains high.

Since the low computational load is a crucial requirement for algorithms used in automatic detection systems, the selection of an alternative edge detection method used as a preprocessing step for fuzzy Hough transform for colonoscopy polyp detection purposes must be tested. Also, it might be possible to find a uniform algorithm for determining threshold domain for edge filtered images.

References

1. Bernal, J., Sánchez, F., Esparrach, G., et al.: WM-DOVA maps for accurate polyp highlighting in colonoscopy: validation vs. saliency maps from physicians. Comput. Med. Imag. Graph **43**(6), 99–111 (2015)
2. Schlemper, R.J., Hirata, I., Dixon, M.F.: The macroscopic classification of early neoplasia of the digestive tract. Endoscopy **34**(2), 163–168 (2002)
3. Bernal, J., Sanchez, F.J., Vilariño, F.: Towards automatic polyp detection with a polyp appearance model. Pattern Recogn. **45**(9), 3166–3182 (2012)
4. Ahmad, O.F., Brandao, P., Sami, S.S., et al.: Artificial intelligence for real-time polyp localization in colonoscopy withdrawal videos. Gastroint. Endosc. **89**(6), AB647 (2019)
5. Sornapudi, S., Meng, F., Yi, S.: Region-based automated localization of colonoscopy and wireless capsule endoscopy polyps. MDPI, Appl. Sci. **9**(12), 2404 (2019)
6. Wittenberg, T., Zobel, P., Rathke, M., Mühldorfer, S.: Computer aided detection of polyps in whitelight-colonoscopy images using deep neural networks. Curr. Dir. Biomed. Eng. **5**(1), 231–234 (2019)
7. Yuji, I., Akira, H., Yoshinori, A., et al.: Automatic detection of polyp using hessian filter and HOG features. In: 19th International Conference in Knowledge Based and Intelligent Information and Engineering Systems. Procedia Comput. Sci. **60**(1), 730–739 (2015)
8. Rácz, I., Horváth, A., Szalai, M., et al.: Digital image processing software for predicting the histology of small colorectal polyps by using narrow-band imaging magnifying colonoscopy. Gastroint. Endosc. **81**(5), AB259 (2015)
9. Georgieva, V.M., Nagy, Sz., Kamenova, E., Horváth, A.: An approach for pit pattern recognition in colonoscopy images. Egypt. Comput. Sci. J. **39**(2), 72–82 (2015)
10. Silva, J., Histace, A., Romain, O., Dray, X., Granado, B.: Toward embedded detection of polyps in WCE images for early diagnosis of colorectal cancer. Int. J. Comput. Assist. Radiol. Surg. **9**(2), 283–293 (2013). https://doi.org/10.1007/s11548-013-0926-3
11. Nagy, Sz., Ismail, R., Sziová, B., Kóczy, L.T.: On classical and fuzzy Hough transform in colonoscopy image processing. In: Proceedings of IEEE AFRICON 2021, Virtual Conference, pp. 124–129. Arusha, Tanzania, 13–15 September (2021)
12. Hough, P.V.C.: Machine analysis of bubble chamber pictures. In: Proceedings of the International Conference on High Energy Accelerators and Instrumentation, HEACC 1959, pp. 554–558. CERN, Geneva, Switzerland, 14–19 September (1959)

13. Han, J.H., Kóczy, L.T., Poston, T.: Fuzzy Hough transform. Pattern Recognit. Lett. **15**(7), 649–658 (1994)
14. Liu, W., Zhang, Z., Li, S., Tao, D.: Road detection by using a generalized Hough transform. MDPI, Remote Sens. **9**(6), 590 (2017)
15. Pugin, E., Zhiznyakov, A., Zakharov, A.: Pipes Localization Method Based on Fuzzy Hough Transform. In: Abraham, A., Kovalev, S., Tarassov, V., Snasel, V., Vasileva, M., Sukhanov, A. (eds.) IITI 2017. AISC, vol. 679, pp. 536–544. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-68321-8_56
16. Nagy, Sz., Kovács, M., Sziová, B., Kóczy, L.T.: Fuzzy Hough transformation in aiding computer tomography based liver diagnosis. In: 2019 IEEE AFRICON, pp. 1–4. Accra, Ghana, 15–17 September (2019)
17. Ismail, R., Nagy, S.: On Metrics Used in Colonoscopy Image Processing for Detection of Colorectal Polyps. In: Kountchev, R., Mironov, R., Li, S. (eds.) New Approaches for Multidimensional Signal Processing. SIST, vol. 216, pp. 137–151. Springer, Singapore (2021). https://doi.org/10.1007/978-981-33-4676-5_10
18. Ismail, R., Nagy, Sz.: Ways of improving of active contour methods in colonoscopy image segmentation. Image Anal. Ster., Original Research Paper **41**(1), 7–23 (2022)
19. Canny, J.: A computational approach to edge detection. IEEE Trans. Pattern Anal. Mach. Intell. **8**(6), 679–698 (1986)
20. Kalbasi, M., Nikmehr, H.: Noise-robust, reconfigurable Canny edge detection and its hardware realization. IEEE Access **8**, 39934–39945 (2020)
21. Chiu, S., Liaw, J., Lin, K.: A fast randomized Hough transform for circle/circular arc recognition. Int. J. Pattern Recognit. Artif. Intell. **24**(3), 457–474 (2010)
22. Pan, L., Chu, W., Saragih, J.M., et al.: Fast and robust circular object detection with probabilistic pairwise voting. IEEE Signal Process. Lett. **18**(11), 639–642 (2011)
23. Barbosa, W.O., Vieira, A.W.: On the improvement of multiple circles detection from images using Hough transform. TEMA (São Carlos) Trends Comput. Appl. Math. **20**(2), 331–342 (2019)
24. Csimadia, G., Nagy, Sz.: The effect of the contrast enhancement processes on the structural entropy of colonoscopic images. In: Proceedings of ICEST 2014, Nis, Serbia June 25–27 (2014)

Development of IoT Indoor Monitoring System for Independent Elderly

Ari Aharari[1(✉)], Jair M. Abe[2], and Kazumi Nakamatsu[3]

[1] Department of Computer and Information Sciences, SOJO University, Kumamoto, Japan
info@ahrary.org
[2] Graduate Program in Production Engineering, Paulista University, Sao Paulo, Brazil
[3] University of Hyogo, Kobe, Japan

Abstract. Due to the increased world population of independent elderly those who live alone, there is an immediate need to develop an intelligent monitoring system at home to support them. On the other hand, the burden on family members and long-term care welfare workers increases due to nursing care at welfare facilities and homes. Therefore, it is required to reduce the burden on individuals, such as physical and mental aspects. The present paper aims to develop an IoT monitoring system for the independent elderly, allowing families to feel at ease even in remote locations.

Keywords: Indoor monitoring system · IoT · Elderly

1 Introduction

The aging population has increased globally faster and is expected to exceed 2 billion by 2050 [1]. The older people of the developed areas are lower than in developing areas. The developing areas filled the most growing share of the world's more aging population (Fig. 1). In 1980, 56% of persons aged 60 years or over were home in the developing areas. It raised over two-thirds of the world's older persons living in developing areas in 2017. The number of elderly (60 years or above) is estimated to increase to 1.7 billion in the developing areas in 2050. The more developed areas are projected to see a 38% increase in the number of elderlies, from 310 million elderly (60 years or above) in 2017 to 427 million in 2050. In 2050, 79% of the world's population aged 60 or over will be living in developing regions.

The majority of the young generation in most developing and developed countries work in big cities far away from their hometown. On the other hand, most parents prefer to live in their own house and hometown instead of in the elderly care homes when getting old. They need to handle their basic daily life to achieve this preference. Sometimes emergency cases happen and need to care and support the elderly in a short timeframe. Some examples are, fall accidentally and cannot get up by themselves, undergo some mental disturbance. Between 28 and 35% of elderly people aged over 65 are fall each year, according to the World Health Organization [2]. These rates are increasing to

R. Kountchev et al. (Eds.): *NAMSP 2022*, SIST 332, vol. 332,, pp. 122–132, 2023.
https://doi.org/10.1007/978-981-19-7842-5_11

32–42% for elderly people aged over 70. In fact, by aging and changes of body biological, fall occurrences are increasing exponentially. Falling is particularly dangerous for elderly persons living alone because it may take more time to receive assistance. Hence, monitoring and gathering real-time information can support the young people or some medical support centers respond quickly to urgent cases and make the appropriate decisions to avoid serious situations.

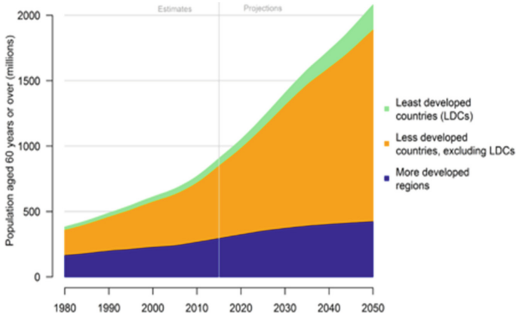


Fig. 1. Number of elderly by development areas [3].

To handle an unexpanded situation, monitoring of elderly persons who are living alone is significant potential in the elderly care research domain.

A most effective way is to monitor the indoor environment using cameras, but most elderly people don't like their daily lives to be under surveillance. In addition, due to the cost of system installation, the camera-based system is limited installed in a few areas in the home. Another solution is a non-intrusive indoor monitoring system using low-cost sensors, such as motion sensors, light sensors, etc. The conventional system prices are differed in a wide range depending on the accuracy and network connection abilities such as Bluetooth, WiFi, 3G/4G/LTE. Due to the high expense of Smart Home Systems, only a few percentages of older persons would be willing to install these high-end sensors systems. Therefore, this paper focuses on developing a low-cost IoT home monitoring system approaching by non-intrusive indoor monitoring system using low-cost sensors.

The rest of this paper is organized as follows. In Sect. 2, we introduce related works. In Sects. 3 and 4, we describe in detail the architecture of the proposed monitoring system. Finally, we conclude our work in Sect. 5.

2 Related Works

The advances of IoT systems have made available efficient, low-cost, low-power miniature devices for remote monitoring applications. Several platforms have been implemented to monitor and support the independent living of elderly adults. For example, the IN LIFE system [4] is a cloud-based platform that provides Ambient Assisted Living (AAL) support to cognitively impaired older people through numerous provided tools and services. The MyLife project [5] supports independence for older people with reduced cognitive function by giving them access to simple and intuitive services that

adapt to their individual needs and wishes. Although both platforms meet several needs posed by cognitive decline, they do not utilize the capabilities provided by the IoT infrastructure. On the other hand, numerous solutions adopt the IoT paradigm for the indoor support of elderly people and mainly focus on human activity recognition.

Recognizing common human activities in real-life settings by extracting knowledge from the data acquired by smart sensors is the goal of human activity recognition [6].

Human activity recognition is a solution that utilizes sensors located in various areas within a smart home, acting as non-intrusive monitoring devices for identifying human behaviors. In particular, a system deploying door sensors, pressure-sensitive mats, float sensors, and temperature sensors is designed by Kasteren et al. [7] for recognizing various living activities in a smart home. A home-based automated system has been implemented in the TAFETA project [8], monitoring the health and well-being of the elderly while remaining unobtrusive by using various types of intelligent sensors in the elderly's home.

A system for recognizing complex activities in a smart home involving contact, motion, tilt, and pressure sensors was proposed by Chen et al. [9]. Zhang et al. [10] also proposed a similar system named "Smarter and Safer Home" that using sensors in homes, acting as non-intrusive monitoring devices for the human behaviour of elderly people.

Other human activity recognition solutions provide activity-monitoring functionalities by using wearable devices for indoor localization. Komai et al. [11] describe an activity monitoring system that utilizes a wearable Bluetooth Low Energy (BLE) beacon device for indoor localization. Popleteev [12] presented an activity tracking and indoor positioning system with a wearable magnet. Similarly, Belmonte-Fernández et al. [13] used a Smart-watch wearable device that acquires the WiFi strength signals of surrounding wireless access points. Finally, [14] Santos et al. have proposed an RFID-based M-health care system using IoT-based connected devices that identifies the position of a health-related item (e.g., elderly) carrying an RFID tag.

Regarding health monitoring, its goal is to allow an individual to closely monitor an elderly person's vital signs, provide feedback for maintaining an optimal health status and create alerts when important measurements are below or over a predefined threshold [15].

Most of the existing commercial home monitoring systems adopt camera devices, e.g., Vivint Smart Home [23], ADT Pulse [22], SimpliSafe home security system [24], Wink Lookout [25], Abode Home Security Starter Kit [26]. Furthermore, they are usually developed as closed systems, and elderly people might become "locked-in," prohibiting opportunities of using products from a different company in the system to suit their needs in terms of affordability and functionality.

Several health-monitoring systems utilize devices placed in the living environment of the elderly and are used by him/her periodically during the day to measure vital signs. Specifically, the system presented by [16] utilizes an electronic blood pressure device that transmits the measured data (i.e., pulse wave) to the data processing center, from where they are accessible in real-time by a doctor. An IoT system that measures saturation peripheral oxygen and pulse rate and transmits them to a cloud-side server was developed by Cao et al. [17].

Another solution is using wearable bio-signal sensors to monitor vital human signs. A new mechanism for uniform biosignals collection from wearables devices and biosignal sensors was introduced by Menychtas et al. [18]. The proposed system is consisting of decision support modules for patient monitoring as well.

A holistic solution for managing Bluetooth biosignal sensors, communication, activity trackers promoting mHealth is proposed in [19]. A monitoring system based on an implantable unit transmitting the measured glucose data to patients' mobile phones is introduced by Ali et al. [20].

A platform for registering patients' vital information and monitoring is demonstrated by Pinto et al. [21]. The system provides mechanisms to trigger alarms in emergencies by using a wristband equipped with numerous sensors.

Tyndall-DMS-Mote [27] is a wireless device that monitors user vital signs indoor/outdoor and needs to carry the device for data collection and a smartphone for local data processing. "Hamon" [28] is a new brand from Mitsufuji manufacturing smart clothing by putting IoT devices into the fabric. The embedded device continuously monitors data, including breath, heart rate, humidity, activities, etc.

Another IoT system based on smartphone and wearable body sensors is proposed by Subasi et al. [29] to monitoring elderly activities. Shende et al. [30] proposed a belt with an embedded fall detector to monitoring the activities and fall detection based on the technology of Global Positioning System (GPS). In most of these solutions, elderly people need to wear devices all the time for monitoring purposes.

A Home Automation System proposed by Ramlee et al. [31] was designed based on environmental sensors and wirelessly connected appliances, e.g., humidity and temperature sensors.

An excellent pilot project has been carried out by IBM equipping hundreds of apartments with IoT sensors (motion detectors, flush-detecting sensors, carbon dioxide, and monoxide sensors) aiming to keep safe elderly people in the home. The collected data was sent to the IBM Watson platform for behavior analysis [32–34].

The present paper aims to develop an IoT monitoring system for the independent elderly, which is not needed to carry or to wear special equipment.

3 IoT Device Architecture

IoT-based applications can be implemented by integrating several technologies such as sensor networks, wireless communications, data processing, and cloud computing. An IoT system is usually represented by combining these technologies, as shown in Fig. 2. We present this architecture as four tiers named Sensing level, Network level, Processing level, Interface level. The sensing level means gathering the environmental and detecting elderly people and sending these data to the processing level for pre-processing. Collecting the data then continues to be transferred to interface level for further analyzing and displaying to the end-users as well as for storing. Data transmissions between levels are implemented by sending level as being illustrated with inter-layers connection lines in Fig. 2. The following sections outline the functions and main elements of each level.

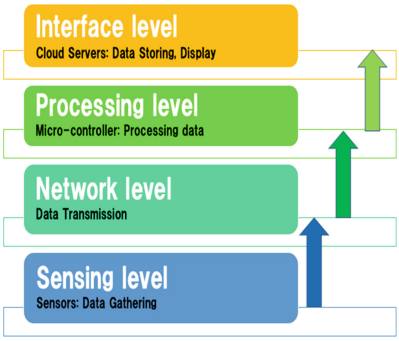


Fig. 2. Number of elderly by development areas.

3.1 Sensing Level

Sensing level involves sensors for gathering the information of elderly people's activities from the indoor environment in our system. More wide range of monitoring devices is capturing vital signs such as body temperature, blood pressure, pulse rate, and respiratory rate [35].

However, other parameters and sensor information are included depending on the application purpose. For example, heart failure patients need monitoring for the following parameters: ECG, Oxygen Saturation (SpO2), heart rate, and weight. In applications that support Ambient Assisted Living for elderly people or the disabled, activity monitoring will be required [36].

3.2 Network Level

Network-level in IoT architecture provides a protocol for things to connect and share data. Also, the network level enables data from existing IT infrastructure to be accessed [37]. The data communication in IoT devices included local and global communication [36]. In elderly people monitoring systems, wireless technology is used for data transmission. Wireless communication is helpful to ensure standardization and compatibility in IoT elderly monitoring systems.

Local network communication between sensing level and processing level is generally implemented by Bluetooth [41–44], or ZigBee [45]. Bluetooth advantages are a low cost, low power consumption technology to transmit data over short distances at 2.4 GHz [48]. ZigBee also offers lower power consumption, but it is not as prevalent as Bluetooth. Some specific communication protocols are also used for elderly monitoring systems, including Radio Frequency Identification (RFID) [46], Near Field Communication (NFC), and Ultra-Wide Bandwidth (UWB) [47]. RFID enables information to be exchanged between two objects—an RFID tag and an RFID reader able to identify, trace and track things within 10 cm to 200 m [39]. NFC works at a high-frequency band at 13.56 MHz, allowing active readers and passive tags or two active readers to communicate with data rate up to 424 kbps and in the range up to 10 cm [40].

3.3 Processing Level

The processing level firstly aggregates data from the sensing level, transfers data to the interface level, and finally processes data. This level consists of processing units and embedded software that apply the computational part of the application [37]. Processing units may be Field Programmable Gate Array (FPGA), System On Chip (SOC), smartphones, microcontrollers, microprocessors, hardware platforms. Hardware platforms such as Arduino, Phidgets, Intel Galileo, Raspberry Pi, Gadgeteer, BeagleBone, Cubieboard, and operating systems such as Contiki, TinyOS, LiteOS, Android, and iOS recently have been developed for running IoT applications [38]. The collected data is processed for further analysis, decision making, generating notifications, and alerts.

3.4 Interface Level

IoT systems connect many physical objects and gather massive data from sensors that need efficient storage. In IoT-based monitoring systems, the collected data from the sensing level is stored for the following analysis. Many cloud services are available for data storage from IoT, such as ThingWorx, OpenIoT, Google Cloud, Amazon, and GENI. Cloud Servers and Physical Servers have three functions at the interface level: storing data, computing, and analyzing data, and displaying data to the end-users. These functions are performed based on cloud computing technology to extract valuable knowledge and trends.

4 Proposed Monitoring System Architecture

The hardware architecture of the proposed monitoring system consists of a specially designed IoT Hub for processing the sensing of elderly people's movement in the indoor environment. The IoT Hub transfers this data to the cloud server, which the User Interface application is later accessed. Figure 3 shows the architecture of the hardware implementation of the proposed monitoring system.

Fig. 3. Hardware architecture of proposed system.

In order to carry out the system without any systematic error, we installed the Watchdog Timer (WDT) in IoT Hub. The function for the WDT is that if the IoT Hub does not become available to receive or transmit the data to the cloud server, the WDT will reset the IoT Hub. The advantage of WDT installation is that future hardware monitoring is not needed, and the IoT Hub can monitor the current condition.

The system configuration of the proposed monitoring system is shown in Fig. 4. The IoT Hub is designed in a compact size that can easily install in an indoor environment.

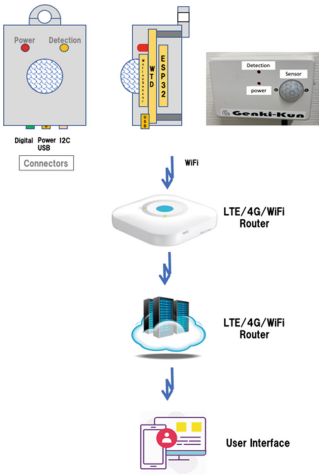


Fig. 4. System configuration of the proposed system.

Additional sensors can be adapted to the IoT Hub through digital or I2C connectors for gathering more specific information.

To monitor the indoor positioning of elderly people, each IoT Hub has installed a device ID linked to the installation location and can easily be set up by the user from the user interface. The proposed system's low cost gives the end-users an advantage for installing several Hubs inside the house to cover monitoring all rooms. In cases if the counter of IN/OUT of some places such as bathroom or shower room is not the same number within specific times, then the alert message will automatically be sent to the initial setup email address.

We deploy IoT Hubs of our monitoring system in a house setting for preliminary validation of the proposed approach, as shown in Fig. 5. Elderly people Activities of Daily Living refers to the daily routine activities mainly in the context of the premises of their home. Therefore, we focused on eating and maintaining continence among five categories of ADLs and installed IoT Hubs in the bathroom and kitchen. We assume that the elderly person uses the bathroom a few times and the kitchen at least two times, concerned they might take their lunch in the day-care house.

An example of the PC user interface of installed IoT Hub is shown in Fig. 6. The counter will reset every day at midnight, and a short record of elderly people's movement will be sent to the end-user email address.

To evaluate the proposed system in long term monitoring, we installed 3 IoT Hubs in 6 elderly people's homes living alone. We started the monitoring on August 2, 2021, and all data are stored in the cloud server.

5 Conclusion

To handle an unexpanded situation, monitoring of elderly persons who are living alone is significant potential in the elderly care research domain.

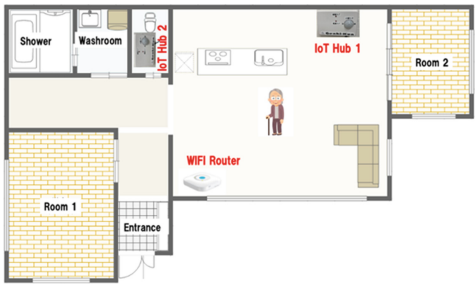


Fig. 5. House setting for preliminary validation of the proposed approach.

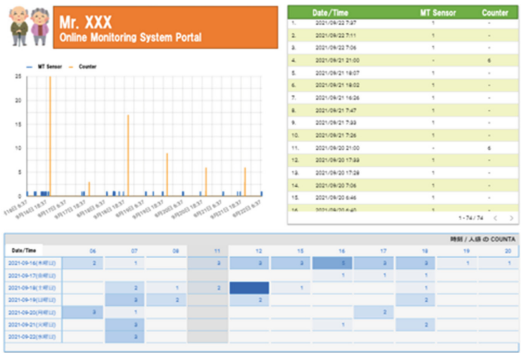


Fig. 6. An example of user interface of the proposed monitoring system.

A most effective way is to monitor the indoor environment using cameras, but most elderly people do not like their daily lives under surveillance. In addition, due to the cost of system installation, the camera-based system is limited installed in a few areas in the home. Another solution is a non-intrusive indoor monitoring system using low-cost sensors, such as motion sensors, light sensors, etc. The conventional system prices are differed in a wide range depending on the accuracy and network connection abilities.

Other human activity recognition solutions provide activity-monitoring functionalities by using wearable devices for indoor localization. Most of these devices need to carry or worn by elderly people, which is less preferred.

The paper proposed an IoT monitoring system for the independent elderly, allowing families to feel at ease even in remote locations. The IoT Hub and connected motion sensors are designed in a compact size and are easy to set up. The proposed system's low cost gives the end-users an advantage for installing several Hubs inside the house to cover monitoring all rooms.

The validation of the proposed system has been evaluated by installing it in 6 elderly houses.

Future works will aim to analyze the data stored in a cloud server and improve the monitoring of the indoor positioning of elderly people.

References

1. World Health Organization: Ageing and health. https://www.who.int/news-room/fact-sheets/detail/ageing-and-health (2018). Accessed 02 July 2022
2. World Health Organization: WHO global report on falls prevention in older age. https://www.who.int/publications/i/item/9789241563536 (2008). Accessed 01 July 2022
3. United Nation: World Population Ageing. https://www.un.org/en/development/desa/population/publications/pdf/ageing/WPA2017_Highlights.pdf (2017). Accessed 01 July 2022
4. Stavrotheodoros, S., Kaklanis, N., Tzovaras, D.: A Personalized Cloud-Based Platform for AAL Support to Cognitively Impaired Elderly People. In: Maglaveras, N., Chouvarda, I., de Carvalho, P. (eds.) Precision Medicine Powered by pHealth and Connected Health. IP, vol. 66, pp. 87–91. Springer, Singapore (2018). https://doi.org/10.1007/978-981-10-7419-6_15
5. MyLife project: http://www.mylife-project.org. Accessed 01 July 2022
6. Kim, E., Helal, S., Cook, D.: Human activity recognition and pattern discovery. IEEE Pervasive Comput. **9**(1), 48–53 (2010)
7. Van Kasteren, T., Englebienne, G., Krose, B.J.: An activity monitoring system for elderly care using generative and discriminative models. Pers. Ubiquit. Comput. **14**(6), 489–498 (2010)
8. Arcelus, A., Jones, M.H. Goubran, R., Knoefel, F.: Integration of smart home technologies in a health monitoring system for the elderly. In: Proceedings of Advanced Information Networking and Applications Workshops 07, pp. 820–825 (2007)
9. Chen, L., Nugent, C.D., Wang, H.: A knowledge-driven approach to activity recognition in smart homes. Knowl. Data Eng., IEEE Trans. **24**(6), 961–974 (2012)
10. Zhang, Q., Su, Y., Yu, P.: Assisting an elderly with early dementia using wireless sensors data in smarter safer home. Serv. Sci. Knowl. Innov. 398–404 (2014)
11. Komai, K., Fujimoto, M., Arakawa, Y., Suwa, H., Kashimoto, Y., Yasumoto, K.: Beacon-based multi-person activity monitoring system for day care center. In: 2016 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops), pp. 1–6 (2016)
12. Popleteev, A.: Activity tracking and indoor positioning with a wearable magnet. In: Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing and Proceedings of the 2015 ACM International Symposium on Wearable Computers, pp. 253–256 (2015)
13. Belmonte-Fernández, Ó., Puertas-Cabedo, A., Torres-Sospedra, J., Montoliu-Colás, R., Trilles-Oliver, S.: An indoor positioning system based on wearables for ambient-assisted living. Sensors **17**(1), 36 (2016)
14. Santos, A., Macedo, J., Costa, A., Nicolau, M.J.: Internet of things and smart objects for M-health monitoring and control. Procedia Technol. **16**, 1351–1360 (2014)
15. Yap, J.H., Jeong, D.U.: Design and implementation of ubiquitous ECG monitoring system by using android tablet. In: Ubiquitous Information Technologies and Applications 214 of Lecture Notes in Electrical Engineering, pp. 269–277. Springer, Berlin, Germany (2013)
16. Wang, J., Sun, S., Zhang, K., Zhang, L., Xing, B., Gao, Z.: Smart blood pressure monitoring system based on internet of things. In: CHI 13 (2013)
17. Cao, G., Liu, J.: An IoT Application: Health Care System with Android Devices. In: Gervasi, O., et al. (eds.) ICCSA 2016. LNCS, vol. 9786, pp. 563–571. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-42085-1_46
18. Menychtas, A., Tsanakas, P., Maglogiannis, I.: Automated integration of wireless biosignal collection devices for patient centered decision-making in point-of-care systems. Healthc. Technol. Lett. **3**(1), 34–40 (2016)
19. Menychtas, A., Padimatos, D., Tsanakas, P., Maglogiannis, I.: On the integration of wear-able sensors in IoT enabled mHealth and quantified-self applications. In: Auer, M., Tsiatsos, T.

(eds.) Interactive Mobile Communication Technologies and Learning. Advances in Intelligent Systems and Computing, p. 725. Springer, Cham (2017)
20. Ali, M., Albasha, L., Al-Nashash, H.: A bluetooth low energy implantable glucose monitoring system. In: Microwave Conference (EuMC), 2011 41st European, pp. 1265–1268. IEEE (2011)
21. Pinto, S., Cabral, J., Gomes, T.: We-care: an IoT-based health care system for elderly people. In: 2017 IEEE International Conference on Industrial Technology (2017)
22. Adt Pulse: https://www.adt.com/pulse. Accessed 01 July 2022
23. Vivint Smarthome: https://www.vivint.com. Accessed 01 July 2022
24. Simplisafe: https://simplisafe.com. Accessed 01 July 2022
25. Wink look out. https://www.wink.com/products/wink-lookout-smart-security-essentials/. Accessed 01 July 2022
26. Abode home security starter kit (2018) review: This entry-level smart home hub continues to evolve. https://www.techhive.com/article/583194/abode-review.html. Accessed 01 July 2022
27. Donoghue, J., Herbert, J., Stack, P.: Remote non-intrusive patient monitoring. In 4th International Conference on Smart Homes and Health Telematics (2006)
28. Hamon: http://www.hamon.tech/en/#products. Accessed 01 July 2022
29. Subasi, A., Radhwan, M., Kurdi, R., Khateeb, K.: IoT based mobile healthcare system for human activity recognition. In IEEE 15th Learning and Technology Conference (2018)
30. Shende, D., Madrewar, S., Dugade, S.: Dementia patient activity monitoring and fall detection using IoT for elderly. Int. J. Trend Sci. Res. Dev. **3**, 363–367 (2019)
31. Ramlee, R., Othman, M.A., Leong, M., Ismail, M., Ranjit, S.: Smart home system using android application. In: International Conference of Information and Communication Technology (2013)
32. IBM Watson IoT Platform: https://www.ibm.com/hk-en/marketplace/internet-of-things-cloud. Accessed 01 July 2022
33. Elder Care Organization: https://www.ibm.com/case-studies/z976639f72075f89. Accessed 01 July 2022
34. AI technology brings innovation to elderly care. https://www.ibm.com/blogs/client-voices/ai-technology-innovation-elderly-care/. Accessed 01 July 2022
35. Ahmed, M.U., Banaee, H., Loutfi, A., Rafael-Palou, X.: Intelligent healthcare services to support health monitoring of elderly. In: Internet of Things, User-Centric IoT, pp. 178–186. Springer International Publishing (2015)
36. Ahmed, M.U., Björkman, M., Čaušević, A., Fotouhi, H., Lindén, M.: An Overview on the Internet of Things for Health Monitoring Systems. In: Mandler, B., et al. (eds.) IoT360 2015. LNICSSITE, vol. 169, pp. 429–436. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-47063-4_44
37. Li, S., Xu, L.D., Zhao, S.: The internet of things: a survey. Inf. Syst. Front. **17**(2), 243–259 (2014). https://doi.org/10.1007/s10796-014-9492-7
38. Al-Fuqaha, A., Guizani, M., Mohammadi, M., Aledhari, M., Ayyash, M.: Internet of things: a survey on enabling technologies, protocols, and applications. IEEE Commun. Surv. Tutor. **17**, 2347–2376 (2015)
39. Want, R.: An introduction to RFID technology. IEEE Pervasive Comput. **5**, 25–33 (2006)
40. Want, R.: Near field communication. IEEE Pervasive Comput. **10**, 4–7 (2011)
41. Suh, M., et al.: WANDA B.: weight and activity with blood pressure monitoring system for heart failure patients. In: 2010 IEEE International Symposium on a World of Wireless Mobile and Multimedia Networks (WoWMoM), pp. 1–6 (2010)
42. Suh, M., Chen, C., Woodbridge, J., Tu, M.K., Kim, J.I., Nahapetian, A., et al.: A remote patient monitoring system for congestive heart failure. J. Med. Syst. **35**, 1165–1179 (2011)

43. Lan, M., Samy, L., Alshurafa, N., Ghasemzadeh, H., Macabasco-OConnell, A., Sarrafzadeh, M.: WANDA: An end-to-end remote health monitoring and analytics system for heart failure patients. In: Conference on Wireless Health, Wireless Health12 (2012)
44. Bisio, I., Lavagetto, F., Marchese, M., Sciarrone, A.: A smartphone-centric platform for remote health monitoring of heart failure. Int. J. Commun. Syst. **28**, 1753–1771 (2015)
45. Zanjal, S.V., Talmale, G.R.: Medicine reminder and monitoring system for secure health using IOT. In: International Conference on Information Security and Privacy, pp. 471–476 (2015)
46. Parida, M., Yang, H.C., Jheng, S.W., Kuo, C.J.: Application of FRID technology for in-house drug management system. In: 15th International Conference on Network-Based Information Systems (NBiS), pp. 577–581 (2012)
47. Fortino, G., Di Fatta, G., Pathan, M., Vasilakos, A.V.: Cloud-assisted body area networks: state-of-the-art and future challenges. Wireless Netw. **20**(7), 1925–1938 (2014). https://doi.org/10.1007/s11276-014-0714-1
48. Fanucci, L., et al.: Sensing devices and sensor signal processing for remote monitoring of vital signs in CHF patients. IEEE Trans. Instrum. Meas. **62**, 553–569 (2013)

Improving the Process of Evaluating User Stories Using the Paraconsistent Annotated Evidential Logic Eτ

Samira Sestari do Nascimento[1(✉)], Jair Minoro Abe[1], Luiz Roberto Forçan[1], Cristina Corrêa de Oliveira[2], Kazumi Nakamatsu[3], and Aharari Ari[4]

[1] Paulista University, São Paulo, Brazil
samira.nascimento3@aluno.unip.br, jair.abe@docente.unip.br
[2] Federal Institute of São Paulo, Bragança Paulista, Brazil
crisolive@ifsp.edu.br
[3] University of Hyogo, Kobe, Japan
akamatu@pf7.so-net.ne.jp
[4] Sojo University, Kumamoto, Japan
aharari@cis.sojo-u.ac.jp

Abstract. Software developers need to be agile to meet users' needs, delivering software on tight, quality deadlines. User history is a technique used in agile methods to elicit requirements. However, this process is performed with the developers and the user, and there may be contradictions between them, resulting in inaccurate metrics. This article presents a model of validation of user history using the Para-analyzer algorithm, based on the Paraconsistent Annotated Evidential Logic Eτ to assist in improving the evaluation, prioritization, and estimation process of user stories. A survey was conducted with a team of developers working with agile methods. The model uses the degrees of favourable and contrary evidence for each INVEST criterion as input variables. The application of this model allows considering extremely relevant issues when it comes to supporting decision-making based on a mathematical model and serving as a support tool for teams, Product Owners, Project Managers, and others. Four user stories were analyzed by nine experts, who evaluated the criteria for each user story. The interpretation of the evaluations performed by the experts was through the global analysis in the unit square of the Cartesian plane, which indicated the degrees of favourable evidence and contrary evidence for the data used. Two stories that could not be developed in a Sprint were verified and, therefore, should be refactored and resubmitted to the opinion of experts. The other two stories had favourable evidence to be used in a Sprint.

Keywords: User stories · Paraconsistent annotated evidential logic Eτ · INVEST

R. Kountchev et al. (Eds.): *NAMSP 2022*, SIST 332, vol. 332,, pp. 133–142, 2023.
https://doi.org/10.1007/978-981-19-7842-5_12

1 Introduction

A requirement is defined as a role, service, or resource to meet a product user's need or demand, as defined by Requirements Engineering, a software engineering subarea. It can be divided into functions, constraints, or business rules regardless of the methodology adopted [1].

In an agile development, user requirements are treated as User Stories (US), which are written requests from the user from the point of view of needs and use [2].

Tasks are prioritized from an available list of software called Product Backlog. The development team (TIME) commits to advancing the cycle in software development, lasting four weeks to run [3].

Due to the constant involvement of the client in the validation and acceptance tests of USs, it becomes a frequent need [4], requiring business knowledge and the domain of the application where the system will be used [4]. Although it does not have a specific input for the validation process, it is essential to narrow the organizational strategy to the corporate culture (Schwaber and Sutherland, 2020). In the validation step, the developer evaluates whether the USs understand, are absent, or replay information to start a sprint, with the certainty that the issues will be resolved by the product owner (PO) or by the customer himself. The project's success depends on the sprint plan, which verifies the quality of the Time and PO, and USs to ensure requirements are correct, documented, and validated [5].

However, the complexity of USs and time determination are not clearly defined and described, and it is considered risky in the project development plan to understand [6, 7]. Another aspect is the collective consensus. Because the analysis of each US is individual, it requires knowledge of the application and the domain of the business and can lead to inconsistencies in a study among the planning participants [4].

The validation process needs to generate a list of issues and actions agreed upon by the client, PO, and TIME. The main problem with this process is that it does not meet the INVEST criteria. It is an acronym for "Independent, Tradeable, Valuable, Estimable, Small or Small, and Testable Stories" [8]. From this list of problems, it needs to create an action plan to define the overall work plan to be executed and agreed upon by all those involved. The various paths of understanding to arrive at results can upset the plan. A Paraconsistent Annotated Evidential Logic Eτ (logic Eτ) evaluates uncertain. Inconsistent data that stakeholders in certain USs do not understand may express a logical contradiction [9], considering the criteria to be logic Eτ assists in decision-making [10].

In this article, the questionnaire results were applied to software development experts in different functions within the company, evaluating USs and measuring evidence through the logic Eτ in conjunction with the INVEST criteria. The model of this study assists in the decision-making process by permitting technical validation and through Para-analyzer algorithms that will be represented in the final analysis of USs.

2 Reasons

2.1 User Stories

User stories (US), in agile development, are a simple way to disbelieve the needs of the product owner because they capture the essential elements using the most widespread format. “As a <role>, I want <goal/desire> so that <benefit>” [11].

They express functional requirements using a stakeholder business language because they represent an agreement between the PO and the developers, forming the basis for development. After all, TIME must understand, estimate, and implement all US [12].

2.2 Invest

The INVEST criterion is an acronym for <I> ndependent, <N> egotiable, <V> aluable, <E> stimable, <S> mall, <T> estable, which should be applied for each US [8], representing an agreement between the PO and TIME to decide whether the US will be assigned in a Sprint.

The main problems encountered in this agreement are communication failures between TIME and the PO [13, 14], and understanding of INVEST [15] from incomplete requirements; formulated or intestable; provided late (scope increase), or not detailed enough to express the smallest significant unit of activity for the user.

Subjectivity in US estimates, which are usually experimentally derived and managed, often without reference to the historical data of the para comparison organization; lack of experience in estimating; or even lack of knowledge of the domain [16], may result in a lack of understanding of the criteria, making it difficult to estimate, as there is no clarity of the outputs that the system should produce and the inputs are not objectively measured [17].

Time estimates for the completion of each task, defined as story points, and for the project total [18] are carried out with the definition of all investment criteria for each US. Although the estimate is important from a commercial point of view, the film is about minimizing errors or failures because the scope of the software project is estimated along with the time and cost [19].

2.3 Paraconsistent Annotated Evidential Logic Eτ

The Paraconsistent Annotated Evidential Logic Eτ belongs to the paraconsistent logic class, not classical, considering the principle of contradiction by obtaining contrary (P) propositions, which are associated with an atomic request, the type of Degree of Favorable Evidence (μ), and the type of Unfavorable Degree of Evidence (λ). The pair (μ, λ), called the annotation constant, where μ and λ [0, 1] e (P) denote a proposition in the usual sense. The annotation μ indicates the degree of favourable evidence, and the λ annotation represents the unfavourable evidence expressed by request (P) τ ∈. Logical states are called extremes and are characterized by internal states, as illustrated in Fig. 1. Non-extreme states are named according to their proximity to extreme logical conditions, as shown in Table 1.

The Para-analyzer algorithm consists of information collected through a research form for decision-making analysis [20]. In data processing, the connectives express favourable and unfavourable opinions about the propositions from the experts participating in the decision-making process. Operators (OR) and (AND) correspond to disjunction and conjunction in classical logic. For example, in Logic Eτ, maximization is considered to be $(\mu 1, \lambda 1)$ OR $(\mu 2, \lambda 2) = (\text{Max}\ \{\mu 1, \mu 2\}; \text{Min}\ \{\lambda 1, \lambda 2\})$ where Max indicates the maximization of real numbers with the standard order and Min indicates minimizing the actual numbers with a legal order. On the other hand, operation AND is described as $(\mu 1, \lambda 1)$ e $(\mu 2, \lambda 2) = (\text{Min}\ \{\mu 1, \mu 2\}; \text{Max}\ \{\lambda 1, \lambda 2\})$ where Min indicates the working of minimizing real numbers with standard order and Max indicates an operation maximizing real numbers with standard order.

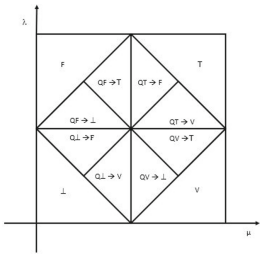


Fig. 1. Representation of the Lattice τ

Table 1. Symbolization of logical states

Extreme states	Symbol	Non-extreme states	Symbol
True	V	Quasi-true tending to inconsistent	$Q_V \rightarrow T$
False	F	Quasi-true tending to paracomplete	$Q_V \rightarrow \perp$
Inconsistent	T	Quasi-false tending to inconsistent	$Q_F \rightarrow T$
Paracomplete	$\perp$	Quasi-false tending to paracomplete	$Q_F \rightarrow \perp$
		Quasi-inconsistent tending to true	$Q_T \rightarrow V$
		Quasi-inconsistent tending to false	$Q_T \rightarrow F$
		Quasi-paracomplete tending to true	$Q_\perp \rightarrow V$
		Quasi-paracomplete tending to false	$Q_\perp \rightarrow F$

3 Case Study

For simplicity, this study is based on a backlog with only four US; selected requirements gathering is carried out in companies. Therefore, they were removed from previous projects, and the respondents were unaware of them (Table 2).

Table 2. User stories

US	User story	Acceptance criteria
A	As a customer, I want to know the price and availability of a product to verify that I want to complete the purchase	The product must have the name and price. The product must be associated with a category and a supplier
B	As an administrator, I want to query product categories to view product categories	Do not list categories that are discontinued
C	As an administrator, I would want to search for products by category; I know the category code and would like to locate it in the system. to learn about your items and to see the products in a well-informed category	Do not list products that are discontinued. You must list all products even if they are out of stock
D	As a customer, I want the system to provide several forms of payment so that I can pay for my order and the system close the order	Customers can select a payment method to be able to pay. When the payment is made with a credit card, validate the operation or restriction with the operator. If everything is ok, proceed with the finalization of the order

Nine participants were selected who are professionals in software development, referred to as "specialists," who work specifically with the Scrum methodology. These professionals from different software companies in the city of São Paulo were divided into three groups, each adopting the position as a grouping criterion. They are called E1, E2,…, and E9 in the database.

The specialists are represented by the letters E1, E2, and E3. The positions are for specialist E1 (Software Architect and Scrum Master), specialist E2 (Senior System Analyst), and specialist E3 (Full System Analyst).

For the INVEST criteria, they were mapped to F1 (independent) factors; F2 (negotiable); F3 (valuable); F4 (estimable); F5 (small); and F6 (testable). Each expert expressed their opinions on each INVEST criteria, pointing to μ and λ criteria for each US. The values of μ and λ were normalized, according to Table 3.

Table 3. Normalization of the values of μ and λ

Degree	Percent (%)	Description
1.00	100	There is no doubt about the evidence
0.75	75	Small doubt regarding evidence
0.50	50	Average doubt in relation to evidence
0.25	25	Low certainty in relation to evidence
0.00	0	Almost no certainty regarding evidence

Given the responses of experts, who attributed the values of μ and λ for each US, according to Table 4, a database was developed with the factors analyzed.

4 Results

After applying the Para-analyzer algorithm, each track receives a diagnosis, which can be: feasible (which is understood as approved); unfeasible (the item is in disagreement with the evaluation factor); Non-Conclusive (the thing requires further evaluation). For each user story, the following results were obtained.

User Story A

The result of the Global Analysis (GA) of the Para-analyzer algorithm (0.58; 0.42) indicates that it is in the "Quasi-True Tending to Inconsistent" state, resulting in no conclusion. The paraconsistent qualitative evaluation demonstrates that US's excellent quality presentation is no longer an absolute truth. Therefore, it is assumed that its quality is insufficient, requiring that new information be obtained and submitted again for analysis by the specialists, as shown in Fig. 2.

User Story B

The result of the GA of the Para-analyzer algorithm (0.96; 0.17) indicates that it is in the "Totally True" state, resulting in viability. Furthermore, the paraconsistent qualitative evaluation demonstrates that US presents an adequate quality. Therefore, it is understood that this US is within a possible quality standard and may be part of a Sprint, as seen in Fig. 3.

User Story C

The result of the GA of the Para-analyzer algorithm (0.96; 0.21) indicates that it is in the "Totally True" state, resulting in a viable. The paraconsistent qualitative evaluation demonstrates that US is feasible. It is understood that this US is within a quality standard and can be part of a sprint to conform to Fig. 4.

User Story D

The result of the GA of the Para-analyzer algorithm (0.59; 0.59) indicates that it is in the "Inconsistent" state, tending to falsity present inconclusive results. The paraconsistent qualitative evaluation demonstrates that it is no longer an absolute truth to conclude that this US has a viable quality. It is understood that this US depends on others and is very large. Therefore, it is assumed that the quality of this US is insufficient, requiring that new information be obtained and submitted for analysis by the specialists (Fig. 5).

5 Conclusion

Current US validation processes do not take into account inconsistency or contradiction. In an actual situation, contradictions appear due to the conditions of the environment in which the requirements were surveyed. These situations of contradiction occur regardless of the will of the software development team members or the company's business area

Table 4. Database formed by assigned by experts analysis

	Group A						Group B						Group C					
	E1		E2		E3		E4		E5		E6		E7		E8		E9	
Factor	μ	λ	μ	λ	μ	λ	μ	λ	μ	λ	μ	λ	μ	λ	μ	λ	μ	λ
F1	0.00	1.00	0.25	1.00	0.50	1.00	0.25	1.00	0.25	0.75	0.25	0.75	0.25	0.75	0.25	1.00	0.25	1.00
F2	1.00	0.00	1.00	0.25	1.00	0.25	0.75	0.25	0.75	0.25	0.75	0.25	1.00	0.25	1.00	0.25	1.00	0.25
F3	1.00	0.00	1.00	0.25	1.00	0.25	1.00	0.25	0.75	0.25	1.00	0.25	0.75	0.25	1.00	0.25	0.75	0.25
F4	0.00	1.00	0.25	0.75	0.25	1.00	0.25	0.75	0.00	1.00	0.25	0.75	0.25	1.00	0.25	0.75	0.25	1.00
F5	0.00	1.00	0.00	1.00	0.25	1.00	0.25	0.75	0.25	0.75	0.25	0.75	0.25	1.00	0.25	0.75	0.25	0.75
F6	0.00	1.00	1.00	0.50	0.75	0.25	1.00	0.25	1.00	0.50	1.00	0.25	1.00	0.50	1.00	0.25	0.75	0.25

Factor	μ and λ	Decision
F1	(0,25; 0,75)	not conclusive
F4	(0,25; 0,75)	not conclusive
F5	(0,25; 0,75)	not conclusive
F2	(0,75; 0)	feasible
F3	(1,00; 0)	feasible
F6	(1,00; 0,25)	feasible

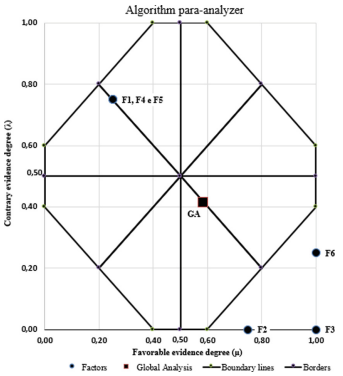


Fig. 2. Analysis of user a story A

Factor	μ and λ	Decision
F1	(1,00; 0,25)	feasible
F3	(1,00; 0,25)	feasible
F4	(1,00; 0)	feasible
F5	(1,00; 0)	feasible
F6	(1,00; 0,25)	feasible
F2	(0,75; 0,25)	not conclusive

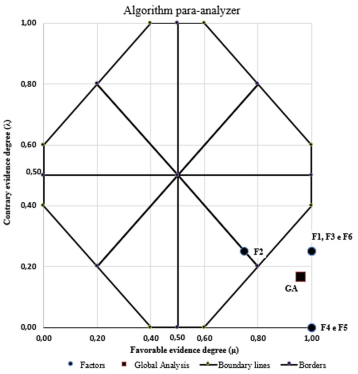


Fig. 3. Analysis of user a story B

Factor	μ and λ	Decision
F1	(1,00; 0,25)	feasible
F3	(1,00; 0,25)	feasible
F4	(1,00; 0,25)	feasible
F5	(1,00; 0,00)	feasible
F6	(1,00; 0,25)	feasible
F2	(0,75; 0,25)	not conclusive

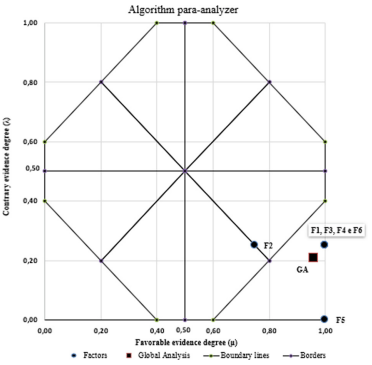


Fig. 4. Analysis of user a story C

team. Therefore, the conflicts are part of the US validation. The bigger involvement of stakeholders, customers, Product Owners, and the development team to resolve the

Factor	μ and λ	Decision
F1	(0,25; 1,00)	infeasible
F5	(0,25; 1,00)	infeasible
F3	(1,00; 0,25)	feasible
F6	(1,00; 0,25)	feasible
F2	(0,75; 0,25)	not conclusive
F4	(0,25; 0,75)	not conclusive

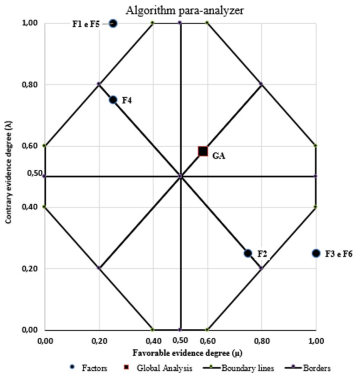


Fig. 5. Analysis of user a story D

client's interests needs, and desires, bigger the levels of conflicts, contradictions, and inconsistencies.

The research presents a model to assist in the decision-making process of evaluation of USs using the Logic Eτ, through the Para-analyzer algorithm, using logical criteria that enable technical validation. The input parameters are established by the experts' opinions, consolidating a collective logic of TIME based on mathematical terms.

The analysis based on these perspectives satisfies experts and stakeholders because the paraconsistent model maximizes all expert opinions and creates a mathematical consensus on these opinions.

This research demonstrates that the concepts of Evidence Noted Paraconsistent Logic Eτ could be used to validate and present perspectives on dealing with situations of uncertainty and inconsistency. These situations or factors are relevant for decision-making and influence the quality and success of an information system.

References

1. Pressman, R.S.: Software engineering: a practitioner's approach, 7 edn. McGraw-Hill Publishing Co. (2011)
2. Chopade M.R.M., Dhavase N.S.: Agile software development: positive and negative user stories. In: 2nd International Conference for Convergence in Technology (I2CT) (2017)
3. Barbosa, R., Silva A.E.A., Moraes, R.: Use of similarity measure to suggest the existence of duplicate user stories in the srum process. In: 46th Annual IEEE/IFIP International Conference on Dependable Systems and Networks Workshop (DSN-W) (2016)
4. Heikkilä, V.T., Damian, D., Lassenius, C., Paasivaara, M.: A mapping study on requirements engineering in agile software development. In: 41st Euromicro Conference on Software Engineering and Advanced Applications (2015)
5. Mateen, A., Abbas K., Akbar, M.A.: Robust approaches, techniques and tools for requirement engineering in agile development. In: IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI) (2017)
6. Hannay, J.E., Benestad H.C., Strand, K.: Agile uncertainty assessment for benefit points and story points. IEEE Softw. **36**, 50–62 (2019)

7. Popli R., Chauahn, N.: Managing uncertainty of story-points in agile software. In: 2nd International Conference on Computing for Sustainable Global Development (INDIACom) (2015)
8. Buglione L., Abran, A.: Improving the user story agile technique using the INVEST criteria. In: 2013 Joint Conference of the 23rd International Workshop on Software Measurement and the 8th International Conference on Software Process and Product Measurement (2013)
9. Carvalho F.R., Abe, J.M.: A Paraconsistent Decision-Making Method, vol. 87. Springer International Publishing (2018)
10. Abe, J.M.: Paraconsistent Intelligent-Based Systems—New Trends in the Applications of Paraconsistency, vol. 94, p. 306. Springer, Switzerland (2015)
11. Sutherland, J.: Scrum: The Art of Doing Twice the Work in Half the Time, 2nd edn. Crown Business Publications, São Paulo (2015)
12. Dalpiaz F., Brinkkemper, S.: Agile requirements engineering with user stories. In: 2018 IEEE 26th International Requirements Engineering Conference (RE) (2018)
13. Hess, A., Diebold P., Seyff, N.: Towards requirements communication and documentation guidelines for agile teams. In: 25th IEEE International Requirements Engineering Conference Workshops (REW) (2017)
14. Liskin, O., Pham, R., Kiesling S., Schneider, K.: Why We Need a Granularity Concept for User Stories (2014)
15. Opazo, J.G., Lopez, C.: Improving user stories: a case study in the chilean banking industry. In: 2018 XLIV Latin American Computer Conference (CLEI) (2018)
16. Kamthan P., Shahmir, N.: Modeling negative user stories is risky business. In: 17th IEEE International Symposium on High Assurance Systems Engineering (HASE) (2016)
17. Bik, N., Lucassen G., Brinkkemper, S.: A reference method for user story requirements in agile systems development. In: 25th IEEE International Requirements Engineering Conference Workshops (REW) (2017)
18. Castillo-Barrera, F.E., Amador-Garcia, M., Perez-Gonzalez H., Martinez-Perez, F.E.: Agile evaluation of the complexity of user stories using the bloom's taxonomy. In: 2017 International Conference on Computational Science and Computational Intelligence (CSCI) (2017)
19. Hayat, F., Rehman, A.U., Arif, K.S., Wahab K., Abbas, M.: The influence of agile methodology (Scrum) on software project management. In: 20th IEEE/ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing (SNPD) (2019)
20. Akama, S. (ed.): Towards Paraconsistent Engineering. ISRL, vol. 110. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-40418-9
21. Sheemar H., Kour, G.: Enhancing user-stories prioritization process in agile environment. In: International Conference on Innovations in Control, Communication and Information Systems (ICICCI) (2017)

Neural Network Algorithm Applied in Electrical Engineering Automation

Yansong Shi, Shiyou Xing, Peijian Wang(✉), and Dai Jiawei

Northeast Electric Power University, Jilin, China
2297915933@qq.com

Abstract. The development of electrical engineering automation as an important symbol of China's modern industrial innovation, industrial labor costs have been effectively controlled, the overall electrical system operation is more stable, the personal safety of internal staff has been guaranteed. According to the analysis of the current economic development trend in the industrial field, electrical engineering automation contains a number of technical concepts, among which the neural network algorithm, as a representative artificial intelligence application algorithm, has been widely used in practical technology exploration and application research. Therefore, on the basis of understanding the current social development trend and the advantages of artificial intelligence technology, according to the application direction of neural network algorithm in electrical engineering automation, taking the fault diagnosis of automatic electrical equipment as an example, the application of neural network algorithm is empirically studied. The final results show that the simulation can not only identify the state feature quantity, but also improve the accuracy of state recognition.

Keywords: Neural Network Algorithm · Electrical Engineering · Automation · Failure to Identify

1 Introduction

Under the background of information age, according to the development of computer technology of artificial intelligence technology is widely used in many fields, it is not only the attention of the society from all walks of life, also in the development of intelligent electrical equipment, to effectively control the cost and the excessive consumption of resources, thus based on the application of artificial intelligence in electrical engineering automation research has certain practical significance. Artificial intelligence technology is mainly used to simulate the extension theory and method about human intelligence. It belongs to the concept of the new era of development of science and technology of the most representative and will be integrated into the electrical engineering automation. On the one hand this can solve the traditional problem such as electrical engineering automation system run unstable, the integrated automation system will become more perfect and convenient in operation; the constructed work environment will be simpler. On the other hand, it can replace mass human efforts, make data collection and analysis

R. Kountchev et al. (Eds.): *NAMSP 2022*, SIST 332, vol. 332,, pp. 143–150, 2023.
https://doi.org/10.1007/978-981-19-7842-5_13

more accurate and effective, and reduce unnecessary intermediate links. In the 1870s, the social and economic development broke out in the second technological revolution marked by the widespread use of electricity, from which human development opened the electric age. Electrical engineering has also been developed in an all-round way under this background. In the initial stage of electrical engineering research, researchers will follow the technical concepts to expand the basic concepts of electrical engineering, such as the knowledge that photons and electronic files belong to the scope of electrical engineering. Under the development of information technology innovation, electrical engineering has gradually entered the stage of automation development. As an important symbol of technological innovation, it has also further accelerated the overall level of China's modern industrial development. From the practical point of view, electrical engineering automation technology as a guide to the steady development of the power system of the basic conditions, involves mechatronics, information network control, computer, power electronics and other technical content [1, 2]. As an application algorithm of reasoning with logic as the core, neural network has been widely used in automatic design system. Especially for the development of electrical engineering automation, the rational use of neural network algorithm can expand the research space of technical system and optimize the automation management level of power engineering. Therefore, in this work, based on the clear neural network algorithm model frame diagram, according to its application direction in electrical engineering automation, the automatic electrical equipment fault diagnosis method based on neural network is studied in practice [3–5].

2 Method

2.1 Neural Network Algorithm

Generally speaking, this application algorithm consists of three layers: the first is the input layer, the second is the hidden layer, and the last is the output layer, where the activation function of the input layer uses the radial basis function, and the output layer chooses the activation function linear function. According to the model diagram analysis of the network algorithm shown in Fig. 1, the input layer contains N neurons, and the input signal vector is $X = [x_1, x_2, ..., x_n]^T$. The hidden layer contains N neuron nodes, and the corresponding output vector is $G = \left[g_1, g_2, ..., g_n\right]^T$. The output layer contains the vector $Y = \left[y_1, y_2, ..., y_n\right]^T$. In the neural network algorithm, the expected output vector is $D = [d_1, d_2, ..., d_n]^T$, the weight matrix between the hidden layer and the output layer should be represented by the symbol Wjk.

Among them, the radial basis function should be studied by Gaussian function, as shown below:

$$G(x) = e^{-\frac{\|x-c\|^2}{\delta^2}} \tag{1}$$

In the above formula, c and δ represent the parameters of the data center width of the radial basis function.

The expression formula of the selected linear function relationship is:

$$y_k = \sum_{j=1}^{q} w_{jk} v_j + b_k, k = 1, 2, ..., n \tag{2}$$

In the above formula b,k represents the threshold vector of neurons.

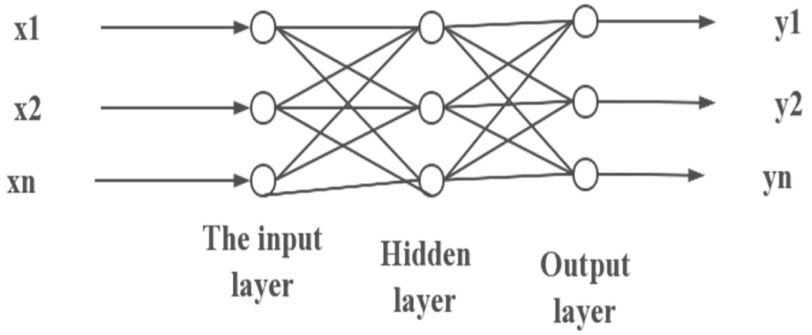


Fig. 1. Model diagram of neural network algorithm.

2.2 Electrical Engineering Automation Based on Neural Network Algorithm

First, fault diagnosis. Algorithm using neural network fault diagnosis of bearing, various parameters have both inside and outside of the circle, rolling state and normal state, the related data signal transmission to the convolutional neural network model to complete data classification, is the current scientific research scholars, combining traditional and convolution neural network algorithm is intelligent diagnosis system constructed a new diagnosis model. The final practice results show that this application method improves the accuracy of bearing diagnosis, and increases 3.7% compared with the traditional support algorithm. At the same time, in order to obtain more accurate induction motor rotor fault, some scholars have developed an optimized neural network fault diagnosis algorithm. The practical application algorithm first uses stator current to obtain fault features, and then identifies related faults according toChaotic particle swa m optimization-Back Propagation neural network algorithm. The final results show that the neural network algorithm based on stator current is more effective than the traditional neural network algorithm, and the accuracy of actual fault identification can be comprehensively improved. In order to deal with the problem of fault diagnosis in multi-battery inverter scientifically, researchers proposed to use deep neural network algorithm to complete fault diagnosis according to relevant literature. The core of the application algorithm is to get the inverter fault data, and then use stack encoding deep neural network algorithm to extract and diagnose the fault, and use MATLAB software to complete the simulation analysis, the final results prove that the neural network algorithm built, effectively improve the effectiveness and accuracy of fault diagnosis. The system understands the most common problems of BP neural network algorithm in automatic fault diagnosis of electrical engineering, which refers to the local optimal problem, which will not only affect the efficiency of diagnosis, but also limit the operation of the construction model. Therefore, researchers put forward a diagnosis model combining adaptive differential evolution algorithm and BP neural network algorithm in practical exploration, and the final empirical results prove that the improved applied algorithm

is more effective than the traditional BP neural network algorithm, and the actual fault diagnosis accuracy is stronger [4–6].

Second, product design. When designing and manufacturing electrical products, neural network algorithm can be used to grasp more precise parameters on the basis of constructing a perfect model. Research scholars, for example, based on permanent magnet synchronous motor nonlinear and multiple parameter change adaptive neural network algorithm is proposed, and its operation in practice using RBF neural network algorithm to comprehensively control the various parameters of the controller, such not only can improve robustness and accuracy of the system operation, can also be combined with the MATLAB software to complete the simulation analysis. Thus the control level of the controller is improved. At the same time, in the treatment of the product design problem of low accuracy, scientific research and academic performance is put forward to in the design model of time-varying data analysis as the core method, this technology will use since sparse coding of unsupervised learning neural network algorithm for processing, thus in promoting electrical product design accuracy at the same time, guarantee the overall design has stability.

Third, control system. The error of synchronous drive system of double motors can be dealt with by using neural network algorithm. In practical exploration, scientists put forward the application algorithm with single neuron PID control as the core. The empirical results show that the designed algorithm is more effective than the traditional control method, which can not only improve the precision of motor synchronous control, but also optimize the anti-interference of the system [7].

2.3 Fault Diagnosis of Automatic Electrical Equipment Based on Neural Network

At present, neural network algorithm as the main form of artificial intelligence network model, BP network refers to a global approximation neural network. In this paper, a typical local approximation neural network is used. According to the topology analysis of RBF neural network as shown in Fig. 2, although there are similarities between RBF neural network and BP network, there are also certain differences.

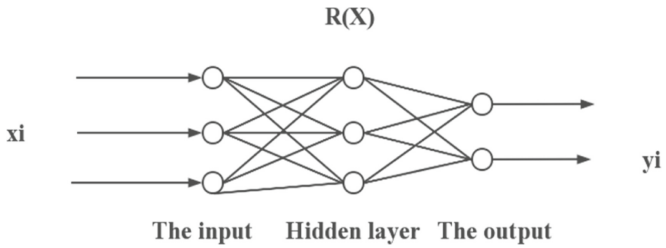


Fig. 2. Topology of RBF neural network.

State recognition algorithm for RBF network to conduct a comprehensive improvement, combined with the concept of credibility clear credible degree recognition results, and on the basis of appropriate decision rules, identification of electrical engineering equipment of the new state, and then see it as a new network model of training sample

input, adjusted already constitute a network weight value and the quality of the jade. After the adjustment, the network can accurately identify the new state categories, so as to ensure the improved state recognition algorithm and have the application function of identifying the new state types. The flow chart of the improved algorithm is as follows (Fig. 3).

3 Result Analysis

Combined with the neural network algorithm model constructed in this paper, it analyzes which state feature quantities can be regarded as input vectors in the state recognition of hV circuit breaker, and thus obtains the common equipment fault codes and causes as shown in Table 1.

This paper constructs the learning samples shown in Table 2 from the statistics and empirical analysis of experts in related fields. According to the simulation results of various faults of the operating mechanism, the influence of different types of faults on electrical machinery is quite different, mainly on the spindle, and the actual Angle curve has obvious changes. In order to quantitatively analyze the difference of spindle Angle curve of various fault types and facilitate the state identification of circuit breaker, the spindle Angle curve can be parameterized.

According to the input vectors described in the table above, the input vectors and state codes in the table are regarded as training samples after initialization. The first 5 input vectors are mainly used to verify the network identification results, and the sixth one is the input vector waiting for identification. Set the threshold of credibility as 0.6, and perform operations according to the algorithm flow chart obtained from the above research, and finally obtain the recognition results as shown in the above table. The study of the table data shows that the input vector as the training sample has no recognition error, and the actual reliability is very high. For the sixth input vector waiting for recognition, the network first proposed the recognition result belonging to the second state type, but the actual credibility was lower than the pre-given credibility threshold, so the recognition result proposed a new state type. After completing the above recognition operations, the application algorithm will adjust the weight value and threshold value, and obtain a new state type recognition function. Sample No. 6 in the above table is regarded as input for subsequent network identification, and the results can be obtained as shown in Table 3.

Combined with the above table analysis results show that the improved algorithm has higher sample recognition accuracy [8, 9].

4 Conclusion

To sum up, this work studies on the basis of the neural network algorithm direction, from the mechanical equipment fault diagnosis of electrical engineering automation, to identify the state of the high voltage circuit breaker based on RBF neural network algorithm application field for empirical research, and the RBF network status recognition algorithm in reference to the related concepts of credibility. The improved application

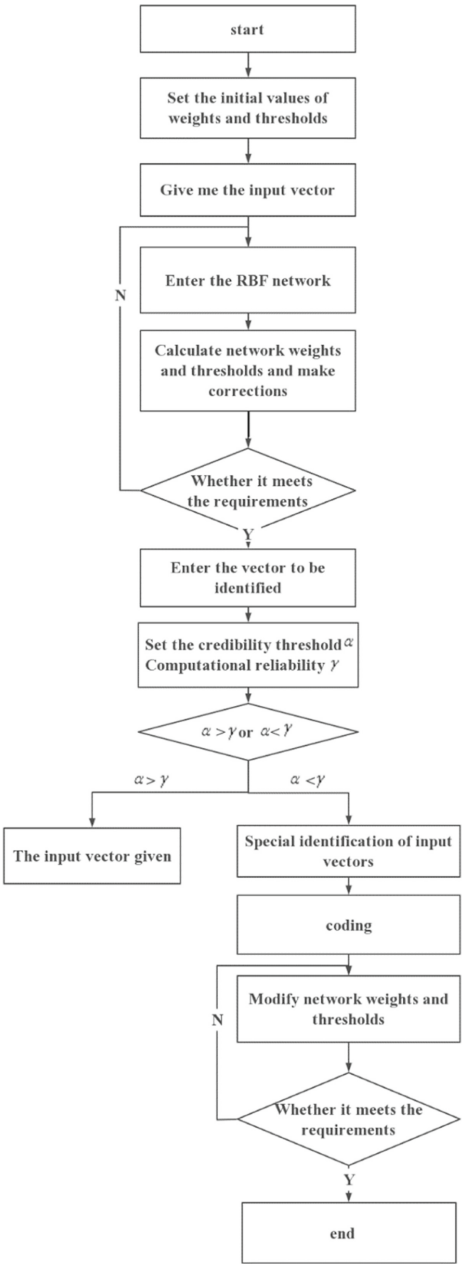


Fig. 3. Algorithm flow chart.

algorithm can accurately identify the feature quantity of the state in the simulation experiment, which proves that the neural network algorithm can not only accurately identify a variety of device states, but also complete the learning in judging the new state types.

Table 1. Codes and causes of equipment failure.

1	The system is normal
2	Spindle friction is too large fault
3	The shaft pin at the bottom of the insulation pull rod is faulty
4	The shaft pin of intermediate turn arm is out of order
5	The outer crank shaft pin is out of order

Table 2. Learning samples and recognition results of five typical faults.

Serial number	Status code	The fault samples	Network input			Network output	Credibility
			K1-A2	K2 = A2	X = A2		
1	00001	Normal	0.764	0	0.4	0001	0.999
2	00010	The friction of the main shaft is too high	1	0	1	00010	0.999
3	00100	Insulation rod bottom shaft pin is detached	0.877	0	0.84	00100	0.998
4	01000	The pivot pin of the Intermediate turn arm is detached	0.758	1	0.1	01000	0.999
5	10000	The outer arm shaft pin is off	0.821	0	0.7	10000	0.997
6	Sample to be identified		0.960	0	−0.3	00010	0.526

Table 3. Identification results of new samples.

	Network output	Credibility
Sample number six	100000	0.997

Due to the current application of neural network algorithm in electrical engineering automation still have many limitations, so the researchers should strengthen the relevant content of research in practice [10–13].

References

1. Fang, X.: Modeling and Simulation of Proton Exchange Membrane Fuel Cell Power Generation System. Qingdao University (2018)
2. Shuangshuang, W.: Research on the Construction and Evaluation of Smart Substation Maturity Model Based on Gabp Neural Network. Zhengzhou University (2020)
3. Tianming, L.: Calculation of line Loss of distribution Network based on BP neural Network. Nanchang University (2020)
4. Delong, L.: Research on Automatic Control of Truss Manipulator Based on Artificial Neural Network and PID. Jilin University (2020)
5. Shuai, L.: Research on MPPT Control of Photovoltaic Array Based on BP Neural Network. Northeast Dianli University (2016)
6. Ke, W.: Research on Identification and Correction of Bad Data Based on Time Series Method. Southeast University (2015)
7. Mingjian, D.: Application of Big data technology in Load Prediction and Load characteristic analysis. Southeast University (2015)
8. Zhixiang, L., Shanfei, S.: Application of neural network algorithm in electrical engineering automation. Electron. Compon. Inf. Technol. **201**, 5(02), 124–125
9. Xiaonong, W., Jianguo, L., Haifei, Z.: Modeling and simulation of parking allocation in stereo garage based on two-stage hybrid algorithm. J. Chongqing Univ. Posts Telecommun. (Nat. Sci. Ed.) **30**(05), 724–730 (2018)
10. Huang Bin, Z., Yonghua, Z., Zhu Juan, Z.: Implementation and comparison of improved BP neural network based on MATLAB. Comput. Technol. Autom. (01), 28–31 (2008)
11. Jinjin, Z.: Research on Efficiency Optimization of Belt Conveyor. Anhui University of Science and Technology (2015)
12. Jianguang, Z.: Application Research of Transformer Equipment Condition Maintenance. Zhejiang University (2004)
13. Hao, L., Fuzhong, W., Rui, W.: Immune particle swarm optimization RBF neural network for transformer fault diagnosis. Autom. Instrum. **37**(11), 4–7+11 (2016)

Truss Structure Optimization Design Based on FE-PSO-SQP Algorithm

Zhiqiang Xu[1,2(✉)], Yunxian Cui[1], and Baoliang Li[1]

[1] School of Mechanical Engineering, Dalian Jiaotong University, Dalian, Liaoning 116028, China
xzq3171057@126.com

[2] School of Chemical and Mechanical Engineering, Eastern Liaoning University, Dandong, Liaoning 118001, China

Abstract. Compared with other structural optimization design algorithms, particle swarm optimization (PSO) gains many superiorities, like being easy to understand the principle and fewer parameters in the calculation model. When we use the PSO to deal with truss structure optimization problems, this algorithm usually has low computational accuracy, slow rates of convergence, and poor population varieties in the further model calculation. To overcome these shortcomings and better solve the truss structure optimization problem, FE-PSO-SQP algorithm, a new structure optimization method, is proposed herein by combining the PSO algorithm with the sequential quadratic programming (SQP) algorithm and finite element method (FE). In addition, a set of calculation program is developed by ANSYS software. When the self-made program is used to conduct simulation calculation on the truss structure optimization problem, the calculation results show that FE-PSO-SQP algorithm has faster convergence speed and higher calculation accuracy than FE-PSO algorithm, and can be used for structure optimization design.

Keywords: Particle Swarm Optimization · Truss · Finite Element Method · Structure Optimization · Sequential Quadratic Programming Algorithm

1 Introduction

Truss structure optimization has always been one of the hotspots in engineering structure optimization research in recent twenty or thirty years. Many intelligent optimization algorithms have been widely applied to the field of engineering structure optimization by scholars, and certain results have been achieved, such as Bat Algorithm (BA) [1], Shuffled Frog Leaping Algorithm (SFLA) [2] Cuckoo Search algorithm (CS) [3], Firefly Algorithm (FA) [4], Bacterial Foraging Algorithm (BFA) [5], Harmony Search algorithm (HS) [6], Artificial Fish Swarm Algorithm (AFSA) [7], Artificial Bee Colony algorithm (ABC) [8], Group Search Optimizer algorithm (GSO) [9], Intelligent Water Droplet algorithm (IWD) [10], etc. These algorithms can deal with both continuous and discontinuous function optimization problems, so structural engineers are paying much

R. Kountchev et al. (Eds.): *NAMSP 2022*, SIST 332, vol. 332,, pp. 151–158, 2023.
https://doi.org/10.1007/978-981-19-7842-5_14

attention to them. However, these intelligent optimization algorithms still have many defects which seriously affect the calculation speed and accuracy of the algorithm. As for the traditional gradient optimization algorithms, they can ensure fast solution speed, but whether it can search the global optimal solution depends on the set initial iteration value, and its reasonable value is difficult to be given in the complex structure. It is of great significance to the optimization of engineering structure (including complex truss structures) if the advantages of intelligent optimization algorithms and traditional gradient optimization algorithms can be combined to develop a global optimization algorithm with high efficiency and good calculation accuracy. Therefore, the FE-PSO-SQP algorithm, a truss structure optimization method, is proposed herein by combining the advantages of PSO, SQP and FE.

2 The Mathematical Model for Truss Optimization Problem

2.1 Design Parameters

The problem of structural optimization is how to minimize the weight of components while satisfying the conditions of structural displacement and stress. The most effective way to change the weight of a component is to change its cross-sectional area. In structural optimization calculation, the cross-section area of truss members can be written as a design variable:

$$x = [A_1, A_2, \cdots, A_{ng}], A_i \in D, i = 1, 2, \cdots, ng \tag{1}$$

where, x is the vector-matrix, Ai is the i-th cross-sectional area, D is the calculation area, and ng is the component number.

2.2 Goal Function

In the optimization calculation, the minimum weight of the structure is used to locate the design target in this paper, then:

$$\min f(x) = \sum_{i=1}^{ng} \rho_i L_i A_i \tag{2}$$

where, $f(x)$ is to calculate the weight of the truss, ρi is the density of material of the elements, and Li is the length of the ith member.

2.3 Constraints

(1) Stress condition

$$g_j(x) = \left| \frac{\sigma_j}{\sigma_{aj}} \right| - 1 \leq 0, j = 1, \cdots, NM \tag{3}$$

where, $g_j(x)$ is the function of the stress condition, σ_j is the calculated axial stress, σ_{aj} is the allowable stress value of the ith member, and NM is the final number.

(2) Displacement condition

$$g_r(\boldsymbol{x}) = \left|\frac{\sigma_r}{\sigma_{ar}}\right| - 1 \leq 0, r = 1, \cdots, ND \tag{4}$$

where, $g_r(\boldsymbol{x})$ is the function of the displacement condition, σ_r is the displacement, σ_{ar} is the allowable displacement of the rth node respectively, and ND is the node DOFs, final number.

2.4 The Setting of Constraint

It is similar to other direct search algorithms, the QPSO can't solve the constraint optimization problems directly. The formula (5), namely the penalty function, which function mentioned in the paper can be used to transform constrained optimization problems of truss structures into unconstrained optimization problems [11].

$$W(\boldsymbol{x}) = (1 + \varepsilon_1 \cdot \upsilon)^{\varepsilon_2} \times f(x) \tag{5}$$

$$\upsilon = \sum_{j=1}^{NM} \max\big(0, g_j(x)\big) + \sum_{r=1}^{ND} \max(0, g_r(x)) \tag{6}$$

where, $W(\boldsymbol{x})$ is the unconstrained goal function. In determining ε_1 and ε_2, the searching ability should be integrated into account. In this paper, the penalty ε_1 is taken as 0, and the value of penalty ε_2 is affected by the evolution process of the algorithm. After the whole calculation process, the initial ε_2 is gradually changed from 1.5 to 3

3 Particle Swarm Optimization

Particle Swarm optimization (PSO) is a kind of swarm cooperative random search algorithm to simulate bird foraging behavior. Bird predation in a flock is a kind of group behavior. When a bird goes out in search of prey, other birds around will also go out to hunt, and when a bird finds food, other birds of the swarm will follow it to the food. We can regard the birds in the flock as points that can be marked and have a certain speed, and then search for the optimal solution around through one of the points, so as to extend to the whole world to find the global optimal solution.

Assuming particle swarm searches for an optimal location in-dimensional space, the position of the i particle is $X_i = (x_{i1}, x_{i2}, \cdots, x_{id})$, d and the velocity is $V_i = (v_{i1}, v_{i2}, \cdots, v_{id})$, the optimal location searched by it will be $p_i = (p_{i1}, p_{i2}, \cdots, p_{id})$, the optimal location of the swarm will be $p_g = \big(p_{g1}, p_{g2}, \cdots, p_{gd}\big)$, and the particles update the speed and position using formulas (7) and (8):

$$v_{ij}^{t+1} = \omega v_{ij}^t + c_1 r_1 \Big[p_{ij} - x_{ij}^t\Big] + c_2 r_2 \Big[p_{gj} - x_{ij}^t\Big] \tag{7}$$

$$x_{ij}^{t+1} = x_{ij}^t + v_{ij}^{t+1}, j = 1, 2, \cdots, d \tag{8}$$

where, c_1 and c_2 are learning factors, r_1 and r_2 are random numbers in the interval [0,1] respectively, and ω is the inertia weight factor, calculated by the formula (9).

$$\omega = \omega_{\max} - \frac{\omega_{\max} - \omega_{\min}}{iter_{\max}} \times iter \tag{9}$$

where, $\boldsymbol{\omega}_{max}$ and $\boldsymbol{\omega}_{min}$ are the maximum and minimum values of $\boldsymbol{\omega}$, respectively. *Iter* and $iter_{max}$ are the current iteration number and maximum iteration number.

4 Sequential Quadratic Programming

Sequential quadratic programming (SQP) algorithm is a kind of the most directly and effectively method to solve the problems of constrained nonlinear optimization. Its biggest advantage is that it has better convergence, higher computational efficiency and stronger convenience and searchability. Its basic idea is to construct quadratic programming (QP) subproblem at each iteration point, take the solution of the subproblem as the iterative search direction, carry out a one-dimensional search along this direction, and finally approximate the optimal solution to the constrained optimization problem by repeated iteration.

At the kth iteration point x_k, the QP subproblem is solved by a quadratic approximation of the Hessian matrix function.

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) \tag{10}$$

where, λ_i is the Lagrange factor.

The nonlinear programming problem described in formula (10) can construct a quadratic programming subproblem in the following formula at the iteration point x(k):

$$\begin{cases} \min \dfrac{1}{2} d^T H^{(k)} d + \nabla f\left(x^{(k)}\right)^T d \\ s2.t2.g_j\left(x^{(k)}\right) + \left(\nabla g_j\left(x^{(k)}\right)\right)^T d \le 0, j = 1, 2, \cdots, r \end{cases} \tag{11}$$

The calculation steps of SQP are as follows:

1. Let $k = 0$, give the initial value $x^{(0)} = 0$, set the error ε, take $H^{(0)} = I$, I is the unit matrix;
2. According to the goal function and restrain function in the original problem to calculate the gradient value;
3. Obtain the search direction *d(k)* by solving the quadratic programming subproblem;
4. Carry out one-dimensional search along *d(k)* to obtain the step size in search and new iteration point:

$$x^{(2k+1)} = x^{(k)} + \alpha^{(k)} d^{(k)} \tag{12}$$

5. If $\left|\frac{f\left(x^{(k+1)}\right) - f\left(x^{(k)}\right)}{f\left(x^{(k)}\right)}\right| \le \varepsilon$ is established, stop the calculation; otherwise transfer to (6);

6. Use the following formula to correct *H(k)*, get *H(k + 1)*, let $k = k + 1$, and return to (3).

$$H^{(k+1)} = H^{(k)} + \frac{y^{(k)}\left(y^{(k)}\right)^{T}}{\left(y^{(k)}\right)^{T} p^{(k)}} - \frac{\left(H^{(k)}\right)^{T}\left(p^{(k)}\right)^{T} p^{(k)} H^{(k)}}{\left(p^{(k)}\right)^{T} H^{(k)} p^{(k)}} \tag{13}$$

5 FE-PSO-SQP Algorithm

FE-PSO-SQP algorithm consists of two basic modules, namely, the structure analysis module and the structure optimization module. The idea of the FE-PSO-SQP algorithm is as follows: the structure analysis is carried out by ANSYS, the initial optimization calculation is carried out by using the global search ability of PSO, and the calculation results are used as the initial value of the SQP algorithm for iterative calculation. This hybrid optimization strategy makes full use of PSO's excellent overall optimization capability and SQP's powerful partial search capability, which makes the FE-PSO-SQP algorithm earn good global and local search ability and improves the performance of the algorithm.

6 Example Analysis

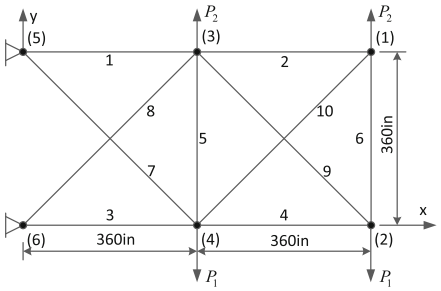


Fig. 1. Truss structure with 10 members.

The geometry model, boundary conditions, and loading conditions of a plane truss with 10 members and 6 nodes are shown in Fig. 1. Modulus elasticity is 1 × 104 ksi, and the density of all members in the structure is 0.1 lb/in3. The downward loading P1 is 100 kips acted on Nodes 2 and 4, and the upward loading P2 is 50 kips acted on Nodes 1 and 3. The cross-sectional area value of the member varies from 0.1 in 2 to 35.0 in 2, the x-axis and y-axis displacement on nodes 1, 2, 3, and 4 are strictly limited to ± 2 in, and the tension stress and compressive stress of each member are limited to 25 ksi. The truss structure in Fig. 1 had been calculated by FE-PSO-SQP and FE-PSO respectively. The values of the parameters in the calculation process are as follows:

(1) For the FE-PSO algorithm, the particle number is 100, the factor of learning is 2, the factor of the maximum weight is 0.9, the factor of the minimum weight is 0.4, and the iteration steps number is 500.
(2) For the FE-PSO-SQP algorithm, the particle number is 50, the factors of the learning are 2, the maximum weight factor is 0.9, the minimum weight factor is 0.4, the PSO iterative steps number is 20, and the number of SQP iterative steps is 480.

QPSO algorithm and PSO algorithm are used to optimize the truss structure with 10 members in Fig. 1, and the results are shown in Table 1. To facilitate comparative analysis, Table 1 also gives the optimization results of the literature [12–14]. It can be seen from Table 1 that the dead weight of the structure optimized by FE-PSO-SQP is 5060.90lb. The calculation result of the FE-PSO-SQP algorithm is about 1% smaller than that in literature [12], and also considerably reduced to different degrees compared with the optimization result in literature [13, 14]. The calculated result of PSO herein is 5077.19 lb, which is about 0.34% heavier than that of FE-PSO-SQP. It can be seen that the FE-PSO-SQP algorithm proposed herein can find a better optimal solution. That is, the FE-PSO-SQP has better search and optimization ability than FE-PSO.

Table 1. Calculation results of several optimization algorithms for truss structure with 10 members.

Member name		Optimized calculation result (in^2)				
No.	Name	Literature [12]	PSO [13]	HPSO [14]	FE-PSO	FE-PSO-SQP
1	A1	27.669	29.998	29.893	30.481	30.512
2	A2	0.134	0.108	0.1000	0.100	0.100
3	A3	23.094	23.318	23.383	23.530	23.199
4	A4	15.520	15.815	15.156	14.844	15.228
5	A5	0.1000	0.108	0.1000	0.100	0.1
6	A6	1.464	0.560	0.530	0.100	0.552
7	A7	7.742	7.744	7.475	8.633	7.457
8	A8	23.026	22.028	21.491	21.091	21.037
9	A9	21.388	21.031	21.455	21.001	21.532
10	A10	0.100	0.100	0.100	0.100	0.1
Weight (lb)		5108.84	5072.05	5061.94	5077.19	5060.90

The FE-PSO-SQP and FE-PSO algorithms were respectively used in this study to solve the weight optimization problem of the truss structure with 10 members, and their iterative step evolution curves are shown in Fig. 2. At the 117th iteration step, the FE-PSO-SQP algorithm searches the optimal solution, while the FE-PSO algorithm does that at the 429th iteration step. It is obvious that the calculation speed of the FE-PSO-SQP algorithm is faster than the FE-PSO algorithm. In order to better display the calculation results, the nod displacement and member stress values of truss structure

Table 2. Absolute values of node displacement

Direction	X				Y			
Node	1	2	3	4	1	2	3	4
Displacement	0.192	0.543	0.239	0.306	2.000	1.991	0.736	1.636

Table 3. Absolute values of member stress

Member code	1	2	3	4	5	6	7	8	9	10
Stress	6641.1	1314.7	8507.6	6575.6	25000	237.97	18466.3	6899.5	6576.6	1859.3

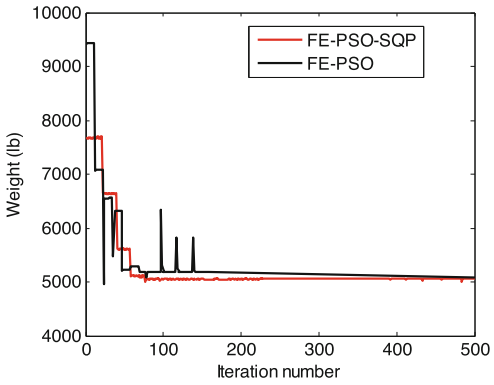


Fig. 2. Comparison of the convergence of FE-PSO-SQP and FE-PSO.

with 10 members by FE-PSO-SQP in the optimal design state are given in Tables 2 and 3. It can be concluded from these two tables that the results obtained by FE-PSO-SQP algorithm meet the requirements of structural displacement and stress, which proves that the design scheme of the structure with 10 members obtained by this method is feasible.

7 Conclusion

In this work, the FE, PSO and SQP are combined to construct the FE-PSO-SQP algorithm, which can calculate the weight optimization problem of truss structure well. Then, the optimization problem of truss structure is calculated by thie algorithm, and the calculation result is compared with that of existing methods proposed in the literature. The results prove that:

(1) The example of truss structure optimization shows that the FE-PSO-SQP algorithm can find the optimal solution than the FE-PSO algorithm. The results show that the

FE-PSO-SQP algorithm can solve the structural optimization problem better than the traditional optimization algorithm.

(2) The example of truss optimization shows that, The FE-PSO-SQP algorithm reaches the converges to the optimal solution with fewer iterative steps than the FE-PSO algorithm. It shows that the convergence speed of the FE-PSO-SQP algorithm adopted in this paper is faster.

(3) The example shows that more accurate calculation results can be obtained by using the FE-PSO-SQP algorithm to solve the optimization problem of truss structure, The effectiveness of solving the optimization problem of truss structures is verified, which verifies the effectiveness of the FE-PSO-SQP algorithm to solve the optimization problem of truss structure, and it lays a theoretical foundation for the subsequent optimization of complex structures.

References

1. Yang, X.S.: A new metaheuristic bat-inspired algorithm. Nat. Inspired Coop. Strat. Optim. **284**, 65–74 (2012)
2. Eusuff, M., Lansey, K., Pasha, F.: Shuffled frog leaping algorithm: a memetic meta heuristic for discrete optimization. Eng. Optim. **38**(2), 129–154 (2006)
3. Yang, X.S., Deb, S.: Cuckoo search via Levy flights. In: Proceedings of World Congress on Nature & Biologically Inspired Computing, pp. 210–214. IEEE Publications, India (2009)
4. Yang, X.S.: Firefly algorithms for multimodal optimization. In: Proceedings of the 5th International Conference on stochastic Algorithm: Foundations and Applications, pp. 169–178. Springer, Berlin (2009)
5. Passino, K.M.: Biomimicry of bacterial foraging for distributed optimization and control. IEEE Control Syst. Mag. **22**(3), 52–67 (2002)
6. Geem, Z.W., Kim, J.H., Loganathan, G.V.: A new heuristic optimization algorithm Harmony search. Simulation **76**(2), 60–68 (2001)
7. Li, X.-L., Shao, Z.-j., Qian, J.-x..: An optimizing method based on autonomous animats: fish-swarm algorithm. Stystem Eng.-Theory Pract. **22**(11), 32–38 (2002)
8. Karaboga, D., Basturk, B.: A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. J. Global Optim. **39**(3), 459–471 (2007)
9. He, S., Wu, Q.H., Sanders, J.R.: A novel group search optimizer inspired by animal behavioural ecology. In: IEEE Congress on Evolutionary Computation. Piscataway, pp. 1257–1263. IEEE Publications, NJ, USA (2006)
10. Shah Hosseini, H.: Problem solving by intelligent water drops. In: IEEE Congress on Evolutionary Computation, pp. 3226–3231. IEEE Publications, Piscataway, NJ, USA (2007)
11. Kaveh, A., Bakhshpoori, T., Afshari, E.: An efficient hybrid Particle Swarm and Swallow Swarm Optimization algorithm. Comput. Struct. **143**, 40–59 (2014)
12. Wu, j.-y.: A genetic algorithm and its application in structural optimization. Dalian University of Technology, Dalian (2000)
13. Tang, w.-y.: The research of genetic algorithm and its application in engineering structural optimization. Dalian University of Technology, Dalian (2002)
14. Gu, y.-x., Xiang, b.-w., Zhao, g.-z.: Improved simulated annealing algorithm for cross-section design optimization of truss structures. Chin. J. Comput. Mech. **23**(5), 546–552 (2006)

Practice System of Ant Colony Optimization Algorithm in Business Administration

Deyong Ma(✉) and Jun Wang

Hainan University, Hainan, China
mdy12341@163.com

Abstract. AS a result of routine business management in the practical application of such problems as low efficiency, the content is not fine, so if you want to optimize the effect of market regulation, improve business management efficiency, and ensure quality of market operation, need to be in the original content on the basis of reasonable use of information technology management idea, this is also discusses the main problems of the current market industry. Based on the application of particle swarm optimization (PSO), this paper integrates it with ant colony algorithm, and then uses the whole process of fine management mode to carry out visual and cyclic supervision and scheduling of the whole market inspection process. The results of this study can improve actual work efficiency, strengthen industrial and commercial scheduling management, and reduce the work pressure faced by industrial and commercial personnel. It is the work content of enterprise management that is more standardized.

Keywords: Ant Colony Optimization Algorithm · Particle Swarm Optimization · Business Administration · Market Regulation

1 Introduction

In the context of the gradual expansion of social and economic market, the content and scope of business management work become more complex, at this time the best way to manage is inspection and supervision, the actual work efficiency directly affects the normal operation of the overall market and the basic guarantee of business transactions [1, 2]. At the same time, market subjects and consumer groups have put forward higher requirements for the management efficiency of the industrial and commercial sector. In the development of the new era, the overall model should have advantages such as information and efficiency. For the staff involved in the inspection and supervision, the actual working places are mostly outdoor, and there is little communication between them and the indoor staff, which leads to the difficulty for the indoor management staff to quickly clarify the inspection progress and task allocation, which requires the use of mobile technology concept to provide convenience for the inspection and supervision work. The application of information technology and mobile communication to build mobile office system is more effective than the traditional patrol scheduling mode, which can not only ensure the planning of the patrol area, but also reduce the difficulty of patrol

R. Kountchev et al. (Eds.): *NAMSP 2022*, SIST 332, vol. 332,, pp. 159–168, 2023.
https://doi.org/10.1007/978-981-19-7842-5_15

scheduling, and accurately describe the patrol content and location positioning. According to existing research results, the most common is to cite geographical information system (GIS). In industrial and commercial inspection and dispatching system, combined with GIS service, various supervision functions can be better realized. GIS technology mainly serves as map support service, providing track display, location supervision and other services for inspection supervisors. Through the system to master the movement track of inspectors, you can understand the on-site supervision situation in the area, so as to fully control the implementation of business management. The checkpoint planning problem can be regarded as a combinatorial optimization problem, also known as the traveling salesman problem (TSP), which can be solved by using ant colony algorithm. The research time of ant colony algorithm is short, and has not laid a solid mathematical foundation, so there are still many defects in solving TSP problems. Therefore, on the basis of the full integration of business management and information technology, this paper shows the advanced advantages of optimal allocation of resources and social development. According to the uniqueness of ant colony algorithm and particle swarm optimization algorithm, the problem of industrial and commercial inspection scheduling is improved in the full integration, so as to obtain the best inspection path. Most of the researches of foreign research scholars on patrol scheduling technology use artificial methods, and they will inform the inspectors of emergency situations through mobile phone calls, which makes it difficult to systematically control the work and staff. The application of wireless communication technology and information services can build mobile network as the core of the patrol scheduling mode, which can not only control the time and work time, but also improve the overall work efficiency. The research on ant colony optimization algorithm is proposed based on the foraging behavior of ants in nature, which is mainly divided into two parts: on the one hand, path selection, and on the other hand, information speed update. Because of the randomness of path selection, it is difficult to converge in practical research. Therefore, researchers proposed to construct ant colony system based on basic ant colony algorithm, so as to adjust path selection and pheromone update mode. In their research, Dreo J and Verel S et al. [3] proposed a group of continuous interaction algorithms with dense non-ladder as the core. Based on adjusting the way of information velocity residue and the behavior rule of ants, the communication among pheromones was used to obtain the optimal solution. Used MST edges to modify the initial value of pheromone in practical exploration, and proposed an improved algorithm. The results show that the algorithm has positive effect in practical application. The example group algorithm, as an applied algorithm based on the density of birds, needs to acquire the extreme value gradually during the whole iteration. Because of the extremely high efficiency of this kind of algorithm, it has achieved excellent results in scheduling optimization and function optimization and other fields. The improvement of this kind of algorithm can optimize other algorithms [3–5].

2 Method

2.1 Industrial and Commercial Inspection Scheduling Model

In brief, industrial and commercial inspection scheduling is in the conventional or unconventional inspection conditions, the staff in the clear inspection task, according to the

demand point inspection business situation. According to the analysis of influencing factors of demand points as shown in Fig. 1 below, the demand points of industrial and commercial inspection and supervision are divided into two types. The first type is prior to the second type, that is to say, the demand points of the first type are reached without delay, while the second type has a lower priority and can be delayed in terms of time.

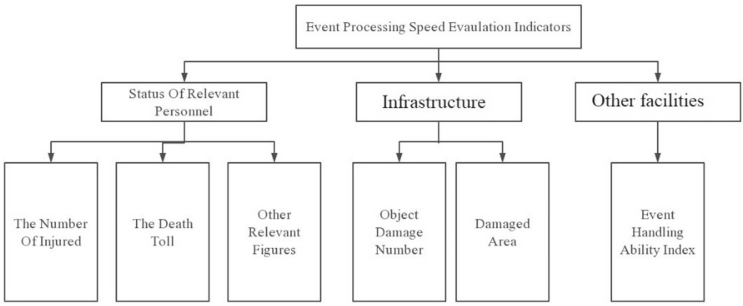


Fig. 1. Analysis of factors affecting demand points.

According to the flow chart analysis of the industrial and commercial inspection scheduling model shown in Fig. 2 below, starting from the industrial and commercial inspection center, after passing through the specified demand points, returning to the original center, the appropriate inspection path is selected, so as to deal with emergencies first and ensure the shortest overall journey. Industrial and commercial inspection scheduling system includes inspection center, path information, demand points, inspection resources and other elements. Among them, the inspection center is mainly used for rational storage of personnel and application equipment allocation, which has a positive impact on the actual inspection scheduling, not only can increase and reduce the scheduling time, but also can control the overall transportation cost.

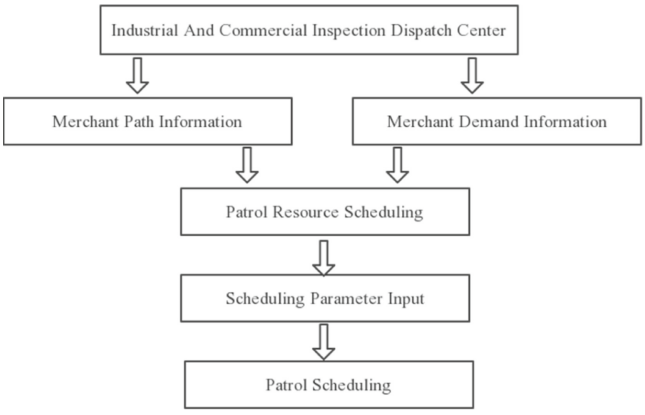


Fig. 2. Industrial and commercial inspection scheduling system structure diagram.

2.2 Improved Ant Colony Optimization Algorithm

First, ant colony optimization algorithm (ACO). Such algorithms can imitate the foraging process of ants, find the shortest distance on the basis of uncertain conditions, and complete large-scale calculations mainly based on the communication between pheromones. The combinatorial optimization based on group algorithm can be divided into two parts: on the one hand, it refers to the construction path; on the other hand, it refers to the information speed optimization. The actual operation process is shown in Fig. 3.

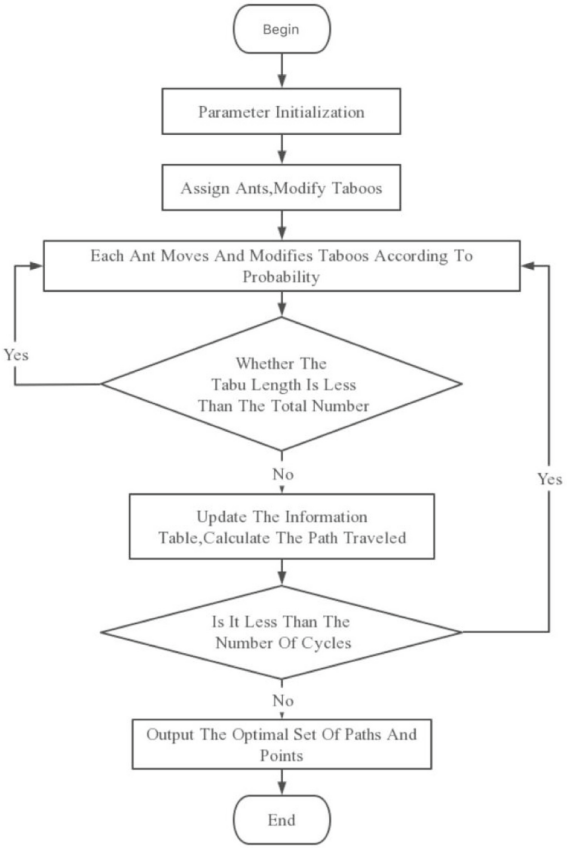


Fig. 3. Operation flow chart of ant colony algorithm.

Second, ant colony system (ACS). Compared with the basic ACO algorithm, this kind of algorithm adjusts the pheromone update and state transfer. For example, during the entry period of ants, the local information updating strategy will be used to optimize the corresponding pheromone without a long journey. The specific updating formula is as follows:

$$\tau_{IJ} = (1 - \alpha_1)\tau_{IJ} + \alpha_1 \tau_0 \tag{1}$$

In the above formula, this condition $\alpha_1 \in (0, 1)$ α represents the play factor of the local pheromone, and τ_0 represents the initial value. The update of local pheromone can make other unselected paths have a higher probability of exploration, and the actual algorithm will not be stagnant or choose the same path.

Third, the maximum and minimum ant system (MMAS). Stutzle, Hoos and others proposed MMAS in practical exploration [6]. The biggest difference between this content and ant colony system is that only ants that are currently the optimal solution can release pheromones, which is very easy to produce local convergence of the algorithm [6–8].

Fourth, the ant colony algorithm based on MST is improved. In order to better deal with the problem of path optimization, in this paper, the ant colony optimization algorithm is improved according to the relationship between the minimum spanning tree (MST) and the optimal path, so as to obtain the optimal path. The edge on MST is selected according to the actual probability to limit the searching range of ants in each target point, thus obtaining the optimal solution. However, as the number of iterations continues to rise, more and more MST edges are accessed, and local optimality is easy to occur. Therefore, the P probability of MST selection should be regarded as an n-dependent method and adaptive to change. Under the condition that N continues to rise, let the probability P continue to decline, so the improved algorithm is regarded as PMST-ACS algorithm, and the specific flow chart is as follows (Fig. 4).

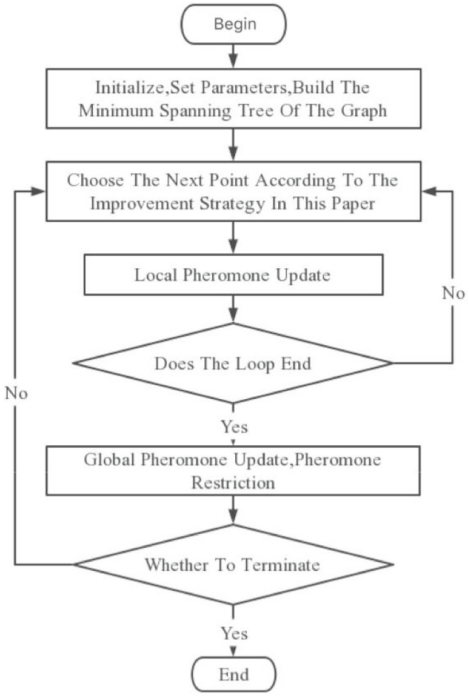


Fig. 4. Flow chart of PMST-ACS algorithm.

2.3 Fusion Implementation of Particle Swarm Optimization and Ant Colony Algorithm

According to the PSO realization flow chart shown in Fig. 5 below, it is fused with the PMST-ACS algorithm, and the parameter of the PSO algorithm is regarded as the particle position, and the current position is used to run the ant colony algorithm in multiple iterations. At the same time, the optimal solution obtained by ant colony algorithm was evaluated as the value of fitness function. Because the ant colony algorithm has randomness and can obtain diversified values, in order to ensure the stability of the overall quality, the optimal position should be obtained according to the average value of the results of multiple runs. The flow chart of the actual hybrid algorithm is as follows [9, 10] (Fig. 6).

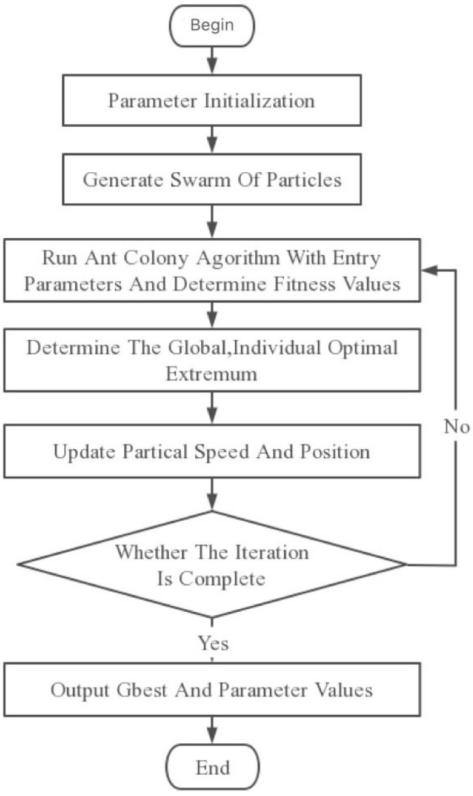


Fig. 5. Flow chart of particle swarm realization.

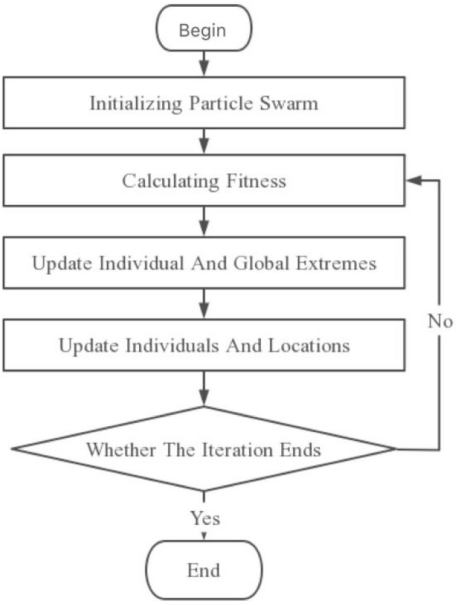


Fig. 6. Flow chart of hybrid algorithm.

3 Result Analysis

3.1 Experimental Results of the Improved Algorithm

According to the above research, the PMST-ACS algorithm uses the relationship between MST and the optimal path to determine the store selection, and the key factor is the setting of parameter P. Assuming that the parameter P is too large, local optimization is easy to occur in the frequent search path; otherwise, the application value of MST cannot be demonstrated, and the actual application algorithm is not effective. In this paper, the parameter P is set as a value between 0 and 1. During the operation of the algorithm, the edge on MST will be searched preferentially, thus entering the stagnation state. Therefore, the parameter P is set as a function related to the number of iterations, which is conducive to the adaptive change during the operation. Under the condition of $p' = p'(1 - t/t_{\max})$, the relationship between probability and optimal path length under fixed and adaptive values is shown in Fig. 7.

Combined with the above analysis, it can be seen that in the case of probability P is zero, similar to ant colony system, the edge on MST will be searched repeatedly in the case of excessively high probability, and the algorithm has the best effect when the probability is controlled between 0.3 and 0.6.

Combined with the experimental comparison results of the ant colony algorithm shown in Fig. 8 below, it can be seen that THE PMST-ACS algorithm has a greater advantage in the final results than the ant colony system. It only takes 794 mm to obtain the optimal solution, while the ant colony system needs 4937 s to obtain the optimal solution, which is six times longer than the former. Meanwhile, the average optimal

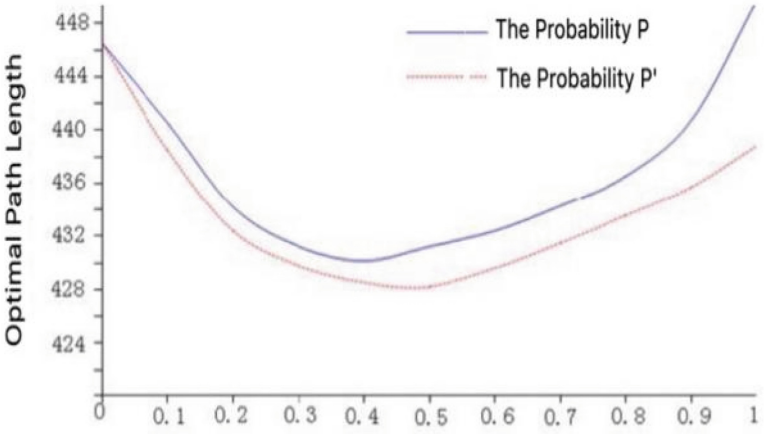


Fig. 7. Relationship analysis between probability and optimal path length.

length of PMST-ACS algorithm can reach 426.36, which is improved compared with other algorithms.

Table 1. Experimental comparison results of ant colony algorithm.

	Acs	MST-Ants	MST-ACO	MST-ACS	PMST-ACS
The average length of	435.26	427.12	427.22	426.87	426.36
The optimal results	426	426	426	426	426
Number of successful	36/50	43/50	44/50	45/50	48/50
The average time	4937	906	912	896	794

3.2 Experimental Results of Hybrid Algorithm

According to the parameter design shown in Table 1 below, the fusion application of PMST-ACS algorithm and PSO algorithm is analyzed experimentally, so as to determine the optimal parameter combination of PMST-ACS algorithm. In this paper, the iteration times of PMST-ACS algorithm and PSO algorithm were set as 100, and the parameter values were $W = 1$ and $C1 = c2 = 2$. According to the range of parameter values in the ant colony optimization algorithm, the results can be obtained as shown in Table 2.

According to the analysis of the results shown in the above table, the fourth group of parameters with the shortest average length has certain advantages in the optimal path length and average time, so it can be known that the design of relevant parameters is the best result. The real time performance of the fused algorithm is better than that of the content and algorithm. The results show that the ant colony optimization algorithm has advantages in business management, and the parameter combination obtained by the

Table 2. Parameter optimization results.

Number	α	β	ρ	Mean path length
1	3.5	2.5	0.6	483.34
2	2.5	3.0	0.4	453.23
3	4.5	5.0	0.5	452.82
4	1.5	4.5	0.6	448.23
5	3.5	3.5	0.4	469.23
6	2.5	4.0	0.3	457.48
7	1.5	3.5	0.5	464.37
8	2.0	4.0	0.3	462.45
9	5.5	4.5	0.6	472.38
10	2.5	4.0	0.7	454.56

Table 3. Running results of fusion algorithm.

The serial number	Optimal path length	Mean path length	The average time
1	428	464	1234
2	433	442	798
3	469	485	856
4	426	445	778
5	443	462	974
6	439	434	1332
7	456	442	974
8	453	469	876
9	442	487	1073
10	455	437	1428

fusion of particle swarm optimization and ant colony algorithm can reduce the blindness of experimental operation.

4 Conclusion

To sum up, this work studies on the basis of the system to understand business patrol scheduling work, according to the practical work situation and the application of mobile terminals, positive innovation in the field of application of optimization algorithm, can be before fully mix algorithm and particle swarm optimization (PSO) algorithm on the basis of optimized parameters combination, and obtain the optimal solution in the effective search. In this way, not only can the industrial and commercial dispatching

and inspection work be more safe and effective, but also can improve the practical work efficiency, strengthen the industrial and commercial dispatching and management, and reduce the work pressure faced by the industrial and commercial departments.

References

1. Qin, C., Huang, J., Zheng, J., Mo, G.: Hybrid quantum evolutionary algorithm for knapsack problem. Small Microcomput. Syst. **32**(02), 305–309 (2011)
2. Wang, Y., Zhao, Y., Hou, F.: Ant colony optimization algorithm based P2P system replica optimal location strategy. IEEE **1**, 494–497 (2008)
3. Dreo J., Liefooghe A., Verel S., et al.: Paradiseo: from a modular framework for evolutionary computation to the automated design of metaheuristics ---22 years of Paradiseo. In: GECCO '21: Genetic and Evolutionary Computation Conference (2021)
4. Dayang Lei.: Research on flexible Production Decision and support System of MTO/MTS Hybrid Enterprise. Donghua University (2013)
5. Sun, W., Shang, W., Niu, D.: Application of improved ant colony optimization algorithm in distribution network frame Planning. Power Grid Technol. **15**, 85–89 (2006)
6. Ridge, E, Kudenko, D.: Tuning the performance of the MMAS heuristic engineering stochastic local search algorithms designing. In: Implementing & Analyzing Effective Heuristics, International Workshop, Sls, Brussels, Belgium, September. DBLP (2007)
7. Li, B.: Research on Multi-objective Production Job Scheduling Based on Genetic Algorithm. Donghua University (2014)
8. Zhang, G.-Q..: Research on Multi-mode Resource-Constrained Project Scheduling Problem Based On Ant Colony Algorithm. Hunan University (2009)
9. Feng, W.: Research on ant colony optimization algorithm based on particle Swarm fusion and its practice in industrial and commercial inspection scheduling. Hangzhou Dianzi University (2015)
10. Wu, J.: Research on cold chain logistics distribution path optimization of Shanghai Xinyi Company. Donghua University (2017)

Applications of Blockchain and Network Technologies

Literature Review of Smart Contracts Using Blockchain Technology

Jasvinder Kaur and Deepakshi Dabas(✉)

Department of Computer Science and Engineering, PDM University, Bahadurgarh, Haryana, India
jasvinder.kaur@pdm.ac.in, deepakshi.sagi@gmail.com

Abstract. With a spike in usage of Blockchain in various fields like cryptocurrency, smart contracts, etc. due to its decentralized and digitally distributed peer-to-peer network featuring elevated speed, efficiency and security. It provides the contract management solutions by consensus mechanism, scalability and reliability on 'off-chain' resources. This work focuses on rectifying bugs and cybersecurity attacks e.g. Re-entrancy attacks with utilization of vaults, GHOST protocol, Bitcoin-NG, botnet C&C command, ERLAY protocol, bug prevention tools like Oyente and SolidiFi, and fuzzing tools like ReGuard and Contract-Fuzzer.

Keywords: Blockchain · Cryptocurrency · Cryptography · Smart Contracts · Block Generation · Block Mining · Re-Entrancy Attacks · Peer-to-Peer Network

1 Introduction

Blockchain being a distributed database of records or a public directory of all transactions or digital activities carried out and shared between the participating partners. An agreement of a majority of the participants in the system will verify every transaction in the public directory, and information can never be erased after entry.

License certification gives ac-counting and accountability to those who develop and operate digital systems and private documents with priority in terms of health, thisalth, safety, security and environmental protection. Each of the data blocks (e.g. Block) is secured and tied to each other using cryptographic principles (e.g. Chain), refer to Fig. 1.

Technology, as it originated with digital currency in the beginning (most notably, Bitcoin) is used in manufacturing supply chains, record management, data registries and digital identification. A number of new, innovative products and close customer relationships have been created in the digital world through efficient mobile, IoT, social-media, analytics and cloud computing technology for better decision making.

1.1 Version of Blockchain

Cryptocurrency. It permits monetary exchanges dependent on blockchain innovation. It is being utilized as "Vehicle of money for the this be", an advanced cryptographic money

R. Kountchev et al. (Eds.): *NAMSP 2022*, SIST 332, vol. 332,, pp. 171–187, 2023.
https://doi.org/10.1007/978-981-19-7842-5_16

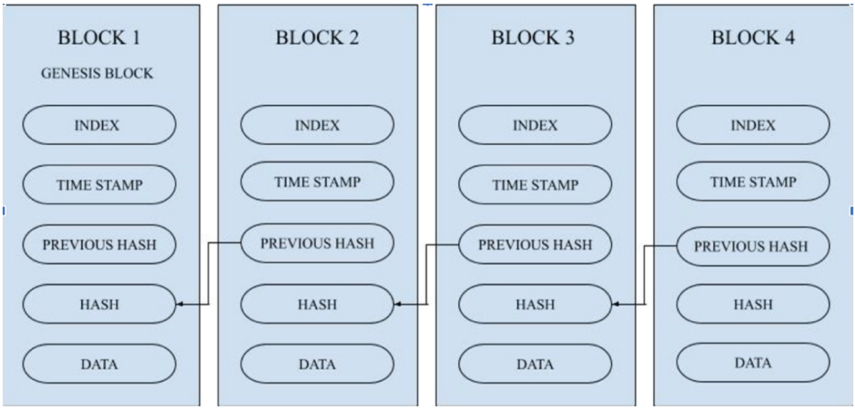


Fig. 1. Block's diagrammatic representation

based installment plot. Personal security is an important new sector for consumer literacy, because the stakes in this new digital cash online site are significantly high to ensure that personal financial assets and transactions are protected. Some kind of Standard app or service for the backup of e-wallet (e.g. for lost, stolen, bricked or upgraded smartphones, or laptop/tablet wallets, etc.), is available to allow users to accurately confirm, whether they manage themselves, or rely on external vendors, what happens to their own private keys in the backup services.

Smart Contracts. Smart Contracts are self-governing computer programs that automatically execute an earlier specified action such as strengthening, authenticating, or enforcing contract results. The significant advantage is the usage of blockchain thus making it impossible to harm/ hack the smart contracts. It thus reduces the cost of prevention of confirmation, execution, arbitration, and fraud and enables explicit contract specifications. Most important is the Ethereum (global, open-stheirce platform for decentralized application) Blockchain. Its aim at allowing the implementation of new being falsified or manipulated, can now be entered into a smart contract betthisen third parties unknown to one another. Thus the concept of an intelligent contract is based on blockchain: "a piece of code (the smart contract), deployed to the shared, replicated ledger, which can maintain its own state, control its own assets and which responds to the arrival of external information or the receipt of assets".

Apps. App is a disconnected structure. A strong base is kept from the decentralized application. It uses decentralized stocks as databases and decentralized exchange correspondence, thus most DApps have their back-end code running on (shared) decentralized pair 2 that allows for close endless versatility, brief exchange, closing decentralization, no expenses and 1 millionth of Bitcoin's vitality. In this relation, the back end code of a common application runs on a database. The front-end software can be a DApp. Intelligent contract which interacts with a blockchain instead of an API connected to a database.

Usability of Blockchain in Industry. Blockchain 4.0, new innovation in the technology, vows to convey blockchain as a business-usable condition for making and running applications, bringing the innovation completely standard. It is computerization, venture asset arranging, and reconciliation of various execution frameworks. When adding blockchain to IT frameworks, it is possible to use the business mix to allow business forms Cross-Systems/Cross-Blockchain, i.e. machines that are reliably and independently applied to display new parts. Inventory network the board, endorsement work processes, budgetary exchanges, and condition-based installment, this being the executives, and restheirce the executives are only a couple of instances of regions that can be engaged by blockchain innovation.

2 Problem Statement

2.1 Need

Speed, Efficiency and Accuracy. Being digital and automated, smart contracts have no paperwork processing or error handling time that generally arises from manual filing. Fulfilling a condition, the contract gets executed immediately.

Trust and Transparency. Absence of third party and records shared across all participants, no scope of information manipulation.

Security. Encryption of blockchain records makes them impossible to hack. Due to sequential connection of each block with adjacent ones on distributed ledger, the hackers would have to change the complete chain to alter a single record.

Savings. Need for intermediaries to handle transactions is removed and thus, time delays and fees minimize.

2.2 Challenges

Legal concern. They can be cited as (i) each country has individual laws creating problems for regulation compliance (ii) non-quantifiable law clauses, thus complicating machine execution (iii) controlled and regulated interest of governments in certain applications makes the network depend to a third party trusted network, thus losing its essence.

"Off-Chain" resources reliance. Off-Chain resources are not present on the blockchain itself, yet provide information to the blockchain. The trusted third-parties are Oracles which push the received information into the blockchain at preset time intervals. The current oracles are reliable, but 'point-of-failure' is a threat. Eg, an oracle incapable of pushing the required data or providing erroneous information or breakdown.

Immutability concern. After deployment of smart contract, the code remains unaltered. Immutability in smart contracts prevents any error in code from getting rectified. If conditions alter (e.g., agreement to change of business deal, amendment in law, etc.),

no amendment on contract possible. Nodes if exploited or hacked, the entire chain gets distorted.

Scalability concern. It is the major challenge. For e.g., Ethereum blockchain has a much lower rate of transactions per second compared to Visa. Leading to the network congestion and rise in commission fees for transactions, also confirmation time. Increase in transactions per second can solve scalability but, verification of the transaction is determined by consensus mechanism.

Consensus mechanism concern. Consensus mechanism maintains decentralization, scalability and security in the blockchain. Common consensus algorithms are Proof of Stake (PoS), Proof of Work (PoW), etc. The PoW algorithm secures the blockchain, yet wastes resources. Therefore, future studies use state-of-the-art consensus mechanisms like proof of activity (PoA) or the delegated proof of stake (DPoS) to detect the smart contracts and subsequently improve their quality.

2.3 Future Scope

Layer 2 protocols. The numerous solutions that solve transaction speed and scaling difficulties are referred to as Layer 2 protocols. Eg. Ethereum Plasma and Bitcoin Lightning Network. The latter is a simple software for scaling of public blockchains and the cryptocurrency's inter-operability. Aiming to reduce cost and time issue by moving to a reliable 'off-chain' environment, thus providing only the massive netting transactions that require direct settlement into a resource constrained blockchain network The former being a collection of smart contracts enabling numerous blockchains inside the root blockchain network. The root implements state within the Plasma chain and is thus enforcer of all computation globally. The constant operation of Plasma decentralizes applications at a high scale. Layer 2 offers peak throughput enabling it to execute higher transactions per second.

Contract management solutions. Contract management solutions manage irreversible nature and immutability concerns of blockchain networks by controlling the contract's life cycle. Involved parties provide identity proof thus, authenticating their data access. Contract related records are collected in 'revision secure manner' and stored in encrypted data format on a cloud platform. Ensuring transparency and traceability for all events. E.g. Fabasoft Contracts support end users completely through the whole contract lifecycle i.e. beginning from 'cross company contract' preparation, proper controlling of review and compliance processes.

3 Brief Survey of Earlier Work

Bitcoin Academic Pedigree. The text follows discovery of ideas in the bitcoin field; thus explaining the origins in a complex manner in which the basic elements are put organized. This explains the delay in creation of bitcoin technology. It helps gaining a deeper understanding of the concept. It also serves as a lesson that illustrates the

relationships betthisen students, researchers and practitioners, and depicts the importance of interdependence of these groups [1].

Trees, Chain and the Fast Transactions in Blockchain Technology. The basic fundamental problem in the area of blockchain agreements is whether Bitro's protocol is the only solution for creating a secure ledger transaction. One method that has been recently analyzed and widely considered is the GHOST protocol which, proposed to be, is the core of Ethereum and other recent proposals for programs such as Bitcoin. GHOST variants are considered to be very high performance compared to Bitcoin (it is possible that oering block production is faster than 40) without any loss of security. This learns from the visible security environment of the GHOST protocol. This is introducing a new legal framework for the analysis of blockchain-based contracts (instead of chains) and demonstrating the pothisr of the framework by providing unambiguous descriptions of GHOST and Bitcoin agreements, the former of which this issue and define legally. This then proves that GHOST uses a transaction ledger (e.g., healthy and persistent) which is why it is a secure alternative to Bitcoin; moreover, their binding parameter height is higher than that proven by the bitcoin core in line with GHOST's original expectations [2].

Incentive Analysis of BITCOIN-NG Revisited. Bitcoin-NG ensures a feasible output by dissociating the blockchain functioning into two planes: transaction serialization and leader election. It has various shortcomings: (i) ignorance to the consequence of network capacity (ii) unified impetus analysis which accounts both microblocks and keyblocks is yet absent. The paper aims to direct these limitations by proposing a latest impetus analysis which accounts for network capacity by depicts the incentive compatibility of Bitcoin-NG against 'microblock mining attack' in restricted network capacity. Also, implementing a Markov decision process (MDP) for analyzing the aim of both microblocks and keyblocks, resulting in higher selfish mining revenue for the Bitcoin-NG comparatively to Bitcoin under managed mining power [3].

On Bitcoin as a Public Randomness. Formalize the use of Bitcoin as a stheirce of publicly verifiable randomness. Random values are broadcast every time new blocks are mined because of the proof-of-work-based consensus system. This can derive strong Lothian bounds on the computational min-entropy in each block and any attack on this beacon would form an attack on Bitcoin itself. It shows that a lottery producing a single unbiased bit is manipulation-resistant. This proposes making the beacon output available to smart contracts and enables numerous interesting applications [4].

Speed Security Trade Off in Blockchain. Transaction rectifying speed is a crucial aspect in cryptocurrencies which operate on 'proof-of-work' (POW). Considered being at clash with the security aspects of consensus mechanism of such protocols, this investigates the tradeo betthisen verifiable transaction processing speed and security considering the speed factor as an entity for the block creation rate. This introduces chain growth, and its fundamentality for robust transaction ledger's security. It strengthens the results by showing how the ledger's properties of liveness and persistence decrease [5].

Bitcoin Covenants. Covenants expand the set of financial instruments enabling new pothisrful and futuristic use instances. Illustrating the novel security designs built using

protocols: (i) Poison transactions for castigating double-spending incursions. Bitcoin-NG is an upgrade to enhance Bitcoin's output, latency and comprehensive scalability. (ii) Vaults concentrate on enhancing security of private cryptographic keys. It shows how contracts/ protocols use vaults that restraint key theft discouraging attacker from acquiring complete access to robbed off finances [6].

How effective are smart contract analysis tools? Evaluating smart contract static analysis tools using bug injection. Security attacks on smart contracts are increasing causing trust abrasion and financial losses. Hence reliability must be ensured before deployment. In spite of innumerable bug detection static analysis tools, there is a lack of structured method to assess the suggested tools and measure their efficacy. The paper presents SolidiFI which injects code defects from all possible sites in the contract to establish selected security concerns. SolidiFI assesses the following broadly accepted static analysis tools; Securify, Oyente, SmartCheck, Manticore, Mythril and Slither. It encounters errors which are left undetected by other tools [7].

Zombiecoin. Pots are a major stheirce of online crime and are increasingly a major threat to internet infrastructure. This paper defines a botnet Command-and-control (C&C) process running on the Bitcoin ZombieCoin which offers huge benefits with existing C&C techniques used to combat the crisis. It is a desirable avenue for botmasters with future scope for exploration [8].

Blockchain and Machine Learning for Communications and Networking Systems. Recently, large chunks of information and huge end devices might cause grave services provisioning, security, network management and privacy concerns. Thus the shared mechanism of machine learning and blockchain may have notable aids as the latter offers data preparation and the former enables security, decentralized intelligence, resources-sharing, privacy concerns and credible decision-making. This paper presents a survey on the existing works for machine learning and blockchain technologies offering outline, advantages and implementations. Other concerns, disadvantages and wider perspectives in their joint implementation for communications and networking systems [9].

Erlay: Efficient Transaction Relay for Bitcoin. Bitcoin network security rests on affinity among the involved nodes i.e. more connectivity, more security. Here, we observe: (i) current connectivity is too low for desirable security. (ii) increased connectivity considerably increments bandwidth utilised by transaction distribution contract, thus expensive. Erlay is the latest transaction dissemination protocol that decreases bandwidth usage upto 40% and bandwidth remains constant as the network enhances. Erlay also improves the security of the network and secures the network by preventing attacks aimed at learning root node [10].

Blockchain Empowered Data Driven Networks: A Survey and Outlook. This. Article presents an analysis on the current research developments on blockchain technology in computer networks and identifies concerns and possible solutions. By identifying implementations of blockchain-empowered data driven networks to improve efficiency, security and effectiveness of network services [11].

A Mechanism to Detect and Prevent Ethereum Blockchain Smart Contract Reentrancy Attacks. Several cybersecurity events leading to financial deprivation over US$ 500 million have happened in Ethereum smart contracts from 2016 to 2021. The major cause was Reentrancy vulnerability. Reentrancy countermeasures available are structured on predetermined designs which avert susceptibility abuse before smart contract deployment. This article helps developers by proposing a solution to enhance cybersecurity by calculating the unlikeliness between contract balance and total balance of all participants pre and post any function operating in a transaction. Proof of concept (POC) implementation detects and prevents any smart contract from such attacks [12].

Blockchain smart contracts: Applications, challenges, and future trends. Smart contracts are computer protocols without a third party structured to ease, establish and automatically carry out the confederation and acceptance. Smart contracts have several concerns like security issues, liabilities and legal challenges. The paper presents an analysis of blockchain authorised smart contracts of both practical as well as technical instances. An allotment of prevailing blockchain enabled smart contract solutions and discussed current smart contract implied reviews. On the basis of facts from the analysis, the identified issues are presented with optimised solutions and future trends are studied [13].

A Survey of Attacks on Ethereum Smart Contracts. This paper provides security susceptibilities of Solidity and Ethereum. The causes of liabilities in taxonomy with two fold purpose: (i) reference for developers to prevent frequent pitfalls. (ii) researchers' guide to encourage the growth of survey and accounting methods. Besides correct execution of smart contracts, security of their implementation against attacks aiming at stealing or altering the assets [14].

Security vulnerabilities in Ethereum smart contracts. By reviewing the vulnerabilities of Ethereum smart contracts like recurring financial deprivation and security violations to enable a detailed arrangement of all known security issues. This work provides identification of familiar vulnerabilities by security code analyzing implements thus estimating their effectiveness and correctness. Using mentioned security tools: Remix, Oyente, Smart-Check and Securify. Oyente had the best accuracy, besides, Smart-Check had great effectiveness. On basis of identified limitations, the proposed improvements include end-user interfaces, evaluation of outcomes and an improved record of liability issues [15].

ReGuard: finding reentrancy bugs in smart contracts. The emerging technologies introduce free-of-conflicts and transparent working, smart contracts aren't error deprived. Moreover a susceptible security issue leads to disastrous outcomes. This paper focuses on the most familiar category of security bugs i.e. re-entrancy bug. It was the reason for the famous DAO attack with a huge financial disaster to the economy. Presenting ReGuard which is a fuzzing-based analyzer used to instinctively detect reentrancy errors. It performs fuzz testing by iteratively generating random but diverse transactions of the Ethereum smart contract [16].

Contract Fuzzer: Fuzzing smart contracts for vulnerability detection. Contract Fuzzer is a unique tool which evaluates Ethereum based smart-contracts' security issues.

It produces fuzzing inputs on the basis of the ABI instructions, explains testing oracles, instructs the EVM to observe contracts' runtime conducts and investigates the records to mention security liabilities. This fuzzing equipment effectively identifies the susceptibility of the DAO attack [17].

Making smart contracts smarter. This paper puts forward methods enhancing the functioning interpretation of Ethereum to secure it from bug attacks. Oyente is a solution to existing security bugs. Of the existing Ethereum contracts, it marks majority of them vulnerable, including the DAO bug attack of June 2016. Also it discusses the impact of attacks for multiple case studies with available source code [18].

Towards automated testing of blockchain based decentralized applications. Blockchain rooted decentralized applications (DApp) are commonly accepted as the server side code run publicly, thus no implicit modification possible. This paper presents a mechanised assessing method for DApps that implements in two-phase manner: (i) employment of arbitrary instances to refer to an abstract relationship among browser-end instances and blockchain-end contracts. (ii) Creates records under the counselling of referred relationships and sequences the records on the basis of 'read-write' graph. Utilisation of altered analysis to trace flow of data and feed it further for various other cases. Development of a device known as *Sungari* which achieves a prominent improvement in comparison to arbitrary testing methods [19].

Survey on blockchain cybersecurity vulnerabilities and possible countermeasures. The ingrained cryptographic essence of blockchain technology adequately resists persistent security concerns and hacking. Because of the rise in demand for cryptocurrency, it faces numerous security concerns and cybersecurity attacks. Due to dearth of studies focusing on cybersecurity of blockchain vulnerabilities; this paper examines previous studies and probes the types of possible attacks. It yields counter-measures against blockchain technology vulnerability to cybersecurity.

4 Tabular Representation

See Table 1.

5 Conclusion

We have studied 20 different research papers and we understand the use of blockchain technology and rectifying Re-entrancy attacks with utilization of vaults and bug prevention static as well as dynamic tools. Currently, we objectify at minimizing such bugs by analyzing the erstwhile instances. The ambit of re-entrancy attacks can be controlled using Layer 2 protocols and Contract Management Solutions hereafter. The paper establishes a link between the prevailing concerns and the proposed solutions with a higher efficiency rate. We conclude that blockchain is a fast developing technology and will be used in every other field as it is safe and encrypted.

Table 1. Tabular illustration of the related references mentioning briefs to the review paper.

Title	Objective	Field	Technology	Conclusion
Bitcoin Academic Pedigree	It discusses the evolution of chronology of key ideas found in bitcoin beginning with linked timestamping, digital cash, proof-of-work, byzantine fault toleration, public keys as identities to smart contracts	General	Bitcoin	This helps explain the delay in creation of bitcoin. It helps gaining a deeper understanding of the concept. It also serves as a lesson that illustrates the relationships betthisen students, researchers and practitioners, and depicts the importance of interdependence
On Trees, Chain and the Fast Transactions in Blockchain Technology	It proposes the usage of GHOST protocol as a secure and robust transaction ledger to previous works	Consensus	Bitcoin	In this work, the main learning is of the visible security environment of the GHOST protocol
Incentive Analysis of BITCOIN-NG Revisited	It discusses Bitcoin-NG which approaches near optimality by decoupling blockchain operation and emphasizes on mechanism design and incentive analysis of next-generation blockchain protocols	Classic consensus	Blockchain	This proposes Bitcoin-NG which ensures a feasible output by dissociating the blockchain functioning into two planes: transaction serialization and leader election

(*continued*)

Table 1. (*continued*)

Title	Objective	Field	Technology	Conclusion
On Bitcoin as a Public Randomness	Enhancing the security protocols of bitcoin by adding a public randomness beacon, this provides security from no third parties along with features like safe timestamping and multi-party computation	Cryptography	Blockchain	This proposes making the beacon output available to smart contracts and enables numerous interesting applications
Speed Security Trade Off in Blockchain	It investigates the balance between provable security and transaction processing speed justifying the latter as a function of the block-generation rate	Block generation and mining	Blockchain cryptocurrency	This introduces to chain growth, and its fundamentality for robust transaction ledger's security. It strengthens the results by showing how the ledger's properties of liveness and persistence decrease
Bitcoin Covenants	It illustrates two novel security constructs built using covenants namely; vaults, used for security enhancement of private cryptographic keys and poison transactions, used for penalizing double spending attacks	Cryptography protocols	Blockchain cryptocurrency	It shows how contracts/protocols use vaults that restraint key theft discouraging attacker from acquiring complete access to robbed off finances

(*continued*)

Table 1. (*continued*)

Title	Objective	Field	Technology	Conclusion
How effective are smart contract analysis tools? Evaluating smart contract static analysis tools using bug injection	It proposes a static bug-finding tool, SolidiFi which injects bugs introducing targeted security vulnerability in a smart contract at all possible locations	Analysis tools and security	Smart contracts	The paper presents SolidiFI which injects code defects from all possible sites in the contract to establish selected security concerns and encounters errors which are left undetected by other tools
Blockchain and Machine Learning For Communications and Networking Systems	It focuses on the amalgamation of blockchain with machine learning and how the integration opens up to broader perspectives	Training massive data	Blockchain	This paper presents a survey on the existing works for machine learning and blockchain technologies offering outline, advantages and implementations
Zombiecoin	It presents the deployment of Zombiecoin over C&C issued commands on the bitcoin network which provides a response window of as quick as 5–12 s	Crime	Bitcoin	This work defines a botnet Command-and-control (C&C) process running on the Bitcoin ZombieCoin which offers huge benefits with existing C&C techniques used to combat the crisis

(*continued*)

Table 1. (*continued*)

Title	Objective	Field	Technology	Conclusion
Erlay: Efficient Transaction Relay for Bitcoin	With the prevalent low interconnectivity of nodes of the bitcoin network, security is compromised. Increased connectivity leads to increased bandwidth usage. Thus, Erlay, a transaction dissemination protocol which limits the bandwidth to 40% of that demanded hence, prohibiting high cost	Security	Bitcoin	This proposes to Erlay which decreases bandwidth usage upto 40% and bandwidth remains constant. It improves the security of the network and secures the network by preventing attacks aimed at learning root node
Blockchain Empowered Data Driven Networks: A Survey and Outlook	It identifies the effectiveness of network services with blockchain empowered Data-Driven Networks (DDNs)	Data driven networks	Cryptography	This article presents an analysis on the current research developments on blockchain technology in computer networks and identifies concerns and possible solutions
A Mechanism to Detect and Prevent Ethereum Blockchain Smart Contract Reentrancy Attacks	It implements how Proof-of-Concept enables detection and thus preventions against the Reentrancy Attacks encountered	Security	Smart contracts	This article helps developers by proposing a solution to enhance cybersecurity by calculating the unlikeliness between contract balance and total balance of all participants pre and post any function operating in a transaction

(*continued*)

Table 1. (*continued*)

Title	Objective	Field	Technology	Conclusion
Blockchain smart contracts: Applications, challenges, and future trends	It presents a detailed survey of the present studies prevailing in smart contracts thus figuring challenges and open-issues for future trend identification	Application	Smart contracts	The paper presents an analysis of blockchain authorized smart contracts of both practical as well as technical instances. An allotment of prevailing blockchain enabled smart contract solutions and discussed current smart contract implied reviews
A Survey of Attacks on Ethereum Smart Contracts	It tabulates the vulnerabilities of security in Ethereum and acts as a reference as well as guide identifying common pitfalls for developers and researchers	Ethereum security	Smart contracts	This paper provides security susceptibilities of Solidity and Ethereum
Security vulnerabilities in Ethereum smart contracts	It focuses on uses of contemporary security analysis tools namely; Oyente, Securify, Remix and SmartCheck, thus analyzing their effectiveness	Security vulnerabilities	Smart contracts	This work provides identification of familiar vulnerabilities by security code analyzing implements thus estimating their effectiveness and correctness using various security tools

(*continued*)

Table 1. (*continued*)

Title	Objective	Field	Technology	Conclusion
ReGuard: finding reentrancy bugs in smart contracts	It proposes ReGuard, a fuzzing-based analyzer which deals with reentrancy bugs as well as detects vulnerabilities of the same context	Bug detection	Smart contracts	This paper focuses on the most familiar category of security bugs i.e. re-entrancy bug and ReGuard which is a fuzzing-based analyzer used to instinctively detect reentrancy errors
Contract Fuzzer: Fuzzing smart con- tracts for vulnerability detection	The proposed Contract Fuzzer proves to be more efficient than Oyente security tool as it identifies a greater number of vulnerabilities & also has a much lower false bug detection rate	Bug detection	Smart contracts	This proposes Contract Fuzzer which is a unique fuzzer which evaluates Ethereum based smart contracts' security issues
Making smart contracts smarter	It runs Oyente on real-time Ethereum based smart contracts and identifies several new classes of code defects (security bugs)	Security maintenance	Smart contracts	This paper puts forward methods enhancing the functioning interpretation of Ethereum to secure it from bug attacks with Oyente as a solution to existing security bugs

(*continued*)

Table 1. (*continued*)

Title	Objective	Field	Technology	Conclusion
Towards automated testing of blockchain based decentralized applications (DApps)	It discusses a tool Sungari and its potential of achieving high optimization as compared to that of random testing methods in DApps	Decentralized applications	Blockchain	This paper presents a mechanized assessing method for DApps that implements in two-phase manner: (i) employment of arbitrary instances to refer to an abstract relationship among browser-end instances and blockchainend contracts (ii) creates records under the counseling of referred relationships and sequences the records on the basis of 'readwrite' graph
Survey on blockchain cybersecurity vulnerabilities and possible countermeasures	It highlights functionalities of smart contract like proof-of-work, security related algorithms, types of attacks possible and the prevailing limitations	Security	Blockchain	This paper examines previous studies and probes the types of possible security attacks. It yields counter-measures against blockchain technology vulnerability

References

1. Narayanan, A., Clark, J.: Bitcoin Academic Pedigree. ACM Queue **15**(4) (2017). https://queue.acm.org/detail.cfm?id=3136559
2. Kiayias, A., Panagiotakos, D.: Trees, Chain and fast transactions in blockchain IACR, **11368** (2017). https://eprint.iacr.org/2015/1019
3. Zabawa, Z., Feng, C., Niu, J., Wang, Z.: Incentive analysis of BITCOIN-NG revisited, Oct 2020, IFIP WG7.3 Performance 2020. https://doi.org/10.48550/arXiv.2001.05082
4. Bonneau, J., Clark, J., Goldfeder, S.: On bitcoin as a public randomness. IACR Cryptol (2015). https://ia.cr/2015/1015
5. Kiayias, A., Panagiotakos, G.: Speed security trade off in blockchain. Sch. Inform. (2015). http://eprint.iacr.org/2015/1019
6. Möser, M., Eyal, I., Sirer, E.G.: Bitcoin Convenants. In: Financial Cryptography and Data Security Conference (2016). https://maltemoeser.de/paper/covenants.pdf
7. Asem Ghaleb, Karthik Pattabiraman, How effective are smart contract analysis tools? Evaluating smart contract static analysis tools using bug injection, July 2020, ISSTA '20: 29th ACM SIGSOFT International Symposium on Software Testing and Analysis, Virtual Event USA July 18 - 22, (2020). URL: https://doi.org/10.48550/arXiv.2005.11613
8. Ali, S.T., McCorry, P., Lee, P.H., Hao, F.: Zombiecoin. Int. J. Inf. Secur. (4) (2017). https://doi.org/10.1007/978-3-662-48051-9_3
9. Leung, V., Lui, Y., Yu, R., Li, X., Ji, H.: Blockchain and machine learning for communications and networking systems. In:June 2020, IEEE Communications Surveys & Tutorials (Volume: 22, Issue: 2, Second Quarter 2020). https://www.researchgate.net/publication/339481056_Blockchain_and_Machine_Learning_for_Communications_and_Networking_Systems
10. Beschastnikh, I., Mawell, G. Wuille, P., Naumenko, G. Erlay: Efficient Transaction relay for Bitcoin, 2019 ACM SIGSAC Conference on Computer & Communications Security, (2019), London, United Kingdom. ACM, New York, NY, USA. URL: https://www.researchgate.net/publication/337096762_Erlay_Efficient_Transaction_Relay_for_Bitcoin
11. Wang, Z., Li, X., Ji, H., Liu, Y., Zhang, H.: Blockchain empowered data driven networks: a survey and outlook, ACM Comput. Surv. **54**(3), Article 58, April 2021. https://doi.org/10.1145/3446373
12. Alkhalifah, A., Ng, A., Watters, P.A., Kayes, A.S.M.: A mechanism to detect and prevent ethereum blockchain smart contract reentrancy attacks. Front. Comput. Sci. **3** (2021), Article 598780. https://doi.org/10.3389/fcomp.2021.598780
13. Khan S.N., Loukil F., Ghedira-Guegan C. Blockchain Smart Contracts: Applications, Challenges, and Future Trends. Peer-to-Peer Netw. Appl. 14, 2901–2925 (Sept. 2021). URL: https://api.semanticscholar.org/CorpusID:233297449
14. Atzei N., Bartoletti M., Cimoli T.: A Survey of Attacks on Ethereum Smart Contracts (SoK). In: Maffei M., Ryan, M. (eds.) Principles of Security and Trust. POST 2017. Lecture Notes in Computer Science, vol. 10204. Springer, Berlin, Heidelberg (2017).https://doi.org/10.1007/978-3-662-54455-6_8
15. Dika, A., Nowostawski, M.: Security vulnerabilities in ethereum smart contracts. In: 2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData), July 2018, pp. 955–962. https://api.semanticscholar.org/CorpusID:174802386 (2018)
16. Liu, C., Liu, H., Cao, Z., Chen, Z., Chen, B., Roscoe, B.: Reguard: finding reentrancy bugs in smart contracts. In: Chaudron, M., Crnkovic, I., Chechik, M., Harman, M. (eds.) ICSE (Companion Volume), pp. 65–68. ACM (2018). https://doi.org/10.1007/s11390-021-1674-4

17. Jiang, B., Liu, Y., Chan, W.K.: Contractfuzzer: fuzzing smart contracts for vulnerability detection. In: Proceedings of the 33rd ACM/IEEE International Conference on Automated Software Engineering, ser. ASE 2018, pp. 259–269. New York, NY, USA, ACM (2018). https://doi.org/10.5281/zenodo.1341421
18. Luu, L., Chu, D.-H., Olickel, H., Saxena, P., Hobor, A.: Making smart contracts smarter. In: Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security, ser. CCS '16, pp. 254–269. New York, NY, USA, ACM (2016)
19. Gao, J., Liu, H., Li, Y., Liu, C., Yang, Z., Li, Q., Guan, Z., Chen, Z.: Towards automated testing of blockchain-based decentralized applications. In: 2019 IEEE/ACM 27th International Conference on Program Comprehension (ICPC), May 2019, pp. 294–299. https://doi.org/10.1109/ICPC.2019.00048 (2019)
20. Hasanova, H., Baek, U.-j., Shin, M.-g., Cho, K., Kim, M.-S.: A survey on blockchain cybersecurity vulnerabilities and possible countermeasures. Int. J. Netw. Manag. **29**(2), e2060 (2019), e2060 NEM-18-0162.R1. https://doi.org/10.1002/nem.2060 (2019)

A Comprehensive Study of 5th Generation Scheduling Algorithms

Pardeep Kumar(✉) and Parvinder Singh

Department of CSE, DCRUST, Murthal, India
{19001901907pardeep,parvindersingh.cse}@dcrustm.org

Abstract. Now a days, the 5th-Generation communication standard for Standalone and Non-Stand-alone mode has been released in "Release 15" to "Release 18". This review paper is written to highlight all the key points of 5th generation communication technology covers core concepts of 5th generation communication network, its architecture for both Standalone and Non-Stand-alone mode, different existing scheduling algorithms used in 5G namely Proportional Fair (PF), Modified least weighted delay first (MLWDF), Exponential Proportional Fair (EX-PF), Frame Level Schedular (FLS), Round Robin (RR) and a detailed review of proposed scheduling algorithms anticipated by different researchers. This article is covering four different sections.

Keywords: 5G · 3GPP · IoT · RB · Scheduling · FEMATO

1 Core of 5th Generation

Data usage is predicted to increase by 30% by 2022, which current technologies such as 3G and 4G will be unable to support. So, the need of the coming generation is a very robust technology which can satisfy the requirements of the users. The 3GPP is now ready with its complete release about 5th generation technology and the world is going to implement 5G communication system. In the coming years the 5th generation is projected to be available throughout the world. Now a days every area like industry, banking sector, ecommerce companies, medical industry, educational institutions and research and development section are demanding for a very high-speed communication system to increase the speed of working. The other aspects like cloud computing and Internet of Things (IoT) will get a boost from 5th generation technology. The 5th generation is a very suitable package to implement IoT [15, 16].

1.1 Requirements for 5G are as Follows

- The 5G provides very high speed as compare to existing LTE and LTE advance technology.
- The 5G provides data rate 5 Gbps in downlink and 750 Mbps in uplink.
- 5G Supports a large number of devices as compare to 4G.

R. Kountchev et al. (Eds.): *NAMSP 2022*, SIST 332, vol. 332,, pp. 188–201, 2023.
https://doi.org/10.1007/978-981-19-7842-5_17

- Battery life is pretty good.
- Based on almost 0 latency. Provide very less latency as compare to 4th generation network.
- One band out of three is totally ready to support Internet of Things.
- Very useful in AI/ML to design automated vehicles like driverless cars etc.

The 3G and 4G were using CDMA and OFDM as the modulation techniques respectively. The CDMA was used in 3rd generation of communication system, but the main disadvantages of CDMA technique was high Intercell Interference abbreviated as ISI, battery drainage was high and high latency. As the time changes and the world moved to next generation known as 4G or LTE (Long Term Evolution). The OFDMA (Orthogonal Frequency Division Multiple Access) was used as modulation technique. This modulation was very robust as compare to the older one and having more advantages over CDMA like low Intercell Interference, less Signal to Noise Ratio (SINR), Ease of Installation, high speed for uplink and downlink and orthogonality to reduce Inter Channel Interference (ICI). But there was a big problem in OFDMA that was the spectral efficiency was not up to the mark. The OFDMA uses Cyclic Prefix (CP) to minimize the Inter Symbol Interference (ISI). That CP was actually a frequency gap between the symbols which causes loss of frequency spectrum. In 4th generation the total loss of bandwidth was around 10% due to CP. But a new modulation technique was invented by different researchers named as FBMA (Filter Band Multi Carrier) which do not uses the Cyclic Prefix (CP) to minimize ISI. So, when both techniques are implemented, the conclusion was FBMC is better than because it provides 10% extra spectral efficiency as compare to OFDMA. Now a days FBMC technique is being used in 5th generation technologies [6, 14–16].

1.2 Services of 5G

Speedy Network: The LTE provides data rate 250 Mbps in downlink and 100 Mbps in uplink, which was a very good speed. But now a days we are moving towards a new generation which require more speed. Upcoming years are of cloud, DL and IoT where speed in Gbps will be required. So 5G will provide downlink speed 10 Gbps and 1 Gbps for uplink [4, 16].

Ultrareliable service in dense areas: 5G is very suitable communication technology to implement IoT. In the coming years according to the demand of users there will be a huge crowd who will use the telecommunication services. So, the 5G is capable to provide a better connectivity to all the user by avoiding denial of service in the highly dense areas like malls, airports etc. [4, 15].

Serving the Distant areas: 5G provides better connectivity for the remote areas. The concept of FEMATO cell is also a very good option in LTE and 5G to provide service in the distant locations, rural areas. Other services like noting down the meter reading of distant places to produce the expense bills, e-health service, the concept of wirelessly connected city, and live CCTV surveillance of the area are possible with this generation communication technology [5, 16].

Amalgamation of Lower power consumption devices: Even though 4G already supports a very big number of low-power user equipments, it still falls short for several applications. As a result, 5G aims to accommodate a huge number of low-power devices, which will be smoothly integrated into commercial 5G mobile networks [15, 16].

Huge/More Capacity: The network traffic is predicted to be increase by 50% after 2021 because IoT is rapidly growing with 5th generation communication technology; thus, the 5th generation communication technology must be able to support the huge number of users while maintaining the highest possible quality of service among the users [10, 11].

Reliable Handover: Handover refers to the transfer of a call from one network to another or within the same network's cell. The current handover scenario is highly problematic because the latency experienced during the transfer is significant, resulting in call drop. As a result, in 5G, an intelligent handover with the shortest possible time is expected during network switching [16].

2 Well Defined Architecture of 5th Generation

2.1 Fifth Generation Works in Two Modes

SA: Standalone Mode: In this mode 5G have its own radio and core and works in a standalone mode. The 3GPP release 16 and onward releases provides this working environment by providing different access and core network.

NSA: Non-Standalone Mode: It is backward compatible with LTE (4G). Uses control signaling of 5th generation networks with the core of 4th generation networks. The 3GPP release 15 provides this mode of working (Fig. 1).

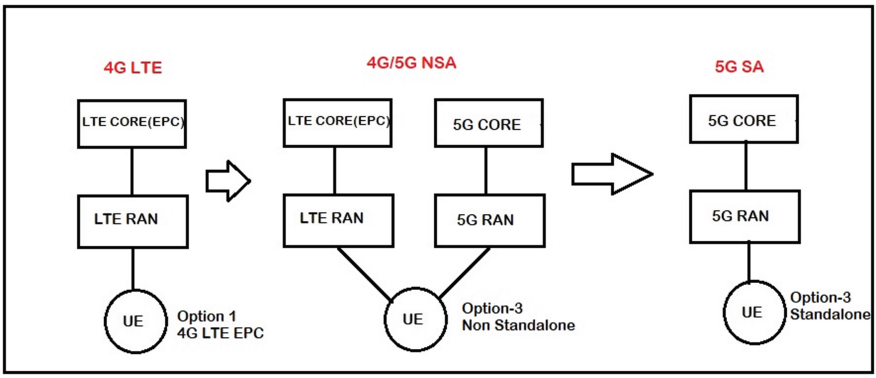


Fig. 1. 5G Architecture (https://www.digi.com/blog/post/5g-network-architecture)

There are three kinds of user applications supported by 5th generation communication technology which are as follows:

- enhanced Mobile Broadband (eMBB)
- enhanced Machine Type Communication (eMTC)
- Ultra-Reliable Low Technologies (URLCC)

The base station which is known as gNodeB in 5th generation technology have a scheduler which is totally responsible to assign the resources to the users connected with that base station under Time Division Duplexing (TDD) or Frequency Division Duplexing (FDD) modulation techniques. A base station allocates the available Resource Blocks (RBs) through TDD over time axis or allocate frequency sub carriers through FDD over the frequency axis to enable diverse 5G use-cases.

The eMBB is a high-speed broadband service which supports a high data rate in both downlink and uplink. So, it requires 100 MHz frequency in the frequency domain and at least 500 Resource Blocks over the time domain to support the proposed speed.

Latency and reliability are crucial for URLLC. It has a moderate bandwidth of up to 5 MHz and a 25-symbol RB sequence for FDD and TDD, respectively. In the frequency and time domains, eMTC uses 1.4 MHz bandwidth and six RBs, respectively. [13, 15, 16].

3 Scheduling Mechanisms and Existing Scheduling Algorithms/techniques

The Radio resource management (RRM) is a management strategy for allocating resources to user equipment (UEs). The RRM enhance the system performance by ensuring that resources are distributed fairly among all users. The scheduling is the mechanism to allocate the resources among all the users.

We apply dynamic scheduling approaches in the fifth generation to make efficient use of resources in both the time and frequency domains. The gNodeB/base station is where these schedulers are deployed. At each Transmission Time Interval (TTI) which is generally 0.5ms in 5th generation networks, the scheduler takes a resource allocation decision for every coming Transmission Time Interval (TTI) and delivers the resources allocation details to the mobile equipment via the PDCCH link. At each TTI, the User/Mobile Equipment computes the CQI value and delivers the Channel Quality Information to the base station's deployed scheduler. This information is called here the information of physical layer. Following receipt of CQI data from UE, the resource schedule makes an allocation decision based on the data and populates the resource block allocation masks. The most appropriate modulation and coding scheme (MCS) is then chosen, which the adaptive modulation and coding (AMC) module will utilize for data transmission. PDCCH provides this information to the user equipment, which includes the assigned resource block as well as the selected modulation and coding scheme. After getting PDCCH payload, the UE learns how to access the physical downlink shared channel (PDSCH) [9, 12, 18, 19].

In this case the QOS (Quality of Service) parameters are not considered. The UE computes the channel quality indicator (CQI) information periodically and transfers it to the gNodeB. The scheduler investigates the CQI information and then estimate the quality of the channel experienced by each user. The scheduler can evaluate the system's feasible throughput once this is known.

3.1 There are Two Types of Scheduling

- Downlink Scheduling
- Uplink Scheduling

Downlink Scheduling: In downlink scheduling the radio resources are scheduled from the base station (gNodeB) to the device also known as equipment.

The downlink scheduling is further divided in 2 types.

- Dynamic Scheduling
- SPS (Semi Persistent Scheduling)

Dynamic Scheduling: In this scheduling the Downlink Control Information is completely responsible to schedule every Physical Downlink Shared Channel on the behalf of available information. [27].

Semi Persistent Scheduling: In this scheduling the Radio Resource Control Information is completely responsible to schedule every Physical Downlink Shared Channel on the behalf of available information. [27].

Uplink Scheduling: In downlink scheduling the User Equipment deals with gNodeB via Physical Uplink Shared Channel.

The uplink scheduling is further divided in 2 types.

- Dynamic Scheduling
- CP (Configured Scheduling)

Dynamic Scheduling: In this scheduling the Downlink Control Information is completely responsible to schedule every Physical Uplink Shared Channel. Decisions are taken over the collected information. [27].

The MAC sublayer's scheduler performs the scheduling algorithms, which determine what will be delivered, when it will be sent, and to or by whom the data will be sent. The scheduler accepts data from a variety of sources. The scheduler has a number of queues for various types of data.

PF Algorithm: Proportional Fairness is abbreviated as PF. This algorithm is not latency sensitive because it is required to transmit non-real time data. This algorithm is a broad time-frequency shared multi-user system algorithm. This is used in 5G to by using Time domain system and Frequency domain system with OFDMA and FBMC modulation techniques. That mixture of TDS and FDS provides a good fairness between fairness of resources and the system throughput among all User equipments. The metric for this algorithm is calculated by combining MT and BET and may be written as [8, 19, 23]:

$$m_{i,k}^{PF} = m_{i,k}^{MT} \cdot m_{i,k}^{BET} = \frac{d_k^i(t)}{\overline{R^i}(t-1)}. \tag{1}$$

M-LWDF Algorithm: (M-LWDF) is an acronym for Modified Largest Weighted Delay First. This scheduling algorithm is proposed to provide delay sensitive data transmission. We can say if we have a real time data to transfer then this algorithm is much suitable for that. At a same time, a large number of users equipments can use this technique for using real time services, each with its own set of quality standards. This algorithm creates a balance between the weighted delay of packet which are going to transfer while considering the efficiency of channel with a priority. The metric is computed as follows for this algorithm [5, 19]:

$$m_{i,k}^{M-LDWF} = \alpha_i D_{HOL,i} \cdot m_{i,k}^{PF} = \alpha_i D_{HOL,i} \cdot \frac{d_k^i(t)}{\overline{R^i}(t-1)}. \tag{2}$$

EXP-PF Algorithm: EXP-PF is a short form of Exponential Proportional Fairness. As the name suggest EXP-PF this technique is proposed by mixing the quality of both Proportional fair algorithm and Exponential function. As discussed earlier the Proportional fair algorithm is most suitable for the non-real time or non-delay sensitive data and the exponential function is used for the real time or delay sensitive data. So, this technique is a full package to deal with both delay sensitive and non-delay sensitive time data. This technique allows the users to transmit multimedia data to be sent using Adaptive Coding Modulation (ACM) and Time Division Multiplexing (TDM) method. This is a combo technique so here both type of users can work: real-time and non-real-time users. This method distinguishes between best effort and real-time data flow. However, a prioritizing system is also employed in this technique. This prioritizing strategy gives real-time service flows a greater priority than non-real-time service flows. The real-time flow measure is computed as follows in this algorithm [13, 19, 23]:

$$m_{i,k}^{EXP-PF} = \exp\left(\frac{\alpha_i D_{HOL,i} - \mathrm{X}}{1+\sqrt{X}}\right) \cdot \frac{d_k^i(t)}{\overline{R^i}(t-1)}. \tag{3}$$

where

$$X = \frac{1}{N_{rt}} \sum_{i=1}^{N_{rt}} \alpha_i D_{HOL,i}. \tag{4}$$

Note: (X represents Chi).

FLS: Frame Level Scheduler. This Schedular is a combination of two scheduling mechanisms stacked over each other. These two stacked techniques are Upper-Level algorithm and lower-level mechanism (Fig. 2).

The FLS is also a delay sensitive algorithm. The FLS is a combination of Delay sensitive algorithm (based upon discreate time linear control) and Proportional fair algorithm. The upper-level algorithm is based upon discrete time linear control which is designed to compute the actual traffic that can be transmitted by the delay sensitive users in a PDU known as frame. The main purpose of framing is to satisfy the delay constraint in this mechanism. The lower level, the PF (Proportional Fair) algorithm is implemented

Fig. 2. FLS schedular

which provides maximum throughput of the system as well as ensure that all the user equipments are feeded with a fair share of the resources [13, 19, 23].

$$u_i(k) = h_i(k) * q_i(k) \tag{5}$$

RR (Round Robin): The RR stands for Round Robin. The round robin algorithm works on preemptive strategy. In this algorithm a time span is fixed for which the RRM allocate the resource to the user. The allocation of resources is done on the behalf of FIFO (First in First Out) Scenario. If a user puts up a request for the resources if the schedular is free and having resources will assign the resource to the requesting UE. This process is repeated among all the user with the same priority. But if a resource is allocated to a user and 2nd user is in wait state then after the completion of time span of user1 the resources are preempted from user 1 and will be assigned to user 2. That process is called switching. The schedular performs this switching in a very speedy manner that the users can't determine that they are working in a RR scenario. This algorithm provides a very good fairness of resources among the users but the main drawback of this algorithm is that this algorithm can't perform its 100% in case of multiuser and high demand scenario [23, 24].

Best CQI Scheduling: As the name suggesting Best CQI scheduling, the scheduling technique which takes CQI (Channel Quality Indicator) in consideration to allocate the resources. As we know when a UE (User Equipment) requests for the resources, that UE computes its CQI at every TTI (Transmission Time Interval) and share that CQI to the gNB. On the behalf of that CQI the schedular allocates the resources to UE. But this scheduling algorithm allocates the recourses to those UE's whose CQI value is highest or very good. That means this algorithm is more suitable for the Cell Center Users.

Because those users who are at cell center always gives perfect or very good CQI value. But those users who are residing at cell edge does not produce good CQI so that scheduling scheme do not consider those requests whose CQI value is less than a threshold value [23, 25].

4 Overview to Related Work

This section covers a review of state-of-the-art scheduling algorithms given for 5th generation networks. The research gaps are traced out for a good future work from the existing work.

4.1 Downlink Scheduling and Resource Allocation for 5G MIMO-Multicarrier: OFDM Versus FBMC/OQAM

Femenias *et al.* Proposed a unifying intersecting layer frame work that have the downlink scheduling and resource allocation (SRA) procedure or algorithm which is builds upon a model of queuing process at an abstraction of data link and physical layer. That procedure can be implemented using any of the modulation technique namely Orthogonal Frequency Division Multiplexing (OFDM) or Filter Bank Multicarrier/Offset quadrature amplitude modulation (FBMC/OQAM). Then a comprehensive comparison is carried out for both modulation formats over different parameters such as Goodput, latency, fairness index and area coverage. Finally, they have concluded that FBMC/OQAM provides better results for all the parameters and especially in spectral efficiency which is improved by 9% (approx.) in FBMC/OQAM in the compression of OFDMA.

The usage of LTE/LTE-Advanced characteristics, as well as the use of an LTE model for simulation, are both responsible for the gap in this research. For instance, a 2×2 MIMO arrangement with a 10 MHz system bandwidth has been deployed. Massive MIMO, which is a configuration of 16×16 and beyond, is present in 5G. 5G uses a minimum bandwidth of 50 MHz. In addition, three scheduling methods, PF, Exponential, and M-LWDF, were examined in this study. The other mechanisms mainly channel-aware algorithms, might be examined, and the comparative analysis will be carried out. The existing access methods can be implemented over OFDM modulation technique by also considering the other parameters which impacts the performance. [14, 26].

4.2 A Novel Scheduling Technique for Improving Cell Edge Performance in 4G/5g Systems

In the cellular networks the users which are near to the cell edge are always suffer from signal problem and interference. At cell edge there is a low SINR because the user is far from the base station or gNodeB. For that Afifi *et al.* [3] developed a novel scheduling algorithm to overcome the problem of cell edge user by providing a good number of resources available at gNodeB. This algorithm provides an ample number of resources to the users that's why they can achieve a good throughput. The proposed technique is an extension of Proportional Fair (PF) algorithm. It is based upon to replace a cell center user by cell edge user at some resources, probabilistically. The cell edge users who will be benefitted by this approach are classified in to a new category of users known as FU's (Featured Users). The proposed technique uses two levels of scheduling. In first level Proportional Fair (PF) scheduling algorithm is applied to check out the proportion of throughput and fairness. Because PF algorithm give result that how much fairly the user is being using the resources and what is the outcome/throughput. After getting the fairness and throughput report from the first level then the proposed algorithm will

decide whether it has to allocate that Resource Block to the chosen PF user to the FU's. The novel scheduling algorithms selects the FU to utilize resource block based on the ratio between the current instantaneous and the accumulated average by the user. After implementation the proposed algorithm has compared with state-of-the-art schedulers and found a conclusion that the proposed algorithm leads to a significant improvement in the average throughput in context of cell edge users. On the other side proposed algorithm results in negligible performance degradation for the cell enter users. [3].

4.3 Inter-cellular Schedular for 5G Wireless Networks

The enhancement of quality-of-service in mobile communication networks is a very typical issue. For that some major inter cell scheduling works are studied by Gueguen *et al.* to propose a new inter-cell scheduling technique named as Inter-cellular Bandwidth Fair Sharing Scheduler (IBFS). The proposed technique supports efficient multimedia services in multiuser 5G environment & balance the load of overloaded cell to improve their quality-of-service. The technique uses a very advance and new resource allocation strategy which is based on a parameter called Mean Cell Delay Outage Ratio (MCDOR). IBFS schedular is having 2 versions:

- IBFSload: Defines the load of a cell means how much data the cell is transmitting and receiving.
- IBFSMCPDOR: Defines that cell which have highest Mean Cell Delay Outage Ratio (MCDOR).

Algorithm

Step 1: Waiting for the scheduling time.

Step 2: For each cell i, update the value of MCPDORi (ith cell MCPDOR Value) for IBFSMCPDOR module, and in the same way update the load of the respective ith cell in IBFSload module.

Step 3: Now in this step on the basic of step 2 values select the cell r (receiver cell) that has highest value of MCPDORi in IBFSMCPDOR and respectively select the highest load value from IBFSload.

Step 4: Find out the neighbors of r cell.

Step 5: For every group g and for every cell k of this group g, which are neighbors of cell r. Here find out the donner cell dg which have highest portion Wg®(k).

Step 6: check is Wg®(dg) = 0 or not.

If 0: No go to step 7.

Step 7: Go to leader donor restitution phase: Restitution form Wg®(dg) a total amount of frequency equal to Qtransfer to the cell r and to all cells gr neighboring dg.

Step 8: Go to other donor cell restitution phase: For every cell neighboring cell r and in the similar group of frequency band dg.

Else Yes.

Step 9: check Wg(k)(k) - Qtransfer > Wmin.

If Yes go to step 10

Step 10: Transfer Phase: Transfer from Wg(k)(k) equal frequency to Qtransfer to cell r and to all the cells gr neighboring k.

Else No go to step 10:

Step 10: End of scheduling and go to step 1.

This proposed scheduling algorithm when implemented and compared with the existing algorithm, it performed very well and produced good results in different-different scenarios. [22].

4.4 A Flexible Scheduling Algorithm for the 5th Generation Networks

From the 15th release to the subsequent releases, the 5G wireless mobile communication standards have been released in stand-alone and non-standalone modes. However, developing an effective and good radio resource management strategy or algorithms to improve network quality of service is a difficult task. Proportional Fair (PF) algorithms are used in current 5G scenarios. While distributing resources and throughput, the PF algorithms consider fairness while assigning resources. This approach focuses on the flow rate as a metric. In this paper Lanlan *et al.* analyzed existing methodologies and presented a new strategy to boost the network's total throughput. The packet delay is used as another priority metric in this technique. Because the existing PF scheduler does not support real-time data transmission, this technique is only suited for non-real-time data transfer. Where latency is a sensitive problem, the delay plays a critical role in transferring time-bound data. The new model allocates resources based on the value of the Channel Quality Index (CQI). The computed CQI value (formulated by the formula) is based on the existing and newly considered parameters supplied back to gNodeB by the UE, and the schedule decides to distribute resources to the user on the basis of that value.

$$\frac{f1(sinri)}{avg\_rate} + \frac{f2(packet_{delay}k - 1)}{DealyTh - f3(packet_{delayk} - 1)} \tag{6}$$

where f1(.), f2(.) and f3(.) are the adjustable functions which are able to adapt the network performance requirements. These functions could be either linear or nonlinear. These functions are referred to as liner functions in the current simulation. In comparison to existing methods or strategies, the simulation yields a fairly good result [2].

4.5 Effective 5G Wireless Downlink Scheduling and Resource Allocation in Cyber-Physical Systems

Vora *et al.* [7] proposed a dynamic programming algorithm with polynomial time complexity for effective downlink scheduling and resource allocation (SRA) by taking channel state and queue state for the fairness index. In this research the SRA is extended for

5G services namely eMTC, URLLC & eMBB. Then the results of SRA and extended SRA (the proposed advanced greedy cross layer algorithm) are compared and concluded that SRA method outperforms the greedy approach by up to 17.24%, 18.1%, 2.5% and 1.5% in terms of average goodput, correlation impact, goodput fairness and delay fairness, respectively. In the case of LTE, that advanced approach outperforms the greedy method by 60%, 2.6% and 1.6% in terms of goodput, goodput fairness and delay fairness compared with tested baseline.

After study this paper as per my observations the main gap in this research the use of LTE tool box. This is a novel technique implemented by the authors is compared with only two scheduling algorithms. To make the results more effective the comparative analysis can be done with many more existing algorithms with more different QoS parameters [11, 20].

4.6 Energy Efficient Scheduling of Small Cells in 5G: A Meta Heuristic Approach

The 5th generation mobile communication technology provides facility to break the large cell into small cells in the ultra-dense areas. Where a large number of devices are working at the same time. So, to provide a very good throughput and QOS (quality of service) the small cells are deputed with a separate small gNodeB/base station which provide a fair scheduling of radio resources among the users. The use of small cells can generate intercell interference. So, to minimize this factor of interference the energy level of singles is considered. The minimization of energy of the signals can be a hurdle between the users and the quality of service. While minimizing the energy level of signals the other side effect is that the users are rapidly increasing and they are demanding a high data rate. But with low energy level signals this is not possible. So, the energy saving approaches are limited while the survey. The Shavo *et al.* Proposed an approach for the scheduling in the small cell by considering the energy levels as the major factor. The authors considered the problem as a non-linear optimization problem and provided a solution as an algorithm which solves the task in a polynomial time. The proposed algorithm is based upon the genetic algorithm. Here the algorithm is divided into 6 sections.

- Defining Chromosome.
- Initializing population and choosing initial solution.
- Calculating fitness value.
- Selecting parent chromosomes.
- Crossover and mutation operation.
- Meta heuristic scheduling.

After the implementation of the proposed algorithm and the results are compared with other state of art algorithms proposed by other researchers known as All Cell ON, ON OFF, Sleep Awake etc. and the performance is evaluated on the behalf of these 6 metrics:

- Activated Cells: The cells who are active and need to be served for completing user's demand.

- Average Energy Expenditure: it defines how much energy is consumed by the small cell in a macro cell.
- Average Switching power: The average time taken by a cell to change its state.
- Aggregated switching delay: The sum of all the small cells switching time from OFF to ON state.
- Average Energy Saving: The average of energy saved from all small cells.
- Average user satisfaction: The %age of those users whose quality of service is perfect [1].

4.7 Green Massive MIMO Scheduling for 5G Traffic with Fairness

The main target of the author in this paper is to design an algorithm which overcomes the trades offs between high energy and bandwidth efficiency, low latency, radio resource fairness and reliability in the wireless networks. Here the author worked on packet scheduling for the downlink channel of massive Multi Input & Multi Output (MIMO) and considered these factors for the implementations. The factors are as follows:

- The packet arrival process.
- The variable packet length.

The proposed algorithm in designed in such a way which is low complex and minimal packet delay. The author's assumptions are that the base station and the user equipments are properly synchronized and working on Time Division Duplexing. Every Time slot is having a scheduling epoch. At the very first epoch the set of users (k) is selected and those users maximize a predefined function. The very first packets also known as HOL (Head of Line) packets are scheduled with the same transmission rate. Every user has its own queue of packets and all the users queue length may be different. So, there is a possibility that packets transfer of some user of set (k) may be complete before other whose packet queue is large. Now the schedular go for the next epoch and find out another user set those will maximize the same predefined function. This process is repeated many times until the time slot get filled with the packets. After that the base station broadcast the resources to the selected users with the time with in the time slot. Like that the packets scheduled are transmitted to the user equipments. [17].

4.8 Survey of Scheduling Schemes in 5G Mobile Communication Systems

Mamode *et al.* reviewed the 5G network architecture in detail, existing scheduling algorithms which are already implemented in 5th generation mobile communication technology and some proposed scheduling algorithms by different researchers based upon the different parameters. The proposed algorithms are studied in such a way that the reader can easily identify the core concepts of the schedulers and their implementations. The research methodology and gaps are also identified during the review which makes other researchers to find out the new opportunity for work [7].

4.9 A Review on Techniques to Improve the Cell Edge Performance for Wireless Network

Mustaffa *et al.* presented a review on those scheduling algorithms which are designed to improve the cell edge performance. Because the normal schedular works on the behalf of equality they provide their 100% to all the users existing the cell. But due to so many factors like Inter-Cell Interference, SINR (Signal to Noise Ratio) etc., the users who are residing at the cell edge do not get the 100% from the base station or schedular. So, the drawback of this seems like call drop, more delay, less spectral efficiency, more packet loss ratio and many more. To overcome these hurdles for the cell edge users the new scheduling approaches are proposed by the researchers which gives more attentions to cell edge users to increase their performance. In this paper the authors studied the papers which are presented after 2015. The approaches like Cooperative Transmission Scheme, Packet Scheduling, Soft Frequency Reuse and Downlink Optimal Power Allocation Scheme who emphasize on the cell edge performance improvement are taken into account [8, 21].

5 Conclusion

In this article a comprehensive study of different scheduling algorithms is carried out. The review is done in different stages. The first stage covers stand-alone, non-standalone architectures and services of 5^{th} generation communication technology. The stage 2 emphasizes on a detailed study of scheduling process and the different scheduling algorithms (existing or proposed). The algorithms work on different parameters like speed, distance, spectral efficiency, signal energy etc. This study points out those areas in scheduling where the improvement is required and the improvement can provide us better results in the form of a new strategy. Our main concern in this study we want to validate the existing and proposed concepts to get something fruitful for the society or researchers.

References

1. Shavo, S., Munna, M., Sarkar, S., Adhikary, T., Razzaque, M., Hassan, M., Aloi, G., Fortino, G.: Energy-efficient scheduling of small cells in 5G: A meta heuristic approach. J. Netw. Comput. Appl. **178**, 1–14 (2021)
2. Li, L., Shao, W., Zhou, X.: A flexible scheduling algorithm for the 5th-generation networks. Intell. Converg. Netw., 101–107 (2021)
3. Afifi, W., El-Moursy, A., Saad, M., Nassar, S., El-Hennaway, H.: A novel scheduling technique for improving cell-edge performance in 4G/5G systems. Ain Shams Eng. J., 1–9 (2020)
4. Mohd R, Huda A.: A Study on packet scheduling algorithms for healthcare contents over fifth generation (5G) mobile cellular network. Int. J. Electron. Telecommun., 729–735 (2020)
5. Abreu, R., Pocovi, G., Jacobsen, T., Centenaro, M., Pedersen, K., Kolding, T.: Scheduling enhancements and performance evaluation of downlink 5G time sensitive communications. IEEE Access, 128106–128115 (2020)
6. Gatti, R., Shankar, S.: Bidirectional Resource Scheduling Algorithm for Advanced Long Term Evolution System. Wiley (2020)

7. Mamo1de, M., Fowdur, T.: Survey of scheduling schemes in 5G mobile communication systems. J. Electr. Eng., Electron., Control. Comput. Sci., 21–30 (2020)
8. Mustaffa, N., Muhamad, W., Anuar, A.: A review on techniques to improve cell edge performance for wireless networks. Intern.Tional J. Adv. Trends Comput. Sci. & Eng., 595–600 (2020)
9. Perdana, D., Sanyoto, A., Bisono, Y.: Performance evaluation and comparison of scheduling algorithms on 5G networks using network simulator. Int. J. Comput. Commun. & Control., 530–539 (2019)
10. Hadar, I., Raviv, L., Leshem, A.: Scheduling for 5G cellular networks with priority and deadline constraints. In: International Conference on the Science of Electrical Engineering (ICSEE) IEEE, pp. 1–5 (2018)
11. Vora, A., Kang, K.: Effective 5G wireless downlink scheduling and resource allocation in Cyber-Physical systems. Technologies (MDPI) (2018)
12. Muller, C., Galaviz, G., Andrade, A., Kaiser, I., Fengler, W.: Evaluation of scheduling Algorithms for 5G mobile systems. In: Springer International Publishing AG, part of Springer Nature, Computer Science and Engineering- Theory and Applications, Studies in Systems, Decision and Control (2018)
13. Garg, K., Mishra, A.: 5G Network Advanced techniques: A Literature Review. Int. J. Eng. Tech. Res. (IJETR) (2018)
14. Femenias, G., Riera-Palou, F., Mestre, A., Olmos, J.: Downlink scheduling and resource allocation for 5G MIMO-multicarrier: OFDM vs FBMC/OQAM. IEEE Access (2017)
15. Kumar, A., Gupta, M.: A review on activities of fifth generation mobile communication system. Alex. Eng. J. (Elsevier) (2017)
16. A White Paper on 5G Architecture. 5GPPP Architecture Working Group, Version-2, pp. 1–140 (2017)
17. Torabzadeh, M.: Green Massive MIMO Scheduling for 5G Traffic with Fairness, pp. 37–42. ICCIS (2017)
18. Kolli, R., Mile, S., Shetty, S., Dixit, S.: Review on 5G wireless technology. Int. J. Adv. Res. Comput. Commun. Eng., 219–223 (2016)
19. Kumar, P., Kumar, S., Dabas, C.: Comparative analysis of downlink scheduling algorithm for a cell affected by interference in LTE networks. Ann. Data Sci., 135–153 (2016)
20. Karimi, A., Pedersen, K., Mahmood, N., Steiner, J., Mogensen, P.: 5G centralized multi-cell scheduling for URLLC: algorithms and system-level performance. IEEE Access, 1–11 (2016)
21. Kela, P., Turkka, J., Costa, M.: Borderless mobility in 5G outdoor ultra-dense networks. IEEE Access, 1462–1476 (2015)
22. Gueguen, C., Ezzaouia, M., Yassin, M.: Inter-cellular scheduler for 5G wireless networks. Elsevier, 113–124 (2015)
23. Gomez, K., Goratti, L., Granelli, F., Rasheed, T.: A comparative study of scheduling disciplines in 5G systems for emergency communications (2014)
24. Hani, U., Samota, K., Singh, S.: A Novel technique of improving 5G networks with round robin scheduling in preamble transmission. IEEE Explor. (2019)
25. Ashfaq, K., Safdar, G., Ur-Rehman, M.: Comparative analysis of scheduling algorithms for radio resource allocation in future communication networks. Peer J. Comput. Sci. (2021)
26. OFDM: httpsmorse.colorado.edu~tlen5510textclasswebch8.html, last accessed 2022/04/06
27. SCHEDULING: https://www.sharetechnote.com/html/5G/5G_Scheduling.html, last accessed 2022/04/22

A Comparative Analysis of Homogeneous and Heterogeneous Protocols to Maximize the Lifetime of Wireless Sensor Network for Precision Agriculture

Neha Neha[1], Jasvinder Kaur[2(✉)], and Banita Banita[3]

[1] PDM University, Bahadurgarh, Haryana, India
[2] PDM University, Bahadurgarh, Haryana, India
jasvinder.kaur@pdm.ac.in
[3] IIAS, Shimla, Himachal Pradesh, India

Abstract. Wireless Sensor Network (WSN) includes a number of Sensor Nodes (SN) which are battery-powered and deployed over a large area in an unattended environment. Various applications of WSN are military, agriculture, monitoring or surveillance etc. As the SN are battery operated and power source cannot be replaced, so to maximize the network lifetime, energy-efficient protocols play a vital role by choosing some low energy path. The use of low energy path repeatedly may lead to exhaustion of battery of SNs in that area that will create energy holes as all the SN are dead in a particular region. So, the protocol must be chosen to prolong the network lifetime by involving all the SNs to participate in data transmission. Clustering proves to be an efficient approach for prolonging the network lifetime. In clustering, the strategy to involve all the SNs will provide efficient and balanced consumption of battery and increase the reliability as well. Various clustering techniques have been proposed over the years but still, there are chances of improvement. Clustering is an efficient approach for precision agriculture because it will provide lots of information to farmers about the field. Heterogeneous protocols are more reliable and energy-efficient than homogeneous protocols. A Heterogeneous Energy Efficient and Reliable Routing (HEERR), which is an advanced version of Distributed Energy Efficient Clustering (DEEC) protocol is proposed and compared with other hierarchical routing techniques. Results revealed that HEERR not only enhanced the network lifetime but also increases the throughput.

Keywords: Wireless · Distributed · Energy Efficient · Clustering · Threshold · Sensor network · Energy Efficient · Hierarchical Routing

R. Kountchev et al. (Eds.): *NAMSP 2022*, SIST 332, vol. 332,, pp. 202–216, 2023.
https://doi.org/10.1007/978-981-19-7842-5_18

1 Introduction

In present, WSN has emerged as a technology in which thousands of nodes are connected to form a large-scale network. WSN consists of devices knows as Sensor Node (SN) that are small, sensing capability, battery-powered, inexpensive, and low-computation capability [1]. These SNs are used for monitoring conditions such as humidity, moisture, temperature, motion, abnormal activities, etc. SNs can be deployed either on the ground or in water to perform applications such as surveillance, intelligence, medical, environmental, and underwater monitoring [2–6]. SN consist of devices such as a microcontroller, sensor, and battery and, radio transceiver [3–10]. SNs are distributed, self-organized, work on dynamic topology [11–16]. The common features of SNs include dynamic topology [17], self-organizing capabilities [18], node mobility [19–22], multi-hop routing [23], broadcasting [24], short-range communication etc. These SNs are connected to form an architecture known as WSN as shown in Fig. 1.

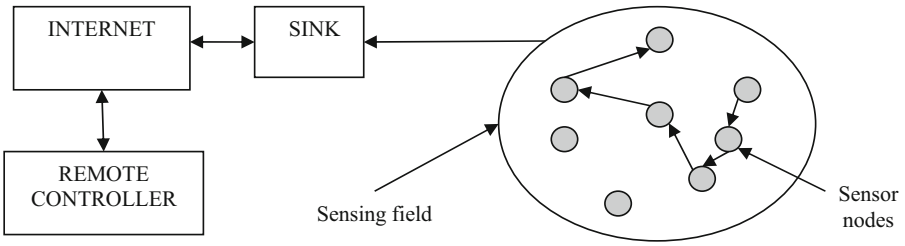


Fig. 1. WSN Architecture

The key challenge is to balance the consumption of batteries and increase reliability [21–30]. But there are other challenges as well that needs to be taken care of. SNs are deployed randomly in an area without any infrastructure and prior knowledge about the topology [27–30]. In these cases, SNs have to self-identify the connectivity and distribution. For example, for surveillance purposes on a battlefield, SNs would be dropped in an area by plane. WSN protocols should be fault-tolerant to accommodate the failure of SN [31–39]. WSN protocols should be dynamic so that they can respond and operate in any number of SNs [17, 36–39]. Protocols must work in such a way that they transmit the data to BS at a specific time to achieve Quality-of-Service (QoS). In this work, we have focused on the application of WSN in agriculture. In agriculture, various data are required periodically as well as critically to monitor the field and crops. Clustering is a good approach to deliver the data from SN to sink in agricultural field. In next section, we have discussed various routing approaches that are using homogeneous i.e. in which initial energy of all the SNs are equal as they are having the same equipment and heterogeneous i.e. in which initial energy of all the SNs are not equal as they are having the different equipment.

2 Clustered Routing Strategies in Wireless Sensor Network

To keep these challenges in mind, researchers worked on various protocols to enhance the network lifetime [1–30]. In clustering, the strategy to involve all the SNs will provide efficient and balanced consumption of battery and increase the reliability as well. The clustering strategy plays an important role for data transmission [25–37]. A few of the routing protocols that are popular in this category are following:

2.1 LEACH Protocol

Low Energy Adaptive Clustering Hierarchy (LEACH) is a hierarchical cluster-based protocol that optimizes the energy consumption through clustering approach. Cluster Leader (CL) will collect that data from neighboring nodes and send to the Base Station (BS).A random selection procedure is used for the election of CL [7] which is given by Eq. (1).

$$\mathrm{Thr}(z) = \begin{cases} \dfrac{\mathrm{Pr}}{1-\mathrm{Pr}.\left[\mathrm{ro}* \bmod \left(\frac{1}{\mathrm{Pr}.}\right)\right]} & \text{if } \in Y \\ 0 & \text{Otherwise} \end{cases} \tag{1}$$

where, Thr: Threshold level

Pr.: Probability.

ro: number of round.

2.2 LEACH-C Protocol

LEACH-C enables low-energy networking in WSN. The only difference between LEACH protocol and LEACH-C protocol is in their Setup phase however the steady-state phase remains ideal in both of them. In LEACH-C cluster formation is performed by the base station (sink).All nodes of the network send their information like location, energy level to the BS [16]. After this, BS calculates the optimal number for SNs that can be CL. Only those SNs can be CL that has sufficient energy. Advantages of this protocol over LEACH are number of CLs in LEACH are not fixed it changes according to round but in LEACH-C BS calculates the number of CLs for every round.

The drawback of LEACH-C is sink requires global knowledge of the network for cluster formation (Fig. 2).

2.3 TEEN Protocol

Threshold-sensitive Energy Efficient sensor Network (TEEN) is a cluster-based hierarchical routing protocol like LEACH i.e. the nodes form clusters and selection of CL for transmission of data to BS. It uses both hierarchical techniques and a data-centric approach. While inside LEACH, absolutely no certain actions are generally driven therefore it is a proactive protocol [17–19].

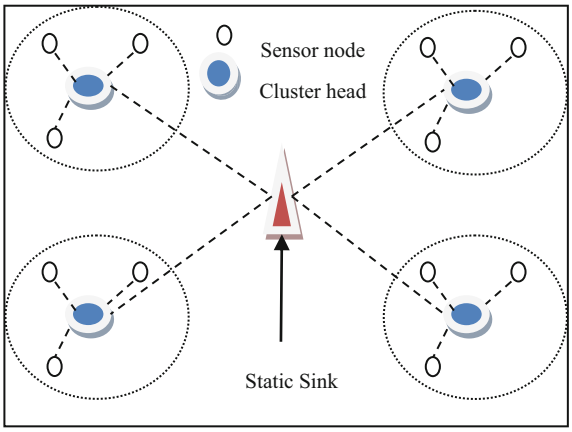


Fig. 2. Centralized Static Sink Approach

2.4 SEP Protocol

Stable Election Protocol (SEP) is a heterogeneous protocol. In heterogeneous protocols the three types of models are used. These models are two levels, three levels, and multilevel heterogeneous model. Discussed routing schemes LEACH, LEACH C, TEEN are advisable only for homogenous sensor networks. For heterogeneity, let Y × Z be the number of advanced SNs where Y is a fraction of the total number of SNs. Let advanced SNs have Z times more energy than the rest of the SNs. So initial energy of each advanced SN in the network is $E_{initial} \times (1 + Z)$. Thus total initial energy of two-level heterogeneous networks could be represented by Eq. (2).

$$E_{total} = E_{initial} \times (1 + Y \times Z) \tag{2}$$

$$Thr(Snl) = \left\{ \begin{array}{cc} \frac{Pnl}{1-Pnl.\left(ro \mod \frac{1}{Pnl}\right)} & if\ Snl \in R' \\ 0 & if\ Snl \in R' \end{array} \right\} \tag{3}$$

$$Thr(Snl) = \left\{ \begin{array}{cc} \frac{Pnl}{1-Pav.\left(ro\ mod \frac{1}{Pav}\right)} & if\ Sav \in R'' \\ 0 & if\ Sav \in R'' \end{array} \right\}. \tag{4}$$

where, pnl is the probability of nodes to become cluster head, ro is the current round, R' is the set of SNs that have not become CLs within the last 1/Pnl rounds, Z" is the set of advanced SNs that have not become CLs within the last 1/*Pav* rounds.

2.5 DEEC Protocol

In Distributed Energy Efficient Clustering (DEEC) [21], higher is the remaining energy, higher will be the chances to become CL. Let $\overline{AvgE}(r)$ denote the average energy at

round rd, which defined as in Eq. (5):

$$\overline{\text{AvgE}}(r) = \frac{\text{total residual energy of all nodes at round ro}}{\text{no of nodes}} \tag{5}$$

$$pi = \left\{ \begin{array}{l} \frac{P_{opt} E_i(r0)}{(1+AM)\bar{E}(r0)} \quad if\ s_i\ is\ the\ normal\ node \\ \frac{P_{opt}(1+AM)E_i(r0)}{(1+AM)\bar{E}(r0)} \quad if\ s_i\ \ is\ \ advanced\ node \end{array} \right\} \tag{6}$$

2.6 EDEEC (Enhanced Distributed Energy Efficient Clustering) Protocol

EDEEC is a modified version of DEEC [23]. EDEEC works as a three-level network based on SNs which are normal, advance, and supernodes. Rest of the work will remain the same apart from the selection of CL which will be done according to Eq. (10).

$$pi = \left\{ \begin{array}{ll} \frac{E_i(rd).P_d}{(1+m'.(ad+b.kd'_o))\bar{E}(rd)} & normal\ node \\ \frac{E_i(rd).P_d(1+ad)}{(1+m'.(ad+b.kd'_o))\bar{E}(rd)} & advanced\ node \\ \frac{E_i(rd).P_d(1+bd)}{(1+m'.(ad+b.kd'_o))\bar{E}(rd)} & super\ node \end{array} \right. \tag{7}$$

where, md' = % of advanced SNs.

P_d = desired probability of CLs.

kd'$_o$ = % of super nodes.

ad = portion of advance SNs.

bd = portion of SNs.

$\bar{E}$(rd) = average energy.

The threshold for CL selection T(CL_j) is given in (8):

$$T(CL_j) = \left\{ \begin{array}{l} \frac{P_d}{\left(1-P_d\left(rd* \bmod \frac{1}{p_d}\right)\right)} \quad if\ p_d \in M' \\ \frac{P_d}{\left(1-P_d\left(rd* \bmod \frac{1}{p_d}\right)\right)} \quad if\ p_d \in M'' \\ \frac{P_d}{\left(1-P_d\left(rd* \bmod \frac{1}{p_d}\right)\right)} \quad if\ p_d \in M''' \end{array} \right. \tag{8}$$

where M', M" &M"' represent group of normal SNs, advanced SNs and super SNs that have not become CLs within the last 1/p_j rounds.

The network average energy can be calculated as:

$$\overline{E}(rd) = \frac{1}{N} E_{total}\left(1 - \frac{rd}{R}\right) \tag{9}$$

R can be calculated as

$$R = \frac{E_{total}}{E_{round}} \tag{10}$$

The total energy of the network E_{total} is calculated by

$$\begin{aligned} E_{total} &= SN \cdot \left(1 - md'\right) \cdot E_i + SN \cdot md' \cdot \left(1 - md'o\right) \cdot (1 + ad) \cdot E_i \\ &+ SN \cdot md' \cdot md'o \cdot E_i \cdot (1 + bd) \\ &= SN \cdot E_i \cdot \left(1 + md' \cdot \left(ad + md'o \cdot bd\right)\right) \end{aligned} \tag{11}$$

where, SN = total number of nodes

E_o = initial energy.

The probability of CL selection for HEERR is given in (12).

$$p_i = \begin{cases} \left(\frac{E_i(rd).P_d}{(1+md' \cdot (ad+bd \cdot md'_o))\bar{E}(rd)}\right) * E_T \ normal\ node \\ \left(\frac{E_i(rd) \cdot P_d(1+a)}{(1+md' \cdot (ad+bd \cdot md'_o))\bar{E}(rd)}\right) * E_T \ advanced\ node \\ \left(\frac{E_i(rd) \cdot P_d(1+b)}{(1+md' \cdot (ad+bd \cdot md'_o))\bar{E}(rd)}\right) * E_T \ super\ node \end{cases} \tag{12}$$

where, E_T is total energy

$$E_{Tx}(B, d) = \begin{cases} B \cdot E_{elec} + B \bullet \in_{fs} \cdot d^2 \ if\ d \leq d_\circ \\ B \cdot E_{elec} + B \cdot \in_{amp} \cdot d^4 \ if\ d > d_\circ \end{cases} \tag{13}$$

Total energy consumed per round is given as,

$$E_{round} = B(2NE_{elec} + NE_{DA} + k \in_{amp} d^4_{toBS} + N \in_{fs} d^2_{toCH}) \tag{14}$$

$$d_{toCH} = \frac{M}{\sqrt{2\pi k}}, d_{toBS} = 0.765\frac{M}{2} \tag{15}$$

$$k = \frac{\sqrt{N}}{\sqrt{2\pi}} \sqrt{\frac{\in_{fs}}{\in_{amp}}} \frac{M}{d^2_{toBS}} \tag{16}$$

where k is the number of clusters.

3 Algorithm for the Simulated Protocols

In this work, various clustered routing protocols have been implemented and following assumption are to be considered.

- The sink will have an unlimited amount of supply.
- The sink position is fixed at the center.
- The SNs are equipped with power control capabilities to vary their transmitted power.
- After a periodic interval, each SN senses the environment and sends the data to CL or BS.
- All SNs are static.

The algorithm for simulated protocols has been shown in Fig. 3. And, the criterion for selecting CL has been shown in Fig. 4.

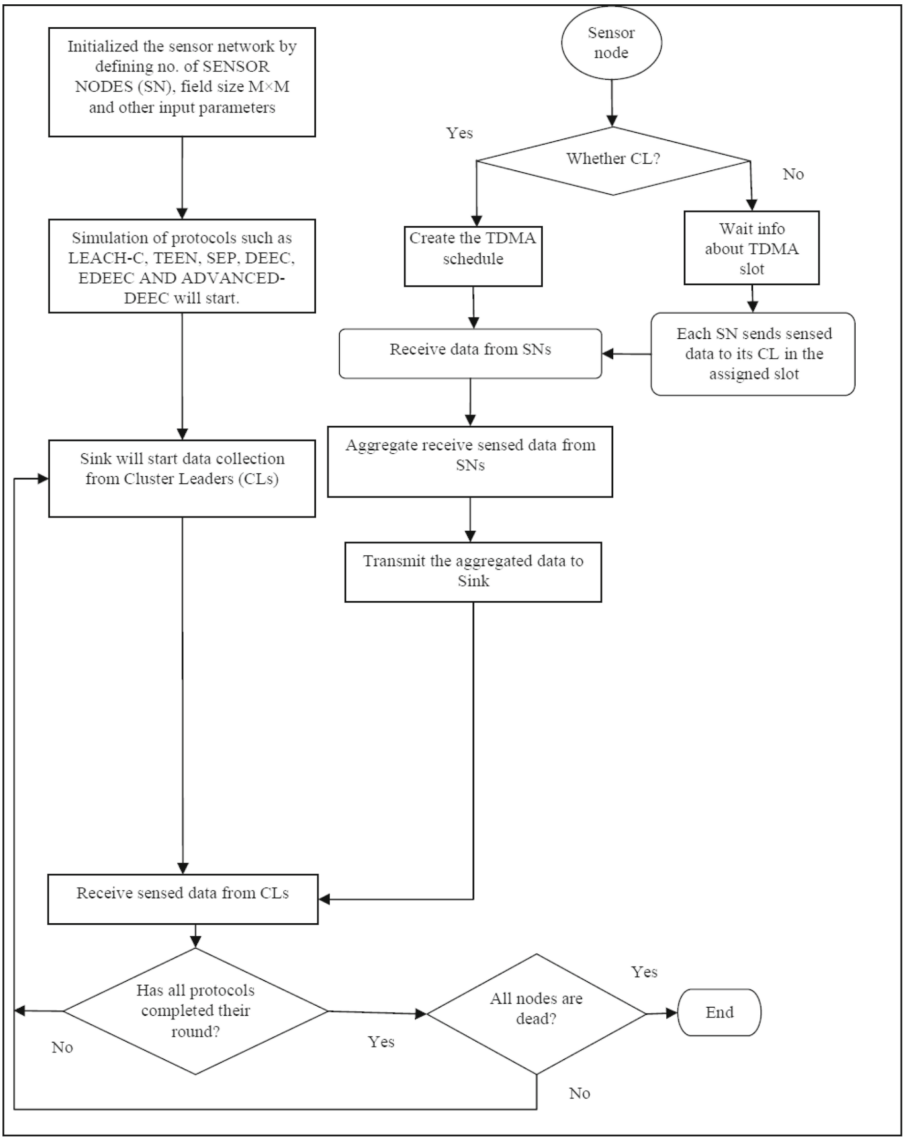


Fig. 3. Framework for simulated protocols

4 Implementation and Results

In this section, simulation and comparison of various protocols LEACH, TEEN, SEP, DEEC and new improved protocols such as E-DEEC and HEERR of the same category have been performed using MATLAB 8.1 based on the parameters like energy efficiency, heterogeneity level, cluster stability, CL selection criteria etc. For this purpose, randomly

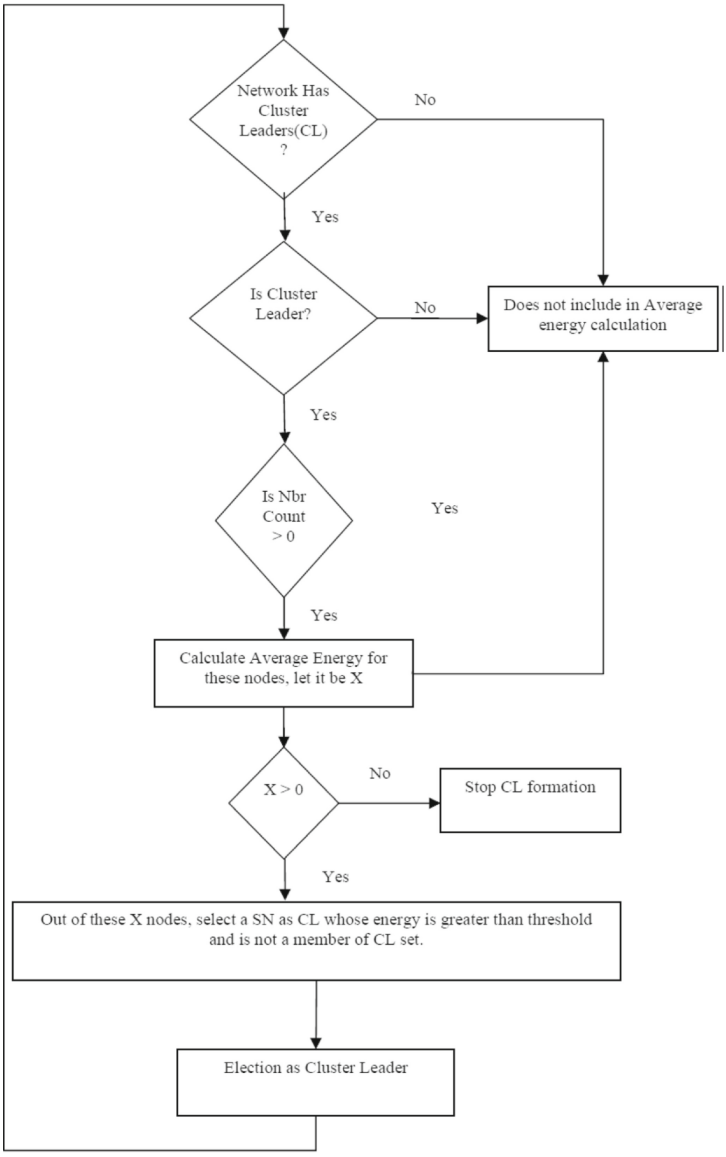


Fig. 4. Cluster Leader Formation by Base Station

distributed WSN consist of 100 SN in a 100 m^2 field is used and centralized BS. Table 1 shows various network parameters.

Figure 5 has shown a plot between SNs alive and the number of rounds of different protocols named LEACH-C, TEEN, DEEC, SEP, E-DEEC,and HEERR Protocols which conclude that in large network area like the agriculture field where we have to include more number of SNs, E-DEEC and HEERR protocols perform well as more number of

Table 1. Network parameters

Parameters	Values
Simulation Area (in meters)	100 ×100
Initial Energy Allotted to SN (in Joules)	0.5
Total no. of SNs	100
E_{TX}	50nJ/bit
E_{RX}	50nJ/bit
E_{DA}	5 nJ/b/message
CL Probability	0.05
Data Packet Size(in bits)	4000
Threshold distance(d_0) (in meters)	87.7
Transmit Amplifier Energy	
E_{FS}	0.0013 pJ/b/m^4
E_{MP}	10pJb/m^2

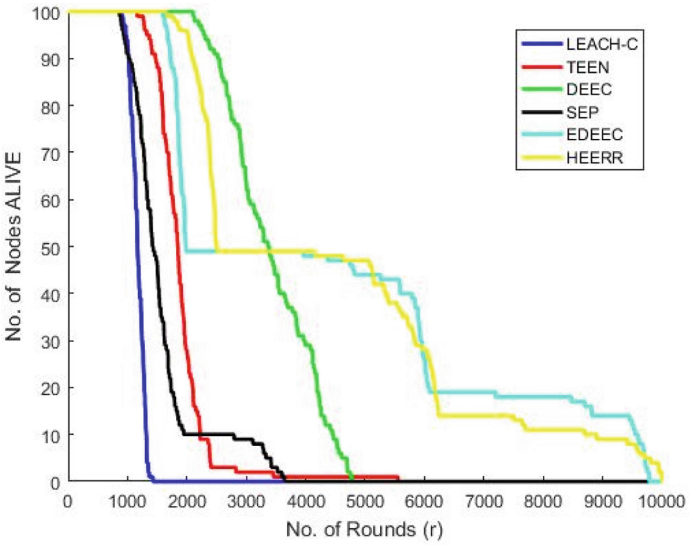


Fig. 5. Comparison of LEACH C, TEEN, DEEC, SEP and Proposed Protocols in terms of nodes alive.

SNs are remain alive at almost all rounds. It has been clear from the figure when all the SNs of other simulated protocols are dead only 50% of SNs are dead for E-DEEC and HEERR.

Refer Fig. 6, it can be observed that E-DEEC and HEERR protocols performs better and showed more stability as compared to other protocols while LEACH-C performs

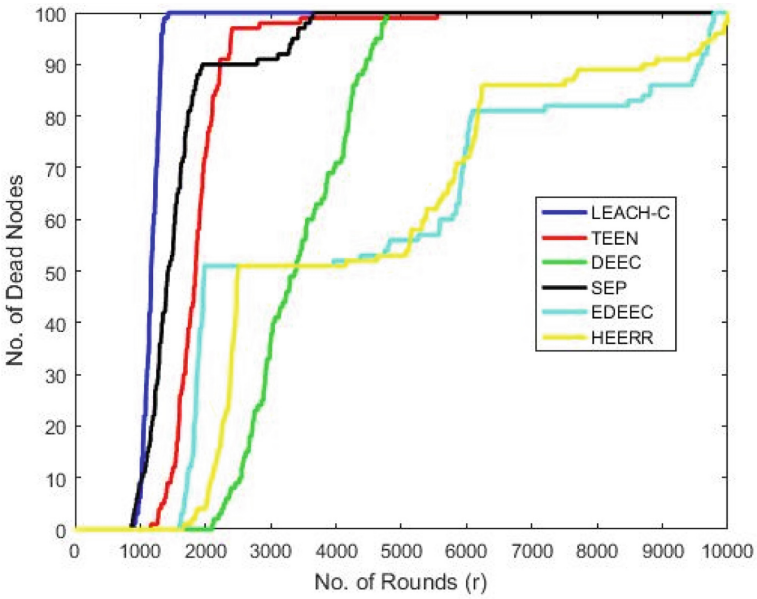


Fig. 6. Comparison of LEACH C, TEEN, DEEC and SEP in terms of nodes dead.

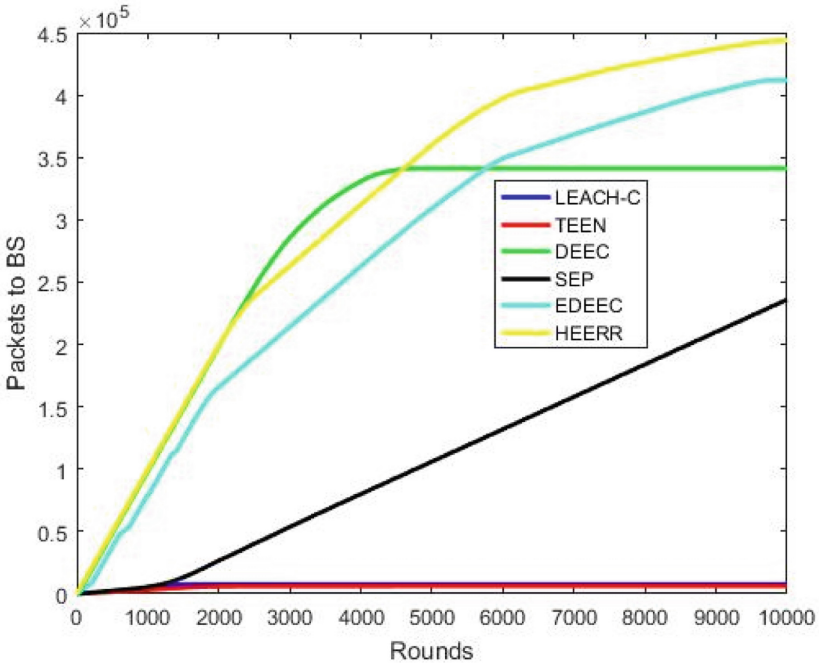


Fig. 7. Comparison of LEACH C, TEEN, DEEC, SEP, E-DEEC and HEERR PROTOCOLS in terms of packets sent to BS

worst. Figure 7 showed information about how many data packets send to sink over the number of rounds. In this case, the HEERR protocol transfers more data from CL to the sink. So PROPOSED strategies are more reliable as compared to LEACH-C, SEP and TEEN. Hierarchical routing protocols have a certain process to choose CHs and have their unique architecture and many other parameters to perform a routing process. Comparison between these protocols based on various parameters like architecture, hop, heterogeneity level, cluster stability etc. as shown in Table 2.

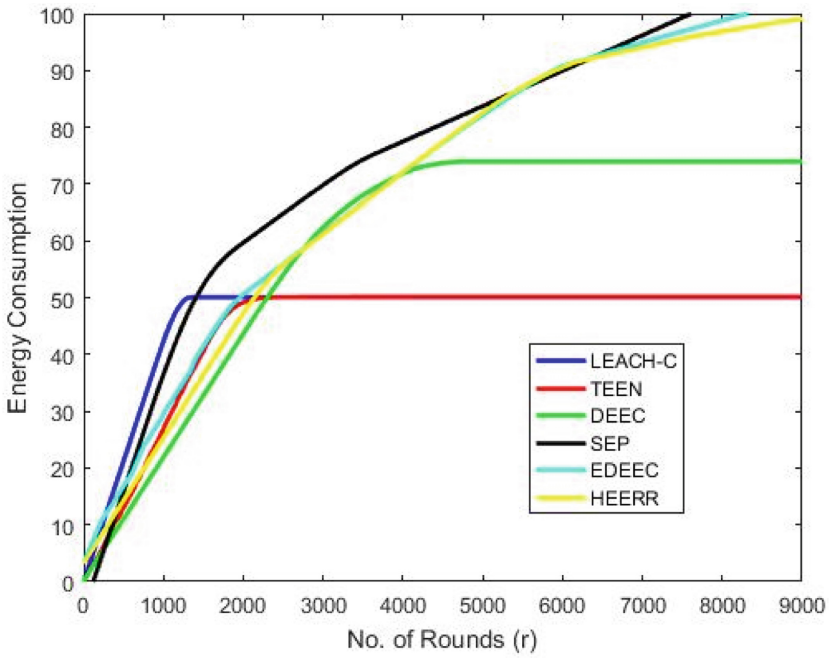


Fig. 8. Comparison of LEACH C, TEEN, DEEC, SEP, E-DEEC and HEERR PROTOCOLS in terms of Energy Consumption

Figure 8 depicts the energy consumption per round in protocols such as LEACH-C, TEEN, DEEC, SEP, EDEEC and HEERR. It can be observed from the figure that heterogeneous protocols has more stability and optimize consumption of battery as compared to homogeneous protocols.

5 Conclusion

Cluster-formation-based routing is the best way to archive energy efficiency goal in hierarchical routing protocols for a large area like agricultural field. In this work, we have simulated and compared protocols such as LEACH-C, TEEN, DEEC, SEP, E-DEEC and HEERR. The performances of such protocols are judged by the simulation results under the various performance metrics. Results showed that E-DEEC and HEERR

Table 2. Comparison of various routing protocol

Performance Criteria	LEACH	LEACH-C	TEEN	SEP	DEEC	E-DEEC	HEERR
Architecture	Distributed	Centralized	Distributed	Distributed	Distributed	Distributed	Distributed
Hop	Single Hop	Single Hop	Multi Hop	Multi Hop	Multi Hop	Multi Hop	Multi Hop
Heterogeneity level	Not present	Not present	Not present	Two level	Multilevel	Multilevel	Multilevel
CL Selection Criterion	Elected rotation-wise by probabilistic approach	Selected by BS w.r.t. nodes energy and distance	Randomly	Based on Initial and Residual Energy	Based on Initial, Residual and Average Energy of the network	Based on Initial, Residual and Average Energy of the network	Based on Initial, Residual and Average Energy of the network
Cluster Stability	Lower	Higher than leach	Very High	Moderate	High	High	High
Global knowledge of network	Not Required	Required	Not Required	Not Required	Not Required	Not Required	Not Required
Energy Efficiency	Very low	Low	Moderate	High	High	High	High

protocols are reliable and energy-efficient than other strategies. Also, it has been cleared that the heterogeneous approach is more reliable and energy efficient as compared to homogeneous approach. From the simulation results, we can conclude that HEERR is reliable because it is sending maximum data packets to the sink as compared to other routing protocols.

References

1. Zeng, D., Dai, Y., Li, F., Sherratt, R.S., Wang, J.: Adversarial learning for distant supervised relation extraction. Comput., Mater. Contin. **55**(1), 121–136 (2018)
2. Shafiq, M., Ashraf, H., Ullah, A., Tahira, S.: Systematic literature review on energy efficient routing schemes in WSN-A survey. Mob. Netw. Appl. **16**, 1–4 (2020)
3. Nakas, C., Kandris, D., Visvardis, G.: Energy efficient routing in wireless sensor networks: a comprehensive survey. Algorithms. **13**(3), 72 (2020)
4. Pan, J.S., Kong, L., Sung, T.W., Tsai, P.W., Snášel, V.: α-Fraction first strategy for hierarchical model in wireless sensor networks. J. Internet Technol. **19**(6), 1717–1726 (2018)
5. Khatri, P., Karunesh K., Gupta R., Panchariya P.: Towards the green analytics: design and development of sustainable drinking water quality monitoring system for Shekhawati region in Rajasthan. MAPAN, 1–15 (2021)
6. Jain, A., Goel, A.K.: Energy efficient fuzzy routing protocol for wireless sensor networks. Wirel. Pers. Commun. **110**(3), 1459–74 (2020)
7. Mishra, D., Sethi, D., Bhattacharya, P.: Modeling and Simulation of A Clustered WSN for Precision Agriculture, Int. J. Comput. Sci., Inf. Technol. Control. Eng. (IJCSITCE) **3**(1/2) (2016)
8. Anand, J., Sethi, D.: Comparative analysis of energy efficient routing in WBAN.In: 2017 3rd International Conference on Computational Intelligence & Communication Technology (CICT), pp. 1–6. https://doi.org/10.1109/CIACT.2017.7977373. (2017)
9. Wang, J., Gao, Y., Yin, X., Li, F. and Kim, H.J.: An enhanced PEGASIS algorithm with mobile sink support for wireless sensor networks. Wirel. Commun. Mob. Comput. (2018)
10. Yao, J., Zhang, K., Yang, Y., Wang, J.: Emergency vehicle route oriented signal coordinated control model with two-level programming. Soft. Comput. **22**(13), 4283–4294 (2017). https://doi.org/10.1007/s00500-017-2826-x
11. Ren, Y., Liu, Y., Ji, S., Sangaiah, A.K., Wang, J.: Incentive mechanism of data storage based on blockchain for wireless sensor networks. Mob. Inf. Syst. (2018)
12. Sethi, S., Bhattacharya, P.P.: A study on energy efficient and reliable data transfer (EERDT) protocol for WBAN. In: Second International Conference on Computational Intelligence & Communication Technology (CICT), pp. 254–258 (2016). https://doi.org/10.1109/CICT.2016.57
13. Sethi, D., Bhattacharya, P.: Revised Multi-Chain PEGASIS for Wireless Sensor Networks. Int. J. Sens.S, Wirel. Commun. Control. **6**(1), 12–17 (2016)
14. Khan, W.A., Abdullah, A., Anisi, M., Javed, I.B.: A comprehensive study of data collection schemes using mobile sinks in wireless sensor networks. Sensors **14**(2), 2510–2548 (2014)
15. Sethi, D., Bhattacharya, P.: A comparative analysis of various mobile sink routing protocols and performance comparison of clustered routing protocols in mobile sink scenario. Majlesi J. Electr. Eng. **12**(3), 11–22 (2018)
16. Zhao, M., Yang, Y., Wang, C.: Mobile data gathering with load balanced clustering and dual data uploading in wireless sensor networks. IEEE Trans. Mob. Comput. **14**(4), 770–785 (2014)

17. Zhu, C., Shuai, W., Han, G., Shu, L., Hongyi, W.: A tree-cluster-based data-gathering algorithm for industrial WSNs with a mobile sink. IEEE Access **3**, 381–396 (2015)
18. Xie, G., Pan, F.: Cluster-based routing for the mobile sink in wireless sensor networks with obstacles. IEEE Access **4**, 2019–2028 (2016)
19. Kumari, Ch Usha, and Tatiparti Padma, Energy-efficient routing protocols for wireless sensor networks. Soft Comput. Signal Process., 377–384. Springer, Singapore (2019)
20. Kumar, R., Dubey, P.K., Zafer, A., Kumar, A., Yadav, S.: Development of Remote Wireless Environmental Conditions Measurement, Monitoring and Recording Device for Metrological and Other Scientific Applications. Mapan **35**(2), 193–199 (2020). https://doi.org/10.1007/s12647-020-00368-3
21. Sasirekha, S., Swamynathan, S.: Cluster-chain mobile agent routing algorithm for efficient data aggregation in wireless sensor network. Journal of Communications and Networks **19**(4), 392–401 (2017)
22. Wang, J., Cao, Y., Li, B., Kim, H.-J., Lee, S.: Particle swarm optimization based clustering algorithm with mobile sink for WSNs. Futur. Gener. Comput. Syst. **76**, 452–457 (2017)
23. Ju, C., Gao, Y., Sangaiah, A.K., Kim, G.-j.: A PSO based energy efficient coverage control algorithm for wireless sensor networks. Comput., Mater. Contin. **56**(3), 433–446 (2018)
24. Wang, J., Yu, G., Wei, L., Arun Kumar, S., & Hye-Jin, K., An improved routing schema with special clustering using PSO algorithm for heterogeneous wireless sensor network. Sensors **19**(3) 2019
25. Deepak, S.: An approach to optimize homogeneous and heterogeneous routing protocols in WSN using sink mobility. Mapan **35**(2), 241–250 (2020)
26. Mohammed, A.-Z., Ahmed, S.M., Sabor, N., Sasaki, S.: Mobile sink-based adaptive immune energy-efficient clustering protocol for improving the lifetime and stability period of wireless sensor networks. IEEE Sens. J. **15**(8), 4576–4586, (2015)
27. Wang, J., Cao, J., Simon Sherratt, R., Hyuk Park, J.: An improved ant colony optimization-based approach with mobile sink for wireless sensor networks. J. Supercomput. **74**(12), 6633–6645 (2018)
28. Tashtarian, F., Moghaddam, M.H.Y., Sohraby, K., Effati, S.: On maximizing the lifetime of wireless sensor networks in event-driven applications with mobile sinks. IEEE Trans. Veh. Technol. **64**(7), 3177–3189 (2014)
29. Varshney, A., et al.: Challenges in Sensors Technology for Industry 4.0 for Futuristic Metrological Applications. Mapan **36**(2), 215–226 (2021). https://doi.org/10.1007/s12647-021-00453-1
30. Gao, Y., Wang, J., Wu, W., Sangaiah, A.K., Lim, S.-J.: A hybrid method for mobile agent moving trajectory scheduling using ACO and PSO in WSNs. Sensors **19**(3), 575 (2019)
31. Wang, J., Zuo, L., Shen, J., Li, B., Lee, S.: Multiple mobile sink-based routing algorithm for data dissemination in wireless sensor networks. Concurrency and Computation: Practice and Experience **27**(10), 2656–2667 (2015)
32. Wang, J., Gao, Y., Liu, W., Wenbing, W., Lim, S.-J.: An asynchronous clustering and mobile data gathering schema based on timer mechanism in wireless sensor networks. Comput. Mater. Contin **58**(3), 711–725 (2019)
33. Banerjee, T., Xie, B., Jun, J.H., Agrawal, D.P.: Increasing lifetime of wireless sensor networks using controllable mobile cluster heads. Wirel. Commun. Mob. Comput. **10**(3), 313–336 (2010)
34. Liew, S.-Y., Tan, C.-K., Gan, M.-L., Goh, H.G.: A fast, adaptive, and energy-efficient data collection protocol in multi-channel-multi-path wireless sensor networks. IEEE Comput. Intell. Mag. **13**(1), 30–40 (2018)
35. Sethi, D., Bhattacharya, P.P.: Artificial neural network based base station localization for energy efficient routing in WSN. Recent. Pat. Comput. Sci., Bentham Science Publisher **9**(3), 248–259 (2016)

36. Liu, Y., Qiong, W., Ting, Z., Yong, T., Fengshan, B., Minglu, J.: An improved energy-efficient routing protocol for wireless sensor networks. Sensors **19**(20), 4579 (2019)
37. Safara, F., et al.: PriNergy: A priority-based energy-efficient routing method for IoT systems. J. Supercomput., 1–18 (2020)
38. Haque, M.E., Baroudi, U.: Dynamic energy efficient routing protocol in wireless sensor networks. Wireless Netw. **26**(5), 3715–3733 (2020). https://doi.org/10.1007/s11276-020-02290-7
39. Daanoune, I., Baghdad, A., Ballouk, A.: An enhanced energy-efficient routing protocol for wireless sensor network. Int. J. Electr. Comput. Eng. (2088–8708);10(5) (2020)

The Integration Development and Upgrading Path of Industry 4.0 Architecture Industrial Engineering Network Driven by Big Data

Hui Li(✉)

Guilin University of Electronic Technology, Guilin, China
yy15904423119@163.com

Abstract. In the overall promotion of network technology, the Internet has gradually changed people's concept of time and space, made cooperation models and development content more diverse and open, and laid a foundation for the integrated development of advanced network forces and industrial systems around the world. In the innovation and development of modern economy, Germany's industry 4.0 strategy and the Industrial Internet model of the United States make the integration development of industrial engineering network driven by big data get the attention of the society, and the manufacturing industry has put forward a new production model and development path. Therefore, this article studies in understanding big background, on the basis of data driven according to the basic connotation of industrial engineering network integration, clear the basic route of industrial architecture engineering network integration, and carries on the empirical research about the mode selection, to enhance the competitive advantage of industrial products, build environment of transformation and upgrading of the new forms.

Keywords: Big Data · Industrial 4.0 · Industrial Engineering · Network Convergence

1 Introduction

Nowadays, as an important industrial power, China ranks first in most of the industrial products manufactured in the world. Although China has not reached the status of an industrial power and the development of all industries presents the phenomenon of overcapacity, after the innovation-driven strategy is put forward, industrial engineering begins to accelerate the pace of industrial development based on system and technological innovation. At this time, the integration development of industrial engineering network driven by big data has laid a foundation for improving the efficiency of China's industrial economy [1–3]. In their empirical research, Wang Hongqiao et al. mainly discussed the development direction of industrialization integration and coal industry informatization under the background of Internet, found in their empirical studies that effective integration can lay a foundation for the realization of personalized production

R. Kountchev et al. (Eds.): *NAMSP 2022*, SIST 332, vol. 332,, pp. 217–224, 2023.
https://doi.org/10.1007/978-981-19-7842-5_19

and manufacturing, collaborative manufacturing logistics, and service-oriented manufacturing enterprises. Nowadays, the economic development of various countries gradually realizes the relationship between "Internet + " and industrial engineering, and puts forward relevant concepts in practical exploration. One refers to the industrial Internet model proposed by the United States, and the other refer to the Industry 4.0 strategy proposed by Germany. Because the manufacturing industry of the United States and Germany is more developed and the actual infrastructure strength is stronger than that of China, the overall development level of the United States and Germany is higher than that of China. Driven by big data and combined with the industry 4.0 framework, China deeply explores the main approaches to the integrated development of industrial engineering network, and actively learns from advanced and mature development experience of foreign countries, which can not only lay a foundation for China's industrial innovation and development, but also obtain more valuable technology systems [4–6]. System to understand the development of industrial engineering network driven by big data fusion the basic definition, found that in the traditional sense of the manufacturing enterprises will use big data, cloud computing, mobile communication technology, such as the Internet, optimization of product research and development production technology, guarantee the rational configuration of production elements, the specific integration innovation system are shown in Fig. 1 below. Regard "Internet + enterprises" as associated manufacturing among manufacturing enterprises, and adjust the production process through integrated research on the condition of grasping changes in market demand and enterprise data manufacturing. The final results show that the basic connotation of "Internet + enterprise" means that the manufacturing subject applies the existing or new Internet thinking mode to innovate and optimize the value chain of current INDUSTRIAL engineering research and development, production and after-sales service, so as to improve the individuation and intelligence of industrial engineering driven by big data. It can be seen that "Internet + Enterprise" has some similarities with Germany's Industry 4.0 strategy and America's industrial Internet. According to the industry 4.0 architecture driven by big data, the development path of industrial engineering network integration is deeply discussed, the importance of industrial intelligence, synergy and high efficiency is clarified, and the network technology and production concept of industrial engineering are continuously transformed and upgraded, so as to truly achieve the expected development goals [7–9].

2 Method

First, we should push forward from downstream industries to upstream industries. According to the industrial engineering industry chain driven by big data, the closer the industry is to the downstream and the end user, the sooner the reform should be completed. Especially in the field of electronic commerce rapid development, the industry chain downstream enterprises gradually transform the traditional production mode to personalization production technology, the industrial chain middle reaches of the industry also followed the downstream manufacturers flexible production and the production of intelligent manufacturing system is put forward, the upstream industry because only provide raw materials, so the relevant production activities and network convergence at

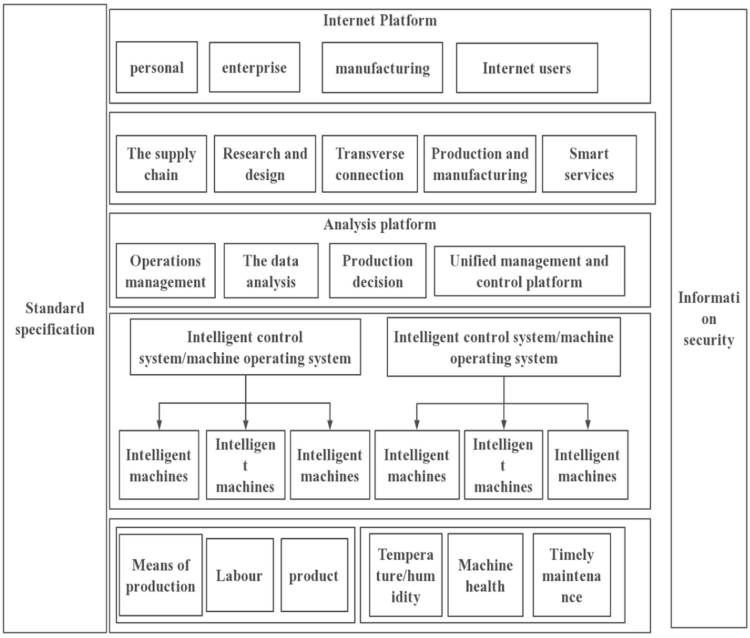


Fig. 1. Innovation system of industrial engineering network fusion.

a slower pace. Driven by big data and influenced by the Industrial Internet in the United States and the Industry 4.0 strategy in Germany, the integration between industrial engineering production activities and network technology is getting faster and faster. The reason for this change is that the production and operation activities at different stages of the supply chain have changed. For raw material suppliers, the downstream industries of the supply chain contain more product types and have a wider integration and development approach with network technology. Nowadays, Internet sales is the main form of early marketing mode innovation, showing a positive advantage in the development of industrial modernization. For example, an enterprise applies network technology to the overall production and circulation of products, and establishes a quality traceability system nationwide. It gets rid of the traditional management mode that only focuses on downstream distribution channels and ignores raw material supply, and makes two-way traceability to the overall operation of the supply chain.

Second, high value to low value link transformation. Generally speaking, the value chain strategy of industrial engineering enterprises contains six links, among which the low value mainly refers to the intermediate link, and the high value is located at both ends of the value chain. China's current industrial engineering development business is in the manufacturing stage, the overall process needs to invest a large number of labor resources, but the economic benefits are very low. According to the analysis of the cuckoo and cantona and others [10], have shown in the Table 1 below industry Internet, 4.0, "Internet + industry" the concept of analysis and comparison, the combination strategy, to upgrade transformation on the whole value chain, intelligent monitoring system is proposed using the Internet of things technology, combined with large data

scientific decision thinking, In order to improve production efficiency at the same time, control labor input. It is important to note that this penetration is not permanent [10, 11].

Table 1. Comparison of concepts.

Concept	Industrial Internet	Industrial 4.0	"Internet plus Industry"
Purpose	Increase productivity and reduce costs	Improve production efficiency and reduce labor	We will optimize resource allocation and enhance innovation capacity
The target	Big data, sensors and human organic combination, break through the boundaries of wisdom and machine, realize industrial production network, intelligent, flexible and service-oriented, upgrade key industrial fields, industrial Internet revolution	To build a highly flexible production model of personalized and digital products and services, improve the competitiveness of German industry, seize the initiative in the new round of industrial revolution, and promote the fourth industrial revolution	The integration of the Internet with the manufacturing industry has been further deepened, and the digitalization, networking and intelligence of the manufacturing industry have been significantly improved
Elements	Three elements: intelligent machines, advanced analytics, and people	Three themes: intelligent products, intelligent factories and intelligent production	Four core: intelligent production, personalized customization, collaborative creation, user service

Third, shift from local to global synergy. For the integrated development of industrial engineering network, it is necessary to reduce the influence of scale efficiency on the composition of enterprise production, and use concentrated scale production to replace the scale and decentralized regional production. In this context, enterprises should independently choose each link to outsource to ensure that product design and production operation can be efficient; The network platform should be used for collaborative services, integrating and researching large quantities of data including the whole life cycle of products, and then feeding back to sales and research and other links to ensure that all links can closely cooperate. Thus, the effective path for the integrated development of industrial engineering network under the industry 4.0 framework can be obtained as shown in Fig. 2 below:

3 Result Analysis

According to the research results of modern scholars of various countries on the integrated development of industrial engineering network, the development of modern industrial engineering driven by big data should be upgraded and optimized based on the

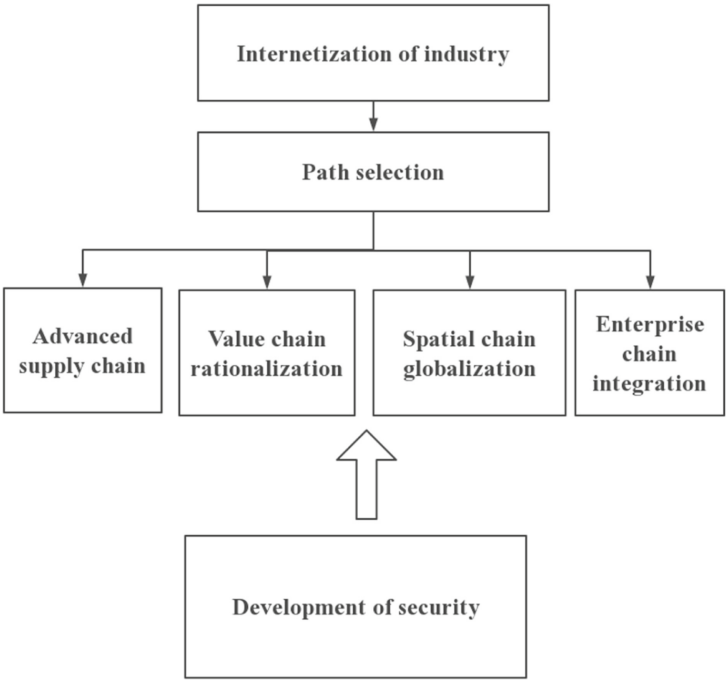


Fig. 2. The effective path of industrial engineering network convergence development under industry 4.0 architecture.

concept of Internet technology on the basis of getting rid of the traditional mode. The industry 4.0 strategy and the industrial Internet concept both prove the importance of this new mode, so the following is a simple analysis of the specific mode selection.

First, big data driven. In the changing market demand, network information technology is more and more mature, industrial engineering product life cycle is shorter and shorter, product production and sales model has also changed. European and American countries will gradually transfer large-scale and large-scale manufacturing production bases to developing regions, while local regions will build personalized and small-batch manufacturing enterprises. Under the guidance of network technology, this personalized service mode will be fully popularized. In this context, the industrial product market can make use of convenient network information technology processing demand, on the basis of sharing resources quickly respond to market changes.

Second, technology spillover. According to the analysis of the interconnection architecture diagram of the industrial Internet shown in Fig. 3 below, the sharing and open performance of network technology enables users to understand the production creation and basic value of products in real time while deeply participating in the consumption and transaction links. Nowadays, some enterprises make full use of the Network to communicate and cooperate with customers in the process of R&D, design and manufacturing. Traditional enterprise thinking also turns to the open and shared network platform.

Under the background of the development of industrial engineering network integration, the development of technology and the integration of a variety of forms to build the ecological service system and combine the user needs to discuss the development and manufacture of industrial products and application model, which not only provides consumers and enterprises to develop more valuable data decision-making information, also according to the whole process of industrial engineering innovation from multiple perspectives.

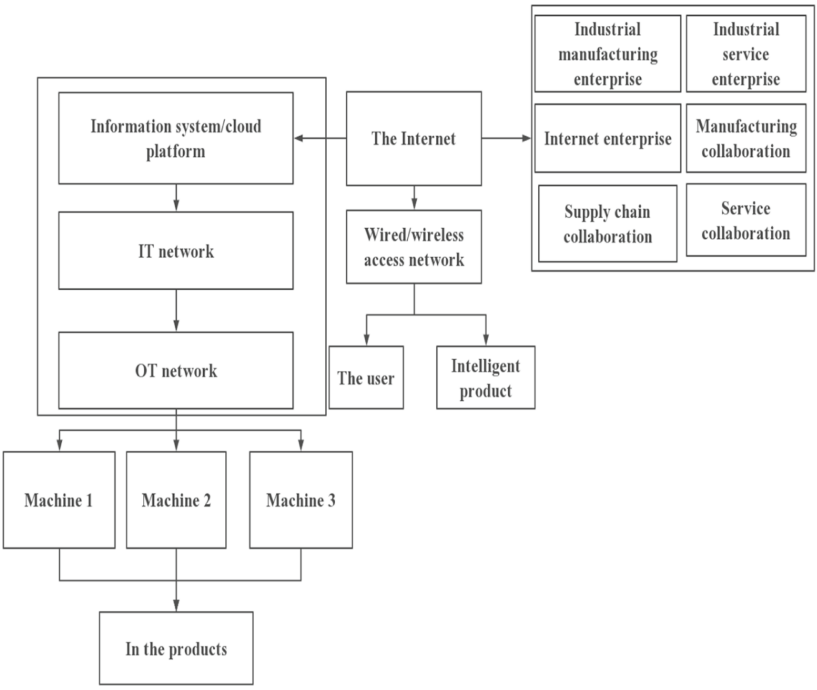


Fig. 3. Interconnection architecture diagram of industrial Internet.

Fourth, fusion lifting type. The service-oriented mode of manufacturing enterprises refers to the transformation of industrial engineering enterprises from the traditional production mode centered on manufacturing to the production process centered on manufacturing, providing services and providing solutions. China's industrial engineering has put forward the servitization model since the 1980s, but the path construction of network convergence is too slow due to the influence of traditional thinking. To understand the relationship between the improvement of network technology efficiency and industrial growth, as shown in Fig. 4 below, it is found that under the influence of big data driving and industry 4.0 architecture, the development approach of industrial engineering network convergence will start from two aspects: On the one hand, on the basis of production specialization, the enterprise's internal and external operation links and network technology are fully integrated to promote the importance of service elements in the innovation collaboration of resource platform. On the other hand, industrial

engineering enterprises should use Internet technology to improve production efficiency and pay attention to creating more business transformation opportunities for service-oriented links, and comprehensively improve the competitive advantages of industrial engineering. The practical results show that the transformation process of industrial manufacturing enterprises effectively improves the return on actual investment, and the application of network core technology provides a new path for the transformation of manufacturing to service-oriented integration. Under this development background, the era of invisible service industry is coming.

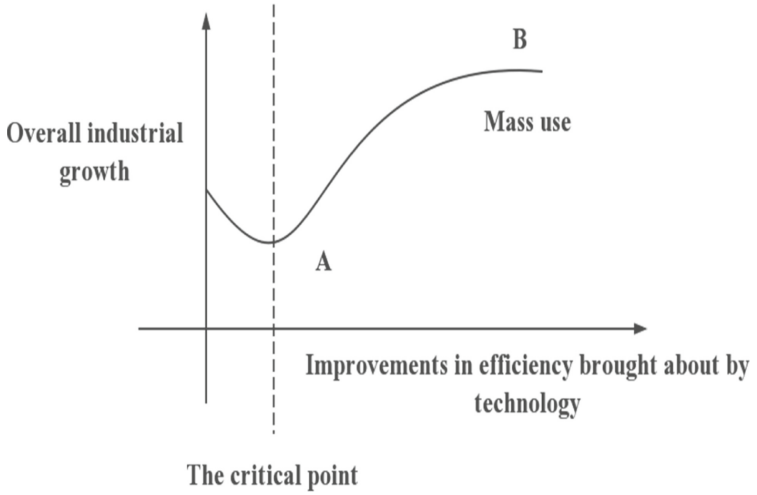


Fig. 4. The relationship between the efficiency improvement of network technology and industrial growth.

4 Conclusion

To sum up, under the background in the era of big data, consumer's personal habits and greater changes have taken place in underlying demand, especially in the "Internet + industry" development mode, the enterprise as the main body of transformation of innovation, to The Times development needs to adjust to the network to industrial infiltration from put forward the strategy of "two" begins, but real integration is only now being fully implemented. Therefore, when choosing the integration mode of industrial engineering, enterprises are required to choose effective countermeasures according to their own transformation direction in the process of continuous innovation. It should be noted that the network integration path of industrial engineering should not only be reflected in the supply and demand chain and value chain, but also be optimized in the enterprise chain and spatial chain. Only all-round integration development can truly achieve the strategic goal of industry 4.0 driven by big data. In this process, governments at all levels should actively mobilize social institutions, such as services, telecommunications department

participation enthusiasm, cooperate fully with network integration development countermeasures, from all walks of life to strengthen the cultivation of professional talents, actively introducing advanced technology experience, only in this way can guarantee industry 4.0 architecture engineering driven by big data can be ordered to carry out the network integration mode.

References

1. Wang, H.: Minggao Ouyang. Transactions of CSICE **20**(1), 36–40 (2002). (in Chinese)
2. Tang, S., Jinhui Yang, Chunye Xu, Yi Yang, Xiaoying Zhou, Yonghua Guo, Hongtao DIao, Zitao Zhang.: Application of industrial data-driven technology in the operation of large and complex natural gas pipeline networks. Natural Gas Industry, 201, 41(09):135–141
3. Zhang, W.T.: Research on data-driven industrial Internet resource adaptation and privacy protection methods. Beijing Jiaotong university (2021). https://doi.org/10.26944/,dcnki.Gbfju.2021.000155
4. Wang, T..: Data-driven industrial information security protection. China Inf. Wkly., 2021-09-13 (024) (2021). https://doi.org/10.28189/n.c.Nkindnjy.2021.000425
5. Wang, T.: Data-driven industrial information security protection. China's Minist. Ind. Inf. Technol. (8), 32–36 (2021). https://doi.org/10.19609/j.carol carroll nki cn10-1299/f 2021.08.004
6. Jiang, Y..: The Impact of Big Data Analysis Capability And Supply Chain Visibility on Mass Customization Capability. Harbin Institute of Technology (2021)
7. Fu, X., Han, J., Cao, K..: Big data-driven intelligent fault diagnosis technology in iron and steel industry. Comput. Meas. Control. **28**(11), 1–6 + 26 (2020). https://doi.org/10.16526/j.carol carroll nki.11-4762/tp. 2020.11.001
8. Sun, L..: Research on "manufacturing + Internet" integration and innovation driven by industrial big data. Electron. Commer. **2020**(9), 19–20 (2020). https://doi.org/10.14011/j.carol carroll nki DZSW. 2020.09.009
9. Xiaohang, J., Yu, W., Zhang, B.: Industrial big data-driven fault prediction and health management. Comput. Integr. Manuf. Syst., 1–27 [2022-02-28]. http://kns.cnki.net/kcms/detail/11.5946.TP.20200814.1703.006.html
10. Perumal, K., Ungati, J.M., Kumar, G., Jain, N., Gaurav, R., Srivastava, P.R.: Test data generation: a hybrid approach using cuckoo and tabu search. Springer, Berlin Heidelberg (2011)
11. Wang, T., Cui, J..: Based on data driven industry information safety protection. J. Inf. Technol. Netw. Secur. **5**(6), 3–5 (2018). https://doi.org/10.19358/j.iSSN.2096-5133.2018.06.001

SAAS Application Prospect Analysis in Hrm and Methods to Upgrade the Contemporary System

Jicong Xue(✉)

University of Shanghai for Science and Technology, Shanghai, China
787768934@qq.com

Abstract. With the soaring of global population and the trend of transitional employment, the amounts of data that human resources practitioners required to process and decode are becoming unprecedently giant. Therefore, the application of big data and artificial technology in Human resources management is an inevitable trend. Based on the previous research papers on the field of Software as a Service, this study uses Methods of Theory of Inventive Problem Solving (in short-termism) to analyze potential innovative methods. Founded that many existed operational modes are no longer applicable due to the drastic changes in the internal and external environment of the enterprise. Currently, the feasible and valuable research direction is to focus on the specific operation module of a certain function.

Keywords: SAAS · Principal Component Regression · TRIZ · Tobit Regression

1 Introduction

Big data tech include diversified components, this article will introduce the study of SAAS mode in HRM to illustrate the functions and potential optimizations of this system to alleviate the work load of HR practitioners.

Relevant academic points are summarized as follow.

1. According to scholar Zhen Zeng [1], the soft ware system design should be based on the employee traits like gender, age, etc. Besides, due to the continual changing of organization's internal and outer environment, the gray scale system should be applied.
2. Scholar Nasreen Nasar [2] believes that the essence of e-HRM is the data organizations have gathered should be pre-processed so as to avert unnecessary negligence and deviation.
3. Scholar Peng Li [3] regards that the criterion of the content firm using big data technology should not merely concentrate on the quantity of data, and instead should value the high-quality products.

R. Kountchev et al. (Eds.): *NAMSP 2022*, SIST 332, vol. 332,, pp. 225–237, 2023.
https://doi.org/10.1007/978-981-19-7842-5_20

2 Contemporary Works of SAAS

The prototype of SAAS is the Enterprise Resource Planning connotation, which is first put forward at 1990s,emphasising putting all business data in a sole system. A series of subsidy policies are implemented to encourage entrepreneurs to view digital resources as vital productive elements to improve working efficiency and optimize the economic schedule. The latest policy released in May, 2021 is the Statistical classification of digital economy. This system confers totally different strategic objects to organizations based on their developing stage. The 4 major objects include fundamental issues solving, informative construction, analysis specific managerial problems and strategy planning. In different developing stage of firms, the managerial traits and potential barricades firms may encountering are usually not same. The specific criterion and indicates could be referred in Fig. 1.

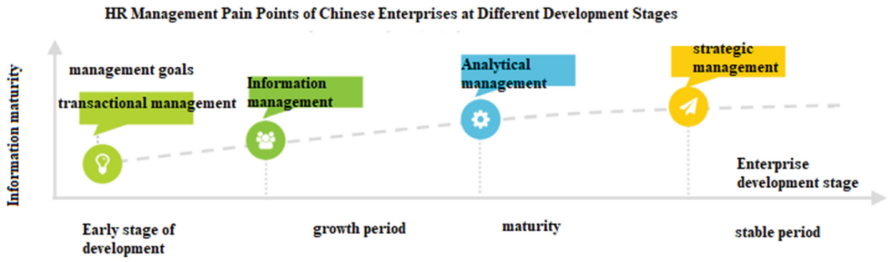


Fig. 1. Managerial key points of enterprises in different development stages

Main reason for firms applying e-HRM is to make a accurate evaluation of the performance of employees and searching for potential employees at the lowest cost. A case in that is firms could use the technology to make a comparison of outcome of employees in same industry. With the application of the 360 degree assessment and Balanced Score Card assessment in SAAS system, a series of evaluation standards like historical standards, standards of excellence are generated in the ERP system. Then supervisors prone to adjust their management strategy based on performance score given by the ERP system rather than the analysis results of financial statements. Reason for this phenomenon is that the work performance data is updated more frequently, thus enable enterprises to better cope with environmental changes. Apart from the managerial reason, firms apply the system so as to earn great reputation and propagate themselves. Consequently, talents are attracted to join in.

3 Data Analysis

3.1 Data Source

This paper get the data from the authoritative investigation institution-iResearch. The specific digits in the report have undergone various tests and have been proved to be reliable. Since the Reliability and Validity is excel to the standard value.

3.2 Marco Analysis

This part include the analysis of digital economy scale and it's share of GDP. The proportion of labor force, rate of workers engaging in the tertiary industry and the CIER rates in CHINA the market scale of cloud computing and the using rate of cloud computing.

(1) digital economy scale and it's share of GDP.

The analytical method I take is Principal component regression, namely, setting the digital market scale as the dependent variable(y) and it's share of GDP as the variable(x).The calculation steps are: ① reconstruct the linear combination, the specific ② delete the variation in the new combination.

Specific equation are presented as follow.

Regression equation:

$$Y = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_p \end{pmatrix} = \begin{pmatrix} \gamma_{11} & \gamma_{12} & \cdots & \gamma_{1p} \\ \gamma_{21} & \gamma_{22} & \cdots & \gamma_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ \gamma_{p1} & \gamma_{p2} & \cdots & \gamma_{pp} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} \tag{1}$$

$$Y_{\mathrm{i}} = \gamma_i^T X, \quad i = 1, \cdots, p \tag{2}$$

Correlation coefficient:

$$\rho(Y_k, X_i) = \frac{\sqrt{\lambda_k}}{\sigma_{\mathrm{ii}}} \gamma_{ki}$$

Spectral decomposition of matrix:

$$\sum_{k=1}^{\mathrm{p}} \rho^2(Y_k, X_i) = 1 \tag{3}$$

Practical statistical data and deep analysis (Table 1).

Table 1. Analysis of variance table

Source of variance	Sum of squares	df	Mean square	F Value	p-Value
Regression	948.4788	1	948.4788	53.7262	0.0007
Surplus	88.2697	5	17.6539		
Total	1036.7486	6	172.7914		

From Table 2, we could figure that the p value is inferior to the standard value 0.05,thus the factor is significantly related.

Table 2. Model fit analysis

R	0.956483
R2	0.914859
Adjusted R2	0.947539

The 3 R indicates in Chart 3 illustrate that the model is well structured with little deviation.

The regression equation: y = 4.717481 + 0.957514x, we could find that the relation between Variable and dependent variable is positive. Therefore, we could make reasonable conclusion that this industry do deserve our investment.

(2) The proportion of labor force, rate of workers engaging in the tertiary industry and the CIER rates in CHINA.

In this part, the big data tech we apply is principal component analysis. The difference from the technology we used in the first data set is that this tech don't require set the dependent variable (Table 3).

Table 3. Correlation coefficient table

Correlation coefficient	X(1)	X(2)	X(3)
X(1)	1	−0.6937	−0.8915
X(2)	−0.4779	1	−0.5804
X(3)	−0.8335	0.1167	1

Variable explanation:

the X(n) n = 1 to 3 refers to rate of workers engaging in the tertiary industry, CIER rate and the proportion of labor force respectively. The sign of digits reflect the negative and positive correlation of variables. We may discern that the X1 has negative correlation with the other two indicates. Therefore, we could make following conclusion: If the overall employment prospect is undergoing continual deteriorating, then the possibility of people engaging in service industry is soaring and the quantity of labor force is escalating.

(3) the market scale of cloud computing and the using rate of cloud computing, in this mode, I'll take the Partial least squares regression method to analysis. The axiom of this method could be referred in Chart 5.

$$X = TP^T + E \tag{4}$$

$$Y = UQ^T + F \tag{5}$$

Among them, X is the estimatation matrix, Y is the Response matrix.

The corresponding regression equation (Table 4):

$$Y = X\tilde{B} + \underset{0}{\tilde{B}} \tag{6}$$

Table 4. Estimated values of model Ti, UI and dependent variation.

t1	u1	y1
−0.8941	−0.8802	54.6206
−0.1858	−0.2071	58.7239
1.0799	1.0873	66.0556

relevant variables listed above (Table 5).

Table 5. Sum of squares of model errors after data standardization

group	y1	R2
c1	0.0007	0.9996

Since the determinable coefficient excel than the normal standard value. Thus, this model could be considered reliable.

The regression result y = 46.081322 + 0.012322x.

Variable interpretation: in this trial, y refers to using rate of cloud computing, x refers to the market scale of cloud computing. Then we could make further conclusion: the two indicates are correlated. If the government want to improve the using rate, investing in fields that improve the quality of cloud computing might have better effect of escalating the using rate.

3.3 Micro Analysis

In this part, I will test the relation between total funding events and vital events of SAAS (the vital one refers to the funding amount which is above 0.1 billion) as well as the market scale and increasing rate of SAAS. This part, I'll take the method of tobit regression. Relevant equation can be referred as follow.

$$\begin{aligned} y* &= \beta' x_i + u_i, \\ y_i* &= y_i \ \ if y_i* > 0, \\ y_i* &= 0 \ \ if y_i* \leq 0. \end{aligned} \tag{7}$$

normal edition

$$y = \max(0, \beta' x_i + u_i) \tag{8}$$

the simplified version.

In this trial, we set increasing rate as the dependent variable, other indicates as independent variables.

Table 6. Outcome of the trial.

	Equation coefficient	Standard error	Wald chi square	P value	95% confidence interval	
β0	81.1685	1.9297	1769.3204	0.0001	77.3863	84.9506
β1	−0.3594	0.0247	212.2834	0.0001	−0.4078	−0.3111
β2	−3.377	0.2119	253.9626	0.0001	−3.7924	−2.9617
β3	−0.5434	0.034	254.9675	0.0001	−0.6101	−0.4767

From Table 6, we could figure that the p value is far below than the standard, thus we could assume that the parameters are significant. The regression equation is y = 81.1685–0.3594x1–3.377x2–0.5434x3.The x1, x2, x3 is correlated while they have negative correlation with y.variable explaination: x1, x2, x3 refers to total events. Vital events, scale respectively. The y refers to the increasing rate.

3.4 Conclusion of Analyzing

In the strategic level, we should realize that priority should give to the improvement of quality of cloud computing rather than the scale. Indeed, we can't deny that improve the scale of cloud computing can also optimize the industrial structure. Since the competition in labor market is still fierce, developing the tertiary industry could be a therapy to solve the embarrassed dilemma. For one side, a well-developed service industry could meet consumers' personalized demand. For the other, the tertiary industry could absorb more surplus labor force since the threshold of this industry is not as high as other industries which require science backgrounds like the semi-conductor industry.

At the executive level, we should first realize that the SAAS industry has entered the stage of diminishing marginal benefits. Solely invest capital can hardly improve the market scale. As is depicted in the below chart, the current stage the SAAS industry at is the second phase. In this stage, though the total production is still rising, the increasing speed, instead are declining gradually. Apparently, the SAAS industry is at the second stage. The two main object for practitioners, especially for mangers are improve the working efficiency so as to reach at the outcome of "MR = MC" and find feasible ways of innovation or upgrading the products (It's prerequisite is that the product is still at the early stage of it's product life circle. If not, managers better to lead the group to invent new products) (Fig. 2).

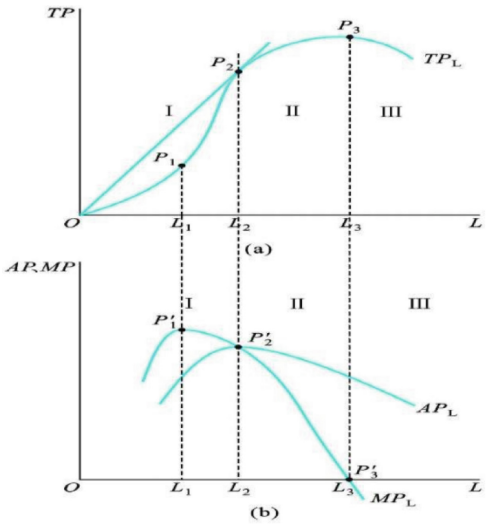


Fig. 2. Relation curve of TP, MP and AP

4 Methods Improvements Which Based on the Previous Study

4.1 General Guidelines of Improved Modes

We could quote the TRIZ conception as our guideline of innovation. The theory which was put forward by Altshuller, giving us various basic traits of innovation and some duplicable measures with almost no threshold, for instance, embed technology from other disciplines into current field, converse multiple dimensions of a specific object into a curve, etc. The essence of this theory is to jump out of the limitations of the disciplines and absorb advanced conceptions of other fields to give birth to a brand-new Interdisciplinary or make astonishing achievement. The HRM has six significant modes, planning of human resource, recruit, training and developing, performance regulation, income regulation and work relation management. This article will innovate current academic methods of SAAS system under the guidance of TRIZ theory.

4.2 Renovation of Existed SAAS Systems or Conceptions

Planning of Human Resource. Scholar Lauren [5] points that we could quote the past data of transactions to estimate the potential business scale of the future and thus establish a simulated sand table system. In this article, we will use the random forest technology to renovate it. Namely, taking the bootstrap method to resample, putting back the duplicated samples to set brand-new sample sets. Then form the new random forest based on the previous outcomes, the form of classification is based on the scores of classification trees. Typical classification trees application in HRM planning could be referred as follow. In the practical operation, we could set and alternate a series of assessment indicates to probe the most significant indicates of estimating the human

resource requirement. Through professional calculating software, we may also acquire some indicates of key deviation. We shall take these indicates to judge the fitting degree of different estimation models, for instance, SVM or Classical discrimination, etc. After testing all the relevant indicates, then we could do the trial of multiple regression analysis. The vital detail of this process is that we might take diversified linear form to ensure our p-value (or f value) is below the standard value and our R^2 is higher than the standard value so that our overall estimation model could be reliable in both parameters and the overall curve. (IDL refers to indirect employees) (Fig. 3).

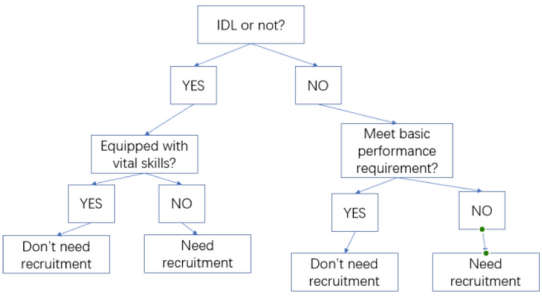


Fig. 3. The Logic diagram of automated employment judgment

Recruitment. Scholar Bo Zhao illustrate that firms could recognize the potential human resources requirement due to the change of payroll list. Besides, the inventory store could also reflect the human resources situation. If highly-skilled employees are well-equipped, then the zero inventory mode may be well implemented. Unnecessary waste and depreciation could be averted. His proposition is mainly relied on the application of ERP system. Actually, this operational mode could be successfully implemented in large enterprises or consortiums. But for small and medium sized firms, they can't afford the high price of system like SAP. Thus, this article will introduce following measures for small-sized firms. To begin with, focus on the exploitation of low code or even no code platform. Since the overall coding level of small-sized firms can't compare with that of listed companies. The custom group of these firms are always changing, thus low-threshold software could cut the training cost and the exploitation cost. Besides, lease the cloud computing system of large firms like Baidu. For one side, they could absorb the advanced conception of data management from using the service. For the other, it could lower the risk of artificial mistakes by the personalized pre-control service from these developer. As a consequent, the firms using them could have less risks of paying a fine as well as propagate themselves.

Training and Developing. Scholar Qi Feng put forward the renovative method of giving training opportunities to employees based on the analysis results of faculty matrix. This article will introduce the game theory to renovate this method. We assume that the effort and passion employees input in working place are based on their opportunities of being promoted (people with high risk are keen on presenting high ability). The below chart will demonstrate the gamble between employee a and b. Set Xi as the risk of employee

A being promoted and set Yi as the risk of other employees being promoted. If Xi > Yi, one can easily estimate that A will exert more effort to be distinct from others to get promoted. If Xi < Yi, A may have two options: (1) work harder to leave a deep impression to his supervisor (2) opt idleness or thwart B to get a outstand performance. Assume every options have the possibilities of 50%, then we could estimate that if we let employees know their probabilities of being promoted. Then, the employees with highest rate of being promoted will have 75% probabilities of paying hard works. The further renovation is that we shall let employees know the quota of promoting every month. The indicates should always modified according to the performance of employees last month.

Performance Regulation**.** Scholar Sadaf Aman [13] regards we could add dimensions of community protection in grassroots staffs' KPI, she emphasis taking both tangible and intangible assessment indicates to encourage employees to take environmental friendly operational modes. The modification measure I'll take is to take Symmetry test. Due to the global economy wave, employees are constantly from different culture background. Thus, they have different motivations of work, thus ethics-oriented objects may not always effective. Therefore, organizations could empower the grassroots HR Specialist the authority to test different staffing combinations. The axiom of this test mode is Correlation test. The equation is

$$r = \frac{\sum_{i=1}^{n} (X_i - \overline{X})(Y_i - \overline{Y})}{\sqrt{\sum_{i=1}^{n} (X_i - \overline{X})^2} \sqrt{\sum_{i=1}^{n} (Y_i - \overline{Y})^2}} \tag{9}$$

We could use the r value to test the extent and direction of diversified variables are related. For those positive odds, we may put employees with these traits in a same team to alleviate contradictions and motivate innovations. For those negative odds, we should try to separate the two group. If conditions are not allowed, we shall train both group to put themselves in others shoes to overcome the cultural shock. Then the follow-up training object is to let them realize the significance of sustainable indicates and mix them into their performance indicates. More important, the weight of the indicates should be moderate, not very tiny. Scholar Ramona Zharfpeykan put forward that firms of different size have different evaluation of GRI standard, the application of BSC is not common in small and medium sized firms. As a consequent, they pay less attention of environmental protecting issues. The driving reason of this phenomenon is that solely listed firms are forced to publish financial information and accept more censorship. In this article, I will combine the method of Intermediary model and 360 degree assessment to renovate it. Due to the division of labor, large amount of employees are expert of their own field, many of them have no basic understanding of other colleagues' work duties or skills required. Thus contradictions may be triggered due to role conflict. Since the inter-discipline employees are rare resource, training one or attracting one this kind of talents are usually costly to firms, especially for firms without affluent cash flow. Therefore, a wise and cheap measure is to extract a grassroots employee, usually the employee with well communicative skills and a great reputation to dedicate in studying the overall

operation structure of the firm. The performance object for this vital position is to coordinate the working process and promote the interpersonal relations of other employees so as to improve the working efficiency and lower the turnover rate. The evaluators of the coordinator will include his supervisor, inferior, colleague as well as the counterpart both inside and outside the firm. The specific operational structure could refer Fig. 4.

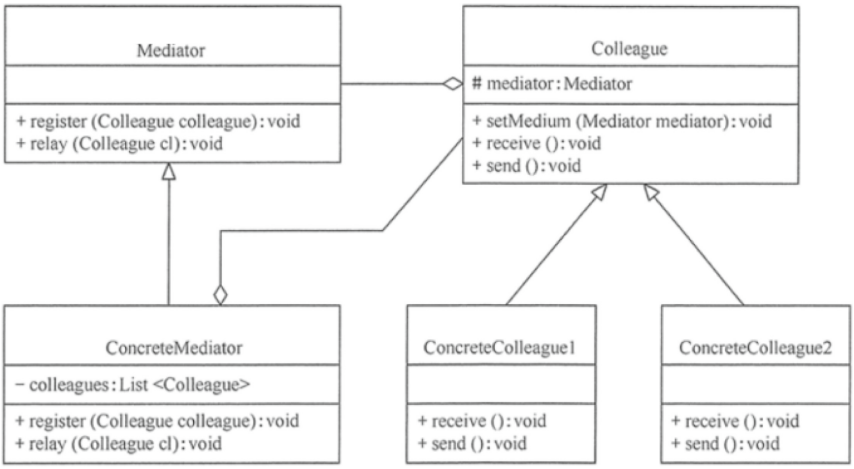


Fig. 4. Operational structure of colleague relation.

Income Regulation. According to scholar Yewande, the high-end talents are suitable for using the mode of paying by the specific service he provided since firms can hardly find their substitutions. Besides, if high-end talents could opt their working content on their will then, we can forecast that large proportion of talents may dedicate in cutting-edge jobs, thus large promotion opportunity may leave to the middle and low end talents. The overall applied technology and market structure of the society may thus improved gradually. This article will apply the technology of substitution effect, we set the salary choice autonomy of the high-end employee as the product A, the quantity of outstand middle-end employees as it's substitution. Since the continual popularization of educational resources, the high-end talents may eventually enrich, therefore it's not wise to permanently enable employees have a high degree of freedom in salary selection. We shall adjust the salary condition based on the whole talent reserve constantly. The SAAS developing duty for HR department is to import the relevant statistical data into the entrepreneurial data-processing system and assess the high end talents reserve of the firm and whole society. The schematic diagram could be referred as follow (Fig. 5).

Scholar Flores regards that mangers could assess the growth potential of employees by analyzing past data and make internal and outer comparison. Besides, the salary raise circle should be vary from employees to employees. The period should based on the score of growth capability evaluation, for those high potential employees, the period should be as short and changeable as possible, for the other, the time should be stable and moderate. Since motivating input in them may bring more paybacks. This article will apply the Endogenous economic growth model to innovate the exist conception.

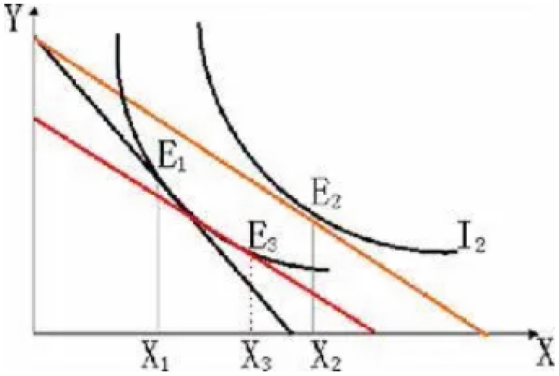


Fig. 5. Indifference curve of Employee allocation

In specific, we could assess the potential by the ability of technology innovation of employees and take factors like economy improvement and government policies into considerations (usually as the variable) to test the significance of these factors to the turnover changing of the firm. The reference diagram of the model can be referred as follow (Fig. 6).

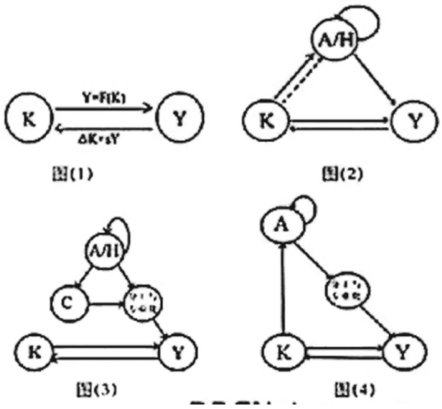


Fig. 6. The adjusted endogenous economic growth model

Work Relation Management. Scholar Gadi Nissim thinks that with the development of cloud technology, employees are no longer need to work at a specific office. They could now work at home to reach a work-life balance. However, the problem is that the work they undertake are more of non-standard and they are having a risk of cooperating with robot colleagues. The Labor Union are losing it's effect. The burgeoning AI-tech Ethics committee are undertaking their role. The low-end employees may face the crisis of being employed. My renovative suggestion to his theory is to pervasively develop and apply the no-code platform. Since consumers usually have different backgrounds and different demands, the fixed programmed setting may not meet their requirement.

Besides, the excessive number of unemployed workers may cause the society to be unrest and sabotage phenomenon may thus triggered. While with the no-code position becoming pervasive, hi-tech firms could absorb large amount of unemployed labor force. It can not only consummate their algorithms but also cultivate their potential consumers. Since employees are incline to choose the products made by their own company rather than that of opponents. Another managerial thought is to choose the potential employee from enthusiast of your products. This kind of employees are more likely to do Organizational citizenship behavior, like helping colleagues, consumers and figure out way to update the technology or overcome flaws of the products. What's more, they are more likely to propagate the product of their firms in their social platform autonomously.

4.3 Conclusion of Renovations Above

The application of TRIZ conception in HR SAAS system design has a huge room for development. The global and domestic statistic data of seeking financing for building SAAS system could help to explain why this field has strong vitality. Current theory on SAAS innovation mainly concentrate on mixing one specific technology with SAAS, while the application of Interdisciplinary like human resource accounting, Physical Economics are still not very common. Thus, we could make reasonable assumption that cross-discipline application in SAAS might be a potential developing trend of this field.

References

1. Zeng, Z., Qi, L.: Engineering M.: "Internet + Artificial Intelligence" human resource information management system construction innovation and research. Math. Probl. Eng. (2021)
2. Clerk Maxwell, J.: A Treatise on Electricity and Magnetism, 3rd ed., vol. 2, pp.68–73. Clarendon, Oxford (1892)
3. Nasar, N., Ray, S., Umer, S., et al.: Design and data analytics of electronic human resource management activities through Internet of Things in an organization. In: Software: Practice and Experience
4. Nicole, R.: Title of paper with only first word capitalized. J. Name Stand. Abbrev., in press
5. Lauren, B., Dimitris, B., Weinstein, A.M., et al.: Prescriptive analytics for human resource planning in the professional services industry. Eur. J. Oper. Res. **272**, S0377221718305708 (2018)
6. Young, M.: The Technical Writer's Handbook. University Science, Mill Valley, CA (1989)
7. Yko, A., Quinn, A., Meaghan L., et al.: Factors that influence specialist physician preferences for fee-for-service and salary-based payment models: a qualitative study. In: Garcia-Diaz, A., Flores, B.E., Noce, R. (eds.) Health Policy, 2021. A computer based heuristic methodology for the development of salary administration guidelines. Omega **24**(5), 583–595 (1996)
8. Goto, H., Hasegawa, Y., Tanaka, M.: Efficient Scheduling Focusing on the Duality of MPL Representatives. In: Proc. IEEE Symp.
9. Computational Intelligence in Scheduling (SCIS 07), IEEE Press, Dec. 2007, pp. 57–64. https://doi.org/10.1109/SCIS.2007.357670
10. Garcia-Diaz, A., Flores, B.E., Noce, R.: A computer based heuristic methodology for the development of salary administration guidelines. Omega **24**(5), 583–595 (1996)
11. Yko, A., Quinn, A., Meaghan L., et al.: Factors that influence specialist physician preferences for fee-for-service and salary-based payment models: a qualitative study. Health Policy (2021)

12. Ramona Zharfpeykana, A., Chris Akroyd, B.: Factors influencing the integration of sustainability indicators into a company's performance management system. ScienceDirect (2021)
13. Aman, S., Seuring, S.: Interestingly it's innovation: Reviewing sustainability performance management in the base of the pyramid (BoP) (2021)
14. Steen, R., Ingvaldsen, G., Patriarca, R.: Engineering resilience in a prison's performance management system. Saf. Sci. (2021)
15. Floros, K., Jrgensen, M.B.: Tracing the future of migrants' labour relations. Experiences of institutionalized migrant precarity in Denmark and Greece. Polit. Geogr. **2020**(March), 102120 (2020)
16. Nissim, G., Simon, T.: The future of labor unions in the age of automation and at the dawn of AI. Technol. Soc. **67** (2021)
17. Gulbransen, K., Thiessen, K., Pidutti, J., et al.: Scoping Review of Best Practice Guidelines for Care in the Labor and Birth Setting of Pregnant Women Who Use Methamphetamines (2021)

All Digital Phase Locked Loop (ADPLL) and Its Blocks—A Comprehensive Knowledge

Lalita Yadav(✉) and Manoj Duhan

ECE Depatment, DCRUS&T, Murthal (Sonepat), Haryana, India
lalitayadav19@gmail.com

Abstract. The Phase-Locked Loop (PLL) is a feedback system used for the synchronization of signals in terms of frequency and phase. It is configured in digital communication, Bluetooth, mobile transmitter, etc. Many analog and mixed techniques have already been proposed with the advancement in time, but design time is much more for that. An ADPLL implements all the major blocks of the PLL in the digital form. The most challenging task is to design a fast locking ADPLL with a small area and low jitter. In this paper, a detailed study of each block of ADPLL architecture is presented and what changes were incurred in these blocks with the advancement in the research and their pros and cons. The comparison of best ADPLL in terms of their parameters and the techniques used in them is discussed.

Keywords: ADPLL · DCO · Biomedical · Signal Processing · Phase and Frequency Detector

1 Introduction

1.1 When the Concept of PLL Come into Being

In 1923, Edward Victor Appleton described the automatic synchronization of triode oscillators, which is the essential element of radio communication. In 1932, Henri de Bellescize described the feedback phenomena for the automatic correction in phase and frequency of the local oscillator in a French journal L'Onde Electrique [1]. In 1969 Signetics introduced a monolithic IC of Phase Locked Loop (PLL), which multiplies the applications of PLL.

PLL is an extremely versatile circuit. It is a closed-loop system that is used for the synchronization of the phase and the frequency of the incoming signal and the reference signal. It is locked when both signals phase and frequency match with each other. PLL used for clock generators, data conversion, and frequency synthesis have low jitter or phase noise [2]. It also found application in the microprocessor, clock recovery circuits, digital signal processing of video signals etc. (Fig. 1).

There are three main blocks of the PLL, namely, phase detector, loop filter, and oscillator. Depending upon the nature of designing of these blocks, PLL is classified into four categories presented in Table 1 [4].

R. Kountchev et al. (Eds.): *NAMSP 2022*, SIST 332, vol. 332,, pp. 238–251, 2023.
https://doi.org/10.1007/978-981-19-7842-5_21

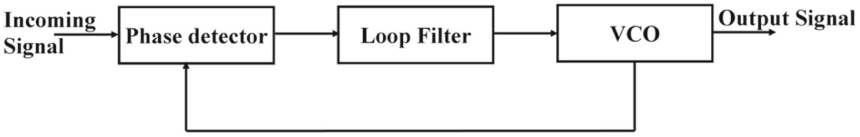


Fig. 1. Basic building blocks of PLL [3]

Table 1. Different types of PLL

Type of PLL	Phase detector	Loop filter	Oscillator
Analog PLL	Analog	Analog	Voltage Controlled Oscillator (VCO)
Digital PLL	Digital	Analog	VCO
All Digital PLL	Digital	Digital	Digitally Controlled Oscillator (DCO)
Software PLL	Software	Software	Software

The hardware PLL can be implemented in discrete or integrated technology, and the software PLL is not discussed here. PLL is a complete analog circuit. The advancement in the technologies led to a reduction in the die size. To build a complex analog system is difficult, and with that it is not possible to achieve the desired accuracy [5]. Analog circuits also have problems with more power consumption, large leakage current, dependency on technological issues, and operating conditions (process, voltage, and temperature) [2, 6]. Hence, analogue circuits are replaced by digital circuits. With the change in technology, the primary issue for the analog PLL is the ample designing time. Then the digital flow has significant advantages of simple tune, higher accuracy, availability of digital route, predictability, low cost, the low power consumption etc. [5, 6].

Hence, ADPLL is considered to be a good option. A manual layout design effort is very difficult during the whole design period which is the long-time. Some synthesized ADPLL were recently proposed to shorten this design time. The custom Application Specific Integrated Circuit (ASIC) design for few hundred quantities is too expensive. Therefore, Field-Programmable Gate Arrays (FPGAs) plays a vital role because they offer the speed, density, and computational power which are otherwise only achievable with ASICs [7]. Hence, for the synthesized ADPLL, FPGA is chosen as the platform but had the problem of poor phase noise, limited frequency tuning resolution, and large reference spurs [2].

The paper is arranged in the following four sections. The second section discusses the ADPLL architecture in detail, and how their main three blocks are introduced till date in detail, the various designing parameters of the ADPLL are discussed in the third section, the fourth section compares the best ADPLL in terms of their performance parameters, which is followed by the conclusion and future work.

2 ADPLL Architecture

To achieve a low voltage operation, low power consumption, and less sensitive to the noise, all-digital implementation presents the great opportunity [8]. It is the circuit that receives a reference input and generate the output signal in phase with the reference signal after performing some control operation over the input signal. Output signal frequency is programmable, rational multiple of a fixed input frequency.

ADPLL has three major blocks: a phase detector, a loop filter, and an oscillator and all the three blocks are digital in nature. The fourth block, i.e., divide by N counter is required only when the frequency of the generated signal is much higher than the reference signal. The basic block diagram of ADPLL is shown in Fig. 2.

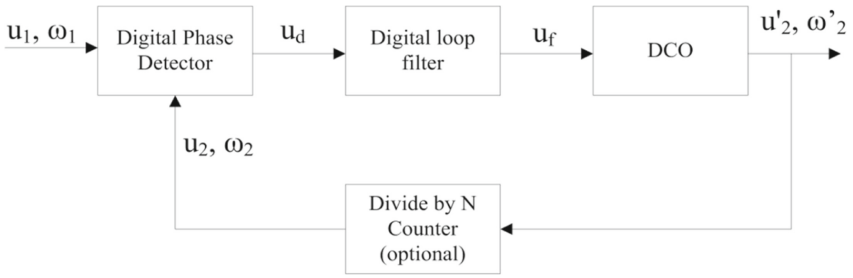


Fig. 2. Basic building blocks of ADPLL [3]

Digital Phase Detector receives an input signal u_1 having frequency of the signal ω_1 and feedback signal u_2 with frequency ω_2 and generates a difference signal u_d. This u_d is the input to the digital loop filter which removes the high frequency component and noise from it and generates a signal u_f which is the filtered output. This u_f is the input to the DCO and generates a high frequency signal from the received input signal. The output signal u_2' is either having the desired frequency, or greater frequency than the desired. If the DCO output signal frequency i.e. ω_2' is greater than the desired output frequency ω_2 then divided by N counter is introduced in the feedback path otherwise it is not required in the circuit.

2.1 Phase Detector

The phase detector (PD) is the comparator that produces pulse width modulated (PWM) signals representing the phase difference between the two signals viz. Reference signal and frequency divider output. It is said to be locked when both signals match with each other in terms of phase and frequency. There are different types of phase detectors available in the literature so far. They are listed below in Table 2, along with their pros and cons.

Table 2. Different types of PD's

S. no.	Author	Types	Pros	Cons	Remarks
1	J. Lee et al. [9]	EX-OR	Good noise rejection	1. Lock to multiple clock frequencies 2. Smaller lock-in range and pull out range	1. Signals should be asymmetrical square waves 2. Used in data and clock recovery circuits
2	Roland E. Best	JK—Flip Flop	Good reliability of data	No cycle of reference signal should be missing	Signal symmetry is not essential
3	J. Lee et al.	Phase and Frequency Detector (PFD)	Large frequency tracking range	1. Poor noise rejection 2. Dead zone problem	Used for clock synchronization and frequency synthesis
4	F. T. Chen et al.	Rotational PFD	Increase the frequency capture range		Used for reference less clock and data recovery circuits
5	Roland E. Best	Double edge-triggered D Flip—Flop	1. Less power dissipation 2. Very fast lock state		
6	Roland E. Best	Zero-crossing PD		1. Output cannot change fast 2. There is a deadband in the output	Input should be analogue. It is basically a comparator circuit
7	Roland E. Best	Hilbert transform PD	Generate two-phase quadrature signals Its gain is one at all frequencies	Architecture is complex	
8	S. H. Lin et al. [10]	Bang Bang PD	Robustness, reduced locked time, and low power consumption	Slow frequency tracking	Also known as single bit TDC

(*continued*)

Table 2. (*continued*)

S. no.	Author	Types	Pros	Cons	Remarks
9	R. B. Staszewski et al. [11]	Arithmetic subtractor based PD	1. Avoids spur introduction in the loop 2. Increasing word length of phase accumulators, increases the dynamic range		
10	W. H. Chen et al. [12]	Latch based PD		Pre-charge time of internal nodes introduce blind zone	
11	Y. S. Tan et al. [13]	Quarter rate linear PD			Suitable for clock data recovery
12	D. Cai et al. [14]	Aperture PD	Low power and low in-band phase noise		Suitable for divider less PLL
13	[15–17]	Flash TDC	The structure is simple	1. Propagation delay of one element limits the time resolution 2. Long delay line consumes large areas, so there is an upper limit of delay elements used 3. Dynamic range is limited by the propagation delay lines and by accumulating effect of mismatch on the propagating edge	

(*continued*)

Table 2. (*continued*)

S. no.	Author	Types	Pros	Cons	Remarks
14	A. K. M. K. Mollah et al.	Ring Delay TDC	1. Occupy less area 2. Mismatch of delay element translates into small cyclic linearity error	1. Time resolution is limited to the delay element 2. The ring cannot be made extremely short	
15	T. E. Rahkonen et al.	Vernier Delay line TDC	1. The structure is simple 2. High resolution compared to flash TDC	1. Power consumption is larger 2. Not very scalable 3. Long delay lines are required to achieve a high dynamic range	
16	[19–21]	Vernier Ring TDC	1. The structure is simple 2. Large dynamic range	1. Two Delay lines are required 2. Thermal noise is more even in the dynamic range of the TDC	
17	R. Szplet et al. [22]	Pulse shrinking delay line TDC	1. It requires only one delay line 2. No Reset is required	Require delay element with asymmetrical propagation delays	
18	J. Borremans et al. [23]	Coarse Fine TDC	1. Power consumption is less 2. Good FoM	Coarse to fine converter gain is to establish which depends on the locking behaviour of the PLL	

(*continued*)

Table 2. (*continued*)

S. no.	Author	Types	Pros	Cons	Remarks
19	[24]	Time amplifying TDC	A very fine time resolution can be achieved	1. Time amplifiers suffer from gain, offset, and linearity problems 2. The large dead time required	
20	Hayun Chung et.al. [25]	Successive Approximation TDC	Achieve sub—gate delay resolution	Mismatch in delays adds non-linearity in the system	
21	M. Z. Straayer et al. [26]	Oversampling TDC	Higher resolution	Charge leakage problem	If technology permits oversampling of signals
22	C. Priyanka et al. [27]	Flash ring TDC	1. Easily implementable and small in area 2. Reconfigurable 3. Power Efficient	Power consumption is limited to toggling of one delay element at each level	

2.2 Loop Filter

This block is used for suppressing the noise and high-frequency component present in the phase detector output and provides a control signal to the DCO. Not all digital phase detectors topologies are compatible with all types of digital loop filters. The inputs of the digital loop filter select which type of phase detector can be used with it [3]. There are two main functions of the filter. First is to determine the stability like how fast the loop achieves lock state, the range over which the loop can achieve lock etc.. Second is limiting the ripple appears in the phase detector output. There are various types of digital loop filter listed in Table 3.

Table 3. Different types of loop filter

S. no.	Author	Type	Pros	Cons	Remarks
1	S. Moorthi et.al. [28]	UP/Down Counter	The design and circuit are very simple	Output pulses do not contain any information like the magnitude of phase error	
2	Roland E. Best	K Counter		It is suited only for EX-OR and JK—FF PDs	Most widely used
3	Roland E. Best	N before M		Non-linear performance	
4	Imran Ali et. al	Digital Filter	Handle low-frequency signals accurately 2. Versatile in processing the signals	1. Complicated architecture 1. Requires a long time to design	

2.3 Digital Controlled Oscillator

It generates a clock signal, and the generated signal is feedback to the input of the phase detector. The comparison of the signals continued until the signal is locked. It is the key component for an ADPLL, which dictates the maximum frequency, the frequency range, and the resolution. For some applications, these features of DCO are the chief considerations. For word signals DCO designing, the main criteria are to provide enough control word resolution and maintain acceptable jitter. There are different types of oscillators listed in Table 4.

3 Design Parameters of the ADPLL

The important design parameters for the ADPLL are [3]:

Table 4. Different types of DCOs

S. no.	Author	Type	Pros	Cons	Remarks
1	M. Song et al. [29]	Path Delay Oscillator	1. Small area 2. Wide tuning range	1. Not suitable for high frequency 2. Phase noise is poor 3. Poor linearity	Cascade many logic gates to form a circular ring oscillator
2	R. Fried [30]	Schmitt Trigger-based current-driven Oscillator		External capacitance is used, which degrades the performance	Formed with Schmitt trigger inverter with a big capacitor and several MOS transistors
3	J. Dunning et al. [31]	Current starved ring Oscillator	1. Generate different frequencies 2. Good linearity	1. The size is large 2. Hardware cost is high	Implemented by controlling the various MOS transistor. Suitable for the microprocessor
4	R. Best	Divide by N Counter	Simple Structure	1. Need high frequency fixed oscillator 2. Jitter design criteria are not good	
5	R. Best	Increment Decrement Counter	Good hold range and lock-in range		Suitable when two input pulses are available
6	R. Best	Waveform Synthesizer	Generate low-frequency signals with high resolution	Generate only sine waves	Suitable for software implementation
7	R. Khalirbaginov et al. [32]	Direct Digital Synthesis DCO		1. Design is complex 2. Requires analogue circuit	Built with phase accumulator, reference tables, and a DAC
8	R. Groza et al.	Combine DAC and ring topology based VCO	Small area and low power consumption	Frequency tuning is trying to achieve by switching the driving strength of the inverter or the load capacitance	Current starved inverter stages are used with variable propagation time w.r.t. control voltage
9	R. Groza et al.	Combine DAC and LC tank based VCO		1. The large area occupied on silicon 2. Poor quality factor at high frequencies	Use passive inductor Frequency tuning is achieved by switching ON/OFF capacitance banks

1. **Frequency Resolution:** It is also called as step size, defines the smallest frequency increment tuning size. It is dependent on the system channel spacing.
2. **Phase noise or Jitter:** It reflects the signal quality. It is the deviation of the signal from its ideal position. Phase noise describes the noise in the phase domain, while jitter describes the noise in the time domain.
3. **Loop Bandwidth:** It indicates the dynamic speed of the feedback loop. It is equal to the natural loop frequency or the frequencies in which the open-loop gain is one**.** The bandwidth of the ADPLL can be changed by setting the parameters of the digital loop filter.
4. **Power Consumption:** It represents the power consumed by the whole circuit. DCO Consumes the maximum power among all the blocks, so it needs to be carefully designed.
5. **Locking time:** It is also known as switching time or settling time. It is the time needed for the PLL to switch the DCO from one frequency to another.

4 Different ADPLLs Architecture and Their Comparison

Synthesized ADPLLs were recently proposed to shorten the design cycle [24]. Synthesized ADPLL are designed using any of the HDL language and its architecture is shown in Fig. 3.

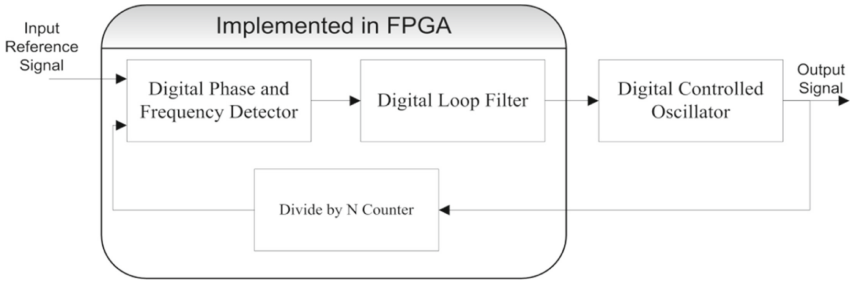


Fig. 3. Synthesized ADPLL architecture

This synthesized ADPLL architecture has the advantage of controlling various features i.e. frequency range, step, jitter etc. which are dependent on DCO designing. From the point of power supply, crosstalk and switching noise it proved to be advantageous [4]. They still have the problem of poor phase noise, limited frequency tuning resolution and large reference spurs. The minimum capacitance of the standard library cells limits the frequency range of the DCO which results in poor jitter performance [2]. The synthesized PLL injection locking technique is adopted in the architecture to reduce the phase noise performance. Injection locking based synthesized ADPLL block diagram is presented in Fig. 4. Injection locking signal frequency is near the free running frequency

of the circuit and it is an external signal. There are different types of injection locking namely first—harmonic, super—harmonic/sub—harmonic, frequency regenerative, edge-injection locking etc. It is costly in terms of silicon area and power consumption.

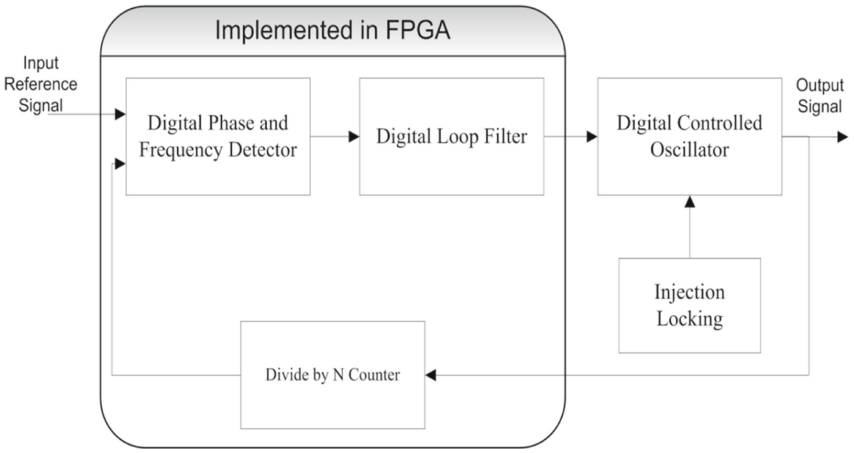


Fig. 4. Injection locking based synthesized ADPLL block diagram

Performance parameters of various ADPLL's are compared in Table 5.

5 Conclusion

In this paper, a detailed progress in research of various blocks used in ADPLL namely phase detector, loop filter, and DCO are presented along with their advantages and disadvantages and their specific applications if any. With the progress in research, the basic architecture of the ADPLL changes to synthesized ADPLL then to injection locking based synthesized ADPLL. Every technique is having its advantage i.e. synthesized ADPLL have a short design period, but they have the problem of poor phase noise and large reference spurs etc. Injection locked synthesized ADPLL improves the phase noise performance of it, but at the cost of the silicon area and power consumption. Depending on the parameter requirement which designer wants to minimize the technique for designing and implementation will be adopted.

Table 5. Comparison of different ADPLL techniques

Author	Techniques	Process	Frequency	Power consumption	Area	Advantages	Applications
Jen Shiun Chiang et al. [34]	Switch tuning DCO Architecture for ADPLL	0.6μm	60 to 400 MHz		923μm* 921μm	Fast frequency locking, easy design, and implementation, reduced hardware cost, good stability	As a clock generator for high-performance microprocessor
R. Groza et al. [33]	DCO based on ring topology and log domain current mode integrators for ADFLL	180nm	90kHz to 3.7 MHz	105μW		The lock is achieved in the frequency domain Jitter performance is good	
Imran Ali	Adaptation technique in the fractional feedback loop of DCO to achieve a highly accurate target frequency	55nm	2.402–2.480 GHz	40 μW	1830μm^2	Ultra-low power and fully synthesizable DPFD	Low energy Bluetooth transceiver
Ahmed Musa et al. [35]	Injection locked dual loop and dual VCO architecture based ADPLL	65nm	Output freq.: 0.5–1.6 GHz Reference Frequency: 40–300 MHz	0.97mW	0.022mm^2	Low jitter, low power, and small area	
J. Yang et al.	Synthesized injection-locked bang-bang PLL	65 nm	Output Frequency—0.45 to 1.8 GHz Reference Frequency—150MHz	1.5mW	0.008 mm2	Best FoM and lowest reference spur	

References

1. De Bellescize, H.: La reception synchrone. L'Onde Electrique **11**, 230–240 (1932)
2. Yang, J., Zhang, Z., Nan QI, Liu, L., Liu J., Wu, N.: A 0.45-to-1.8 GHz synthesized injection-locked bang-bang phase-locked loop with fine frequency tuning circuits. Sci. China Inf. Sci. **62**(6), 1–16 (2019)
3. Best, R.E.: Phase-Locked Loops: Design, Simulation & Applications, 4th edn. McGraw-Hill Professional Engineering (1999)
4. Introduction to FPGA based ADPLLs. Silicon Labs (2011)
5. Chen, F.T., et al.: A 10-Gb/s low jitter single-loop clock and data recovery circuit with rotational phase frequency detector. IEEE Trans. Circuits Syst. I Regul. Pap. **61**(11), 3278–3287 (2014)
6. Imran, A., Oh, S.J., Abbasizadeh, H., Rikan, B.S., Rehman, M.R.U., Lee, D.S. Lee, K. Y.: An Ultra low power fully synthesizable digital phase and frequency detector for ADPLL. In: Applications in 55nm CMOS Technology, International Conference on emerging trends, IEEE (2017)
7. Kumm, M., Klingbeil, H. Zipf, P.: An FPGA-based linear all-digital phase-locked loop. IEEE Trans. Circuits & Syst. – I: Regul. Pap., **57**(9), 2487–97 (2010)
8. Chung, C. C., Sheng D., Chen, C. H. An all-digital phase-locked loop compiler with liberty timing files, IEEE International Symposium on VLSI Design, Automation and Test (2014)
9. Lee, J., Liu, M.: A 20-Gb/s burst-mode clock and data recovery circuit using injection-locking technique. IEEE J. Solid-State Circuits **43**(3), 619–630 (2008)
10. Lin, S.-H., Liu, S.-I.: Full-rate bang-bang phase/frequency detectors for unilateral continuous-rate CDRs. IEEE Trans. Circuits Syst. I, Reg. Pap. **55**(12), 1214–1218 (2008)
11. Staszewski, R.B., et al.: ALL – digital PLL and transmitter for mobile phones. IEEE J. Solid-State Circuits **40**(12), 2469–2482 (2005)
12. Chen, W.H., Inerowicz, M.E., Jung, B.: Phase Frequency Detector With Minimal Blind Zone for Fast Frequency Acquisition. IEEE Trans. Circuits Syst. II Express Briefs **57**(12), 936–940 (2010)
13. Tan, Y.S., Yeo, K.S., Boon, C.C., Do, M.A.: A dual-loop clock and data recovery circuit with compact quarter-rate CMOS linear phase detector. IEEE Trans. Circuits Syst. I Regul. Pap. **59**(6), 1156–1167 (2012)
14. Cai, D., et al.: A dividerless PLL with low power and low reference spur by aperture-phase detector and phase-to-analog converter. IEEE Trans. Circuits Syst. I Regul. Pap. **60**(1), 37–50 (2013)
15. Van Veldhoven, R.H.M.: Analog to digital conversion circuit and method (2011)
16. Yu, J., Dai, F.F., Jaeger, R.C.: A 12-bit vernier ring time-to-digital converter in 0.13um CMOS technology. IEEE J. Solid-State Circuits **45**(4), 830–842 (2010)
17. Rahkonen, T.E., Kostamovaara, J.T.: The use of stabilized CMOS delay lines for digitization of short time intervals. IEEE J. Solid-State Circuits **28**(8), 887–894 (1993)
18. Mollah, A.K.M.K., Rosales, R., Tabatabaei, S., Cicalo J., Ivanov, A.: Design of a tunable differential ring oscillator with short start-up and switching transients. IEEE Trans. Circuits Syst. – Part I: Regul. Pap. **54**(12), 2669–2682 (2007)
19. Mandai, S., Iizuka, T., Nakura, T., Ikeda M., Asada, K.: Time-to-digital converter based on time difference amplifier with non - linearity calibration. In: Proc. Of ESSCIRC, 266–269 (2010)
20. Lee, M., Abidi, A.A.: A 9 b, 1.25 ps resolution coarse-fine time-to-digital converter in 90 nm CMOS that amplifies a time residue. IEEE J. Solid-State Circuits **43**(4), 769–777 (2008)
21. Chen, P., Chen, C.-C., Zheng, J.-C., Shen, Y.-S.: A PVT insensitive Vernier-based time-to-digital converter with extended input range and high accuracy. IEEE Trans. Nucl. Sci. **54**(2), 294–302 (2007)

22. Szplet, R., Klepacki, K.: An FPGA integrated time-to-digital converter based on two-stage pulse shrinking. IEEE Trans. Instrum. Meas. **59**(6), 1663–1670 (2010)
23. Borremans, J., Vengattaramane, K., Giannini, V., Debaillie, B., Van Thillo, W., Craninckx, J.: A 86 MHz–12 GHz digital—intensive PLL for software-defined radios, using a 6 fj/step TDC in 40nm digital CMOS. IEEE J. Solid-State Circuits **45**(10), 2116–2129 (2010)
24. Park, Y., Wentzloff, D.D, An All Digital PLL synthesized from a digital standard cell library in 65 nm CMOS, In Proc. Of IEEE Custom Integrated Circuits Conference (CICC), (2011)
25. Chung, H., Ishikuro, H., Kuroda, T.: A 10-bit 80 ms/s decision—select successive approximation TDC in 65-nm CMOS. IEEE J. Solid-State Circuits **47**(5), 1232–1241 (2012)
26. Straayer, M.Z., Perrott, M.H.: A multipath gated ring oscillator TDC with first-order noise shaping. IEEE J. Solid-State Circuits **44**(4), 1089–1098 (2009)
27. Priyanka, C., Latha, P.: Design and implementation of time to digital converter. In: IEEE Sponsored 2nd International Conference on Innovations in Information Embedded and Communication Systems ICIIECS'15
28. Moorthi, S., Meganathan, D., Janarthanan, D., Kumar, P., Perinbam, J.R.P.: Low jitter all digital phase locked loop based clock generator for high speed system on-chip applications. Int. J. Electron. **96**(11), 1183–1189 (2009)
29. Song, M., Sung I., Pamarti, S., Kim, C.: A 2.4 GHz 0.1-fref-bandwidth all-digital phase-locked loop with delay-cell-less TDC. IEEE Trans. Circuits Syst. – I: Regul. Pap., **60**(12), 3145–3151(2013)
30. Fried, R.: Low-power digital PLL with one cycle frequency lock –in time for clock syntheses up to 100 MHz using 32,768 Hz reference clock. Proc. IEEE **84**, 291–294 (1996)
31. Dunning, J., Gareia, G., Lindberg, J. Nuckolls, E.: An all-digital phase-locked loop with 50 – cycle lock time suitable for high-performance microprocessors. IEEE J. Solis – State Circuits **30**, 412–422 (1995)
32. Khalirbaginov, R., Alexandr H.: Novel HDL Design of Digital Controlled Oscillator for ADPLL, pp. 1678–1682. IEEE (2019)
33. Groza, R., Potarniche, I.A., Kirei, B.S., Topa, M.D.: Digitally controlled oscillator for all digital frequency locked loops. Rom.Ian J. Inf. Sci. & Technol. **21**(1), 3–17 (2018)
34. Chiang, J.S.:, Chen, K.Y.: A 3.3V All digital Phase-locked loop with small DCO hardware and fast phase lock. In: IEEE International Symposium on Circuits and Systems (1998)
35. Musa, A., Deng, W., Siriburanon, T., Miyahara, M., Okada, K., Matsuzawa, A.: A compact low-power and low-jitter dual-loop injection locked PLL using all-digital PVT calibration. IEEE J. Solid State Circuits **49**(1), 50–60 (2014)

On Realization of Smart Logistic Warehouse Management with Internet of Things

Mingyang Liu(✉)

Construction and Installation Engineering Corporation, Fangzheng Forestry Bureau, Heilongjiang, China
hqqguet@guet.edu.cn

Abstract. Internet of Things (IOT) platforms can offer effective solutions in various domains. Logistic warehouse management could be facilitated with the IOT. Warehouse management activities ranging from critical equipment positioning, driver guidance, automated and accurate stock counting, pedestrian safety, shipment validation, staff tracking to consolidated view and overall system manipulation can be realized and optimized with IOT. The Cloud solutions applied in IOT improve the system mobility and integrate many plug-play peripherals into an organic entity. Also it is possible to run the whole system intelligently in the private network with or without internet access. Application Programming Interface (API) ranging from Message Queuing Telemetry Transport (MQTT), Constrained Application Protocol (COAP) to Lightweight Machine to Machine (LWM2M) are applicable and compatible for plug-play peripherals or newly developed applications. The IOT platform, with the employment of both software and hardware entities, ensures that the whole warehouse functions automatically and intelligently. Also, this is the prevailing technology applied in the logistics sector. In a word, within the new world of IOT and Radio Frequency Identification (RFID), business equipment, devices and processes are creating new efficiency, revenue and opportunities for the warehouse management.

Keywords: Realization · Smart · Warehouse Management · IOT

1 Introduction

Warehouse management is an essential sector in the supply chain. And it serves as sourcing, storing, producing, packaging and distributing goods. But complexities of warehouse management often lead to low efficiency in operation, waste of time, labor force and money. So an effective management is crucial for minimizing cost and optimizing the whole supply chain. Core warehouse management functions are receiving, stacking, stockpile, tallying and packaging, and handling goods [1]. And affiliated warehouse management functions can bring added value, including labor force planning, tracking, and terminal scheduling. Cloud severs, IOT communicative devices stamped with bar code, and RFID formulate the transactional foundation for a warehouse management system. The system ensures efficiencies of all work and accurate transmission of information in near real time through all procedures, and creates efficient, stream-lined operations.

R. Kountchev et al. (Eds.): *NAMSP 2022*, SIST 332, vol. 332,, pp. 252–263, 2023.
https://doi.org/10.1007/978-981-19-7842-5_22

2 A Glance at IOT and Smart Warehouse Management

2.1 Internet of Things

IOT employs multi-layer technology to ensure the normal and integral functions of an organic straightforward provisioning, management, and automation of connected peripherals within the Internet of Things system. With multiple connectivity solutions, enterprise-grade surveillance system, and vast data processing capacity, it combines diverse hardware to the cloud. For customers, an IOT platform bears ready-to-use properties that greatly expands applications of connected peripherals and realizes scalability and cross-device compatibility.

2.2 Features of Smart Warehouse Management

IOT-based management system is actually an inventory application that can be used in many domains, such as logistics, tourism, hospitals, oil & gas, transportation etc. In a warehouse, the system offers an accurate surveillance of the whole inventory process through the whole supply chain and improves the efficiency of warehouse operations including receiving, stacking, stockpile, tallying and packaging, handling, etc. With data flow, all labor force activities and equipment operations are traceable and visible. As time goes by, further rectification can be made for the system and it becomes more adaptable to different use cases according to specific situations [2]. The system consists of a few both co-related and independent modules, either handheld or autonomous, connecting the warehouse and the supply chain seamlessly (Fig. 1).

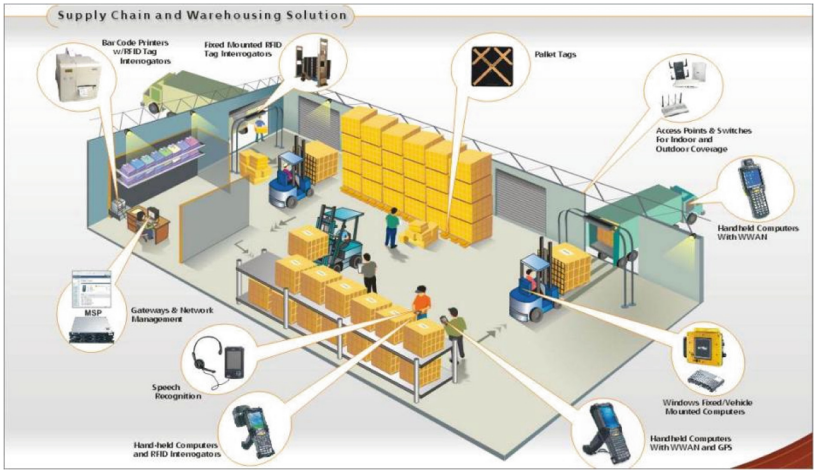


Fig. 1. IOT sketch in the warehouse management

With IOT, a typical smart warehouse management system should fulfill the functions in the following so as to make the whole system operate efficiently, automatically and economically.

Timely alarms. This module gives notifications and/or prevention about any unauthorized handling of inventory items. Alarms can be sent in different perspectives: texting E-mails or messages to staffs, blowing buzzers, flashing red lights, closing doors, noting the case through cameras.

Freight status notification. This module takes a quick glance at what is coming up at the receiving gate and notifies the staff involved to deal with the freight.

Freight inspection. The module closely monitors all operations and functions when receiving freight. Any misfunction or exceptions occurring through the receiving can be rectified. The staff will be granted with login authorities to manage the system [3]. At the end of the inspection process, each item will be stamped with an identification, either with a Bar-code Label or an RFID Tag, to record all relevant information about the item.

Stacking. After cargoes are inspected and labeled, they will be routed to the Stacking Queue. The module ensures that the freight are accurately put away at their own designated position according to different item sort-out commands. If a deviation or error occurs, the system will give an alarm for handmade correction from the staff.

Picking. This module puts goods issue orders in queue and will manage the picking process automatically or manually. This system makes a picking process highly efficient for large quantities of items in huge and complicated warehouse circumstances. A novel algorithm will be run when the system picks any certain items.

Shipping. This module deals with the process before transferring goods to freight companies. It testifies that the goods shipped should comply with orders. If a mistake is made in any previous process the system will recognize it in this phase.

Inventory tallying. In this module both quick stock count and inventory audit of the whole inventory are carried out. And stock counts can be conducted periodically, daily or monthly. Based on item properties, the database can sort out all items and make reports on their status at real time. Inventory audit can give data about found, lost or mis-putaway items. This information helps to boost inventory circulation and improve the whole warehouse management.

Misplaced location. In the module, misplaced items are located and the capacity of the warehouse can be fully exerted. This tool plays an important in cleaning up, tallying and picking.

Inventory positioning. This module helps to locate the where-about of any missing goods. Data flow, visible records and labels guide us to the cargo we want. When the staff approaches close to the target, the reader buzzes and detailed descriptions about the item will be demonstrated on a nearby screen [4]. From the modules, we get powerful management advantages from reducing inventory spending in finished items, raw materials and articles being processed.

Meanwhile, within the system, all information presented in whatever forms, can be recognized and transmitted among all devices. All tools drive a future-proofed e-commerce fulfillment strategy, via a ready-to-use, out-of-the-box software platform. Warehousing businesses leverage the automation execution capabilities, streamline inventory management and drive more efficient order fulfillment workflows. Major properties of a smart warehouse management system can be demonstrated from the following perspectives:

Connectability. Connectability focuses on transmitting messages through the cloud and all plug-play peripherals. It defines how the hardware are connected to the cloud to carry out different performances [5]. The whole system supports open IOT protocols for plug-play devices, Message Queuing Telemetry Transport (MQTT), Constrained Application Protocol (COAP) and so on. It is flexible and tolerant enough and can comply with all existing IOT protocols. The platform can establish new applications that work out under any type of persistent or intermittent IOT connections.

Hardware monitoring and supervision. The platform contains series of electronic and digital hardware, including tangible things, equipment, other entities and so on. Attributes are also assigned to all store devices and they define and describe detailed information on characteristics of all hardware. In this way, simple data processing, or more complex, structural hardware, including a serial of outer peripherals, along with their own properties can be integrated into an organic system and function smoothly. Any device that connects to the system has to offer qualifications or credentials, including pre-alloted keys, tokens, register or login specifications, certificates, etc. [6]. The system monitors and supervises all devices throughout their whole life-cycle, from the provisioning and connectivity events at the very beginning to software upgrades and final decommissioning.

Data Receiving. Under all circumstances, protocols help to collect data from all related peripherals. The protocol guarantees reliable data transmission with response codes, which clarifies the whole procedures of data processing in the system. All device data collected by the system will be distributed to multiple processing thoroughfares. If any error occurs during the data processing, disk damage, or processor overburdening, corresponding alarm will be sent to the device for further adjustment. In this regard, the device always deals with all data with correct and real-time instructions. Data can be varied from structured to unstructured, simple digital numbers or text, or key-value graphs, arrays, charts or nested objects.

Data analytics. The system defines standardized and strict properties to data collection adapters that ensures data transmission to all databases or data analytic platforms. Because of the highly modular architecture, any modification to all existing adapters can be easily conducted. Any raw, unstructured data will be processed and turned into well-structured time series for easy and swift analytics, structural analysis, visualization, and so on.

Configuration management. Configuration management is important because it controls the routine device functions, deals with data processing parameters, ensures analytics, prescribes flagging, and so on. The configuration management guarantees that all functions of the platform can be realized smoothly regardless of all arbitrary data structures [7]. Thus, the configuration data can be processed easily within the system, either historical configuration data or pending delivery can be traced accurately. An integration of software and hardware management in the warehouse system can be seen in Fig. 2.

3 The Realization of Smart Warehouse Management with IOT

With an integral application of Enterprise Resource Planning (ERP), Warehouse Management System (WMS), Radio Frequency Identification (RFID), and other self-developed software and peripherals, a typical IOT-based warehouse management system

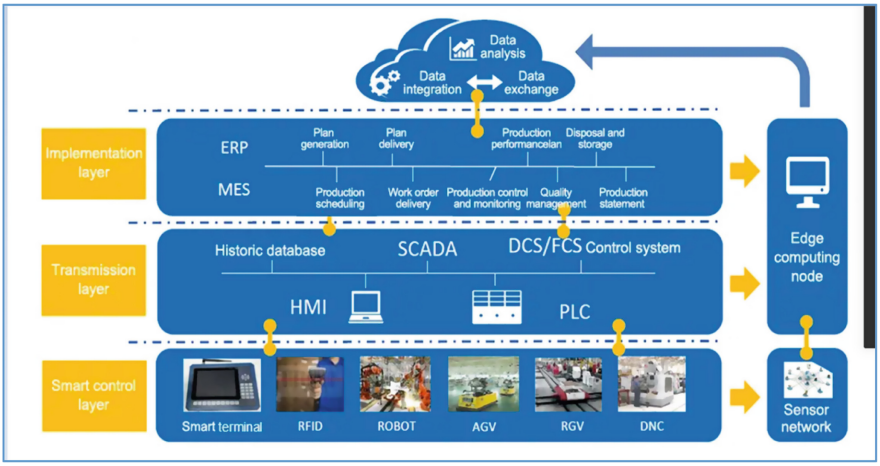


Fig. 2. Data processing in smart warehouse management

becomes more customized and user-friendly to different use cases according to specific work situations. The system consists of a few both co-related and independent modules, either handheld or autonomous, combining the inventory system and the whole supply chain seamlessly.

3.1 Automated and Accurate Stock Counting

At the beginning, the following functions can be fulfilled: real time inventory management, accurate positioning of all items and their united containers. All goods are counted and audited easily within the system, with readers on vehicles, tags or labels at all items, optical devices and handheld computers and interrogators, and so on. A typical use case can be seen in Fig. 3.

Fig. 3. Automated stock counting.

3.2 Critical Equipment Positioning in Real Time

The warehouse system enables complete warehouse visibility without manual scanning of goods. The solution is an all-in-one mapping and positioning system that lets workers accurately track vehicles and goods movements [8]. It enables highly accurate and repeatable indoor and outdoor vehicle location under rapidly changing environments where localization accuracy, quality and consistency are critical. The Radio Frequency Identification (RFID) and real time location system technologies facilitate warehouse management and offer real time monitoring of all process: accurate location, system operation status and reporting, which can be shown in Fig. 4.

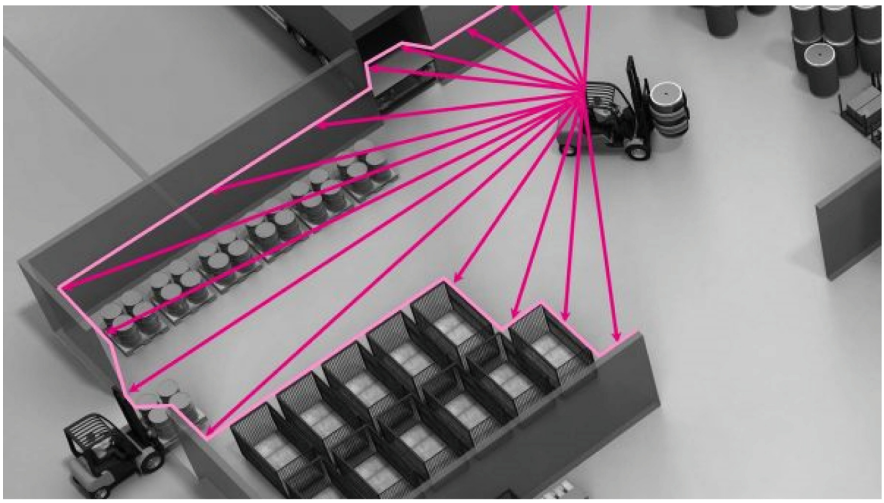


Fig. 4. Real time positioning.

3.3 Pedestrian Safety

The whole plant or warehouse is a workplace, and even the most automated conveyor system needs safe access. The safety department, the training department and plant layout need to be involved. Separating and defining work areas and walkways as well as providing drop-off protection or collision for interior loading and rail docks has long been a challenge in industrial plants and facilities. The IOT can reduce risk and prevent mistakes or accidents which bring about losses in the warehouse by early detection [9]. Sensors in the warehouse can monitor temperature, moisture, and movements of equipment and pedestrians. Warehouse workers or visitors wearing active tags, once detected by the forklift or moving machine, the system activates a stop or slow down of the engine. In this way, timely alarm can be given to both pedestrian and vehicles. Whether autonomous or manual, the IOT system synchronizes all data in an easy-to-access network, optimizing the inventory control, labor force planning. The IOT is the overarching concept that informs how to control all the moving parts of the warehouse, which can be seen in Fig. 5.

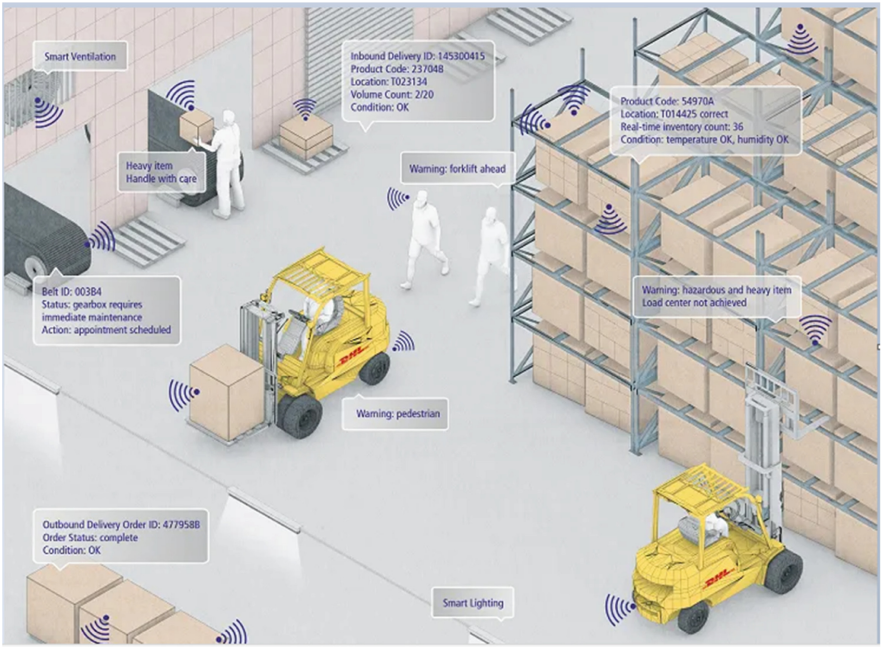


Fig. 5. Pedestrian safety.

3.4 Shipment Validation

Avoiding misplaced cargo on wrong vehicles can prevent loss in time and money, and the system can help us trace every item and guarantee their precise load on a pallet, in a container or a truck [10]. The system equipped with RFID readers and Portal Gates can check it all and automatically and report it to us. Because the smart RFID gates know the shipment planning, we can prevent two common mis-operations: “mistaken load” or “missing goods”. A typical use case can be seen in Fig. 6.

Fig. 6. Shipment validation

3.5 Robots Empower Intelligent Warehouse Management

Technical improvements in sensors and software allow autonomous robots to identify and avoid human workers or other obstacles and to travel to their destination by the shortest route, rather than being limited to defined robot pathways. The nimble robots are instructed to carry parts or finished goods in the warehouse, providing faster delivery and lower costs [12]. In a warehouse facility or distribution center, robots drive around picking up packages and delivering them to stations for storing or processing. For the package-sorting scenario, a Central Scheduler sends commands to robots to pick up packages from the loading station and deliver them to another unloading station. The Robot Controller deploys the trajectory based on the positions of these stations, and generates velocity commands for the robot [13]. These commands are fed to the Plant, which contains a differential-drive robot model for executing the velocity commands and returning ground-truth poses of the robot. The poses are fed back to the scheduler and controller for tracking the robot status (Fig. 7).

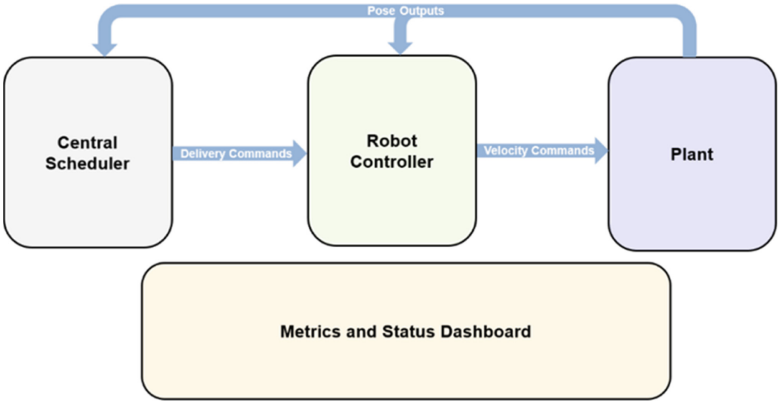


Fig. 7. Sketch of Robotics System Toolbox

The Central Scheduler adopts a workflow chart to deal with package allocation to the robots from the package dispenser [14]. Each robot carries one package at one time and is instructed to start from the loading to an unloading station according to the required position for each package. The scheduler traces the status of all packages and robots. On basis of robot poses, the scheduler sends stop commands to one robot if it detects any imminent collision [15] (Fig. 8).

An array of robot controllers is generated for each robot by The Robot Controller (Fig. 9).

The Plant system adopts the Differential Drive Kinematic Model block to model the movement of each robot.

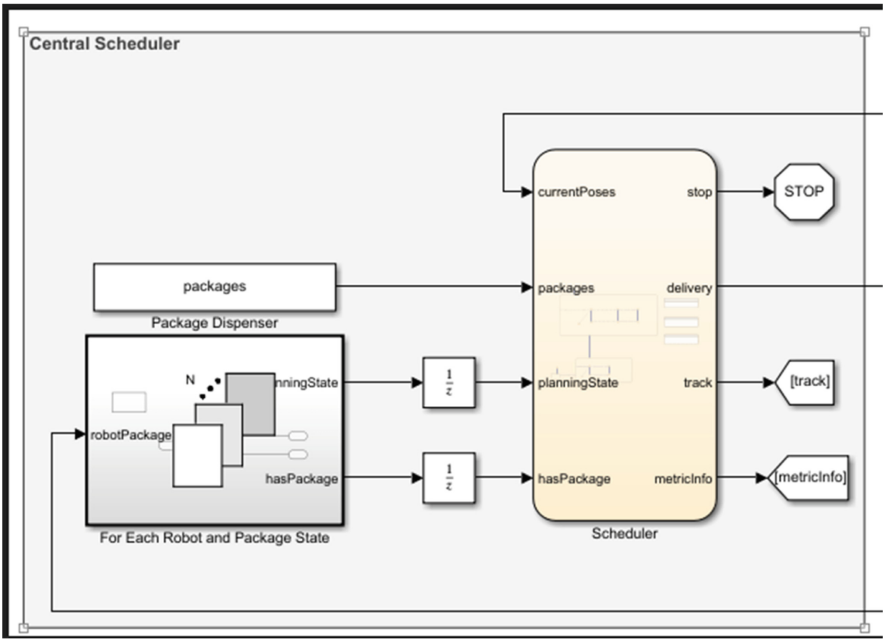


Fig. 8. The mechanism of Central Scheduler

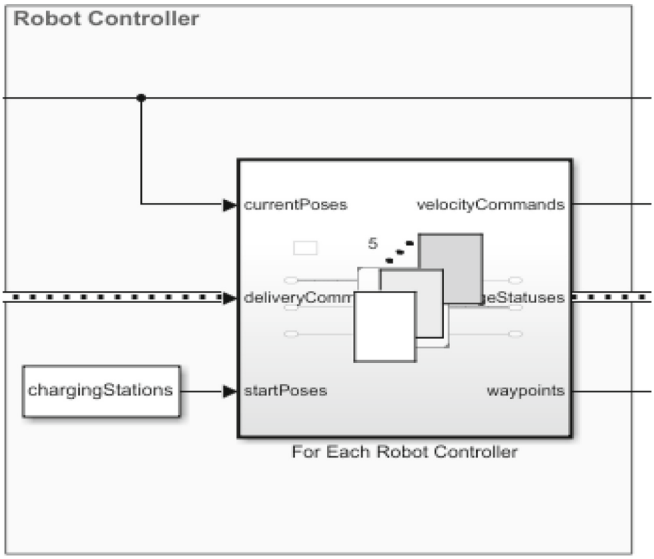
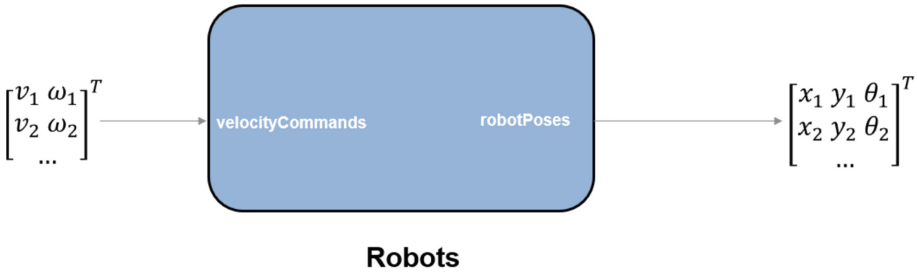


Fig. 9. Array of robot controllers



A logical type matrix, logical map indicates the occupancy map of the warehouse. The facility includes obstacles representing walls, shelves, and all processing stations. Also loading, unloading, and charging stations are given in xy-coordinates [16].

3.6 Open IOT Protocols for Building Event-Based Workflows

Rule engine is powerful so all complex rule chains can be generated to deal with information conveyed through devices and make applications feasible under various use cases and environments.

Message-The platform can handled all incoming information from all possible peripherals and events under different rules.

Rule Node--All incoming information can be processed through filtration, transformation or commands can be executed under various circumstances.

3.7 Conclusion

The IOT technology will make a prominent and profound impact on all sectors of industries. For an intelligent warehouse system in logistics, the integral architecture and accurate information communication are very important. The IOT-based system can handle a complicated supply chain network by combining cloud servers, Radio Frequency Identification (RFID), devices and sensors, tags, bar codes, mobility Wifi, surveillance and other smart connected utilities into an organic whole, which is equally applicable and feasible by the developers and business stakeholders [17]. Whether we are requiring excellent cargo storage or swift handling solutions, optimization of inventory and warehouse operations are crucial elements to be considered. An IOT-based warehouse management system offers a right answer for this logistic sector. Work tasks from receiving, stacking, stockpile, tallying and packaging, and handling and so on are intelligently and automatically conducted. Under the IOT platform, effective and efficient operations are characterized by the following: increases in precision and rapid progress of order processing, decrease of misoperations in all process, decline in costs of inventory management, improvement in traceability and efficiency, optimized handling of freight, reduction over “dead-head” process through the whole management. As a result, intelligent and automatic management is demonstrated in full play within the IOT system: products are staged, stored according to motion velocity; slotting storage locations are conducted based on actual usage record then fast moving products becomes feasible; shelf life supervision is carried out swiftly-tracks expiration and intelligently identifies

non-conforming inventory operations; industrial standard bar-codes such as 2D symbology are used; mobile solution is realized-easy use of hand held scanners, tablets and smart phones to execute transactions and track activity; real-time scheduling of operations including picking, putaway, order staging, dock and container management are monitored; Enterprise Resource Planning (ERP) system integration, embedded Electronic data interchange (EDI), accounting, sales orders, Material Requirement Planning (MRP) and shipping management work both independently and integrally [18].

References

1. Biabani, M.: An Energy-effificient evolutionary clustering technique for disaster management in IOT networks. Sensors **20**, 2647 (2020)
2. Baykasoglu, A., Subulan, K., Karaslan, F.S.: A new fuzzy linear assignment method for multi-attribute decision making with an application to spare parts inventory classification. Appl. Soft Comput. **42**, 1–17 (2016)
3. Fernndez-Carams, T.M.: An intelligent power outlet system for the smart home of the internet of things. Int. J. Distrib. Sens. Netw. **11**(11), 214,805 (2015). https://doi.org/10.1155/2015/214805
4. Mourtzis, D., Vlachou, E.: A cloud-based cyber-physical system for adaptive shop-flfloor scheduling and condition-based maintenance. J. Manuf. Syst. **47**, 179–198. [CrossRef] (2018)
5. Solic, P., Radic, J., Rozic, N.: On software defined radio based implementation of RFID tag in next generation mobiles. IEEE Trans. Consumer Electron. **58**, 1051–1055 (2012)
6. Kim, H.M., Laskowski, M., Zargham, M., et al.: Token economics in real life: cryptocurrency and incentives design for Insolar's blockchain network. Computer **54**(1), 70–80 (2021)
7. Bordel, B., Alcarria, R., Robles, T.: Denial of chain: evaluation and prediction of a novel cyberattack in Blockchain-supported systems. Futur. Gener. Comput. Syst. **116**(11), 426–439 (2021)
8. Rolinck, M., Gellrich, S., Bode, C., et al.: A concept for blockchain-based LCA and its application in the context of aircraft MRO. Procedia CIRP **98**(14), 394–399 (2021)
9. Zhang, X., Liu, S.: Action mechanism and model of cross-border e-commerce green supply chain based on customer behavior. Math. Probl. Eng. **2021**(3), 1–11 (2021)
10. Prajapati, D., Zhou, F., Zhang, M., et al.: Sustainable logistics network design for multi-products delivery operations in b2b e-commerce platform. Sadhana **46**(2), 100 (2021)
11. Kong. X., Zhong, R.Y., Zhao, Z., et al. Cyber physical ecommerce logistics system: An implementation case in Hong Kong. Comput. & Ind. Eng. 2020, 139(Jan.), 106170.1–106170.15
12. Liang, Y., Yao, L.V.: Analysis of the pattern and geopolitical impacts of the agricultural trade between China and ASEAN countries. Resources Science **31**(10), 1654–1661 (2009)
13. McFarlane, D., Sarma, S., Chirn, J.L., Wong, C., Ashton, K.: Auto ID systems and intelligent manufacturing control. Eng. Appl. Artif. Intel. **16**(4), 365–376 (2003)
14. Rubrico, J., Ota, J., Higashi, T., Tamura, H.: Scheduling multiple agents for picking products in a warehouse. In: IEEE International Conference on Robotics and Automation, pp. 1438–1443 (2006)
15. Rubrico, J.I., Higashi, T., Tamura, H., Ota, J.: Online rescheduling of multiple picking agents for warehouse management. Robot. Cim.-Int. Manuf. **27**(1), 62–71 (2011)
16. Koster, R.D., Le-Duc, T., Roodbergen, K.J.: Design and control of warehouse order picking: a literature review. Eur. J. Oper. Res. **182**(2), 481–501 (2007)

17. McFarlane, D., Parlikad, A., Neely, A., Thorne, A.: A framework for distributed intelli-gent automation systems developments. In: IFAC Proceedings Volumes (IFAC-Papers Online), pp. 758–763 (2012)
18. McFarlane, D., Giannikas, V., Wong, A.C.Y., Harrison, M.: Product Intelligence in industrial control: theory and practice. Annu. Rev. Control. **37** (2013)

Development and Design of Intelligent Traction System for Bulk Cargo Terminal

Xingtian Zhang[1], Maoyuan Mou[2(✉)], and Bingfeng Hu[3]

[1] CRRC Qingdao Sifang Rolling Stock Research Institute Co.Ltd, Qingdao, China
[2] JARI Industrial Control Technology Co.Ltd, Qingdao, China
mmyuan0119@163.com
[3] Ningbo Water Environment Group Co, Ltd, Ningbo, China

Abstract. At present, the bulk cargo wharf in the port mostly adopts winch iron ox traction loading, which has high failure rate, poor safety, low transportation efficiency and high maintenance cost in the later stage. In view of the current situation and problems of winch iron ox traction loading, an integrated unmanned intelligent traction loading system is innovatively proposed. The system adopts the structural design of high-performance automatic tractor, combines the development of key technologies such as sliding contact power supply, vehicle number identification, decoupling control, continuous and accurate loading control, and bridging control, and builds a high-quality communication network architecture to solve the defects of the traditional iron ox traction loading process, Realize fully automatic unmanned loading operation. This study can effectively improve the bulk cargo shipping capacity of the port area, reduce the potential safety hazards in the loading area, and achieve the expected effect of energy saving, cost reduction and efficiency increase in the overall operation process.

Keywords: Bulk Cargo Terminal · Continuous Loading · Traction System · Unmanned

1 Introduction

In recent years, more and more old port areas have improved railway transportation capacity through upgrading [1, 2]. Because the iron ox traction system used the steel wire rope guide hook to tow the vehicle for loading, and the end weight is tensioned, the steel wire rope is seriously worn and easy to trip from the guide pulley during the traction process, the maintenance cost is high, which seriously damages the overall performance of the vehicle and affects the safe transportation. In order to meet the requirements of modern railway management and the needs of safe production of port bulk cargo terminals [3, 4], an intelligent traction system suitable for bulk cargo terminal is innovatively designed.

R. Kountchev et al. (Eds.): *NAMSP 2022*, SIST 332, vol. 332,, pp. 264–273, 2023.
https://doi.org/10.1007/978-981-19-7842-5_23

2 Technical Parameters

2.1 Operating Environmental Conditions

Maximum altitude: 2000 m.

Ambient temperature: 20 °C to +45 °C.

Average relative humidity: 95%

Gauge: 1435 mm.

2.2 Main Technical Parameters

The main technical parameters of intelligent traction system are shown in Table 1.

Table 1. Table of technical parameters.

Item	Value
Traction weight	5000 t
Driving mode	Motor-driven
Maximum speed	21 m/min
Working speed	9.5 m/min
Voltage	380 AC
Power	260 kw
Vehicle weight	200 t
Vehicle size	10 × 3 × 4.2 m

3 System Composition

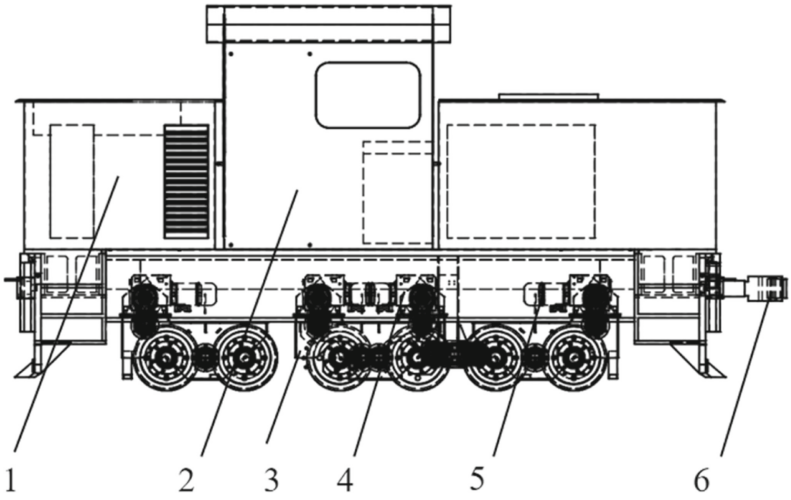
The intelligent traction system is composed of tractor, power supply system, control system, communication system and other systems. In terms of system function, it basically realizes the automatic unmanned loading vehicle, and the process meets the operation requirements of high efficiency and high precision (Table 2).

4 Design of Intelligent Tractor

The intelligent tractor is composed of vehicle body, traveling mechanism, control system, power system, braking system, coupler device, etc. Flexible operation and convenient operation (Fig. 1).

Table 2. Comparison of application and economy between intelligent scheme and traditional scheme.

Item	Traditional locomotive	"Iron ox" traction	Intelligent traction
Efficiency	Middle	Low	High
Accuracy	Middle	Low	High
Control difficulty	Middle	High	Middle
Running cost	High	Middle	High
Energy consumption	High	Low	Low
Pollution	Middle	Low	Low
Security	High	Low	High



1-vehicle body; 2-control system; 3-traveling mechanism; 4-braking system;5-power system; 6-coupler device

Fig. 1. Schematic diagram of intelligent tractor.

4.1 Vehicle Body

The vehicle body is the key component of the tractor and the installation carrier of other components. The car body is welded with all steel, the middle beam bearing steel structure, and the stiffener plate is added at the key stressed parts, which is the main stressed part of the tractor. The vehicle body is composed of underframe, cab, hood, counterweight and auxiliary equipment group.

4.2 Traveling Mechanism

The traveling mechanism carries all the weight of the traction vehicle body and its accessories. The overall structure adopts modular design, which has good universality.

4.3 Dynamic System

The power is driven by 6 variable-frequency motors to realize two-way drive, and the AC variable-frequency speed regulation and soft start scheme are adopted. Frequency converter is a combination of frequency conversion technology and microelectronics technology. It controls the power control equipment of AC motor by changing the frequency mode of motor working power supply [5, 6].

4.4 Brake System

The tractor braking system is mainly composed of JZ-7 air brake, hydraulic motor, air compressor, air cylinder, etc. The braking device adopts JZ-7 air brake used by railway locomotive, which is composed of automatic brake valve, relay valve, equalizing air cylinder, etc. to apply air braking to the tractor.

4.5 Manual Control System

It is used when the central control of the loading building switches the loading mode to manual control. Control system the control system is located in the cab, and the direction handle of the driver controller has forward, 0 position and backward functions. The speed control handle of the driver controller is set with 0 ~ 4 gears. When the tractor is in traction condition, the gear of the speed control handle corresponds to the different set speeds of the tractor, which is the speed control mode [7, 8].

4.6 Coupler

The coupler device is composed of standard gondola coupler, guide mechanism and electric lifting device. The electric lifting device is composed of lead screw, motor, coupling, etc. the electric lifting device is controlled to rise or fall by operating the front and rear coupler lifting buttons in the cab, so that the coupler height can be automatically adjusted within a certain range.

5 Application of Power Supply System

In order to meet the loading operation environment of bulk cargo wharf and the operation requirements of tractor, the power supply system adopts sliding contact line power supply (Fig. 2 and Table 3).

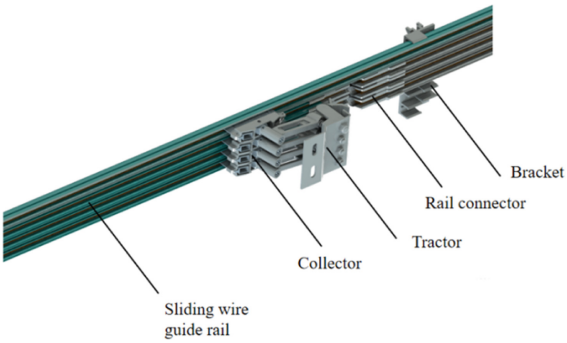


Fig. 2. Sliding contact wire collector structure diagram

Tabel 3. Technical parameters of power supply system.

Item	Value
Type	Unipolar combination
Conductor material	aluminum alloy
Cross sectional area	500 mm^2
Size	32 × 42
Maximum current	700 A
Installation length	2 km

6 Development and Implementation of Control System

The control system is the core of the whole intelligent traction system. It controls the traveling position of the tractor driving carriage to achieve the dynamic response relationship corresponding to the material flow and realize the functional requirements of continuous loading in the loading building and intelligent traction collaborative control. In order to achieve this goal, three measures are taken in the control scheme of the system, namely: train number identification control, hook and uncouple control, continuous loading and crossover control, etc., the overall control flow chart is as follows (Fig. 3).

6.1 Continuous Loading Control

Traction motor Coupling Control. The tractor has 12 wheels in total, including 6 driving wheels driven by high-power motor and 6 driven wheels driven without power. The driving wheel and driven wheel are equipped with encoders. The driving wheel encoder acts as wheel speed feedback and acts on motor synchronization; The driven wheel encoder is used as vehicle position feedback to reflect the actual speed and position of the vehicle.

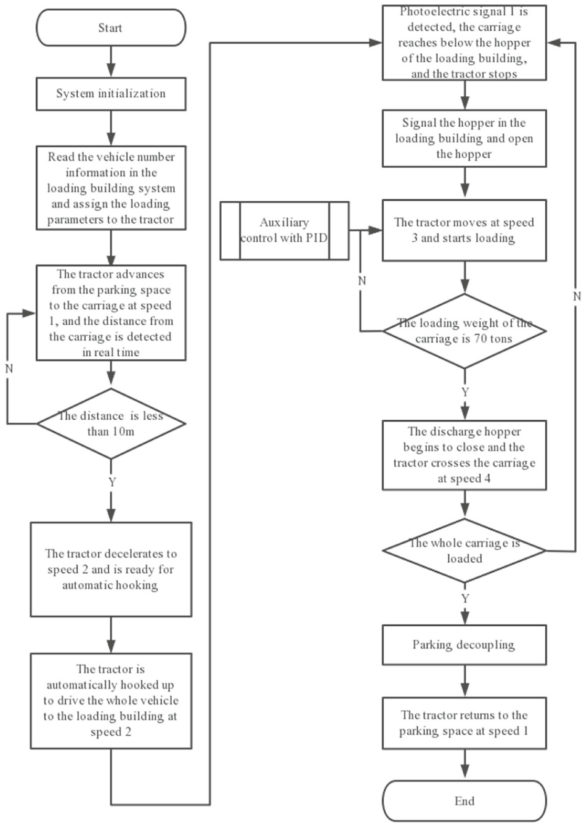


Fig. 3. Flow chart of tractor automatic loading operation control system.

PID Control of Traction Speed. When the loading building controls the traction system for loading, it is necessary to control the speed of the tractor in order to ensure uniform loading and no materials. When the speed of the material conveyor belt in the loading building is certain, the actual speed value of the tractor motor is used as the PID feedback signal, the speed value of the tractor is calculated through the cumulative loading capacity of the load cell as the PID set value input, and the PID calculated output value (control value) is used to adjust the traction speed to achieve the purpose of uniform loading [9–11].

6.2 Adjacent Compartment Management

During loading operation, there is a gap between carriages. When the gap runs below the loading hopper, it is necessary to close the loading hopper to prevent materials from leaking out of the carriage, causing waste and affecting the driving of vehicles. With the cooperation of photoelectric sensor and load cell, when the cumulative loading of load cell is close to the set value, the loading hopper is controlled to close. When the loading hopper is closed, the carriage just reaches the crossover. When the photoelectric sensor

detects that the next carriage reaches the loading position, control the loading hopper to open and start loading again (Figs. 4 and 5).

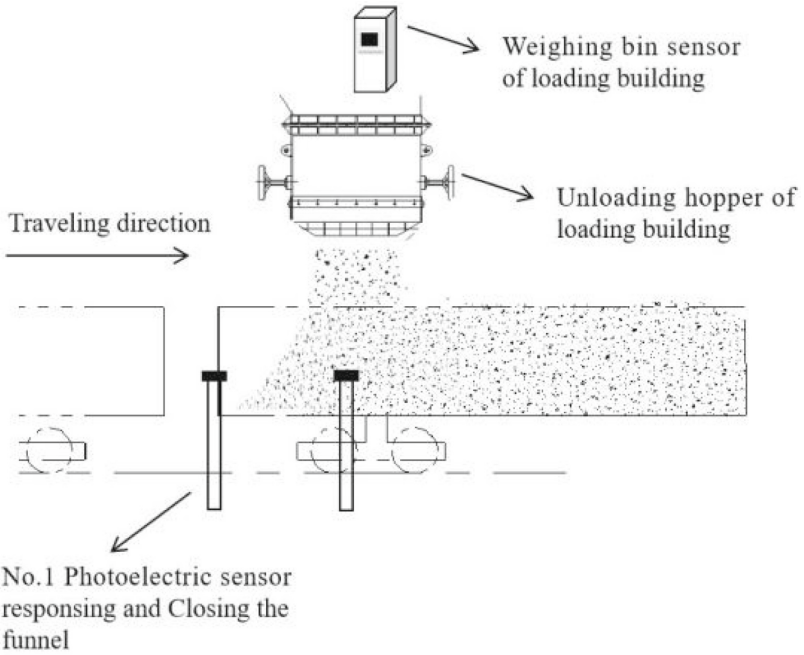


Fig. 4. Schematic diagram of controlling before carriage jumper.

7 Construction of Communication System

The communication system is divided into two parts: the communication between the upper computer and the control system of the loading building, and the communication between the upper computer and the local PLC system of the tractor. By means of 5G technology and wireless relay, the interactive transmission of information among the three is realized. At the same time, is detected the communication status. If communication interruption is detected, immediately control unit commands the charging hopper to close and the tractor to stop, and sends an alarm to the on-site personnel to ensure the safety of on-site personnel and equipment (Fig. 6).

8 Conclusions

The design of the system solves the safety hidden danger of “iron ox” traction in bulk cargo wharf, and better ensures the good running state of vehicles. It has improved the safety performance of loading operation, improved the production efficiency of car loading, reduced the cost of maintenance and use, and advanced the system’s performance

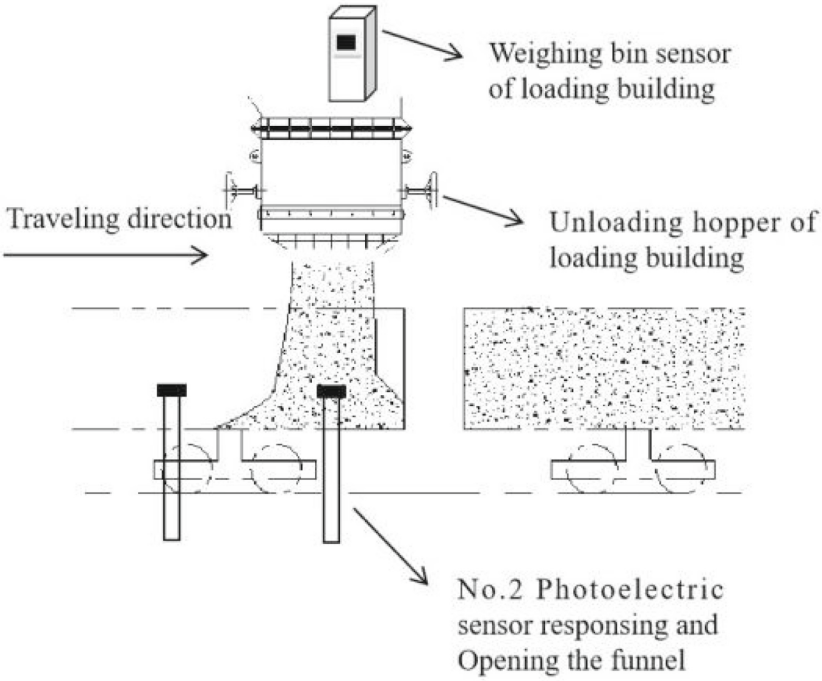


Fig. 5. Schematic diagram of controlling after carriage jumper.

and safety and reliability, and achieved the leading level in China. It has outstanding economic and social benefits and obvious popularization and application value, which has laid a foundation for the development of enterprise railway transportation.

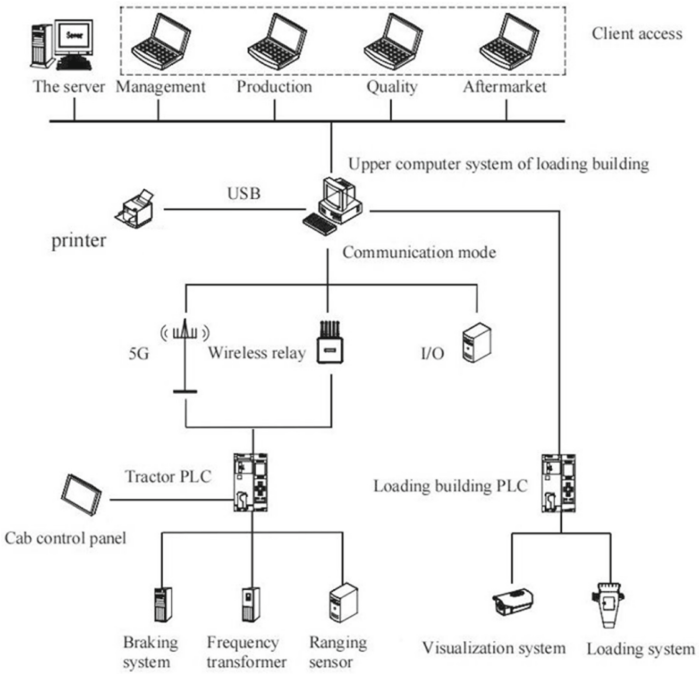


Fig. 6. Communication network architecture.

References

1. Wen, S., Zeng, X., Lang, S.: Suggestions on the development of intelligent handling of bulk cargo terminals. Port Technol. (1), 47–49 (2019)
2. Xingming, L., Meng, Y., Hao, L., et al.: Problems and Countermeasures in loading and unloading operation of automated bulk cargo terminal. Water Transp. Eng. **6**, 204–207 (2018)
3. Yuelin, X., Ronglin, L.: Design and implementation of coal mine railway loading traction system. Coal Mine Mach. **8**, 19–21 (2014)
4. Yin, Q., Yang, X.: Stability analysis of endless rope continuous tractor system. Coal Min. Mach. **32**(6), 98–100 (2011)
5. Yingchun, L., Xi, B., Haifeng, Z., Zhengang, Z.: Research and application of transmission mechanism of coarse crushing lower long ore feeding belt conveyor. China Molybdenum Ind. **43**(06), 30–32 (2019)
6. Qu, X.: Research and application of new track changing iron cattle in Coal Mine Railway Rapid Loading System. Coal Mine Mach. **33**(6), 181–182 (2012)
7. Niu, C., Zhang, J.: Development of four-wheel drive road and railway tractor. Mech. Eng. Autom. (2), 113–115 (2020)
8. Liu, X., Ma, C.: Application of iron bull shunting system in railway freight unloading. Soda Ash Ind. (3), 21–23 (2016)
9. Zhu, X.: A complete set of integrated full-function automatic loading system technology and its application. Port Sci. Technol. (3), 3–6 (2014)

10. Li, J., Zhou, Y., Ren, J., Xue, P.: Control system design of fully automatic mobile flat loading machine. Port Load. Unloading (6), 19–23 (2018)
11. Zhang, H.: Design of automatic electrical control system for continuous loader. Hoisting Transp. Mach. (5), 13–16 (2014)

Author Index

R. Kountchev et al. (Eds.): *NAMSP 2022*, SIST 332, vol. 332,, pp. 275–276, 2023.
https://doi.org/10.1007/978-981-19-7842-5