

Peter Fleming  
Nalinaksh Vyas  
Saeid Sanei  
Kalyanmoy Deb  
*Editors*

# Emerging Trends in Electrical, Electronic and Communications Engineering

Proceedings of the First International  
Conference on Electrical, Electronic and  
Communications Engineering (ELECOM  
2016), Bagatelle, Mauritius, November  
25–27, 2016

# Lecture Notes in Electrical Engineering

Volume 416

## Board of Series editors

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

### *About this Series*

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

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

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

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

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

Peter Fleming · Nalinaksh Vyas  
Saeid Sanei · Kalyanmoy Deb  
Editors

# Emerging Trends in Electrical, Electronic and Communications Engineering

Proceedings of the First International Conference  
on Electrical, Electronic and Communications  
Engineering (ELECOM 2016),  
Bagatelle, Mauritius, November 25–27, 2016

 Springer

*Editors*

Peter Fleming  
ACSE  
University of Sheffield  
Sheffield  
UK

Saeid Sanei  
Department of Computer Science  
University of Surrey  
Guildford  
UK

Nalinaksh Vyas  
Department of Mechanical Engineering  
Indian Institute of Technology  
Kanpur  
India

Kalyanmoy Deb  
Michigan State University  
East Lansing, MI  
USA

ISSN 1876-1100

ISSN 1876-1119 (electronic)

Lecture Notes in Electrical Engineering

ISBN 978-3-319-52170-1

ISBN 978-3-319-52171-8 (eBook)

DOI 10.1007/978-3-319-52171-8

Library of Congress Control Number: 2016963329

© Springer International Publishing AG 2017

This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.

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

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

Printed on acid-free paper

This Springer imprint is published by Springer Nature

The registered company is Springer International Publishing AG

The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

# Preface

The first International Conference on Emerging Trends in Electrical, Electronic and Communications Engineering (ELECOM 2016) was held in Mauritius during 25–27 November 2016. The main theme of this conference was “Advances in Research Through Innovative Technologies” which covered a wide spectrum of theoretical and applied issues in electrical, electronic, communications engineering and computing/IT. The main objective of this conference was to provide a platform for researchers to interact and present high-calibre research. The participants shared their expertise, discussed areas for collaboration and presented their latest findings and innovations.

ELECOM 2016 was organised by the University of Mauritius, a leading tertiary education institution. The event received positive responses from a multitude of stakeholders including public and private sectors, local and international universities and research institutions. Twenty eight papers were accepted following a rigorous blind review process and all papers were screened using Turnitin. In addition to the accepted papers, the present proceeding includes three invited papers by prominent researchers. There were three keynote addresses by the following experts: Prof. Michael Faulkner from the Victoria University gave a keynote address on the fifth generation frontier, mm-waves; Prof. Miloud Bessafi from University of Reunion Island gave a keynote address on the high spatial resolution mapping of the monthly climatology of global solar radiation observed in Mauritius; Prof. Krishna K. Busawon from the Northumbria University gave a keynote address on the design and implementation of chaos based communication systems.

This book contains a selection of revised manuscripts presented at ELECOM 2016 within the four main conference themes, which are electrical engineering, electronic engineering, communications engineering and computing/IT. This book covers various aspects of advanced electrical and communications engineering and computing.

The conference chairs wish to acknowledge the University of Mauritius for successfully organising the conference and, the Mauritius Research Council, Tertiary Education Commission and ABB Group - Mauritius for sponsoring this event.

November 2016

Robert T.F. Ah King  
Tulsi Pawan Fowdur

# Message from Honorary Chair

It is my great pleasure to welcome you to the first International Conference on Emerging Trends in Electrical, Electronic and Communications Engineering (ELECOM 2016), held in Mauritius, a volcanic island of lagoons and palm-fringed beaches with coral reefs surrounding most of the coastline. Mauritius is a multi-cultural nation, known for its dreaming beaches, lagoons and reefs with colourful mountainous and rainforests, waterfalls, hiking trails and wildlife yet carrying a long history of cultural and social changes.

The conference, held during November 25–27, 2016, is a premier forum for the exchange of ideas, open and direct discussion in the fields of electrical, electronics and communication engineering as well as computing/IT, and its related disciplines. It attracted high-quality and impact research articles in science and technology which were peer reviewed by eminent scholars. Therefore, these proceedings include rigorously peer-reviewed articles submitted by research scholars from all over the world reflecting the conference value as an international event. We were also pleased to see the broad range of subjects addressed in these submissions which covered the conference themes.

I wish to use this opportunity to express my appreciation to the excellent Organising Team, Program Committee Members and reviewers for their hard work and valuable contribution. Certainly, ELECOM would not have taken place without such an outstanding effort.

In addition, the conference was graced by outstanding keynote speeches delivered by international expert guest speakers.

I trust that this conference has started on a very solid foundation and indeed the forthcoming editions of this conference will have an even greater impact while significantly extend their reach to the international community in the related fields.

Saeid Sanei  
Honorary Chair, ELECON 2016



# Proceedings Editors



**Professor Peter Fleming** Peter Fleming is Professor of Industrial Systems and Control in the Department of Automatic Control and Systems Engineering at the University of Sheffield, UK. His control and systems engineering research interests include control system design, system health monitoring, multi-criteria decision-making, optimisation and scheduling, and applications of e-science.

He has over 400 research publications, including six books, and his research interests have led to the development of close links with a variety of industries in sectors such as automotive, aerospace, energy, food processing, pharmaceuticals and manufacturing.

He is a Fellow of the Royal Academy of Engineering, a Fellow of the International Federation of Automatic Control, a Fellow of the Institution of Engineering Technology, a Fellow of the Institute of Measurement and Control, and is Editor-in-Chief of *International Journal of Systems Science*. He was Director of the Rolls-Royce University Technology Centre for Control and Systems Engineering (1993–2012).

Further details may be found at [http://www.shef.ac.uk/acse/staff/peter\\_fleming](http://www.shef.ac.uk/acse/staff/peter_fleming).



**Dr. Saeid Sanei** Saeid Sanei received his PhD in Signal and Processing from Imperial College London, UK. He has been a member of academic staff in Iran, Singapore, and the UK. He has published three books in applications of signal processing to biomedical signals particularly electroencephalographs, a number of book chapters, and over 340 papers in peer-reviewed journals and conference proceedings. His research interest is in nonlinear and adaptive signal processing, cooperative learning, multi-way, multimodal, and multichannel signal processing, blind source separation, compressive sensing, and deep neural networks with

applications to biomedical data, audio and speech, biometrics, body sensor networking, and communications. He has served as Associate Editor for the IEEE Signal Processing Letters, IEEE Signal Processing Magazine, and Journal of Computational Intelligence and Neuroscience, and a Guest Editor for many other journals. Dr Sanei is a recipient of a number of best paper awards in brain-computer interfacing, EEG analysis, and other signal processing applications. Currently, he is with University of Surrey, UK, and a Visiting Academic to Imperial College London, UK.



**Professor Kalyanmoy Deb** Kalyanmoy Deb is Koenig Endowed Chair Professor at the Department of Electrical and Computer Engineering in Michigan State University (MSU), East Lansing, USA. He also holds a professor position at the Department of Computer Science and Engineering, and at the Department of Mechanical Engineering in MSU.

Professor Deb's main research interests are in genetic and evolutionary optimisation algorithms and their application in optimisation, modelling, and machine learning. He is largely known for his seminal research in developing and applying evolutionary multi-objective optimisation. His Computational Optimization and Innovation (COIN) Laboratory (<http://www.coin-laboratory.com>) at Michigan State University promotes research in these areas. He consults with various industries and software companies.

Professor Deb was awarded the prestigious 'Infosys Prize' in 2012, 'TWAS Prize' in Engineering Sciences in 2012, 'CajAstur Mamdani Prize' in 2011, 'JC Bose National Fellowship' in 2011, 'Distinguished Alumni Award' from IIT Kharagpur in 2011, 'Edgeworth-Pareto' award in 2008, Shanti Swarup Bhatnagar Prize in Engineering Sciences in 2005, 'Thomson Citation Laureate Award' from Thompson Reuters. Recently, he has been awarded a Honorary Doctorate from University of Jyväskylä, Finland. His 2002 IEEE-TEC NSGA-II paper is judged as

the Most Highly Cited paper and a Current Classic by Thomson Reuters having more than 20,000+ citations. He is a fellow of IEEE, ASME, Indian National Science Academy (INSA), Indian National Academy of Engineering (INAE), Indian Academy of Sciences (IASc), and International Society of Genetic and Evolutionary Computation (ISGEC). He has written two textbooks on optimisation and more than 435 international journal and conference research papers with Google Scholar citations of 90,000+ with h-index of 100. He is in the editorial board on 20 major international journals. More information about his research can be found from <http://www.egr.msu.edu/~kdeb>.



**Professor Nalinaksh S. Vyas** Professor Nalinaksh S. Vyas has been at the Indian Institute of Technology Kanpur since 1987 and is currently the Chairman of the Technology Mission for Indian Railways (TMIR), announced by the Government of India in March 2015. He obtained his B. Tech. in Mechanical Engineering from IIT Bombay (1980) and M.Tech. (1983) and Ph.D. from IIT Delhi (1986).

At IIT Kanpur, Prof. Vyas has served as the Head of Mechanical Engineering Department, Nuclear Engineering Department, Centre for Mechatronics and the Innovation Laboratory. He has also been a Visiting Professor at Virginia Tech, USA; INSA Lyon, France; Lulea University Sweden and National Chung Cheng University. He also served as the Vice Chancellor of Rajasthan Technical University from 2013 till 2015. He has supervised more than ninety doctoral and masters students and has nearly one hundred international publications to his credit.

His research interests lie in the areas of turbomachinery dynamics, nonlinear parameter estimation, instrumentation and integrated health monitoring of machinery. He has executed more than thirty major projects for organisations like the Aeronautical Research & Development Board, Department of Science & Technology, Indian Space Research Organisation, Ministry of Railways, TATA Consultancy Services, Larson & Toubro, Scooters India, HAL, NAL, GTRE among others.

Professor Vyas has been the National Coordinator for the Technology Mission on Railway Safety, Govt of India; Chairman, Automotive Parc, National Program on Smart Matls & Structures, Govt of India; Project Coordinator, Nano-Satellite, JUGNU Project with ISRO; Consortium Leader, Automotive Electronic Stability Program, Core Group on Automotive Research, TIFAC, DST; Member, Expert Task Force on IVHM of LCA (Light Combat Aircraft); Editor (mid-Asia), International Journal of Engineering Asset Management; Editor, ISSS Journal (International Society of Smart Systems); Editor, Advances in Vibration Engineering; Member, Indo-US Task Force on Embedded Systems; Founding Director, International Society on Asset Management, Australia; Member, Working

Group on Energy Research, Planning Commission, Govt of India. Currently, he is also on the Board of Governors, IIT Jodhpur and on the Executive Council of the AICTE.

Professor Vyas has been featured by India Today (2010), as one of the 20 Innovators Changing Our Lives. He has been awarded the Awadh Samman, 2009 and the Madhavrao Scindia Leadership Award, 2012.

## Plenary Speakers



**Professor Michael Faulkner** Michael Faulkner received a B.Sc.(Eng) from London University, UK, and a Ph.D. (1993) from the University of Technology, Sydney, Australia. He is Professor in Telecommunications at Victoria University, Melbourne, Australia. He co-founded and led the Centre for Telecommunications and Micro-Electronics (CTME) 2002–2008. Professor Faulkner has worked in amplifier linearisation and low energy (green) wireless transmitters. He is now an expert in the use of signal processing to correct for radio frequency circuit imperfections and received a prize for an article on this

topic in the IET Proceedings on Communications. Professor Faulkner has been involved in standardisation and commercialisation activities in the IEEE802.11 (WLAN) space and is well connected with the wireless industry. He has supervised research projects in radio propagation measurements (wideband channel sounding, direction of arrival, etc), MIMO, transceiver algorithms, architectures and circuits, physical layer signal processing, and modulation. His research interests cover all areas of wireless system design and his current activities are focused on cognitive radio, flexible transceiver design and mm-waves for future wireless systems. He has authored or co-authored over 100 publications.

## Keynote Speech: The 5G Frontier; mm-Waves

With the introduction of the fourth generation (4G) of wireless equipment almost complete, the focus of the research community has switched to the fifth generation, targeted for commercialisation in 2020. Increased data rates, a renewed focus on the Internet-of-Things and the scarcity of spectrum will force operators into higher frequency bands despite deteriorating performance in terms of coverage. The new

mm-wave bands under consideration offer both the opportunity for wider bandwidths and the challenge of providing the coverage. Repeaters might be necessary to extend coverage zones. A number of research organisations are doing measurements to better understand how the mm-wave bands behave in different environments. The presentation will describe the mm-wave measurement program currently underway at Victoria University, which aims to identify performance issues under local conditions.



**Professor Miloud Bessafi** Miloud Bessafi is Full First Class Professor of Meteorology and Physics of Environment at the University of Reunion Island, and member of LE2P laboratory. From 1991 to 1998, he was a member of LPA (Laboratory of Physics of Atmosphere) at University of Reunion Island working on Tropospheric ozone, Stratosphere-Troposphere exchange, Aerosol, and Rainfall. From 1998 to 2003, he was a member of CRC (Tropical Cyclone Research Group) at Météo-France (DIRRE) working on Tropical Cyclone Forecast and on tropical equatorial waves.

Since 2004, Prof. Bessafi is a member of LE2P at University of Reunion Island working on solar energy, mapping of solar irradiation, solar variability in southwest Indian Ocean and climate change. He was First Vice-Dean of the faculty of science and technology from 2007 to 2012. Since 2010, Prof. Bessafi is Co-Investigator in French Research Program ANR blanche « cyclobulle » LE2P (EA-4079)-LOMA (UMR CNRS/Université de Bordeaux).

From 2011 to 2013 he was a member of Regional Committee of Africa Thorpex, Coordinator of the research activities of East Africa and an active member in French Research Program VOASSI (Variabilité Océan-Atmosphère du Secteur Sud-Ouest de l’océan Indien). In 2013, he was the Principal Investigator of Regional project GEOSUN (GisEment sOLAire : meSures, evalUation et coopération) and AXIS\_OI (Analyse des Corrélations de l’Irradiance Solaire dans la zone sud-ouest de l’Océan Indien).

Professor Bessafi has published in topics related to atmospheric physics, aerosols, tropical meteorology, and renewable energy. He has been the advisor for 14 doctoral students. He also contributed as a scientific reviewer for the Journal of Geophysical Research, Journal of Climate, Weather and Forecasting Review, Journal of Atmospheric Science, Monthly Weather Review, and Theoretical and Applied Climatology. Professor Bessafi’s scientific expertise was sought for National Science Foundation Identification program: Pro ATM - GEO/ATM - Climate & Large-Scale Dynamics in August 2008. He was an invited guest at Meteorological Institute of University of Hamburg during June–July 2006.

## **Keynote Speech: High Spatial Resolution Mapping of the Monthly Climatology of Global Solar Radiation Observed in Mauritius**

Over the past few years, the development of photovoltaic energy in Mauritius has become a necessity in the context of the sustainable approach. The exploitation of renewable energy and in particular solar energy is a reality for the island. In the context of climate change adaptation, expansion of related industries is expected in the medium and long term for the island. As such, solar power through photovoltaic farms will be an important vector of energy production for Mauritius. However, knowledge of the solar radiation on the island both spatially and temporally are still poorly mastered. To our knowledge, there was no work on the spatial distribution of high resolution solar radiation for Mauritius. A downscaling approach for estimating the monthly solar radiation with a resolution of 90 m on the satellite observation based CM-SAF with 3 km spatial resolution will be presented. The methodology used took in account the effects of terrain on observation and anisotropy of solar radiation effects. Finally, the results of the downscaling method with a presentation of the mapping of the monthly climatology of global solar radiation observed in Mauritius will be shown.



**Professor Krishna K. Busawon** Krishna Busawon is a Professor in Control Systems Engineering and is currently the head of Nonlinear Control research group in the Faculty of Engineering and Environment at Northumbria University. He obtained his first degree in Mathematics and Fundamental Sciences from University of St-Etienne in 1989. He then pursued his studies at the University of Lyon where he obtained his BEng and MSc Degree in Electrical Engineering in 1990 and 1991, respectively. He went on to continue his postgraduate studies in the same university and consequently he obtained his MPhil

and PhD degree in Control Systems Engineering in 1992 and 1996, respectively. After his PhD he was appointed as a Research Fellow at Simon Fraser University in 1997. He then joined the University of Nuevo León in Mexico where he worked as a Lecturer in the Department of Mechanical and Electrical Engineering (FIME). In the year 2000, he joined Northumbria University where he was appointed as a Senior Lecturer in the School of Computing, Engineering and Sciences. Later in 2006, he became a Reader in Control Systems Engineering and a Professor since February 2013 at the same university.

His research interest lies mainly in the area of mathematical modelling, nonlinear control and observer design, fault detection and isolation with application to various engineering disciplines such as mechanical, power, communication and biotechnological systems. He has published over 160 research papers in his area of research.

## **Keynote Speech: Design and Implementation of Chaos Based Communication Systems**

In this talk, we shall give an overview of chaos based communication systems. Unlike traditional communication systems, chaos based communication systems or simply chaotic communication uses chaotic oscillators to transmit and receive messages. It is shown that such communication systems provide an extra level of security for data transmission. However, it also bring along some serious difficulties with regards to message recovery. In this talk, we shall also discuss the various available techniques for data recovery as well as the possibility of employing standard modulation techniques in the context of chaotic communication.

# **Committee**

## **Honorary Chair**

Saeid Sanei

University of Surrey, UK

## **Organising Committee**

### **General Chair**

Robert T.F. Ah King

University of Mauritius, Mauritius

### **General Co-chair**

Tulsi Pawan Fowdur

University of Mauritius, Mauritius

### **Secretary**

Baby Gobin

University of Mauritius, Mauritius

## **Technical Programme Committee (TPC)**

### **TPC Chair**

S.Z. Sayed Hassen

University of Mauritius, Mauritius



**Co-chairs**

M.I. Jahmeerbacus                      University of Mauritius, Mauritius  
M.A. Hosany                              University of Mauritius, Mauritius

**Website and Advertisement Committee****Chair**

Anshu Prakash Murdan                      University of Mauritius, Mauritius

**Co-chair**

Rajeshree Ramjug-Ballgobin                      University of Mauritius, Mauritius

**Members**

Oree Vishwamitra                      University of Mauritius, Mauritius  
Baby Gobin                              University of Mauritius, Mauritius  
Nuzhah Gooda Sahib                      University of Mauritius, Mauritius  
Visham Hurbungs                      University of Mauritius, Mauritius

**Finance and Registration Committee****Chair**

Vandana Bassoo                      University of Mauritius, Mauritius

**Co-chair**

Bhimsen Rajkumarsingh                      University of Mauritius, Mauritius

## **Members**

Oree Vishwamitra                      University of Mauritius, Mauritius  
Chooramun Nitish                      University of Mauritius, Mauritius

## **Paper Submissions Committee**

### **Chair**

Visham Hurbungs                      University of Mauritius, Mauritius

### **Co-chair**

Vidasha Ramanarain Seetohul      University of Mauritius, Mauritius

## **Logistics Committee**

### **Chair**

Rameshwar Ashvin Jugurnauth      University of Mauritius, Mauritius

### **Co-chair**

Heman Shamachurn                      University of Mauritius, Mauritius

## **Members**

Abdel Khoodaruth                      University of Mauritius, Mauritius  
Santaram Venkannah                      University of Mauritius, Mauritius

## Proceedings Committee

### Chair

Nitish Chooramun                      University of Mauritius, Mauritius

### Co-chair

Avinash Mungur                      University of Mauritius, Mauritius

### Member

Vidasha Ramanarain Seetohul      University of Mauritius, Mauritius

## International Technical Program Committee Members

Peter Fleming	University of Sheffield, UK
Nalinaksh Vyas	Indian Institute of Technology Kanpur, India
Kalyanmoy Deb	Michigan State University, USA
Poompat Saengudomlert	Bangkok University, Thailand
Zoran Bojkovic	University of Belgrade, Serbia
Chandra Sekhar Paidimarry	Osmania University, Telangana, India
Keshav Dahal	University of the West of Scotland, UK
Santosh Nagaraj	San Diego State University, USA
Cheng-Hsiung Hsieh	Chaoyang University of Technology, Taiwan
Nikos Mastorakis	Technical University of Sofia, Bulgaria, Hungary
K.V. Mahendra Prashanth	SJB Institute of Technology, Bangalore, India
João Duro	University of Sheffield, UK

## List of Trackchairs and Reviewers

### Trackchairs

Robert T.F. Ah King                      University of Mauritius

Yasdeo Bissessur	University of Mauritius
Bhimsen Rajkumarsingh	University of Mauritius
Avinash Mungur	University of Mauritius
Vishwamitra Oree	University of Mauritius
Santaram Venkannah	University of Mauritius
Mussawir Hosany	University of Mauritius
Nuzhah Gooda Sahib	University of Mauritius, Mauritius
M. Iqbal Jahmeerbacus	University of Mauritius

## Reviewers

Glenn Bright	University of Kwazulu-Natal, South Africa
Andre Nel	University of Johannesburg
Janez Trontelj	University of Ljubljana, Slovenia
S. Poompat	Bangkok University, Thailand
Xuemin Chen	Texas Southern University, USA
Tulsi Pawan Fowdur	University of Mauritius
Sarjoosing Goolaup	Senior Research Fellow, Nanyang Technological University, Singapore
Girish Bekaroo	Middlesex University, Mauritius branch campus
Vandana Bassoo	University of Mauritius
Rameshwar Ashvin Jugurnauth	University of Mauritius
G. Ramsawock	University of Mauritius
Pablo Orduña	DeustoTech - Deusto Institute of Technology, Spain
Kehinde Awodele	University of Cape Town, South Africa
Ayman Faza	Princess Sumaya University for Technology, (PSUT), Jordan
Nitish Chooramun	University of Mauritius
Aatish Chiniyah	University of Mauritius
Razvi Doomun	University of Mauritius
Amal Bholah	CEO of Smart Health Ltd
Sudha Cheerkoot-Jalim	University of Mauritius
Rajeshree Ramjug-Ballgobin	University of Mauritius
Muhammad Yaasir Khodabacchus	Senior IT Consultant at PricewaterhouseCoopers
Vidasha Ramnarain-Seetohul	University of Mauritius
Nassirah Laloo	University of Technology, Mauritius
Raj Moloo	University of Mauritius
Leckraj Nagowah	University of Mauritius
Shehzad Jaunbuccus	University of Mauritius
Roopesh Kevin Sungkur	University of Mauritius
Yogesh Beeharry	University of Mauritius
Heman Shamachurn	University of Mauritius
Jing-Min Wang	St. John's University, Taiwan

## Sponsors



**Mauritius Research Council**

<http://www.mrc.org.mu/>



**ABB Group – Mauritius**

<http://www.abb.mu/>



**Tertiary Education Commission,  
Mauritius**

<http://tec.intnet.mu>

# Contents

## Invited Papers

<b>Survey on Models and Methodology for Emergency Relief and Staff Scheduling</b> . . . . .	3
Bhupesh Kumar Mishra, Thepparit Sinthamrongruk, Zeeshan Pervez, and Keshav Dahal	
<b>A Technology Vision of the Fifth Generation (5G) Wireless Mobile Networks</b> . . . . .	25
Zoran Bojkovic and Dragorad Milovanovic	
<b>An Overview of Cloud RAN: Architecture, Issues and Future Directions</b> . . . . .	44
Meruyert Makhanbet, Xuewei Zhang, Hui Gao, and Himlal A. Suraweera	
<b>Electrical and Electronic Engineering</b>	
<b>Resonance and Impedance Matching for an Experimental Low-Frequency Wireless Power Transfer System</b> . . . . .	63
Raaju Hoolaus and Yasdeo Bissessur	
<b>Impact of Introducing Small Scale Distributed Generation on Technical Losses in a Secondary Distribution Network</b> . . . . .	71
Ismaël Adam Essackjee and Robert T.F. Ah King	
<b>Noise Measurement and Analysis in a Power Line Communication Channel</b> . . . . .	81
B. Rajkumarsingh and B.N. Sokappadu	
<b>Using the Development of the Information and Communications Technology Sector to Enhance Teaching in Electrical Engineering</b> . . . . .	94
Bhimsen Rajkumarsingh and Sarjoosing Goolaup	

<b>Quantifying the Pumping Energy Loss Associated with Different Types of Leak in a Piping System</b> . . . . .	107
L. Latchoomun, D. Mawooa, Robert T.F. Ah King, K. Busawon, and R. Binns	
<b>A Power Flow Control Scheme for a Photovoltaic to a Low Voltage Microgrid System</b> . . . . .	119
S.Z. Sayed Hassen, M.I. Jahmeerbacus, K. Sewraj, and M.S. Ruhomaun	
<b>Performance Analysis of a Hybrid Algorithm for Power Loss Reduction by Distribution Network Reconfiguration</b> . . . . .	132
Sarah Marappa Naiken and Robert T.F. Ah King	
<b>Voltage Stability Maximization by Distribution Network Reconfiguration Using a Hybrid Algorithm</b> . . . . .	144
Robert T.F. Ah King and Sarah Marappa Naiken	
<b>Design and Implementation of a Smart Dual Axis Solar Tracker with an Anti-theft Alarm Mechanism</b> . . . . .	154
Anshu Prakash Murdan, Rameshwar Jugurnauth, and Ravishwara Rakesh Nirsimloo	
<b>A Low-Cost Autonomous Cleaning System for Photovoltaic Arrays</b> . . . . .	165
Mohammad Fardeen Islam, Vishwamitra Oree, and Anshu Prakash Murdan	
<b>Voltage Control of a Power System with Exciter and Generator Saturation</b> . . . . .	175
R. Ramjug-Ballgobin and S.G. Calchand	
<b>Estimation of Solar Photovoltaic Parameters Using Pattern Search Algorithm</b> . . . . .	184
M. Derick, C. Rani, M. Rajesh, K. Busawon, and R. Binns	
<b>An Improved Reconstruction Method for Compressively Sampled Magnetic Resonance Images Using Adaptive Gaussian Denoising</b> . . . . .	192
Henry Kiragu, George Kamucha, and Elijah Mwangi	
<b>A Novel Implementation of FPGA Based Finite Difference Time Domain (FDTD) Technique for Two Dimensional Objects</b> . . . . .	201
Srinivasa Rao Gandham, Kartheek Bodireddy, Boya Pradeep Kumar, and Chandra Sekhar Paidimarry	
<b>An Iterative Back-Projection Technique for Single Image Super Resolution with Natural Texture Preservation</b> . . . . .	210
Boniface M. Ngocho and Elijah Mwangi	
<b>Development of Active Acoustic Noise Cancellation Using the Professional Audio Development Kit (PADK) Featuring the TMS 320C6727 DSP</b> . . . . .	220
Sajaad Boodoo, Yasdeo Bissessur, and Roshun Paurobally	

**Cooperative Spectrum Sensing in the DSA: Simulation of Spectrum Sensing Time Consumptions by Cognitive Radio Secondary Users. . . . .** 230  
 Masiala Mavungu and A.L. Nel

**Communication Engineering and Computing**

**Large Scale Fading Pre-coder for Massive MIMO Without Cell Cooperation. . . . .** 239  
 Tedros Salih, Elijah Mwangi, and Kibet Langat

**Performance Analysis of Symmetric and Asymmetric LTE Turbo Codes with Prioritisation and Regression Based Scaling . . . . .** 252  
 Y. Beeharry, Tulsi Pawan Fowdur, and K.M.S. Soyjaudah

**Wireless Body Area Network System Architecture for Real-Time Diabetes Monitoring . . . . .** 262  
 Geshwaree Huzooree, Kavi Kumar Khedo, and Noorjehan Joonas

**A Secure Device and Service Discovery Protocol for Proximity Based Social Networks (PBSNs) . . . . .** 272  
 Asslinah Mocktoolah, Krishiv Askoolum, and Kavi Kumar Khedo

**Performance of Unequal Error Protection Schemes for Audio Transmission Over ADSL with Reed Solomon and Turbo Codes. . . . .** 283  
 Prateema Ragpot, Tulsi Pawan Fowdur, and K.M.S. Soyjaudah

**Performance Analysis of Link Adaptation with MIMO and Varying Channel Estimation Schemes . . . . .** 293  
 Maryam Imran Sheik Mamode, and Tulsi Pawan Fowdur

**A Context-Aware Mobile Learning System Using Dynamic Content Adaptation for Personalized Learning. . . . .** 305  
 Brita Curum, Nigel Chellapermal, and Kavi Kumar Khedo

**Towards Improving the Security of Low-Interaction Honeypots: Insights from a Comparative Analysis. . . . .** 314  
 Abubakar Zakari, Abdulmalik Ahmad Lawan, and Girish Bekaroo

**Implementation of Driver Drowsiness Detection Application in Mauritius. . . . .** 322  
 Karishma Ramodhine and Shireen Panchoo

**A Hybrid Three-Phased Approach in Requirement Elicitation . . . . .** 331  
 Abubakar Zakari, Abdulmalik Ahmad Lawan, and Girish Bekaroo

**A Framework to Reduce the Testing Time of a Mobile Application Using an Automation Tool . . . . .** 341  
 V. Hurbungs, B. Dookheea, and Y.K. Suttroogun

**Author Index. . . . .** 353



## **Invited Papers**

# Survey on Models and Methodology for Emergency Relief and Staff Scheduling

Bhupesh Kumar Mishra<sup>(✉)</sup>, Thepparit Sinthamrongruk,  
Zeeshan Pervez, and Keshav Dahal

School of Engineering and Computing, Artificial Intelligence,  
Visual Communication and Networks (AVCN) Research Centre,  
University of the West of Scotland, Paisley, UK  
{bhupesh.mishra, thepparit.sinthamrongruk,  
zeeshan.pervez, keshav.dahal}@uws.ac.uk

**Abstract.** Decision support is required for effective planning on all kinds of scheduling scenarios. The stochastic scenarios and uncertainty in demands make the scheduling task complex. Multiple objectives in terms of cost, timing window, priorities and travel routes are the driving factors in the scheduling task. These objectives are often associated with given constraints like time, cost, resource limit etc. To meet all these objectives with the given constraints, it requires effective scheduling methods. Among different application areas of scheduling, emergency relief and staff scheduling are two domains which present major challenges for the scheduling research. These two areas provide analogy with many other areas of scheduling. Issues like finding appropriate locations and establishing them in appropriate group, discovering effective path for routing and making efficient plan for distribution and servicing are major challenges for these two and related scheduling cases. This paper covers a survey study on some of the recent papers of these areas that highlights the problem formulations, technologies, methods and algorithms applied. It provides a literature review on technologies and algorithms applied in the area of emergency case relief scheduling and staff scheduling.

**Keywords:** Resource scheduling · Staff scheduling · Emergency relief

## 1 Introduction

Scheduling systems are used to manage and optimize access to service providers. Several factors occur that affect the performance of scheduling systems including availability of resources, services, time variability, service preferences, information availability, level of the scheduling. For each scheduling environment relevant plans, ranging from a set of rules to the real-time responses are described that guide schedulers. There are many private and public organizations which provide relief distribution or staff services to the people. Each of those may have different scenarios, objectives, interests, capacity, and expertise but have common objective of scheduling optimization [35]. Scheduling service starts with requirements. In most of the cases, the requirements are in forms of estimation, mostly done by experts. The schedulers are

typically equipped with resources having specific time windows [36]. Schedulers must handle volume flexibility, delivery flexibility, and supply system flexibility for effective and efficient management [1]. Schedulers have to work in such way that they can handle critical and challenging issues like minimizing unsatisfied demands in case for goods or services allocations to seekers [7, 8]. Major objectives and constraints during the scheduling are to minimize cost, unsatisfied demand, travel time, and rational resource distribution. Satisfying these objectives are the major challenges in scheduling task. To meet the objectives, there are some constraints that make the job complex. Some of the very commonly faced constraints are stochastic supply and demand, resource availability, and vehicles routes.

Conventionally, planners are likely to create schedule for services and resources manually conducting to the issue related to travel distance, travel costs and resource utilization. All services and resources cannot be accomplished within their specified time windows according to insufficient manual schedule [37] since there are multiple objectives. Traditionally, solving such a problem consists of converting multi objectives into a single objective function [19]. The desired goal is to find the best solution that minimizes or maximizes this single objective while retaining the constraints of the problem. Compared to a single objective problems multi-objective problems are more difficult to solve, because there is no unique solution; rather, there is a set of acceptable trade-off optimal solutions. The solution algorithms of such problems should be shifted from exact to heuristic or meta-heuristic due to the complexity of the problem. To solve these multi-objective problems, heuristic methods are generally used. Some of commonly applied approaches are as: Nearest Neighbour Heuristic, Genetic Algorithm, Memetic Algorithms, and Particle Swarm optimization. These were developed with different algorithms and techniques but they have the same target to minimize the total cost and maximize preferences for both resource and staff scheduling.

There are numerous areas where scheduling is required for effective planning, decision making and task implementation. Among several applications of scheduling optimization, we choose two areas to review in this paper: emergency relief scheduling and staff scheduling. These two scheduling problems characterise the many features of the general scheduling and present challenges for the researcher. Several techniques, approaches, models and algorithms have been applied for scheduling solutions that make the best use of the corresponding objectives and constraints. In case of emergency such as in disaster relief operation, efficient scheduling plan reduces the severity of the impact of the emergency. The optimized scheduling of relief item distribution not only helps decision makers to plan effectively, but also to implement timely. For the case of staff scheduling, it requires improved schedule plan so that there is minimization of cost, time and maximization in the service satisfaction level. It needs optimized scheduling to balance between the skill level and time window. Scheduling task has number of issues. Appropriate resource centre location, transportation path and proper allocation and distribution planning are among the issues discussed in most of the cases. Each issue has impact on the efficiency of the scheduling decision making.

This paper surveys the recent papers on these two areas. The survey focused on the key issues of scheduling task in terms of pre-positioning, transport and routing and resource scheduling. We focus on the solution techniques and algorithms that are used to improve and optimize scheduling task. The rest of the paper is organized as follows.

In Sect. 2, problems associated with scheduling task are discussed. In Sect. 3, different techniques, models and algorithms are considered. Analysis and discussion is accomplished in Sect. 4. Conclusion and future research issues are discussed in Sect. 5.

## 2 Problem Classification and Models

There are many sources of uncertainty in scheduling that may affect coordination efforts. The nature, the location, timing, intensity of sudden needs, population characteristics and pre-existing regional infrastructure such as communications, transportations are the major concerns in the scheduling. The scheduling may also be challenged by the counter-active problem such as oversupply, unwanted and unused resources for longer period. Among many issues of scheduling, we review some of the key issues in scheduling managements and are categorised into following sub groups.

### 2.1 Pre-positioning, Warehousing and Clustering

One of the major goals of emergency relief scheduling effort is to provide services or resources as effectively and as timely as possible. In order to achieve this goal, several types of supplies, scheduling plans and decision making are essential during the distribution period. Pre-positioning of service centres at strategic locations is obligatory so that the commodities are available when they are needed [10, 11]. Where to preposition supplies in preparation for a scheduling is one of the key issues. Developing an effective pre-positioning of resource centres are challenging because of the uncertainty occurrence of demand locations and their magnitudes. The location and capacities of the service providers are key components in managing response efforts after an event. There can be a single resource distribution centre or number of such centres. A resource network is a collection of multiple resource centres, characterized by connecting link and storing capacity, built to address specific services or resource demands. The network needs to be designed to optimize timely distribution of the items. All the resource centres need to be modifiable to meet the optimum demand-supply when distribution starts. Based on the presented assessment parameters and constraints, resource network optimization methods could be developed. A resource network may exhibit different behaviours in different stages of scheduling.

Designing the network with appropriate location and storing capacity is challenging since network structure dynamically evolves [9]. There can be different alternatives to supply centre locations. If supplies are located closer, it can allow for faster delivery of supplies [12] but the major threat is that the supplies may be in a risky location for the cases such as if any the disaster occurs. If supply centre is located far, it increases the transport cost and delay in distribution. Considering these risks, determining the optimal location, stocking quantity and the total expected costs associated with delivering to a demand point from a supply point is required. Supply points at different locations may have different probabilities of being shattered and may offer different travel costs/time to serve the demand so in the network of resources centres cooperation and collaboration is required for optimum result.

## 2.2 Transportation and Connecting Network

Scheduling models that are able to accurately interact and represent the reality may be more desirable. In some scenarios, information about accurate need is hard to obtain, and a stochastic modelling approach can be useful to represent those scenarios. Transportation and connecting networks design depends upon location and also the specific challenges of resource distribution. Transportation and routing problems are very difficult to solve. The problem's difficulty increases as the model's level of detail increases [4]. Trying to solve all issues at once, stochastic data, heterogeneous vehicle fleet, in a multi-period and multi-commodity network context, the resulting model will be complex to solve. The resulting model may not be optimum solution when looking for fast and efficient solutions. Therefore choosing the factors that are best suited to the context and then establishing hypotheses on the other features is required to simplify the model.

In most of the scenarios, transportation capacities often exceed the available capacity significantly and vehicles depart from and arrive to warehouses with full load. Therefore, it is not quite possible to modify the vehicles route whenever new information arrives. One of the issues during the transportation is the last mile distribution problem [5] that arises in case of emergency response which involves delivery from local distribution centres or from central warehouses to a population in need. It included both delivery and pickup functions, and call it "the last mile delivery and pickup problem", where the last mile delivery is concerned with materials transported from warehouses to affected locations. For the logistics distribution, not only finding the best path is enough for vehicle routing but it is also important to find the capacity of the path along with the vehicle [6].

The vehicle routing problem can be further sub-categorized as path finding and network flow capacity. Both the sub problems have its own constraints that need to be considered while finding optimum method for logistic support routing. An essential issue is capacity feasibility because of the vehicle capacity fluctuate throughout the route path and also the assignment of the multiple types of load to available vehicles with different capacity and type. These parameters keep changing dynamically during the disaster relief system. Another challenge in the path selection is condition of the path that also keeps changing with respect to time in scenarios such as emergency cases [8]. Path condition is a time varying function that depends on the parameters of the path description. Objective of cost minimization varies with the path conditions with varying time. For transportation, one task is to design the transport network and another is to optimize the route selection for a specific task [15]. In planning and scheduling problems route scheduling is required to find effective routes for offering service to the demand points with the shortest path.

## 2.3 Resource Distribution/Scheduling

Distribution of commodities and services in the disaster struck areas with quick response is a critical issue for emergency relief. Resource distribution is challenging because of its stochastic nature of occurrence and inadequate information. A critical and

challenging component of distribution is the allocation of resources to the beneficiaries [2, 3]. The major objectives and constraints during the resource distribution or scheduling are broadly classified as:

- i. **Minimize cost:** Minimize the running cost of resource distribution task. This mainly includes travel cost and inventory costs and other operational costs. Operational cost of emergency distribution is different than the business operational cost because the decisions taken during the emergency are more diverse.
- ii. **Minimize unsatisfied demand:** Demand in most of cases is stochastic and varies rapidly with time. For affective distribution operation within demand areas, minimization of unsatisfied demand is required.
- iii. **Minimize travel time:** Timely arrival of commodities is required. To achieve this, the shortest path or the minimum time for travel is desired.

Above listed objectives are the major challenges in distribution job but it also has some constraints that make the job complex. Some of very commonly faced constraints are:

- i. **Stochastic supply:** The quantity of goods available at different supply centres with the network for distribution is varying and in most cases that may not meet the exact demand in distribution phase.
- ii. **Stochastic demand:** The amount and type of commodities and services request vary with time and also that vary from one location to another location.
- iii. **Vehicle Availability:** Types of vehicles availability, capacity, cost, travel time varies. Also there can be single or multiple dispatch point in the distribution network.
- iv. **Vehicles routes:** Finding the shortest feasible route from distribution node to end node is another constraints for distribution implementation.

Timely distribution with minimization of cost is required during distribution process [19, 20]. Apart from these two objectives, another objective during the distribution of minimization on unfair distribution to all points [20]. Minimization in unfair distribution is as higher as possible is essential but cannot reach to the level of 100% is challenging since the demand within the network is stochastic in nature.

Workforce resource is also one of the most difficult issues associated with constraints in case of service scheduling, even when they have only a single criterion and homogeneous skills [53]. This problem also concerns about when and how the workers should be employed with their skills. Human resources consist of different task skills [54]. They can be categorised into either hierarchical task or categorical task skills. Under hierarchical task skills, each staff has different efficiency and quality of work where staff with a lower skill level can do less task than staff with a higher skill level. Under categorical task skills there is no difference on performance level. Each staff group has different skills that they can perform in the group. Another challenge for resource distribution is kinds of service to be scheduled. The scheduling model should incorporate the time slot or assignment of each individual worker task. Some service requires only single type resources where as some requires multiple types.

### 3 Methodologies

In general, real world scheduling problems are multi-objective. For instance, some cases require minimum cost, some require high priority service first, some requires minimum time duration where as some may requires minimum utility utilization such as human resource and vehicles. Most of scheduling scenarios are multi-objective optimizations related to a set of results or solutions called Pareto-optimal solution. In the wider sense, no other solutions in the search space are better when all objectives are considered [55]. During the last decade, a number of multi-objective algorithms have been proposed including classical and intelligent approaches. The Classical approach comprises of converting the multi- objective problem into a single problem which can be solved using traditional scalar-valued techniques comprising weighted aggregation, goal programming,  $\epsilon$ -Constraint, and discussion on classical method. The second approach is directly a tool to investigate the Pareto-front using Artificial Intelligent (AI)/computing algorithms. The results will be generated in many iteration of the AI search to create and then is analysed in different effective methods. However, the problem of this approach is how to choose to the Pareto-optimal set and how to keep up the convergence of the result.

There are many popular tools which can determine the Pareto capable solution. Vector Evaluated Genetic Algorithm (VEGA) separates the population into several equalled sub-groups according to the number of objective functions to find each objective. Multi Objective Genetic Algorithm (MOGA) is a simple and proficient technique based on Pareto-Based Approaches and Niching Mechanism in [56] to create the possibility of answer called “non-dominated individual” which was developed to the Non-dominated Sorting Genetic Algorithm (NSGA). Moreover, Niche Pareto Genetic algorithm (NPGA) and (SPEA) were proposed to handle with multi-objective problems. The former is an algorithm which desires to categorize classification layers in [57]. The latter uses an external archive to maintain the non-dominated solutions found during the evolution. Candidate solutions are compared to the archive [58]. In the next sub-section, a number of methods and algorithms applied for different issues are analysed.

#### 3.1 Pre-positioning, Warehousing and Clustering

The first stage challenge in scheduling task is to identification of appropriate location of the distribution centre. Also, the appropriate capacities in terms of their storage and availability resources need to be identified. There can be a single or number of distributed locations for resource distribution. Evolution of the resource network can be considered using a discrete-time model to deal with the uncertainty and dynamism. Hu et al. [9] proposed complex resource network based on a multi-agent system (MAS). They developed a framework for resource network with the state diagram describing the various states of each resource distribution centre of the network at discrete time interval. Each resource centre was a unit of storage, where the resources were stored in units or blocks depending on their types and properties. Their network was designed for small scale scheduling case for fire disaster that may be applicable for larger scale.

Although in the large scale demand, there are many number of distribution centres in the network so network optimization is required. The model also considers gradual decrease in resources availability at distribution centres. But, in most of the long run distribution centres, there occurs in and out of resources from the centre so value of a centre changes over time but not necessarily decrease every time.

Stochastic programming, a general purpose technique, deals with uncertainty in the input values to optimization models and was used for pre-positioning of resource centre. Rawls et al. [10] had applied a two-stage stochastic mixed integer program (SMIP) that provides response pre-positioning strategy for emergency threats. The SMIP considered uncertainty in demand for the stocked supplies as well as uncertainty regarding transportation network availability after an event. A heuristic algorithm, Lagrangian L-shaped method (LLSM), was developed to solve large-scale instances of the problem to reduce the computational complexity. The model combined facility locations and sizes of storage facilities. It also considered the amounts of various materials stocked in each facility, decisions on stocking levels for emergency supplies, and distribution of those supplies to multiple demand locations after an event with uncertainty about demand. The model used a set of discrete scenarios with probabilities to represent potential locations and magnitudes of the uncertain events. A bundled commodity was defined that includes specific elements in fixed proportion but was still handled by the algorithm as a single commodity. It did not change the structure of the formulation but changed the decisions to acquire, stock and transport units, commodity costs, storage requirements and other related parameters that reflect the aggregation. Finding appropriate location for storing commodities with sensitivity analysis was applied to show how different parameters impact stocking levels and costs.

Campbell et al. [11] showed how cost model can be used to select the single best supply point location from a discrete set of choices. Apart from this it also included how it can be embedded within existing location algorithms to choose multiple supply points taking the constraints on standard inventory purchasing, restocking costs, salvage values and delivery costs. Identifying the supply centres, the algorithm had considered the cost of assigning each demand point to each of the supply points and allocating each supply centre to the demand point where it could be served at the lowest cost. Model could be used to select the single best supply point location from a discrete set of choices but the algorithm failed to identify more than one distinct supply points in case if the demand is not fulfilled by single supply centre. Barzinpour and Esmaeili [12] developed multi-objective mixed-integer linear programming model for preparation of disaster logistics scheduling based on demand area population and damage severity. They applied goal programming approach to prioritize objectives in order to have the least deviation from goals. For the planning, worst-case scenario was considered for need estimation and preparation. The model showed improvement in quality of solutions when collaborative and cooperative between sub-regions were used.

Heuristic algorithm was designed to solve emergency resource allocation problem based on an integer mathematical programming and network optimization. Zhang et al. [13] modifies the solutions of the linear programming with constraints of multiple resources by setting priorities of preference for each demand location with certain possibilities. The algorithm was designed to efficiently satisfy the constraints and



objective function for scheduling resource distribution by setting allocation priorities according to their probabilities of occurrence. The local search technique was employed to assign the emergency resources to those points based on the set priorities. Mete et al. [14] proposed a two stages stochastic programming model to select the storage locations of medical supplies. During first stage, warehouse selection and inventory decisions was done with the objective of cost minimization. During second stage, transportation plans and demand satisfaction decisions were decided which represents the amount of medical supply to be delivered from warehouse to hospital with the minimization on the total transportation duration. The model was limited only to land transportation and used a set of predetermined routes that may not be fully applicable in all scenarios since the condition of the route changes dynamically and uncertain in most of the scheduling cases.

Aghamohammadi et al. [30] applied heuristic method based on two nested genetic algorithm for optimal allocation of the distribution centre in the form of medical centre for resource scheduling in case of emergency. It used Geographic Information System (GIS) for positioning data. Among the two nested genetic algorithm, one is applied for location part on the basis of number of population need to be serviced by the medical service centre and the next one is applied for resource scheduling optimization by reducing the operation time.

### 3.2 Transportation and Network Flow

Sending out resources from distribution centre to demand point cannot be effective without considering the connecting path, available vehicles and their capacity, especially if there is more than one demand points need to be in consideration for scheduling plan. Ozdamar et al. [5] used hierarchical cluster and route procedure (HOGCR) for coordinating vehicle routing in large-scale post-disaster distribution and evacuation activities. The HOGCR utilized an efficient network flow model in a hierarchical “cluster first, route second” approach that produced feasible solutions. It offered flexibility in adjusting the solution quality by imposing different runtime restrictions. HOGCR approach divided the demand centres in the relief network into geographically dense clusters using the k-means clustering algorithm and then solved the top-level routing problem. When a cluster had more demand centres than desired its routing problem was not solved immediately, but it was sub-divided further into smaller sub-clusters. Finding efficient and stable cluster in its hierarchy cluster centre requires number of iteration with correct parameters. In the case of emergency, there is very high chance of incorrectness that makes the approach not very effective for early stage of distribution. The initial procedure of the route Scheduling begins by partitioning entire area into sub-areas using clustering.

Generally, clustering algorithm can be divided into two major branches: Hierarchy and Partition [61]. The partition clustering illustrates outstanding results compared with the Hierarchy technique. Typically, partition clustering technique has a simple and easy concept to implement by attempts to find to centre of clusters or sections call “centroid” by minimizing an objective function. Simple K-means or modification of k-means was applied such as K-mean such as Kd-tree, K-mean++, Fuzzy c-means [62–64] were

applied for clustering. After that, the nearest-neighbour heuristic is used for all clusters by finding the unrouted customer locations “nearest” in terms of a measure of Euclidean distance between its centroid and location. At every repeating loop, the heuristic search gathers the unrouted customer locations into the nearest cluster depended upon its centroids until each cluster reaches the maximum number of members. Thus, using the K-mean in the initial phase is very effective for solving scheduling problem. Although, the K-mean, an iterative technique, is used widely; it has several disadvantages such as difficult to know the number of actual clusters, an iterative technique, a higher chance to fit in bad local optima [65].

Yi et al. [6] applied a meta-heuristic of ant colony optimization (ACO) for solving the logistics problem arising in relief distribution scheduling activities by decomposing the original logistics problem into two phases of decision making. The first phase builds stochastic vehicle paths under the guidance of pheromone trails while a network flow based solver was developed in the second phase for the assignment between different types of vehicle flows and commodities. With the increase of dimensionality and covering space this approaches might not be the highly efficient techniques. A modified Dijkstra algorithm was designed by Yuan et al. [7] to solve the time-varied shortest path problem for routing. A mathematical model was built for path selection in the real-time effect of disaster for emergency logistics management with the objective to minimize total travel time along the path from distribution point to the relief distribution centre. A path selection model for emergency logistics management was built based on path complexity consideration having objectives to minimize total travel time along the path and to minimize the complexity of the path. The complexity of the path was modelled as the total number of arcs included in a selected path. An ant colony optimization algorithm was applied to solve the multi-objective optimization model. The algorithm converts the multi-objective path selection model into a single-objective model. The model considered travel speed under disaster conditions that is different from normal conditions travel speed. The assumption of the model that speeds on each arc of the network will decrease with the extension of disasters in time and space is not valid for all cases. The conditions of the arc need to be updated with varying time and hence the speeds and travel times can be calculated.

For solving problem of emergency transportation scheduling in the relief supply chains, a multi-objective fuzzy optimization was applied by Zheng et al. [8] with the consideration of three transportation modes: air, rail, and road. They applied a cooperative optimization method that divides the integrated problem into a set of sub-components and evolved the sub-solutions concurrently. These sub-solutions were brought together to construct complete solutions. The arrival time at the relief centre was defined by fuzzy number to cope with the uncertainty of travel time of vehicles because of variation of environment. Among the three means of transportation, the highest priority items assigned to air transportation then after to railways and then to road transportation. For the solution, the integrated problems were divided into a set of sub-components and applied multi objective tabu search. After then multi objective genetic algorithm was applied to optimize sub solutions for transportation task allocation and resource allocation. The major focuses were on the optimization of the task allocation plan, resource allocation plan and delivery scheduling and vehicle routing plans where the items priorities varied with time and areas.

Ahn and Ramakrishna [23] applied genetic algorithm for routing. The model used variable-length chromosomes with population-sizing equation based on the gambler's ruin model with the major objective to minimize the cost associated with the path. For similar scenarios, Nagata et al. [24] used a penalty-based edge assembly memetic algorithm for vehicle routing with time window. The model calculated time window violation and hence the sum of the penalties in the route for vehicle routing. Vidal et al. [25] also applied time window for vehicle routing. They used hybrid genetic algorithm with adaptive diversity management to minimize time and cost during scheduling. Zidi et al. [28] applied multi-agent approach for vehicle scheduling. The model used guided genetic algorithm that has multiple agents in the forms of system agent, information manager agent, forecast agent, distribution agents and others. It first dealt with the emergency planning with set of request then it concerned with the contingency planning. The contingency plan addressed the issues of new requests or any breakdowns in the scheduled task implementation. It applied local searching for contingency management. The major focus of the schedule was to maximize the number of saved people in case of emergency while also minimizing the cost of the rescue operations. To minimize the transportation service delay during scheduling task, Özdamar and Yi [33] applied constructive heuristic model to find the feasible acceptable solutions. The feasible efficient routes were examined based on greedy neighbourhood search based on the vehicle's utilities.

Zhi-Hua Hu [15] proposed a multi-objective integer linear programming model to build the path selection for container supply chain in the context of emergency relief. A scheduling framework was proposed for the container multimodal transport emergency relief that was modelled by immune concepts for optimization. Affinity measures were designed to represent the complex relations among the components. Based on the affinity model, a decision process of emergency relief was proposed considering the characteristics of container. The main objective of optimizing the multimodal transport was to choose optimal combination of transport means by minimizing the cost and satisfying the time constraints. It used binary value 0/1 for transportation link that is either the fully conditioned link or no link condition. For transportation scheduling, the connecting link conditions are one of the constraints that need to be considered while finding the effective shortest cost path in the transportation network.

### 3.3 Relief Resource Distribution/Scheduling

In general, scheduling of resources in most of the cases is close to the supply chain management. The primary objective and actions are similar but the major change or challenge is uncertainty and nature of the information.

On mathematical programming methods, Felici and Gentile [44] established an integer programming model that maximizes the total satisfaction of the scheduling of staff with use of positive weights for shifts. Moreover, Bard and Purnomo [59] proposed mathematical programming model for minimizing the penalty in scheduling of staff members violating the preferences. They adopted the column generation scheme to solve the problem in which many conflicting factors guided the decision process. M'Hallah, R. and A. Alkhabbaz [48] used a mixed-integer programming model for the

staff scheduling problem applied in a health-care unit. They adopted a case study as Kuwaiti Health Care Unit for nurse scheduling for this model. The result showed the better schedule than the nurse manually planned scheduling. Smet et al. [49] solved the scheduling problem using a generic mathematical model, which is related to the common elements used in earlier similar works and some limitations that are usually ignored in the medical centres. Fan et al. [43] used a binary integer programming model to plan a practical solution of management in nursing timetable. The purposes are to maximize all nurses' satisfaction after considering seven shifts of both 8.5 and 12.5 h a day and some hard and soft restriction. The model also covered preferences of each staff for creating schedule of selection.

Given the fundamental differences between supply chains scheduling and emergency resources scheduling, supply chain coordination mechanisms might not be feasible or practical for relief scheduling [1]. Different methods had been applied for solving these issues. Researchers have applied mathematical approaches and computational methods such as fuzzy logic, genetic algorithm for solving such problem during different phases of scheduling. Camacho-Vallejo et al. [19] applied a bi-level mathematical programming model applied where problem was modelled with two different levels of pre-set hierarchy for decision making to get optimum result. The model assumed decision-makers in the upper level called the leader and in the lower level called the follower with set of variables, constraints and corresponding objective functions at each level. The leader made decision based on the actions and the follower reacts accordingly. Doing so, the follower optimized its objective with considering the decision made by the leader. The Upper level chose the means of transportation and the speed of distribution of goods that minimized the total response time for delivering resources and the lower level chose the storage centre with minimized shipping cost. Bi-level model then reduced into a nonlinear single-level mathematical model and hence problem was linearized as a mixed integer programming problem. The model used Chile earthquake (2010) as a case study where it found solutions with lower cost and time for distribution of emergency logistics.

Tzeng et al. [20] designed a relief-distribution model with the multi-objective programming method for relief delivery systems. It focused on the objective to minimize cost, time and unfair distribution during the planning period. For the maximization in fair distribution, scheduling model considered even distribution to the demand points regardless of cost by defining the upper boundary limit of the resource demands among all of the relief demand points in each period of distribution. The model assumed need of the people and current connected road network was only considered for the distribution. Wright, P.D. and S. Mahar [51] considered staff's preferences to deal with a centralizing scheduling problem with nurses from multiple units such as surgical, medical, and intermediate cardiac units in a hospital using an integer programming process. The method involves in two examining objectives: minimize cost of scheduling and maximize preference with satisfying the constraints of the staff's availability and preferences.

Solution approach used in business supply chain help to identify the major issues that can be used during the other scheduling management. Solution, such as an agent-based micro-simulation framework proposed by Roorda et al. [16] represented the diversity of roles and functions of actors in the resource scheduling system. Diverse

actors involved in the production and distribution of goods illustrates the major issues of consideration for affective distribution system. Two-stage approach was used to solve the response problem. At first stage, based on primary information, stochastic estimation of transportation capacities, supply availabilities and demand were identified. During second stage, actual values were used as they were revealed. Due to uncertainties associated with the information and a lack of supporting resources pre- and post- operation coordination activities were compared to enhance the effectiveness of the scheduling outcome. Two-stage, decisions of multisource relief supply and relief distribution, demand chain based dynamic optimization model was formulated in the specified emergency logistics co-distribution framework.

Sheu [21, 22] in his works proposed a time-varying relief demand associated with each affected area. The demand was predicted using a short term dynamic relief demand forecast model that mainly relied on the number of survivals trapped in the affected areas without receiving any rescue aids and may change over time upon updating the conditions. The model focused on group-based relief distribution having a composite weighted multi-objective optimization to deal with the problem of distribution. It clustered multi-type reliefs from multiple urgent relief distribution centres to multiple affected-area groups that maximized the time-varying relief demand fill rate and minimized the time-varying distribution costs.

Genetic Algorithm (GA) is also one of the effective approaches that were applied for solving scheduling problems. Genetic algorithm has been proven effective for solving optimization problems in various fields. However, GAs essentially uses generation succession to search for optimal solutions. Chou et al. [18] applied a Biological-based Genetic Algorithm (BGA) for improving the solution performance and execution time with their optimal solution search capability. It used the highest and the lowest fitness values to normalize chromosome population fitness values between 0 and 1. Non-linear fitness values were adjusted by increasing their fitness values with superior chromosomes and eliminated inferior genetic groups by decreasing their fitness values. Higher prioritized areas retained superior chromosomes for use in further crossover and mutation operations. The method increased chromosome population diversity and moved beyond local optimal solutions by using multi point search and imitating the biological phenomenon of migration instead of using mutation. Guided search direction was applied where fitness values were generated from a fitness function instead of from differential gradient data. The computational cost of the BGA was 50% lower than those of the other corresponding algorithms. This approach had a limited use of previous search in the selection, crossover, and mutation operations.

Lin et al. [26] proposed a heuristic genetic algorithm approach for scheduling that was designed for logistics model for delivering prioritized items. The model was designed with multi-objective that considered multiple items, vehicles, time periods. It applied weights on the objectives. By applying the weights, the multi objective problem was converted into a single objective problem that minimized the total unsatisfied demands, total travel time for all vehicles and tours. Chunguang et al. [27] developed distribution scheduling where genetic algorithm was applied for optimizing the travel route and also determining the number of vehicles being used for relief item. The models targets to total number of vehicles being used and also to minimize the distribution time. The major limitation of the model was that only one kind of resource

was considered for scheduling. With the objectives of minimization of unsatisfied demand, delay in service and transportation cost, Chang et al. [34] proposed greedy search based multi objective genetic algorithm for resources scheduling. The model dynamically adjusted distribution scheduled from various distribution points as per the requirements from various points. The algorithm adjusted the distribution of available resources by generating various feasible schedules. In addition, it also assembled multiple routing schedules for each form of vehicles in accordance with the resources required.

Gonzalez et al. [32] applied both fuzzy logic and genetic algorithm for resource distribution. The model found the cost-minimized transportation schedule without deviating from supply constraints. It used fuzzy sets to represent initial information related to cost, demand and other variables. The scheduling task was solved by applying genetic algorithm with fuzzy fitness value for solution evaluation. To improve the flexibility in schedule planning and high success rate, D'Uffizi et al. [29] applied discrete event simulation that support in decision making during the planning for different activities. In the model, travelling time over the defined routes was supplied for each vehicle. Relief scheduling was designed on first-in-first-out basis where both preemption and non-preemption scenarios were considered for distribution policy. Lin et al. [39] used genetic algorithm with an immigrant scheme for staff scheduling. The model applied balancing between the preferred work shifts and intervals among the staff to generate efficient scheduling solutions.

It is hard to predict the time unit precisely and also transport network can dynamically change during the time varying scenarios. Considering the uncertainty, Sheu et al. [17] proposed fuzzy processing times and weights based on expert opinion of the tasks in emergency scheduling modelling with multiple objectives. In some scenarios such as in disasters, the corresponding shortest paths in the transportation network could vary with time if there are damaged points on the centres connecting network. Any path through that point is blocked thus cannot be used for computing the shortest paths. Multiple permutations and ranking measures based on disaster geography were applied for effective scheduling. For effective decision making during scheduling implementation, Baky [31] used Fuzzy Goal Programming (FPG) to minimize the degree of dissatisfaction of decision makers. It applied single decision makers at top level and multiple decision makers at lower level where the algorithm was extended to solve bi-level multi objective programming. Iteratively the re-evaluation was performed to increase the satisfaction level in decision making. The model applied membership functions to the objective functions and to the decision vectors in the upper level. The FPG was used to achieve the maximum degree of membership goals by minimizing their deviational variables and hence obtained the most satisfactory solution for decision makers. Coordination and cooperation between upper and lower level decision making is required for effective implementation. Topaloglu, S. and H. Selim [50] proposed a fuzzy goal programming model where multi-objective integer programming model was applied in a view of changeable factors that affect scheduling timetable and preferences considering the staff scheduling. The model was also able to cope with the uncertainties associated with target value of the management and staff preferences.

To minimize the unsatisfied demand at the receiving end, Yi et al. [6] proposed a meta-heuristic of ant colony optimization for the distribution scheduling. The scheduling model consisted of two phase decision making. The First phase dealt with vehicle route finding whereas the second phase dealt with the scheduling of vehicles and resources. Todorovic and Petrovic [60] proposed a staff scheduling with preference using a bee colony optimization which is able to eliminate some ineffective plan from the neighbourhood solution. The model performed scheduling in two phases. In first phase constructive search is applied where unscheduled shifts are assigned to available staff. In the second phase, the model applied local search to optimize the model and hence improve in the quality of scheduling.

Maenhout and Vanhoucke [47] created a population based Evolutionary algorithm to support planning for nurse allocation to the different departments in hospital under conditions of staff appointment, nursing shift in particular ward policies and personality traits. They applied immune cells to calculate objective function value for different objective functions. The objective functions with acceptable total objective function are selected. Bai et al. [41] proposed a hybrid evolutionary algorithm with the local search associated with simulated annealing hyper-heuristic with the better result. The model improved ability to handle constraints during scheduling. The algorithm showed enhancement in the performance level of evolutionary method by hybridizing with simulated annealing. Ray et al. [57] illustrated evolutionary algorithm for multi objective optimization where as Ngatchou et al. [58] applied pareto multi objective optimization.

Constantino et al. [42] proposed a new multiple assignment problems based deterministic heuristic algorithm to solve a nurse scheduling problem consisting of two phases. In first phase, constructive phase, complete schedule was developed by implementing successive assignment of task each day in the planning. In the second phase, improving phase, task assignments were resolved for better scheduling result. Hadwan et al. [46] proposed a Harmony Search Algorithm, population-based meta-heuristic algorithm, for the nurse scheduling problem for a hospital in Malaysia. The algorithm improved solutions iteratively based on good candidate solutions from the initial population by applying a stochastic random search on the solution with a number of alterations. Algorithm evolved upon the harmony memory and solutions were updated during the evolution. The test result showed better performance than the basic genetic algorithm. Wright and Vanhoucke [52] utilized heuristic procedure and exact solution for scheduling. Depending on the duration the method was selected. They chose exact solution approach if the planning length is less than 7 days and for higher than 7 days they applied meta-heuristic. The model utilised a multi-start heuristic to select combination of nurses randomly by adopting individual schedules. After then a random permutation was generated for those selected nurses. Moreover, an evolutionary algorithm and a branch-and-price approach were combined to reschedule planning to solve violated constraints for nurses from external department.

Gao and Lin [45] developed a mathematical model to overcome the scheduling problem that also covered different aspects and targets such as minimized cost and maximized efficiency. The goal of research was to find the maximum happiness level while nurses are working considering hospital regulation and then they used classical Particle Swarm Optimization to solve the problem to decrease the time-consuming

manual scheduling. Akjiratikar et al. [37] applied particle swarm optimization to schedule home care workers. They applied heuristic assignment scheme for scheduling. The heuristic assignment was mainly considered for transforming the continuous particle swarm optimization algorithm into the discrete job schedule.

## 4 Analysis and Discussion

We analysed several recent papers in the area of scheduling mainly focused on logistic distribution and staff scheduling. Distribution centre location and transportation and routing issues have impact on scheduling. Major objective and methodologies applied for the scheduling plans are surveyed. Tables 1, 2, 3 and 4 shows the major objectives and methodology applied categorically in the area of positioning the service centres, transport and routing, logistic distribution and staff management scheduling. Most of the scheduling scenarios come with multi-objective like minimization in time, cost delay in service and maximization in performance level. Optimization is required to enhance the effectiveness of scheduling. To plan better scheduling, intelligent methods have been applied apart from the mathematical approach. Intelligent approach is highly suitable for non-linear and multi-objective optimizations task. In some cases hybrid approaches have also been applied.

**Table 1.** Methodologies and objective used with focused on pre-positioning, warehousing and clustering

Author	Major objective(s)	Methodology
Zhi-Hua Hu and Zhao-Han Sheng (2015)	Rescue time optimization	Multi-agent system
Carmen G. Rawls and Mark A. Turnquist (2010)	Minimize the expected costs over all scenarios	Two-stage stochastic mixed integer program
Ann Melissa Campbell and Philip C. Jones (2011)	Minimize the sum of costs of serving each demand point from its closest supply point	Heuristic approach
F. Barzinpour and V. Esmaeili (2014)	Minimize cost	Multi-objective mixed-integer linear programming
Jiang-Hua Zhang et al. (2012)	Minimize cost of dispatching time	Integer mathematical programming and heuristic algorithm
Huseyin Onur Mete and Zelda B. Zabinsky	Minimize operating cost of warehouse, transport duration and unmet penalty	Two-stage stochastic programming model
Hossein Aghamommadi	Minimize the time of relief operation and the number of fatalities	Heuristic method based on two nested genetic algorithms



**Table 2.** Methodologies and objective used with focused on transportation and connecting network

Author	Major objective	Methodology
Linnet Özdamar and Onur Demir (2012)	Minimize the estimated total travel time and promotes efficient vehicle utilization	Mathematical optimization
Yuan Yuan and Dingwei Wang (2009)	Minimize total travel time along a path	Ant colony optimization
Yu-Jun Zheng and Hai-Feng Ling (2013)	Minimize the total time delay	Multi objective fuzzy optimization
Zhi-Hua Hu (2011)	Minimize cost	Integer linear programming
Chang Wook Ahn and R.S. Ramakrishna (2002)	Minimize cost associated with path	Genetic algorithm
Thibaut Vidal et al. (2013)	Minimize cost and time	Hybrid genetic algorithm with adaptive diversity
Yuichi Nagata et al. (2010)	Minimize cost and time	Penalty-based edge assembly memetic algorithm
K. Zidi et al. (2013)	Maximize the number of saved people and minimize the costs of the rescue operation	Multi-agents approach using a guided genetic algorithm
Enrique LoÁpez Gonza Ález and Miguel A. Rodrõ Águez Ferna Ández (2000)	Find the cheapest transporting schedule	Fuzzy system and genetic algorithm
Linnet Özdamar and Wei Yi (2008)	Minimize total service delay	Greedy neighbourhood search
Kergosien Y. et al. (2009)	Minimize travel distance	Integer linear programming

**Table 3.** Methodologies and objective used with focused on relief resource scheduling

Author	Major objective	Methodology
Jiuh-Biing Sheu et al. (2005)	Minimize distance cost and travel time between source and destination	Hybrid fuzzy-optimization
Yu-Jun Zheng et al. (2015)	Minimize the total weighted waiting time of the tasks	Biogeography-based optimization and fuzzy system
Jui-Sheng Chou (2014)	Minimize delay	Biological based genetic algorithm

*(continued)*

**Table 3.** (continued)

Author	Major objective	Methodology
Jos_e-Fernando Camacho-Vallejo et al. (2015)	Minimize the shipping costs and time	Bi-level mathematical programming
Gwo-Hshiung Tzeng et al. (2007)	Minimizing the total cost, travel time and maximize the minimal satisfaction during the planning period	Multi-objective programming
Jiuh-Biing Sheu (2007)	Maximize the time-varying relief demand fill rate and minimize the time-varying distribution costs	Hybrid fuzzy clustering-optimization
Jiuh-Biing Sheu (2010)	Improve the performance of relief-demand management	Multi-source data fusion, fuzzy clustering, multi-criteria decision making
Yen-Hung Lin et al. (2011)	Minimize total unsatisfied demand, total travel time	Heuristic approaches of genetic algorithm and integer programming
Antonio D'Uffizi et al. (2015)	Improve flexibility and relief success rates	Discrete event simulation
Wei Yi and Arun Kumar (2007)	Minimize the weighted sum of unsatisfied demand	Meta-heuristic of ant colony optimization
Chang Chunguang et al. (2010)	Minimize task completion time and vehicle count	Genetic Algorithm
Ibrahim A. Baky (2009)	Minimize the group regret of degree of satisfactions of all the DMs	Fuzzy goal programming and multi-objective linear programming
Fu-Sheng Chang et al. (2014)	Minimize unsatisfied demand for resources, time to delivery, and transportation costs	Greedy-search-based multi-objective genetic algorithm

**Table 4.** Methodologies and objective used with focused on staff management

Author	Major objective	Methodology
Chananes Akjiratikar et al. (2007)	Optimum home care workers scheduling	Particle swarm optimization and heuristic assignment scheme
Mankowska and Dorota Slawa (2014)	Optimum daily planning of health care services	Mathematical modelling
Chun-Cheng Lin et al. (2015)	Balanced preferred work shifts and day-off of nursing staff	Genetic algorithm with an immigrant scheme
Chun-Cheng Lin et al. (2015)	Maximize the total satisfaction of the nursing staff with their preference rights and interests	Memetic algorithm with recovery scheme

(continued)

**Table 4.** (continued)

Author	Major objective	Methodology
Bai, R. et al. (2010)	Improve the constraint handling ability,	Penalty based genetic algorithm and simulated annealing hyper-heuristics
Constantino, A.A. et al. (2013)	Maximize the satisfaction of nurses' preferences and minimize the violation of soft constraints.	Multi assignment heuristic algorithm
Fan, N. et al. (2013)	Balance the preferences of the employees	Integer programming
Felici, G. and C. Gentile (2004)	Maximise staff satisfaction	Integer programming
Gao, S.C. and C.W. Lin	Minimise cost and balance nurses' happiness	Particle Swarm Optimization (PSO)
Hadwan, M. et al. (2013)	Allocate the required workload to the available staff	Harmony search algorithm
Maenhout, B. and M. Vanhoucke (2013)	Cope with schedule disruptions for the staff shift	Artificial immune system
M'Hallah, R. and A. Alkhabbaz (2013)	Cope with real time context dependent constraints	Mixed integer program
Topaloglu, S. and H. Selim (2010)	Treat uncertainties in the target value of the hospital management and nurses preference	Fuzzy goal programming
Wright, P.D. and S. Mahar (2013)	Minimize costs and reduced overtime result for multiple units	Heuristic algorithm
Wright, P.D. and M. Vanhoucke (2013)	Understanding in the consequences and outcomes of various personnel re-rostering characteristics and strategies.	Artificial immune system heuristic procedure
Cai, X., & Li, K.N (2000)	Minimize the total cost for assigning manpower	Genetic algorithm with multi-criteria optimization
Nikola Todorovic and Sanja Petrovic (2013)	Handle with the nurse rostering problem by minimizing the use of outside nurses	Bee colony optimization algorithm

## 5 Conclusion and Future Directions

Distribution of resources and allocation of staff requires a careful planning. Scheduling provides the enhanced and efficient way of planning. Each scheduling scenario has its own objectives and constraints in terms of cost, delay time, resource limitation. Meeting all constraints at a time is complex therefore optimization is required for effective scheduling. Different techniques, models and algorithms are applied for scheduling optimization including mathematical and computational approaches.

For multi-objective scheduling scenarios computational intelligent approaches have been used effectively in diverse ways which includes fuzzy logic, genetic algorithm, heuristic methods, swarm optimization, ant colony optimization, agent system. This survey paper covers approaches applied by many authors in recent times for scheduling in the area of relief distribution and staff scheduling. Finding appropriate locations or aligning them into appropriate groups has been applied in different ways to enhance the effectiveness of the scheduling plan. It is analysed through many papers that finding appropriate path and vehicles for resource transfer also play role in scheduling effectiveness. For this, wide ranges of approaches were applied. Cautious planning of distribution of resources, careful setting up of staff allocation can be done using different methods. Hybrid methods can be applied to optimize the scheduling performance. The paper provides the methods and techniques that are applied in the area of relief and staff scheduling along with the limitations.

Surveying many papers exposed some the issues for further research in the area of resource and staff scheduling. Improvement in the performance of demand management is required for the effective, timely and fair distribution. The nature of demand is stochastic and also there is uncertainty of requirements in most of the scenarios so forecasting mechanism needs to be more dynamic and exact. An appropriate grouping and priority identification of the demand is needed for the optimum scheduling result. Also, inclusion of the transportation path, vehicle availability, capacity and preferences are desirable in scheduling. Alteration in methods and algorithm is required for multi-objective scheduling plan.

**Acknowledgement.** The first and second authors are currently pursuing their PhD at the University of the West of Scotland under the Erasmus Mundus SmartLink scholarship.

## References

1. Beamon, B.M., Balcik, B.: Performance measurement in humanitarian relief chains. *Int. J. Public Sect. Manag.* **21**(1), 4–25 (2008)
2. de la Torre, L.E., Dolinskaya, I.S., Smilowitz, K.R.: Disaster relief routing: integrating research and practice. *Socio-Econ. Plann. Sci.* **46**, 88–97 (2012)
3. Altay, N., Green III, W.G.: OR/MS research in disaster operations management. *Eur. J. Oper. Res.* **175**, 475–493 (2006)
4. Anaya-Arenas, A.M., Renaud, J., Ruiz, A.: Relief distribution networks: a systematic review. *Ann. Oper. Res.* **223**, 53–79 (2014)
5. Özdamar, L., Demir, O.: A hierarchical clustering and routing procedure for large scale disaster relief logistics planning. *Transp. Res. Part E* **48**, 591–602 (2012)
6. Yi, W., Kumar, A.: Ant colony optimization for disaster relief operations. *Transp. Res. Part E* **43**, 660–672 (2007)
7. Yuan, Y., Wang, D.: Path selection model and algorithm for emergency logistics management. *Comput. Ind. Eng.* **56**, 1081–1094 (2009)
8. Zheng, Y.-J., Ling, H.-F.: Emergency transportation planning in disaster relief supply chain management: a cooperative fuzzy optimization approach. *Soft. Comput.* **17**, 1301–1314 (2013)

9. Zhi-Hua, H., Sheng, Z.-H.: Disaster spread simulation and rescue time optimization in a resource network. *Inf. Sci.* **298**, 118–135 (2015)
10. Rawls, C.G., Turnquist, M.A.: Pre-positioning of emergency supplies for disaster response. *Transp. Res. Part B* **44**, 521–534 (2010)
11. Campbell, A.M., Jones, P.C.: Prepositioning supplies in preparation for disasters. *Eur. J. Oper. Res.* **209**, 156–165 (2011)
12. Barzinpour, F., Esmaili, V.: A multi-objective relief chain location distribution model for urban disaster management. *Int. J. Adv. Manuf. Technol.* **70**, 1291–1302 (2014)
13. Zhang, J.-H., Li, J., Liu, Z.-P.: Multiple-resource and multiple-depot emergency response problem considering secondary disasters. *Expert Syst. Appl.* **39**, 11066–11071 (2012)
14. Mete, H.O., Zabinsky, Z.B.: Stochastic optimization of medical supply location and distribution in disaster management. *Int. J. Prod. Econ.* **126**, 76–84 (2010)
15. Zhi-Hua, H.: A container multimodal transportation scheduling approach based on immune affinity model for emergency relief. *Expert Syst. Appl.* **38**, 2632–2639 (2011)
16. Roorda, M.J., Cavalcante, R., McCabe, S., Kwan, H.: A conceptual framework for agent-based modelling of logistics services. *Transp. Res. Part E* **46**, 8–31 (2010)
17. Jiu-BingSheu, Y.-H.C., Lan, L.W.: A novel model for quick response to disaster relief distribution. *Proc. East. Asia Soc. Transp. Stud.* **5**, 2454–2462 (2005)
18. Chou, J.-S., Tsai, C.-F., Chen, Z.-Y., Sun, M.-H.: Biological-based genetic algorithms for optimized disaster response resource allocation. *Comput. Ind. Eng.* **74**, 52–67 (2014)
19. Camacho-Vallejo, J.F., Gonz\_alez-Rodríguez, E., Almaguer, F.J., Gonz\_alez-Ramírez, R.G.: A bi-level optimization model for aid distribution after the occurrence of a disaster. *J. Cleaner Prod.* **105**, 134–145 (2015)
20. Tzeng, G.H., Cheng, H.J., Huang, T.D.: Multi-objective optimal planning for designing relief delivery systems. *Transp. Res. Part E* **43**, 673–686 (2007)
21. Sheu, J.-B.: An emergency logistics distribution approach for quick response to urgent relief demand in disasters. *Transp. Res. Part E* **43**(2007), 687–709 (2007)
22. Sheu, J.-B.: Dynamic relief-demand management for emergency logistics operations under large-scale disasters. *Transp. Res. Part E* **46**, 1–17 (2010)
23. Ahn, C.W., Ramakrishna, R.S.: A genetic algorithm for shortest path routing problem and the sizing of populations. *IEEE Trans. Evol. Comput.* **6**(6), 566–579 (2002)
24. Nagata, Y., Braysy, O., Dullaert, W.: A penalty-based edge assembly memetic algorithm for the vehicle routing problem with time windows. *Comput. Oper. Res.* **37**, 724–737 (2010)
25. Vidal, T., Crainic, T.G., Gendreau, M., Prins, C.: A hybrid genetic algorithm with adaptive diversity management for a large class of vehicle routing problems with time –windows. *Comput. Oper. Res.* **40**, 475–489 (2013)
26. Lin, Y.-H., Batta, R., Rogerson, P.A., Blatt, A., Flanigan, M.: A logistics model for emergency supply of critical items in the aftermath of a disaster. *Socio-Econ. Plann. Sci.* **45**, 132–145 (2011)
27. Chunguang, C., Xiang, M., Xiaoyu, S., Bo, G.: Emergency goods scheduling model and algorithm during initial stage of disaster relief. *Int. Conf. Logistics Syst. Intell. Manag.* **3**, 1518–1521 (2010)
28. Zidi, K., Mguis, F., Borne, P., Ghedira, K.: Distributed genetic algorithm for disaster relief planning. *Int. J. Comput. Commun.* **8**(5), 769–783 (2013)
29. D’Uffizi, A., Simonetti, M., Stecca, G., Confessore, G.: A simulation study of logistics for disaster relief operations. *Procedia CIRP* **33**, 157–162 (2015)
30. Aghamohammadi, H., Mesgari, M.S., Molaei, D., Aghamohammadi, H.: Development a heuristic method to locate and allocate the medical centres to minimize the earthquake relief operation time. *Iran. J. Publ. Health* **42**(1), 63–71 (2013)

31. Baky, I.A.: Fuzzy goal programming algorithm for solving decentralized bi-level multi-objective programming problems. *Fuzzy Sets Syst.* **160**, 2701–2713 (2009)
32. LoÁpez GonzaÁlez, E., RodrÃ³guez FernaÁndez, M.A.: Genetic optimisation of a fuzzy distribution model. *Int. J. Phys. Distrib. Logistics Manag.* **30**(7/8), 681–696 (2000)
33. Özdamar, L., Yi, W.: Greedy neighborhood search for disaster relief and evacuation logistics. In: *IEEE Intelligent Systems*, pp. 541–1672 (2008)
34. Chang, F.-S., Jain-Shing, W., Lee, C.-N., Shen, H.-C.: Greedy-search-based multi-objective genetic algorithm for emergency logistics. *Expert Syst. Appl.* **41**, 2947–2956 (2014)
35. Kergosien, Y., Lenté, C., Billaut, J. C.: An extended multiple travelling salesman problem. In: *4th Multidisciplinary International Conference on Scheduling: Theory and Applications* (2009)
36. Mankowska, D.S., Meisel, F., Bierwirth, C.: The home health care routing and scheduling problem with interdependent services. *Health Care Manag. Scipp.* **17**, 15 (2014)
37. Akjiratikar, C., Yenradee, P., Drake, P.R.: PSO-based algorithm for home care worker scheduling in the UK. *Comput. Ind. Eng.* **53**(4), 559–583 (2007)
38. Begur, S.V., Miller, D.M., Weaver, J.R.: An integrated spatial DSS for scheduling and routing home-health-care nurses. *Interfaces* **27**(4), 35–48 (1997)
39. Lin, C.-C., Kang, J.-R., Chiang, D.-J., Chen, C.-L.: Nurse scheduling with joint normalized shift and day-off preference satisfaction using a genetic algorithm with immigrant scheme. *Int. J. Distrib. Sensor Netw.* **2015**, 1–10 (2015)
40. Lin, C.-C., Kang, J.-R., Hsu, T.-H.: A memetic algorithm with recovery scheme for nurse preference scheduling. *J. Ind. Prod. Eng.* **32**(2), 83–95 (2015)
41. Bai, R., Burke, E.K., Kendall, G., Li, J., McCollum, B.: A hybrid evolutionary approach to the nurse rostering problem. *IEEE Trans. Evol. Comput.* **14**, 580–590 (2010)
42. Constantino, A.A., Dario, L.S., de Melo, E.L., de Mendonça, C.F.X., Rizzato, D.B., Romão, W.: A heuristic algorithm based on multi-assignment procedures for nurse scheduling. *Ann. Oper. Res.* **218**, 165–183 (2013)
43. Fan, N., Mujahid, S., Zhang, J., Georgiev, P., Papajorgji, P., Steponavice, I., Neugaard, B., Pardalos, P.M.: Nurse scheduling problem- an integer programming model with a practical application. In: Pardalos, P.M., Georgiev, P.G., Papajorgji, P., Neugaard, B. (eds.) *Systems Analysis Tools for Better Health Care Delivery*. Springer Optimization and Its Applications, vol. 74, pp. 65–98. Springer, New York (2013)
44. Felici, G., Gentile, C.: A polyhedral approach for the staff rostering problem. *Manag. Sci.* **50**, 381–393 (2004)
45. Gao, S.C., Lin, C.W.: Particle swarm optimization based nurses' shift scheduling. In: *Proceedings of the Institute of Industrial Engineers Asian Conference*, pp. 775–782 (2013)
46. Hadwan, M., Ayob, M., Sabar, N.R., Qu, R.: A harmony search algorithm for nurse rostering problems. *Inf. Sci.* **233**, 126–140 (2013)
47. Maenhout, B., Vanhoucke, M.: An artificial immune system based approach for solving the nurse re-rostering problem. In: *Proceedings of 13th European Conference on Evolutionary Computation in Combinatorial Optimization*, pp. 97–108 (2013)
48. M'Hallah, R., Alkhabbaz, A.: Scheduling of nurses: a case study of a Kuwaiti health care unit. *Oper. Res. Health Care* **2**, 1–19 (2013)
49. Smet, P., De Causmaecker, P., Bilgin, B., Berghe, G.V.: Nurse rostering: a complex example of personnel scheduling with perspectives. *Autom. Sched. Plann. Stud. Comput. Intell.* **505**, 129–153 (2013)
50. Topaloglu, S., Selim, H.: Nurse scheduling using fuzzy modeling approach. *Fuzzy Sets Syst.* **161**, 1543–1563 (2010)
51. Wright, P.D., Mahar, S.: Centralized nurse scheduling to simultaneously improve schedule cost and nurse satisfaction. *Omega* **41**(6), 1042–1052 (2013)

52. Wright, P.D., Vanhoucke, M.: Reconstructing nurse schedules: computational insights in the problem size parameters. *Omega* **41**, 903–918 (2013)
53. Cai, X., Li, K.N.: A genetic algorithm for scheduling staff of mixed skills under multi-criteria. *Eur. J. Oper. Res.* **125**(2), 359–369 (2000)
54. De Bruecker, P., Van den Bergh, J., Beliën, J., Demeulemeester, E.: Workforce planning incorporating skills: state of the art. *J. Oper. Res.* **243**(1), 1–16 (2015)
55. Deb, K., Pratap, A., Agarwal, S., Meyarivan, T.: A fast and elitist multiobjective genetic algorithm: NSGA-II. *Evol. Comput. IEEE Trans. Evol. Comput.* **6**(2), 182–197 (2002)
56. Fonseca, C.M., Fleming, P.J.: Multiobjective optimization and multiple constraints handling with evolutionary algorithms—Part II. *IEEE Trans. Syst. Man Cybern. Part A Syst. Hum.* **28**(1), 26–37 (1998)
57. Ray, T., Tai, K., Seow, C.: Multiobjective design optimization by an evolutionary algorithm. *Eng. Optim.* **33**(3), 399–424 (2001)
58. Ngatchou, P., Zarei, A., El-Sharkawi, M.A.: Pareto multi objective optimization. In: *Proceedings of the 13th International Conference on Intelligent Systems Application to Power Systems* (2005)
59. Bard, J.F., Purnomo, H.W.: Preference scheduling for nurses using column generation. *Eur. J. Oper. Res.* **164**(2), 510–534 (2005)
60. Todorovic, N., Petrovic, S.: Bee colony optimization algorithm for nurse rostering. *IEEE Trans. Syst. Man Cybern. Syst.* **43**(2), 467–473 (2013)
61. Jain, A.K.: Data clustering: 50 years beyond K-means. *Pattern Recogn. Lett.* **31**(8), 651–666 (2010)
62. Redmond, S.J., Heneghan, C.: A method for initialising the K-means clustering algorithm using kd-trees. *Pattern Recogn. Lett.* **28**, 8 (2007)
63. Arthur, D., Vassilvitskii, S.: k-means++: the advantages of careful seeding. In: *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms* (2007)
64. Stetco, A., Zeng, X.-J., Keane, J.: Fuzzy C-means++: fuzzy C-means with effective seeding initialization. *Expert Syst. Appl.* **42**(21), 7541–7548 (2015)
65. Celebi, M.E.: Improving the performance of k-means for color quantization. *Image Vis. Comput.* **29**(4), 260–271 (2011)

# A Technology Vision of the Fifth Generation (5G) Wireless Mobile Networks

Zoran Bojkovic<sup>(✉)</sup> and Dragorad Milovanovic

University of Belgrade,  
Studentski Trg 1, 11000 Belgrade, Republic of Serbia  
z.bojkovic@yahoo.com,  
dragoam@gmail.com

**Abstract.** 5G is now the next generation of wireless communication systems. The general demonstration of one technology vision for fifth generation mobile networks is presented in this work. The vision of 5G technology provides guidance for the definition of requirements, architecture, and other aspects. By expanding the performances limits of mobile networks, it is necessary for the 5G to include a flexible designing which can optimize network utilization with large range of examples for partnership and business models. The 5G should include flexibility to optimize the network usage by design. Modular network functions with the ability of on demand deployment and scaling capabilities need to be included in the architecture. The purpose of this inclusion would be to handle the accommodation of various demands in a cost-effective and agile manner. Direct communication in D2D services is one of the new 5G networks characteristics. Advanced technologies such as massive MIMO and millimeter-wave radio systems have significant impact on design of cellular architecture. It is crucial for 5G technology to maintain a massive traffic volume owing to increase efficiency of radio link and density of cells. Also, it is necessary for the network to be transformed in a cloud architecture, which coordinates radio resources of multi-radio access, and inter-cell interference during network deployment. Research trends in 5G tell us that using an aggregation of technologies, it is possible to realize numerous goals. Our use case study for the perspective period post-2020 shows extremely broad variety of applications and their attributes performance.

**Keywords:** Wireless technology · Research activities · Standardization process · Network densification · Cloud architecture · Use case study · 5G requirements

## 1 Introduction

The fifth generation (5G) is the next step in the evolution of mobile communication. 5G represents a key component in the vision of unlimited access to information and data sharing. The next generation is positioned to fulfill the requirements (large increment in traffic volume/density and connectivity, including multi-layer densification) in a broad range of use cases. Industry demands 10x higher data rate for users and 1000x more capacity. Additionally, devices directly communicate (D2D) to each other, including vehicle-to-vehicle (V2V), machine-to-machine (M2M), etc. All impose different



requirements to be optimized in mobile broadband data access [1, 2]. The long battery lifetime, high reliability, massive number of devices, very low response time, require further improvements. Reducing the power consumption in cellular network is particularly a challenge because a simultaneous increase in peak data rates and capacity is needed. There is a requirement for 5G to achieve a spatio-temporal consistent user experience in a highly heterogeneous environment consisting of multiple access technologies, multilayer networks, several types of different devices with different types of user interactions, etc. The requirements can be grouped along with identification of system devices, user, business demands, network management and service enhancement. Mobile network operators are under the stress of continuously incremental demand for higher data rates, larger network capacity, high spectral, and energy efficiency [3]. Technology trends provide an insight about the approach to be taken in order to limit the gap which is present between the expected and existing capabilities.

Many research trends in 5G have demonstrated that a lot of goals can be achieved with the aggregation of technologies. Since initial support of the European Commission to the research projects in early 2013, the pursue of innovative solutions for 5G has begun worldwide. Standardization bodies share the vision that target year for the initial commercialization of 5G networks is 2020. The final goal is to achieve seamless communications between machine-to-machine (M2M), human-to-human (H2H), and human-to-machine (H2M). However, it is expected that 5G will be a collision of different radio access technologies (RATs), differently sized network tiers, backhaul connections, transmit powers, and number of heterogeneous and smart devices [4].

The structure of this overview paper is as follows. We start with an outlook of wireless technology evolution, continue with 5G research activities and standardization processes. The emphasis is on key advantages of 5G technology, such as massive MIMO, mm-wave radio-technology, device-to-device (D2D) communication as well as heterogeneous cloud radio-access networks. Next, network densification impact on spectral efficiency gain and transmit power reduction are analyzed. The presentation concludes use case study of selected applications.

## 2 Wireless Technology Evolution and Research Activities

The technology of wireless phones is dated back to the middle of the 20<sup>th</sup> century and they were not considered to be portable handset. The usage and most probably the manufacturing of the first ever wireless phones as well was in Japan in 1979. About the same time, Nordic Mobile telephone (NMT) was also working towards the development of similar wireless phone technologies [5]. A few years later, Motorola mobiles in USA made the wireless phone technology.

The first generation (1G) mobile networks was introduced in 1970. The systems were referred to as cellular due to the method by which the signals were handed off between stations. Cell phone signals were based on analog system transmission. The global mobile phone market grew by 30–50 percent annually. The number of subscribers worldwide reached approximately 20 million by 1990 [6]. Some of the most popular standards for 1G systems were NMT, AMPS (Advanced Mobile Phone System), TACS (Total Access Communication Systems).

2G phones were introduced in the early 1990s. The GSM (Global Systems for Mobile Communications) network offered digital modulation to improve voice quality and coverage, together with additional data services, such as paging, faxes, text messages and voice mail. GSM networks deployed time division multiple access (TDMA) and code division multiple access (CDMA) technologies [7]. An 2.5G intermediary phase uses the GPRS (General Packet Radio Service) which delivers packet-switched data capabilities to existing GSM networks. The Internet becomes popular and the importance of Internet Protocol (IP) in packet-switching increased. Phones start supporting Web browsing, multimedia services and streaming. EDGE network increases data rate from to 20–40 Kbps, with 171,2 Kbps as a peak value.

3G phones were introduced in 1998. Globally standardized 3G service UMTS (Universal Mobile Telecommunication System) sustain higher data rates up to 2 Mbps and open the way to Internet applications. Various mobile devices compatible with a set of standards could be used throughout the world. One possibility could be the use of UMTS which allows for internet access from any location together with global roaming. An 3.5G intermediary phase towards the road of 4G generation supports higher throughput HSPA (High Speed Packet Access) technology with the peak speed of 14,4 Mbps.

4G technology is commercially available since 2006. It provides rates of transmission up to 20 Mbps and simultaneously Quality of Service (QoS) which allows prioritizing traffic according to the type of application. In order to keep pace with the demand for data access by several services, the speeds for 4G are being further increased. Also, worldwide roaming becomes reality. A wider bandwidth is provided to vehicles and devices moving at high speeds within the access area by 4G networks. Long Term Evaluation (LTE) is the latest wireless cellular system since 2009, which comes with a new OFDMA (Orthogonal Frequency Division Multiplexing) air interface and an All-IP network. LTE-Advanced is deeply optimized for wireless data throughout all the layers of its protocol stack, unlike HSPA (High Speed Packet Access) technology which is constrained to operate within 3G networks. It promises several improvements in performance in addition to the increased data rate. This achievement is due to the support provided by technologies such as coordinated multi-point enhanced multi-antenna capabilities, carrier aggregation, relaying and improved inter-cell interference coordination for LTE cells as well as heterogeneous networks (HetNets) and small cells supporting aggressive spectrum spatial reuse.

Designing, dimensioning and optimization of telecommunications and ICT infrastructure have changed substation only over the recent years in line with changes towards 4G/5G wireless multiservice networks, networks convergence, and mobile communications as well as. A substantial increase in the amount and complexity of issues which were previously tackled by engineers and theoreticians follows each generation of technology. However, the essential tasks in developing engineering tools and algorithms for the design, analysis dimensioning and optimization of 5G wireless networks, remain the same:

- (a) develop new technologies for increasing radio network capacity,
- (b) evaluation and determination of the relationship between the QoS/QoE parameters and the parameters characterizing traffic sources (services),

- (c) control and optimize the usage of radio network resources, and
- (d) enhance the capabilities of data transport, transmission, and reception between end users and the core network.

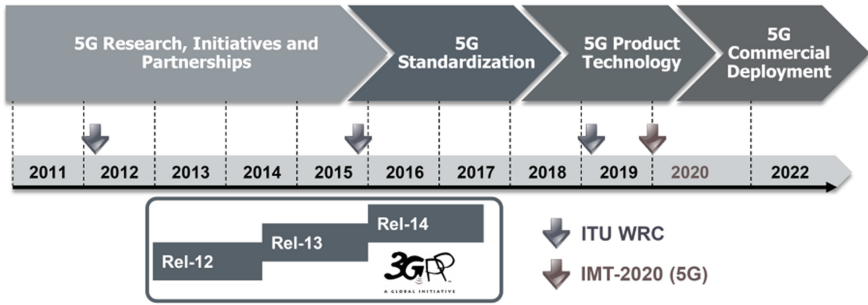
The move to 5G has a greater focus on applications as well as a novel approach to operator and vendor business models. This will derive a further set of dimensions to simulation, measurement, and validation with a unique emphasis on software and applications relating to network performance. Today, 5G wireless technology holds a place amongst the biggest research areas within both industry and academia. A number of technology components which aim to achieve ambitious goals are being revealed through research. Some examples of the most important 5G research initiatives and activities include:

- Research projects funded by EU as part of the 7<sup>th</sup> Framework Program (FP7) started in September 2012,
  - *METIS* research project as part of the FP7 started in November 2012,
  - *Horizon2020* is seven years EU FP8 research and innovative program (2014–2020),
  - *5GPP* infrastructure public-private partnership project is a joint initiative between the ICT industry and the European Commission.
- *5G Innovative Centre*: 5G research center in the UK started in November 2013,
- China: IMT-2020 and Future Forum (February 2013)
- Taiwan: Office Science and technology working with the National Science Council on a print for 5G development in 2014.
- *5G Forum*: Korean industry-academy-R&D cooperation system, established in May 2013,
- Japan: ARIB established as a new working group (2020 and beyond) in September 2013.
- USA: Several universities - led research projects sponsored by key industry players.

Some of the most important 5G standardization processes include:

- 3rd Generation Partnership Project (3GPP) started its work on 5G in March 2015 and defined a timeline for the standardization.
- ITU Radiocommunication Sector (ITU-R) setting the stage for 5G research activities in early 2012, started a program to develop next-generation IMT-2020 cellular systems.

3GPP joined together seven global telecommunications SDOs (ETSI, ARIB, TTC, CCSA, TTA, TSDSI, ATIS), and has made provision for their members to have access to a stable environment producing Specifications and Reports which define the technologies. The creation and organization of the standards for a number of mobile communication systems including 2G, 3G, HSPA, LTE and 5G has emerged from this collaboration. In 2015, 3GPP began working on four technical reports outlining the new services and markets technology enablers, based on potential 5G requirements (Fig. 1). The completion of these reports were attained in June 2016 and used as input for R15 release, set for the standardization of the first phase of 5G requirements by June 2017.



**Fig. 1.** 3GPP and ITU timeline for 5G: research, standardization, and technology.

ITU-R Working Party 5D has finalized its timeline towards IMT-2020. The finalization of ITU-R’s version of 5G mobile broadband and connected society (Table 1) was done in September 2015.

**Table 1.** ITU standardization process for defining 5G technology.

Vision		Defining the technology	
2012–2015	2016–2017	2018–2019	2019–2020
Development plan	Spectrum arrangements	Proposals	Spectrum arrangements
Market/services view	Technical performance requirements	Evaluation	Decision & radio framework
Technology/research kickoff	Evaluation criteria	Consensus building	Detailed IMT-2020 radio interface specifications
Vision & framework	Invitations for proposals	CPM report	Future enhancement/update plan & process
Name IMT-2020	Sharing study parameters	Sharing study reports	
<6 GHz spectrum view	Sharing studie		
>6 GHz technical view			
Process optimization			

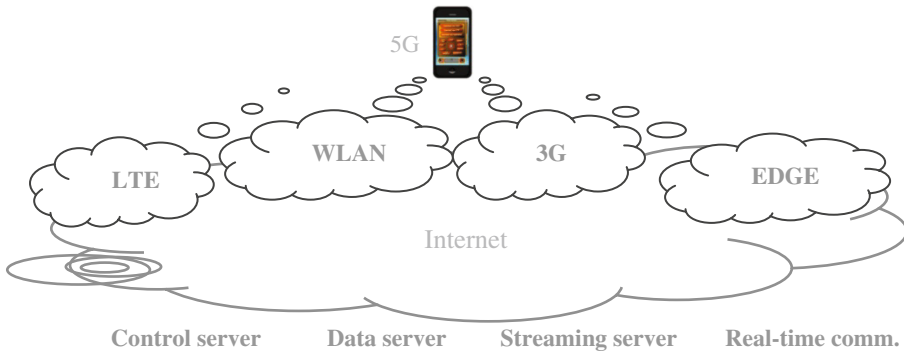
### 3 Technology Components

In what follows key advantage of 5G technologies has to be pointed out such as:

- A large bi-directional bandwidth with high resolution is offered by the technology.
- Containment of billing interfaces with advanced features.

- Provision of supervision tools for high-end subscribers.
- Massive data (in Giga bit) broadcasting.
- Highly performing transporter class gateway.
- Statistics about traffic with high accuracy.
- Option for the management of bandwidth from remote sites.
- Very high speeds for both download and upload.
- Improved and enhanced connectivity across the globe.

Block scheme for designing of All-IP mobile network is presented in Fig. 2.



**Fig. 2.** 5G network model of All-IP system for wireless and mobile networks interoperability.

The achievement of the following ambitious goals is aimed by a number of technology components:

- massive (large-scale) multiple input- multiple output (MIMO),
- millimeter wave (mmWave),
- network densification,
- heterogeneous dense networks,
- heterogeneous cloud radio access networks,
- direct device to device (D2D) communications in the inband and outband form.

5G has diverse requirements, however, not all of these will need to be satisfied at once because different applications make different demands on system performance. The question often arises when discussing candidate technology components is what about the increase of the data rate. As a conclusion, the required increase in data rate could be achieved through:

- densification of the network to an extreme level with the aim to improve the area spectral efficiency, which means more nodes per unit area and frequency
- making better usage of the 5 GHz unlicensed spectrum by moving towards millimeter-wave (mmWave) spectrum with the aim for increased bandwidth
- making use of the advances in Multiple-Input Multiple-Output (MIMO) techniques to obtain an increased spectral efficiency.

When planning the future cellular system capacity, we based on the well-known *Shannon* theory of mathematical communication model, as an example of a cellular system, where following relation can be applied [8]:

$$R \leq C = m(B/n) \log_2(I + S/(S + N)) \quad (1)$$

This means that the throughput per user  $R$  is upper bounded by capacity  $C$  of an additive white *Gaussian* noise channel. The parameter  $m$ , integer spatial multiplexing factor, denotes the number of spatial streams between a base station (BS) and user equipment (UE),  $B$  denotes the BS signal bandwidth, while the load factor – integer parameter  $n$ , denotes the number of users sharing the given BS. Finally,  $S$  denotes the desired signal power, while  $I$  is the interference power and  $N$  represents the noise power at the receiver end.

### 3.1 Benefits and Shortcomings of Massive MIMO

A cellular base station (BS) serves a multiple number of single-antenna terminals over identical time-frequency intervals. Efficient time division duplexing (TDD) based on reverse-link pilots enables the BS to estimate the reciprocal channels. A 5G system uses massive MIMO technology with even hundreds of BS antennas. Massive MIMO relies on phase-coherent but computationally very simple processing of the signal from all antennas. The more BS antennas is equipped with, the more possible signal paths and the better the performance, on the other hand, more antennas increases complexity of signal processing and energy consumption. Nevertheless, the advantages of massive MIMO is so immense that the effect of thermal noise becomes negligible, while the system performance can only be limited by the pilot pollution [9]. High-order multiuser transmission and improved SNR provide increase in spectral efficiency of the system. This can support high-order multiuser communication. The same time and frequency resources are shared by a large number of users without any significant interference with each other [10]. Some specific benefits of a massive MIMO system are [11]:

- A 10 times or more increase in capacity is obtained by Massive MIMO with a simultaneous 100 times improvement in the efficiency of radiated energy.
- The aggressive spatial multiplexing causes an increase in capacity.
- The fundamental principle that makes the significant increase in energy efficiency possible is that with a large number of antennas, energy can be focused with extreme sharpness into small regions.
- It can be implemented using low-power and inexpensive components.
- Using the flow of large numbers and beamforming in order to avoid fading tips, a significant reduction of latency on the radio interface is enabled on the system.
- Multiple access layer is simplified because each subcarrier will have substantially the same channel gain.
- The whole bandwidth can be allocated to each terminal, thereby making redundant most of the physical layer control signaling.
- These systems offer a lot of possibilities to eliminate harmful signal.

Before incorporating massive MIMO in the 5G cellular architecture [12], recent research trends have to be pointed out. For example, beamforming requires a large amount of channel state information which is not convenient for the downlink. Massive MIMO shortcoming is as follows:

- It may not be practical for FDD but due to the channel reciprocity, it can be used in TDD systems [13].
- Massive MIMO lack pilot contamination from other cells with high transmit power, and suffers from thermal noise otherwise [11].

### 3.2 Millimeter-Wave Technology

Bandwidth expansion is an obvious way to increase the throughput. Mobile communication systems use sub-3 GHz spectrum which is becoming increasingly crowded as the mobile traffic demands grow. On the other hand, a vast amount of spectrum in the 3–300 GHz range remains underutilized. mmWave frequencies of 28 GHz and 38 GHz are studied to obtain a proper understand of their propagation characteristics in 5G systems [14].

Two main features of the mmWave technology are identified as large amounts of bandwidth, which enable a very high throughput coverage, and very small wavelength which enables the deployment of several large antennas in a specific area. The main challenges for mmWave communications include large path loss (especially with non-line-of-site propagation), signal blocking/absorption by various objects in the environment, and low transmission power capability of current amplifiers.

Signal attenuation can be combated using large antenna arrays driven by smart beam selection tracking algorithms [15]. With the high density of mmWave BSs, the cost to connect every BS via a wired infrastructure can be very high. One solution is to allow some mmWave BSs to connect to the backhaul via other mmWave BSs. Due to large beamforming gains, the mmWave inter-BS backhaul link can be deployed in the same frequency as the mmWave access link. Cost-effective and low-latency solutions for wireless backhaul will be essential for supporting the envisaged densification in high capacity 5G networks.

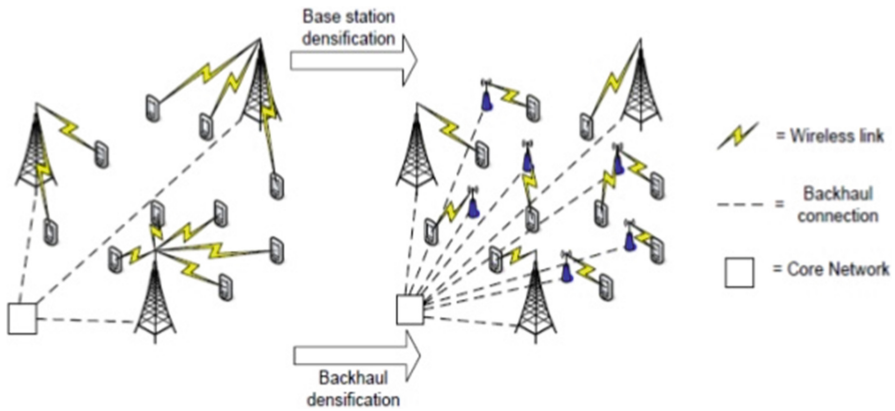
### 3.3 Network Densification

Reduction of the cell size is one of the solutions available to operators for growing data rates demands. The question often arises is how to increase the spatial efficiency through higher frequency range. The answer is to reduce directly the cell size. As for reducing the transmit power, it can be done if the propagation power loss is lower. On the other hand, coverage is improved by deploying indoor cells under the assumption of serving overflow traffic from macro cells when required.

Cell site density needs to be drastically increased to accommodate large volume of traffic within a small geographical area. In the first three generation of cellular networks, cell layouts had rather homogeneous topology, i.e., system base stations had the

same configurations for transmitting power and antenna gain. Base stations were more or less equally distanced. Cell splitting can be viewed as offloading traffic from macro base stations to low-power nodes, which help fill the coverage holes in a homogeneous layout [16]. In order to provide a convenient coverage as well as improved spectral and energy efficiency, heterogeneous networks (HetNets) can be considered as a perspective solution. HetNets represent the concurrent operation of different technologies, as well as various BSs classes (i.e., macro, pico, and femto). It introduces centralized or distributed processing, coverage and capacity, inter-cell coordination, and improved performance at cell edge.

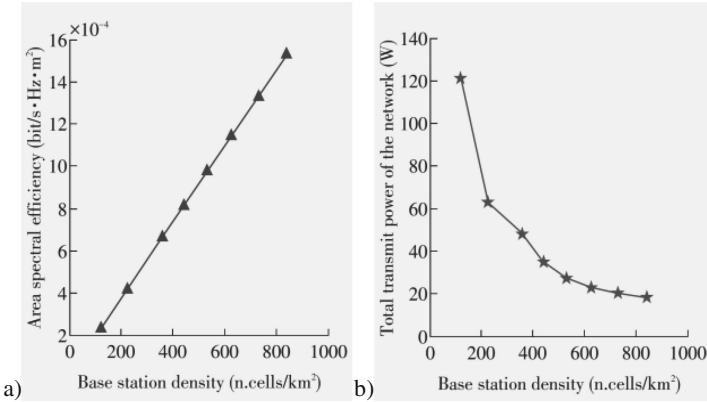
Base stations become smaller and more numerous and more users locate within the same spectrum. As for base station classification, widely spread backhaul network is required to support it. Hence, the number of backhaul links will increase along with the number of base stations. Network and backhaul densification are shown in Fig. 3.



**Fig. 3.** Network and backhaul densification (backhaul connections can be either wired or wireless) [16].

Two advantages of increasing the BS density are shown in Fig. 4. The decrease of total transmission power of the network while maintaining linear area spectral efficiency (ASE) gain is shown in Fig. 4a. And the reduction in total transmission power as base station density increases is shown in Fig. 4b. Therefore, a higher throughput and reduced overall radiated power by the base station antennas is enabled by the increase of the base station density. The reduction in the overall transmit power may have positive effects for reducing the aggregate interference in primary network which share spectrum with a secondary system of small cells [18]. It is useful in future scenarios based on shared access schemes LSA/ASA in which small-cell networks exploit new spectrum-sharing opportunities.





**Fig. 4.** Reduction in total transmit power of the network: (a) area spectral efficiency gain vs. base station density, (b) transmit power reduction vs. base station density [17].

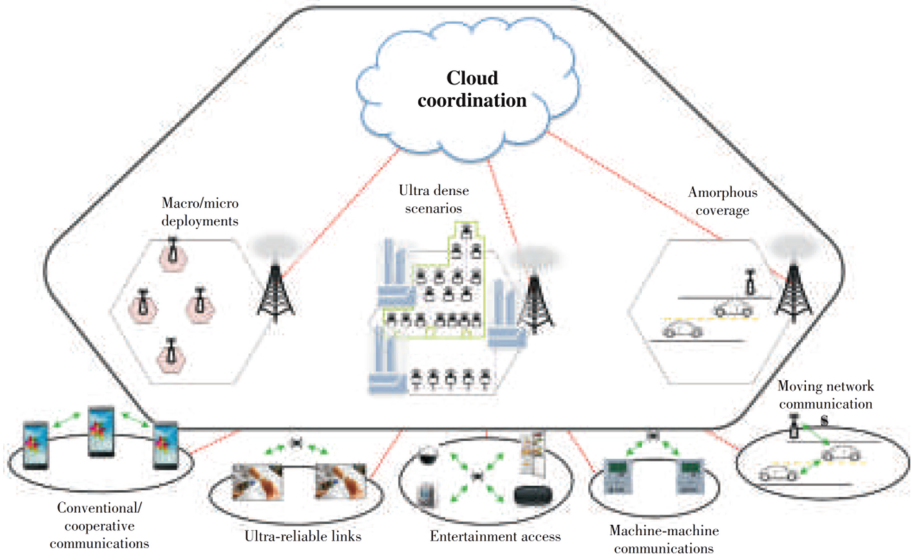
### 3.4 Heterogeneous Cloud Radio Access Network

5G networks are expected to support multiple radio access technologies (RATs) with overlapping coverage creating rich opportunities for combining and aggregating capacity. In order to realize this, 5G systems need to support an integrated virtual access network. Single end-to-end network architecture enables joint management and simultaneous use of radio resources to significantly improve network capacity and reliability of wireless link.

Cloud architecture is a solution for coordinating different categories of network and radio resources in a seamless and transparent method. 5G networks are a mix of new network components and existing systems as well as RATs of a non-cellular kind. Within each generation of networks, the allocated spectra may be different, depending on country, operator, or year of deployment. Highly licensed and unlicensed spectrum may be included in radio resources. A cloud architecture for unified coordination of network access in converged multi-RAT network is shown in the Fig. 5.

H-CRAN (Heterogeneous Cloud Radio Access Network) based 5G systems are traffic-driven and user-centric. The H-CRAN centralized server platform receives and processes the baseband signals from several hundred BSs when small cells are connected to macrocells with low-latency high-rate backhaul [19, 20]. The evolution of H-CRAN includes even more advanced techniques such as joint resource allocation across multiple RATs and simultaneous processing of multiple users signals to further increase capacity of 5G systems. Network is characterized by attributes of both cloud computing and software defined functionality. In order to improve spectral and energy efficiency [21], four key H-CRAN functionalities need to be implemented:

- The self-organizing functionality of H-CRAN automatically configures and optimizes the traffic, front-haul, and radio resources, regulating operations of the management plane without human interaction.



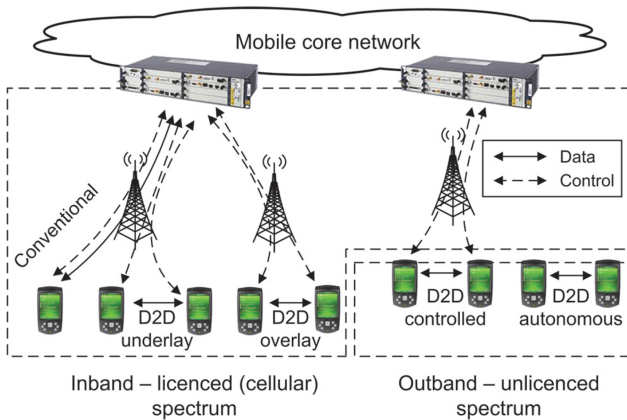
**Fig. 5.** 5G unified cloud architecture: network access using cloud-coordination [16].

- The radio resource cloudization technology decreases the inter-tier interference and improve reuse of the radio resources.
- The cognitive processing technique is used in the overlaid scenario to avoid inter-tier interference and make small cells work cooperatively when the overall system load is not high.
- The big data mining functionality bring in machine learning (ML) technology into large-scale cooperative signal processing and intelligent networking. Adaptive ML techniques prepare 5G system for the future Internet/IoT traffic properties and user demands.

### 3.5 Direct Device-to-Device (D2D) Communication

The main goal for direct device-to-device communications (D2D) is to improve the overall spectral efficiency for wireless systems. User terminals are in a position to form a direct link without the influence of BSs and core networks. Also, a special case of the densification of network is D2D communication. D2D communication helps increase the density of low-power nodes (LPN) with wireless backhaul. A device acts as a LPN for unicasting, multicasting or broadcasting traffic directly to the user without being routed through the network [10]. In 5G networks, the system capacity is expected to be increased by D2D, especially for big outdoor events in dense urban environments. This is especially useful for proximity services where sharing and exchanging local information by users in the vicinity is of concern.

D2D communications are divided into two classes: inband and outband, as shown in Fig. 6 [22]. Inband D2D communications occur in licensed spectrum. It means that cellular spectrum is used for both D2D and cellular links. The motivation for choosing inband communications is high reliable control over licensed spectrum. Inband communications are further divided into underlay and overlay categories. In the case of underlay approach, cellular and D2D communication share the same radio resources. On the other hand, in overlay communication, cellular resources are dedicated to D2D links. Inband D2D communication is advantageous in the sense that underlay concept increases the cellular spectral efficiency by exploiting the spatial diversity. Cellular spectrum is controlled by BS. Thus, the QoS provisioning is not a challenging issue. On the other hand, cellular resources might be wasted in overlay D2D concept, while the interference management among D2D and cellular transmission in underlay is challenging [23]. Inband power control and interference management solutions usually explore complex resource allocation methods. The main criteria of choice is higher spectral efficiency.



**Fig. 6.** Two classes of direct device-to-device communications: inband and outband.

The major motivation in using outband D2D communication is the elimination of the interface between direct and cellular lines. An extra interface is required by unlicensed spectrum and is usually adopted in two environments: WLAN and WPAN wireless networks. Outband D2D communications are further divided into autonomous and controlled communication. In autonomous outband communications, the cellular network controls all the communication but leaves the D2D communication to the users, while in controlled outband communication, the management of a second interface is under the cellular network. It means that the secondary interface is not under cellular control. Since outband communication does not occur in the cellular spectrum, there is no interference issue as in inband. However, only mobile devices with two different interfaces (WiFi and LTE) can use outband D2D. In that way users can have simultaneous maintaining D2D and cellular communication. The significant benefit of outband communications are absence of interference between cellular and

D2D users. It should be added that there is no need to dedicate cellular resources to D2D spectrum like in overlay inband approach [15]. Outband D2D communications have some disadvantages. The BS does not control the interference in unlicensed spectrum. Also, the efficient power management between two wireless interfaces is crucial because the power consumption of the device can be increased.

## 4 Use Case Study

5G supports evolution of established mobile broadband use cases encompassing various applications with variable performance attributes. Video applications which are delay-sensitive evolve to ultra-low latency. Best effort applications evolve to reliable and ultra-reliable. Vehicular high speed entertainment applications evolve to mobility on demand for connected objects. Furthermore, use cases will be delivered across a wide range of devices and across a fully heterogeneous environment. The use case study is an input to determine requirements and defining the technology base in the architecture for 5G. The use cases are a tool to ensure the comprehension of the level of flexibility required in 5G. NGMN (Next generation mobile network) alliance has come up with twenty-five use cases for 5G [24–26]. The eight groups with representative examples and imposed requirements are shown in Table 2.

**Table 2.** The 5G use case study and resulting requirements on technology.

Use cases	Representative examples	5G requirements
Broadband access in dense area	Pervasive video	Extremely high resolution video, end-to-end latency and data rate, concurrently active connections
	Smart office	Ultra-high bandwidth-intensive processing, instant communication by video
	Operator cloud services	Higher QoE as well as seamless interworking across clouds, networks, and devices
	Video/photo sharing	Ultra-high connection density, low latency, and high data rate
Broadband access everywhere	50+Mbps	Minimum user data rate and not a single user's theoretical peak rate
	Ultra-low cost networks	Low-cost deployment and operation of networks infrastructure and terminals
High user mobility	High speed train	Passengers satisfied with the service (e.g. up to 1000) at a speed of 500 km/h

(continued)

**Table 2.** (continued)

Use cases	Representative examples	5G requirements
	Remote computing	Robust communication links with low latencies together with high availability
	Moving hot spots	Real-time, Non-stationary, and dynamic provision of capacity
	3D connectivity and aircrafts	Typical routes have an altitude of up to 12 km, sporting event live services
Massive Internet of Things	Smart clothes	Overall management of the number of devices and sensors
	Sensor networks	High density of devices in a common communication and interworking framework
	Mobile video surveillance	Secured and highly reliable network with instant interaction and the right performance
Extreme real-time communications	Tactile internet	Audio and/or visual feedback and tactile control signal and in real-time within sub-millisecond
Lifeline communications	Natural disaster	Robust communications in efficient network and user terminal energy consumptions
Ultra-reliable communications	Automated traffic control	Ultra-low end-to-end latency communication with high reliability, and high scalability
	Collaborative robots	Low latency and high reliability of underlying control network
	eHealth	Reserve/prioritize capacity, security, authentication management, identity, and privacy
	Remote object manipulation	Very strict requirements in terms of security, latency, and reliability
	3D connectivity drones	Remote control system
Broadcast-like services	News and information	Real-time or non-real time services having a wide distribution focusing on either address-space focused (many end-users) or geo-location
	Local/regional/national/continental services	

#### 4.1 Broadband Access in Dense Areas

This group focuses on service availability in densely-populated areas where for each square kilometer, there are thousands of people who live and work. An increasingly significant role will be played by three-dimensional (3D) services, and multi-user interaction. An essential aspect at the network close to the user is context recognition in ensuring delivery of consistent and personalized services to the customers. The following use cases are included in this group.

- **Pervasive video** with extremely high resolution in person-to-group communication or person-to-person in everyday life will have much more advanced capabilities.
- **Smart office** scenario in where ultra-high bandwidth-intensive applications are required by hundreds of users, instant communication by video as well as vast amount of data processing in a cloud.
- **Operator cloud** support the future value-added services which need for higher QoE as well as seamless interworking across networks, and devices.
- **HD video/photo sharing** in open-air gathering/stadium requires ultra-high connection density, low latency, and high data rate combined altogether.

#### 4.2 Broadband Access Everywhere

A minimum guaranteed data rate is needed everywhere for consistent user experience. Cost of infrastructure deployment is a key factor in further development of services. This group includes the two most important cases.

- **50+Mbps everywhere** broadband access is considered as the minimum user data rate in the coverage area, even at cell edges. The target value between 50–100 Mbps will be indicative depending upon the 5G technology evolution.
- **Ultra-low cost networks** deployment in scarcely populated areas to offer Internet access and provide the ability for new businesses to develop in underserved areas.

#### 4.3 Higher User Mobility

The degree of mobility required depends upon the specific traffic devices. For example, vehicle demands accessing the Internet, enhanced connectivity for in-vehicle entertainment, autonomous driving, safety and vehicle diagnosis and, enhanced navigation through instant and real-time information. The following use cases are included in this group.

- **High speed train** reaches 500 km/h with passengers using high-quality mobile Internet for interaction, information, entertainment or office-like applications.
- **Remote computing** in transportation industry enables the ease of vehicle maintenance and offers novel services to customers in public transport.

- **Moving hot spots** is becoming a challenge in 5G which complements the stationary mode of planning in capacity and incorporates non-stationary, real-time, and dynamic provision of capacity.
- **3D connectivity** is implemented in civil aviation commercial services and 3D sporting event live services.

#### 4.4 Massive Internet of Things (IoT)

IoT involve the interconnection of huge number of devices such as actuators, sensors, and cameras with a wide range of demands and characteristics. The use case includes/long-range/low-cost/low-power mobile type communication (MTC) and broadband with similar characteristics closer to human type of communication (HTC). The cases are as follows.

- **Smart clothes** involve a number of low-power, ultra-light, waterproof sensors integrated in people's clothing. These sensors measure various environmental and health attributes like body temperature, heart rate, blood pressure, breathing rate and volume, etc. The key challenge is the overall management of the number of devices, as well as the data and applications.
- **Sensor networks** require very high density and low-cost devices with high battery life. The aggregation of all smart services a common communications and inter-working framework is challenging task.
- **Mobile video surveillance** may evolve to be available on traffic devices, as well as safety and security personal for monitoring targeted area, specific events, etc. These applications require a high reliable and secure network with the right performance and instant interaction with back-end and remote systems.

#### 4.5 Extremely Real-Time Communications

Use cases with strong demands for real-time interactions may require extremely high reliability, throughput, mobility, etc. Remote computing, with strict latency requirement, need highly available and robust communication links.

- **Tactile Internet** is a system where humans wirelessly control real and virtual objects in tactile interaction with audio/visual feedback. The main challenge is the expectation of real-time reaction to be within sub-milliseconds.

#### 4.6 Lifeline Communications

These use cases include new applications for authority-to-authority communication, emerging disaster relief and prediction. Lifeline communications require the ability to support traffic surges and a very high level of availability.

- **Natural disaster** requires robust communications in case of earthquakes, tsunamis, floods, hurricanes, etc. Basic communications (voice, text messages) are needed in the disaster area in order to signal location/presence of survivors. Several days of user terminal operation should be supported.

#### 4.7 Ultra-Reliable Communications

These use cases include applications requiring extremely low latencies and involving significant growth in remote control and operation. The following use cases are included.

- **Automated traffic control and driving** in advanced applications for safety which reduce the road accidents and improve traffic efficiency, require high reliability, low latency and high scalability 5G networks.
- **Collaborative robots and control network for robots** in applications with diverse tasks in different environment, require an underlying control network with very low latency and high reliability. For many robotics scenarios, a round-trip reaction time of less than 1 ms is anticipated.
- **E-Health** mobile applications of remote health monitoring include immediate and automatic surveillance of patients. The application is life critical and the system requires reserve/priorities capacity for the related communications, including out of coverage warnings. For each device, identity, privacy, security and authentication management must be ensured.
- **Remote object manipulation** in application like remote surgery has very strict requirements in terms of security, latency, and reliability.
- **3D connectivity drones** require terrestrial and up-in-the-air locations coverage in logistics application.
- **Public safety** requires enhanced and secure communications which include real-time video and sending high-quality pictures. Also, one of the main challenges is to ensure reliable communications over the entire footprint of the emergency services. Priority over the traffic is also required together with an ability for direct communications between devices and high security.

#### 4.8 Broadcast-like Services

Real-time or non-real-time services are characterized by having wide distributions which can be either geo-location focused or address-space focused with many end-users. These services may distribute content as being currently done, but also provides a feedback channel for interactive services or acknowledgment information.

- **News and information** include receiving text/pictures, audio and video, everywhere and as soon as events happen.
- **Local broadcast-like services** is active at a cell level with a reach of 1 to 20 km (advertisements, festivals, fairs, congress/convention, stadium services as well as local emergency services).



- **Regional broadcast-like services** are required within 1 to 100 km (communication of traffic jam information, emergency warnings).
- **National broadcast-like services** are interesting as a substitute or complementary to broadcast services for radio or television. Industries will benefit from national broadcast like services to upgrade/distribution of firmware.

## 5 Conclusion

Fifth generation mobile systems aggregate best technologies and have an extraordinary capability to support software. The routing and switching provide high connectivity. Also, Internet access to nodes are distributed and can be deployed with wireless/wired network connections. The advanced systems such as massive antennas, millimeter-wave radio systems, and direct communication have significant impact on new heterogeneous network architecture. Cloud computing becomes a promising solution for high energy and spectral efficiency. However, cloud virtualization of communication hardware and software impose stress on architecture and protocols in software defined radio networks.

5G systems become adapted to big traffic load, which is fundamental for service ubiquity and to sustain a massive number of connections. The three requirements for 5G networks are always services presence and availability, a massive number of connections and energy efficiency. It is of importance to focus improvements in the following areas: network capacity, consistent user experience, flexibility, efficiency, innovation. 5G vision of these technologies provides guidance for the definition of architecture, requirements, and other aspects.

Consistent user experience is the driving force of 5G. Researcher in the field of wireless and mobile operators will come with innovative ways of converging networks, devices, and services. Versatile and intelligent 5G networks are able to support applications for great benefits of individuals and organizations. In that way, 5G systems will be comprehensive and able to penetrate in many aspects of people everyday life.

## References

1. Hu, F.: Opportunities in 5G Networks: A Research and Development Perspective. CRC Press, Boca Raton (2016)
2. Dahlman, E., et al.: 5G wireless access: requirements and realization. *IEEE Commun. Mag.* **52**(12), 42–47 (2014)
3. Osseiran, A., et al.: Scenario for 5G mobile and wireless communications: the vision of the METIS project. *IEEE Commun. Mag.* **52**(5), 26–35 (2014)
4. Hossain, E., et al.: Evolution towards 5G multi-tier cellular wireless networks: an interference management perspective. *IEEE Wirel. Commun.* **21**(3), 118–127 (2014)
5. Santhi, K.R., et al.: Goals of the true broad band's wireless next wave (4G–5G). In: *Proceedings of IEEE Vehicular Technology Conference, VTC 2003*, vol. 4, pp. 2317–2321 (2003)

6. Ramnarayan, N., et al.: A new generation wireless mobile network – 5G. *Int. J. Comput. Appl.* **70**(2), 26–29 (2013)
7. Gupta, A., Jha, R.K.: A survey of 5G network: architecture and emerging technologies. In: *IEEE Access Special Section: Recent advances in Software Defined Networking for 5G Network*, vol. 3, pp. 1206–1232 (2015)
8. Shannon, C.E.: A mathematical theory of communication. *Bell Syst. Tech. J.* **27**, 623–656 (1948)
9. Marzetta, T.L.: Non-cooperative cellular wireless with unlimited number of base station antennas. *IEEE Trans. Wirel. Commun.* **9**(11), 3590–3600 (2010)
10. Bhushan, N., et al.: Network densification: the dominant theme for wireless evolution into 5G. *IEEE Commun. Mag.* **52**(2), 82–89 (2014)
11. Larson, E., et al.: Massive MIMO for next generation wireless systems. *IEEE Commun. Mag.* **52**(2), 186–195 (2014)
12. Chin, W.H., Fan, Z., Haines, R.: Emerging technologies and research challenges for 5G wireless networks. *IEEE Wirel. Commun.* **21**(2), 106–112 (2014)
13. Choi, J., Love, D.J., Madhow, U.: Limited feedback in massive MIMO systems: exploiting channel correlations via noncoherent trellis-coded quantization. In: *Proceedings of Conference on Information Sciences, CISS 2013*, pp. 1–6 (2013)
14. Rapaport, T.S., et al.: Millimeter wave mobile communications for 5G cellular: it will work! *IEEE Access* **1**(2), 335–349 (2013)
15. Bojkovic, Z., Bakmaz, B., Bakmaz, M.: Recent trends in emerging technologies towards 5G networks. In: *Proceedings of Advances in Circuits and Systems, Signal Processing and Telecommunications, CSST 2015*, pp. 137–143 (2015)
16. Yuen, Y., WuZhao, X.: 5G vision, scenarios and enabling technologies. *ZTE Commun.* **13**(1), 3–10 (2015)
17. Marchetti, N.: Towards 5<sup>th</sup> generation wireless communication systems. *ZTE Commun.* **13**(1), 11–19 (2015)
18. Galliotto, C., Marchetti, N., Doyle, L.: The role of the total transmit power on the linear area spectral efficiency gain of cell-splitting. *IEEE Commun. Lett.* **17**(12), 2256–2259 (2013)
19. Peng, M., et al.: Heterogeneous cloud radio access networks: a new perspective for enhancing spectral and energy efficiency. *IEEE Wirel. Commun.* **21**(6), 126–135 (2015)
20. Peng, M., et al.: System architecture and key technology for 5G heterogeneous cloud radio access networks. *IEEE Netw.* **29**(2), 6–14 (2015)
21. Bojkovic, Z., Bakmaz, M., Bakmaz, B.: Research challenges for 5G cellular architecture. In: *Proceedings of TELSIKS 2015*, pp. 215–222 (2015)
22. Mumtaz, S., Mohammed, K., Hug, S., Rodriguez, J.: Direct mobile-to-mobile communications: paradigm for 5G. *IEEE Wirel. Commun.* **21**(5), 14–23 (2014)
23. Feng, D., et al.: Device-to-device communications underlying cellular networks. *IEEE Trans. Commun.* **61**(8), 3541–3551 (2013)
24. Agiwal, M., et al.: Next generation 5G wireless networks: a comprehensive survey. *IEEE Commun. Surv. Tutor.* **18**(3), 1617–1655 (2016)
25. Boccardi, F., et al.: Five disruptive technology for 5G. *IEEE Commun. Mag.* **52**(2), 74–80 (2014)
26. NGMN Alliance: 5G Initiative White paper, February 2015

# An Overview of Cloud RAN: Architecture, Issues and Future Directions

Meruyert Makhanbet<sup>1</sup>, Xuewei Zhang<sup>1</sup>, Hui Gao<sup>1(✉)</sup>,  
and Himal A. Suraweera<sup>2</sup>

<sup>1</sup> School of Information and Communication Engineering,  
Beijing University of Posts and Telecommunications, Beijing 100876, China  
mika.makhanbet@outlook.com,  
{zhangxw, huigao}@bupt.edu.cn

<sup>2</sup> Department of Electrical and Electronic Engineering,  
University of Peradeniya, Peradeniya 20400, Sri Lanka  
himal@ee.pdn.ac.lk

**Abstract.** Cloud radio access network (C-RAN) has been considered as one of the enabling network architectures towards the implementation of fifth generation (5G) wireless systems. Combining with advanced radio, wireless and computing techniques, C-RAN provides great potential to improve the network capacity, spectrum efficiency, energy efficiency and operational flexibility. Therefore, C-RAN has attracted considerable attention from both academia and industry. In this paper, we review the architecture and some of the key physical (PHY) layer signal processing issues of C-RAN. The functional split and limited-capacity fronthaul impose PHY layer issues and signal processing opportunities, e.g., channel state information acquisition, effective compress-and-forward methods and intelligent resource allocation schemes. Moreover, some emerging 5G techniques which can be supported by or integrated with C-RAN architecture, such as software-defined networking, heterogeneous networking, communication at millimeter wave frequencies and full-duplex radio are discussed. The challenges and future research directions are also presented, where mobile edge computing and caching, multi-dimensional resource management and PHY layer security and energy-efficient designs are highlighted.

**Keywords:** Cloud RAN · 5G networks · Fronthaul · HetNet · mmWave · Full-duplex · Mobile edge computing

## 1 Introduction

The volume of the wireless data traffic is increasing rapidly due to the unprecedented popularity of smartphones, tablets and machine-centric devices [1–3]. The explosive growth of wireless data traffic cannot be satisfied by traditional cellular system due to pretty high per-bit cost and scarce spectrum resources. Confronted with such unprecedented traffic growth, mobile operators are looking for new technologies to meet the diverse and increasing demands of users with guaranteed quality-of-service (QoS) and quality-of-experience (QoE) [4]. However, it is a challenging task to improve the key performance metrics of the network under limited resource and cost.

The network throughput, service coverage, spectrum efficiency (SE) as well as energy efficiency (EE) should be jointly considered and optimized. Aiming to address these pertinent issues, dense deployment of access points (APs) or base stations (BSs) can be considered [5], which makes it proximal for any device to access the network. Such AP/BS densification compensates the path loss of the wireless channel and has the potential to improve the network throughput and EE. Meanwhile, full frequency reuse can be adopted to improve the SE with advanced physical (PHY) layer techniques, such as multiuser multiple-input multiple-output (MIMO) and massive MIMO [6, 7]. As the density of the AP/BS increases, co-channel interference may also increase [8], which calls the need for proper interference management to maintain high network throughput.

From a system's perspective, advances of the radio access network (RAN) itself may only contribute to a fraction of the performance gains to meet the requirement of ever-increasing data demands. Note that there have been heterogeneous networks (HetNets), which can be deployed with different available spectrum resources and service coverage capabilities [9]. However, cooperation among different HetNets is neither easy nor smooth under concurrent network architectures and protocol stacks. Such inefficiency in heterogeneous resource utilization also indicates inefficient usage of infrastructure. Therefore, integrating these heterogeneous network components, spectrum resources, APs/BSs to offer flexible and reliable on-demand services are of great practical and commercial interest. Such functional integration may require the evolution of the whole network, not only the RAN itself. Cloud computation techniques have been widely and successfully used in the area of computation/data networks [10]. Recently, the concepts of software-defined network (SDN) [11–13] and network function virtualization (NFV) [14] have been introduced to wireless communication networks. The basic idea behind these concepts is to abstract different physical resources, components and functional entities into a logic resource pool, then different resources can be centrally orchestrated to meet the specific demands according to flexible protocols [15]. Therefore, APs/BSs in RAN can be abstracted as antennas with different coverage abilities and signal processing abilities. Assuming centralized control and processing, the seemingly interfering APs/BSs can together form an MIMO network in RAN, where AP/BS densifications may be beneficial and interference-free. In the core network, signal processing and control can be done at the central unit (CU) and the heterogeneous resources can be flexibly scheduled and used without any barrier to meet the required demands. To this end, Cloud-RAN (C-RAN) architecture has attracted significant research attention from both academia and industry [2, 3, 6]. It has been regarded as a promising technology towards implementing the fifth generation (5G) wireless systems [16].

In this survey paper, we review the general architecture of C-RAN with special focus on PHY layer signal processing aspects. In particular, we first review the functional split solutions in C-RAN, which partially move the baseband signal processing from the conventional AP/BS to the cloud server. It is noted that the capacity of the fronthaul needs to be carefully considered, and some key PHY signal processing techniques are introduced, namely channel state information (CSI) acquisition, compression and transmission, and resource management. The possible integration of other promising 5G techniques such as HetNets, SDN, mmWave and full-duplex with

C-RAN infrastructure is also discussed. These techniques are expected to offer more physical spectrum resources, infrastructure resources, elastic networks and flexible protocols. Finally, we envision several challenges and future directions associated with C-RAN design.

The rest of this paper is organized as follows. In Sect. 2, C-RAN architecture in terms of key components and functional split is introduced. In Sect. 3 emerging PHY layer signal processing approaches for C-RAN are presented. In Sect. 4 we discuss some promising 5G techniques that can be supported and integrated with C-RAN. Some open research challenges and promising future directions are outlined in Sect. 5. Section 6 concludes this paper.

## 2 C-RAN Architecture

In this section, the architecture of C-RAN is presented with detailed explanations on the key components and the options of functional split. It is noted that the PHY layer and some upper layer functionalities are all located in the BS within the conventional RANs, which imposes high cost for network deployment and upgrading. Different from conventional RANs, C-RAN simplifies the BS by moving a significant part of its functionalities to the cloud server, namely the baseband unit (BBU). Therefore the deployment of a larger number of APs, namely the remote radio heads (RRHs), is possible and cost-effective. Such a new paradigm of network architecture is considered to be the foundation of the future wireless communication systems.

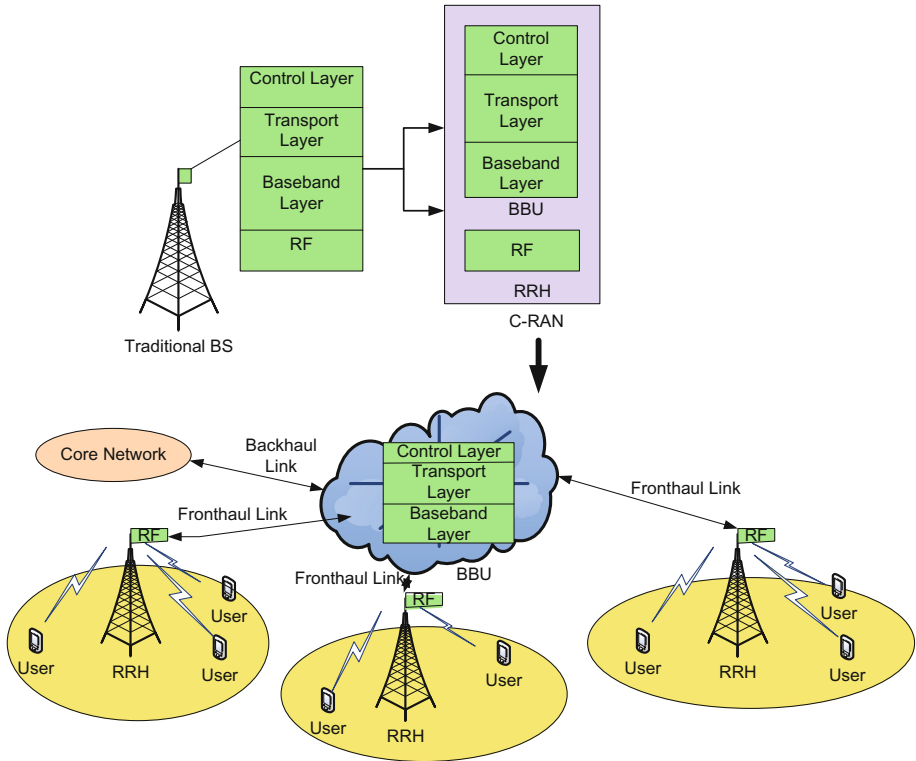
### 2.1 Key Components

Figure 1 shows the architecture of C-RAN, where the BBUs are in the cloud server, and the RRHs are geographically separated from BBUs. In addition, BBU and RRH are connected via a fronthaul link, while BBU and core network are connected via a backhaul link [17].

**BBU:** The BBU is enabled by cloud computing, which achieves flexible spectrum management and advanced network coordination [4]. Moreover, BBU can handle a significant part of baseband signal processing of the whole network, it also controls the signaling to RRHs. Compared to a traditional BS, the joint signal processing across a larger coverage area can be done at the BBU side in a centralized and soft way [18], which bears the potential to mitigate interference and improve the performance.

**RRH:** The RRH is mainly responsible for the radio frequency function and some simple signal processing. The deployment of RRH can provide seamless connection, especially in the hot spot areas. Moreover, densely deployed RRHs can offer stronger coverage with high data rate.

**Fronthaul link:** The fronthaul links can use wired or wireless medium based on the application scenarios, and they are capacity-limited in general. Within the two-hop C-RAN structure, the limits of fronthaul should be carefully considered for practical system designs.



**Fig. 1.** C-RAN architecture.

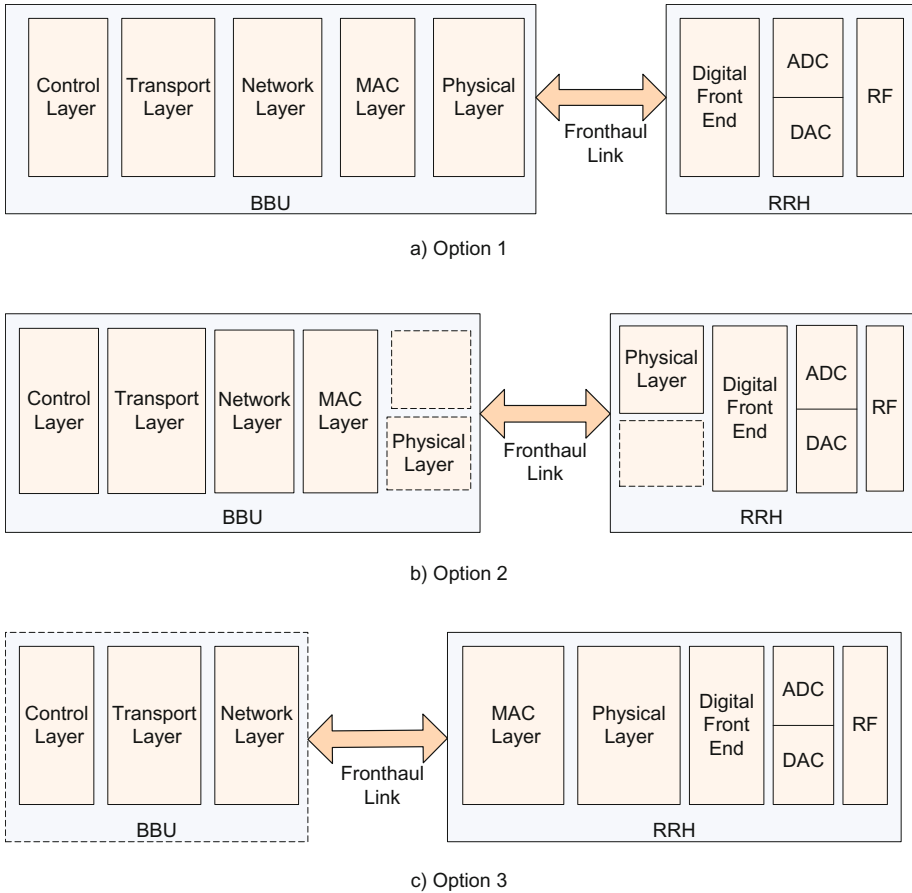
It is worth pointing out that there are prominent advantages of the C-RAN architecture. For example, C-RAN can provide larger bandwidth by integrating heterogeneous spectrum resources, which increases the service coverage area of users. In addition, the centralized management of BBU can decrease the capital and operation expenditures. Potential gains of a large number of RRHs can be achieved by using the cooperative processing and advanced MIMO techniques.

## 2.2 Functional Split

The centralization of the baseband signal processing in C-RAN architecture improves the flexibility of network coordination, which also enables high cooperative processing gains [19]. However, there exist tradeoffs among the split option, fronthaul capacity and signal processing complexities. In this subsection, we present three possible options of functional split between BBU and RRH, and briefly remark on the tradeoffs.

The first option is shown in Fig. 2(a), which is proposed for the initial C-RAN architecture [1]. Almost all PHY layer functionalities are moved to BBU. Meanwhile, RRH acts as a simple relay with RF, ADC/DAC and digital front-end. In this option,

BBU and RRH are connected by the common public radio interface (CPRI) [2]. This centralized PHY architecture may achieve the highest cooperative processing gain. However, forwarding I/Q samples via fronthaul links requires very high transmission bandwidth. Aiming to limit the fronthaul bandwidth, I/Q samples should be compressed at BBU and RRH [19, 20]. The second option is shown in Fig. 2(b), where the baseband processing is partially centralized and some PHY layer processing is still reserved at RRH. This functional split can significantly reduce the transmission bandwidth within the fronthaul, and also achieves high cooperative processing gains.



**Fig. 2.** Functional split between BBU and RRH.

However, the implementation of cooperative processing becomes complicated due to the distributed deployment of PHY functionalities [18]. This architecture aims to strike a balance between fronthaul capacity and signal processing complexity. The last option is shown in Fig. 2(c), where all functions of PHY layer are moved to the RRH.

In this architecture, the transmission bandwidth of fronthaul link is reduced to the maximum medium access control (MAC) layer throughput, which imposes the minimum fronthaul bandwidth requirement as compared to the previous two options. The price paid for this reduction is the increased scheduling delay in the fronthaul link, which may degrade system performance and network throughput [1], while the benefit of this architecture is the saving of power consumption at BBU and higher flexibility to support radio resource allocation towards users [17].

To sum up, dividing baseband signal processing between BBU and RRH gives several options to deploy the C-RAN with different fronthaul capacity limits. However, without suitable PHY layer signal processing, it is challenging to realize cost-effective deployment of C-RANs.

### 3 PHY Layer Signal Processing

In this section, we discuss signal processing approaches in the PHY layer, where CSI acquisition, data compression and transmission, and resource management are focused as shown in Fig. 3. It is noted that accurate CSI is the prerequisite of many advanced data transmission schemes, where both training-based and blind schemes can be used. In addition, the massive RRHs and the limited capacity of the fronthaul impose new challenges for CSI acquisition and data transmission. Therefore, advanced compression schemes are necessary to address these issues. Finally, radio resource allocation is discussed, which improves system performance.

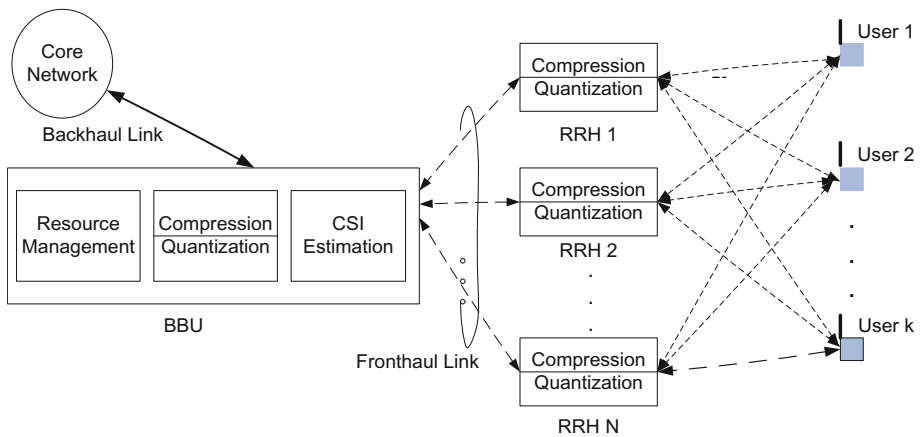


Fig. 3. PHY layer signal processing for uplink and downlink of C-RAN.

#### 3.1 CSI Acquisition

CSI acquisition is the first step for data transmission in C-RAN. It is noted that C-RAN consists of two-hop links, namely the access links between RRHs and users and the



fronthaul links between BBUs and RRHs. To this end, the channel estimation methods are similar to those of the relay channels while the constraints of the fronthaul and massive RRHs should be carefully considered. Saving overhead for CSI acquisition in C-RAN is necessary and important. Considering the limited capacity of fronthaul links, the transmission overhead of training sequences should not be very large, and the limited resources devoted to CSI estimation shall be fully utilized. For uplink transmission in C-RAN, the joint channel estimation and CSI compression transmission schemes have been proposed in the literature. The BSs compress the received data and training signals and transmit them to BBU, then BBU estimates the channel and decodes the data [22, 23]. These schemes can significantly reduce the transmit overhead of training sequences and relieve the severe burden of fronthaul links. Moreover, sparse channel estimation has been proposed in [24] based on the principle of compressed sensing, where the inherent sparse nature of C-RAN can be utilized to save system overhead. It is worth pointing out that as C-RAN evolves, more factors should be taken into account in the channel estimation, e.g., the increasing number of channel parameters, the heterogeneity of APs and the delay of fronthaul links. Therefore, fast, scalable and cost-effective channel estimation methods should be developed to meet the demands of C-RANs.

### 3.2 Compression and Transmission

A high amount of traffic including training sequences and data must be transmitted on the fronthaul links. However, the capacity of fronthaul link is in general limited, which serves as the main bottleneck to fully achieve large-scale cooperative processing gains in C-RANs. To this end, various compression and quantization methods for uplink and downlink transmission have been proposed in the literature [21, 24–26]. For uplink transmission, signals from users are quantized at the RRHs and sent to BBU for further signal processing. This is known as the compress-and-forward strategy in the literature [21]. BBU needs to extract the useful CSI from the quantized and compressed samples while performing detection and decoding. For downlink transmission, BBU performs joint compression and encoding, and then the compressed samples are forwarded to the RRHs using fronthaul links. Moreover, for multi-user transmission BBU can perform joint quantization of multiple sources, which is known as the multivariate quantization [24]. Compared to the conventional point-to-point quantization, multivariate quantization can reduce the negative impact of quantization errors, which distinctly improves the quality of signals transmitted to users. Table 1 compares various relevant compression methods for C-RAN, which has been considered in [2, 4, 21, 25].

### 3.3 Resource Management

Resource management is an important issue to fulfill the potential of C-RAN to achieve higher SE and EE. Although conventional resource allocation schemes can be borrowed to improve SE in C-RAN, EE-oriented optimization calls for new designs. Recently, there has been some works on the methods to improve EE in C-RAN [9, 27–29].

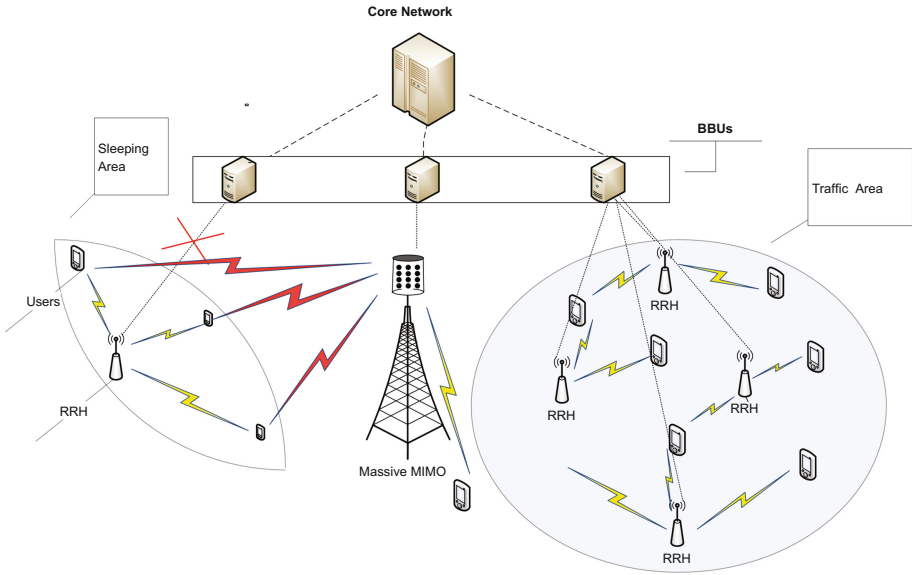
**Table 1.** Compression methods for C-RAN uplink and downlink transmission.

Name of the compression methods	Communication-Oriented Optimization	Complexity	Compression ratio	Number of RRH antennas
Compression using scheduling side information	No	Low	High compression ratio	Single antenna
Wyner-Ziv lossy compression	No	High	High compression ratio, support a larger number of antennas	Single antenna
Adaptive compression	Minimize the number of bits under the BLER constraint	Low	Adaptive compression ratio, exploit the cross-correlation of signals	Single antenna
Spatial compression	Maximize the minimum SINR by delicately designing the power allocation	Low	Significant gain over the conventional quantize-and-forward	Multiple antennas

For example, energy saving can be achieved by switching off a subset of RRHs when the network traffic load is not high. To this end, the rest RRHs should collaboratively serve all users including users in the BS sleeping area as shown in Fig. 4. Results show that the BS sleeping method [29] reduces the total power consumption. Although promising, the selection of the sleeping RRHs is a non-trivial task; the related optimization problem is complicated and requires network coordination. On the other hand, the radio resource in C-RAN is multi-dimensional. The RRH clustering, transmit beamforming, user association and power allocation can be jointly optimized to meet different objectives in terms of SE, EE and delay [28]. To this end, efficient resource allocation schemes should be developed; and the implementations of these algorithms are also influenced by the functional split of the underlying network architecture.

## 4 C-RAN and 5G Technologies

In this section, we discuss flexible network configuration of C-RAN, which can be combined with other candidate 5G technologies. As mentioned earlier, next generation of wireless communication networks call for significant innovation in both PHY layer techniques in the RAN as well as the entire network architecture. In order to exploit potential gains, C-RAN should cooperate with other advanced technologies, e.g., HetNets, SDN, mmWave and full-duplex radio, as illustrated in Fig. 5. Integrated with SDN framework, the BBU is placed at the control layer, while RRH is placed at the infrastructure layer. Infrastructure layer of SDN is often operated with HetNets, which includes low power nodes (LPNs) and high power nodes (HPNs). Furthermore, new resources such as mmWave spectrum can be used to further improve the network



**Fig. 4.** Resource management in C-RAN.

throughput. Finally, full-duplex operation can be embedded into C-RAN architecture to further improve the system SE.

#### 4.1 HetNets and C-RAN

In this subsection, we briefly show the advantages of implementing C-RAN into HetNets to enable more flexible utilization of heterogeneous resources. Low power nodes, such as pico BS, femto BS and small cell BS, are components to increase the capacity of cellular networks in the hot spot areas [30]. Low power nodes can cooperate with high power nodes, such as macro BSs or micro BSs. The prominent advantage of HetNets is to support high data rates in hotspot areas. However, the handover among different networks/APs/BSs may not be smooth, and the interference management is in general challenging. The 5G system is expected to solve this problem with the help of C-RAN. The combination of HetNet and C-RAN results in a new paradigm of RAN, namely the heterogeneous cloud radio access network (H-CRAN) [31], which is shown in Fig. 5. Such network architecture enjoys several advantages. Firstly, deploying C-RAN to support HetNet can achieve seamless connection. Moreover, compared to the traditional wireless cellular network, H-CRAN can provide significant SE and EE by using large-scale cooperative processing and resource management. Finally, the collaboration between HetNets and C-RAN facilitate effective mobility management to achieve a smooth handover.

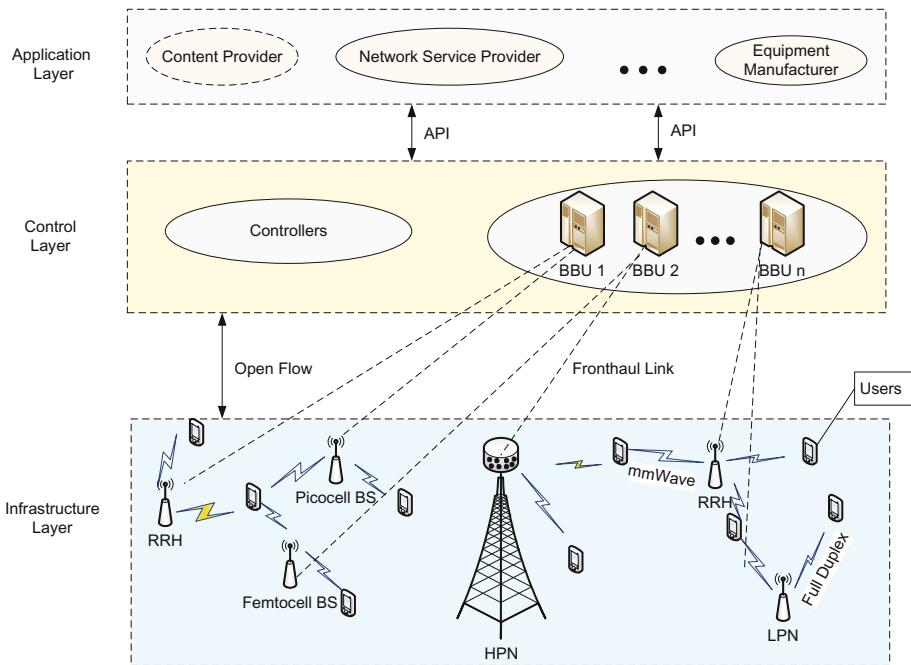


Fig. 5. C-RAN and some of the candidate 5G technologies.

## 4.2 SDN and C-RAN

SDN is a cost-effective, adaptable and manageable architecture, which decouples network control and forwarding functionality [17]. It is convenient for the operators to upgrade software separately from hardware via centralized controllers in the control layer. The first architecture of SDN was proposed only for wired communications. Unfortunately, the original conceptual SDN framework is not readily implementable in wireless communication networks. A solution of this problem is to combine SDN with C-RAN, i.e., SDN-CRAN [31]. SDN consists of an application layer, a control layer and an infrastructure layer. In SDN-CRAN architecture, the BBU is operated in the control layer and RRHs are in the infrastructure layer as shown in Fig. 5. The control layer is the core part of SDN architecture. The functions of data traffic offloading, caching, storage, mobility management are all controlled by BBU with the help of SDN.

## 4.3 mmWave and C-RAN

Millimeter wave (mmWave) is a strong candidate technology for 5G systems [32–35], and it can be used to empower C-RAN. First, the mmWave spectrum spans from 3 GHz to 300 GHz, which offers abundant spectrum resources to enable a huge transmission bandwidth. In C-RAN, mmWave can not only be utilized to improve the capacities of the access links between RRH and user, but also improve the wireless fronthaul capacity as

well. Secondly, the propagation characteristics of mmWave [35] facilitate the dense deployment of RRHs and also simplify the interference management. The high attenuation of mmWave indicates its application in line-of-sight (LoS) environment, which matches well to the layout of C-RAN. The RRHs are deployed to be near to users, where LoS scenarios are predominant. Meanwhile, the strong path-loss of mmWave will cause less interference leakage; therefore, the location-aware and local interference managements may be useful in C-RAN, which in turn relieve the heavy burden of large-scale network coordination and fronthauling. Finally, mmWave can enable cost-efficient MIMO techniques to further improve the SE and EE in C-RAN [6]. Because the wave-length of mmWave frequencies is small, the technology also enables to build very compact and low-cost antenna arrays with large number of elements [32]. When RRHs are equipped with such antenna arrays, advanced directional beamforming can be dynamically designed to improve the SE and EE in C-RAN.

#### 4.4 Full-Duplex and C-RAN

Full duplex communications can realize the simultaneous transmission and reception on the same frequency resource; therefore it is possible to achieve the goal of doubling the SE in future wireless systems [36]. In C-RAN, full-duplex radio technique can be implemented at the RRHs [37]. When employing full-duplex communications, the major challenge is to combat with the loopback interference (LI), which is leaked from the output to the input side of transceiver. There are several antenna domain, analog domain and digital domain approaches for suppressing the LI. In a full-duplex enabled C-RAN, LI can be naturally mitigated due to the path loss among the distributed RRHs [38, 39]. Moreover, to improve the transmit data rate and mitigate LI in a full duplex C-RAN, it is worthwhile to investigate schemes to properly design the RRH association and beamforming. The combination of full-duplex and other types of duplexing models in C-RAN also deserves further research attention.

## 5 Challenges and Future Directions

In this section, challenges and interesting future directions for C-RAN are briefly discussed. Specifically, mobile edge computing and caching, multi-dimensional resource management, and PHY layer security of C-RAN are highlighted.

### 5.1 Mobile Edge Computing and Caching

Aiming to offload the heavy burden of the fully centralized C-RAN, mobile edge computing (MEC) has been recently proposed to offer computing capabilities and resources to the edge devices in the RAN [47]. In particular, some new elements are deployed at the BS, RRH and users, which can perform local signal processing and storage functions. Employing mobile edge computing, some popular contents such as videos or social media can be locally cached and processed by the edge devices; the delivery and service latency towards the users can be significantly reduced, improving

the QoE of users [39]. To this end, a possible evolution of C-RAN has been proposed, namely the fog radio access network (F-RAN) [40]. The main idea of F-RAN is to remove some signal processing functions from the central cloud to the geographically distributed RRHs. A convenient realization of F-RAN is to equip the RRHs with limited cache storage so these RRHs can pre-fetch the popular content videos during the off-peak hours. Compared to the conventional C-RAN, F-RAN can significantly relieve the serious traffic burden of fronthaul links to effectively avoid the network congestion and reduce the service delay. In this scenario, the major challenge is to effectively utilization the limited storage capacity and to increase the probability that the requirement of users can be satisfied locally to improve the QoS of users.

## 5.2 Multi-dimensional Resource Management

C-RAN encompasses multi-dimensional resources, including radio resources, computation resources, storage resources and power in general [41–43]. Noting the MEC and caching have been introduced as new functional elements of C-RAN, the resource management and allocation is more challenging and complicated to meet diverse requirements, objectives and constraints. For example, the radio resources should be jointly considered with the computational and storage capabilities to enable high quality video service with mobile edge computing, while the capacity constraints of the fronthaul and the delay constraints for the fetch-and-forward transmission should be carefully taken into account in C-RAN. Although the conceptual model of SDN and NFV suggests resource abstraction in a global view, the large-scale and fast resource slicing is still very challenging due to its inherent complexity, protocol barrier and stringent delay constraints. Moreover, the integration of multi-dimensional heterogeneous resources is challenging, such as the mobility management and cooperation among HetNets [43]. Aiming at large-scale and multi-dimensional resource management, the game theoretical approach [44] as well as distributed learning based schemes [45] may be useful in conjunction with mobile edge computing. Moreover, big data and advanced data mining techniques [46] can be introduced in C-RAN to model and predict the user's behavior, contents popularity and service diversities and so on, which will enable more intelligent and fast resource management.

## 5.3 PHY Layer Security of C-RAN

Due to the openness of the radio mode in C-RAN, exchanged information becomes more vulnerable to eavesdropping. The broadcast nature of the wireless transmission medium, the relay-like RRH functions and the cloud environment all make the secure information transmission challenging in C-RAN. Focusing on the PHY layer, recently PHY layer security has been proposed as a viable technique to secure wireless communications [48, 49] against eavesdropping and attacking. Different from the conventional cryptographic techniques in the higher layers, PHY layer security exploits the characteristics of the wireless channels to secure message transmission. In some network settings, such as the conventional signal-cell, multi-cell and HetNets, PHY layer security

can improve secrecy rates [48, 49]. In C-RAN, one of the unique challenges to apply PHY layer security is the smart coordination of the ubiquitous interference. As the density of RRH increases, mutual interference may become stronger and the interplay among useful signal, interference signal and jamming signal becomes more complicated. To this end, advanced collaborative MIMO signal processing can be used, where RRHs can be clustered to jointly perform secrecy beamforming [50, 51] to combat eavesdropping and interference can be judiciously utilized to improve the secrecy performance [52].

## 6 Conclusion

In this paper, we discussed fundamental concepts, key techniques, and challenges of the C-RAN architecture in the context of emerging 5G wireless communication systems. We reviewed key components of C-RAN, functional split between the BBU and RRH as well as the signal processing approaches for uplink and downlink transmission. To effectively utilize the limited capacity of fronthaul links, compression and quantization methods for C-RAN uplink and downlink transmission were investigated. In addition, resource management schemes for energy saving were discussed to meet the requirements of green communications. To exploit potential performance gains, SDN, HetNets and mmWave and full-duplex communications can be combined with C-RAN to increase the SE, EE and network throughput. Finally, we presented some challenges and future directions, where mobile edge computing and caching, multi-dimensional resource management and wireless security are major concerns for further evolution of the C-RAN architecture.

**Acknowledgement.** This work was supported in part by the National Natural Science Foundation of China (Grant 61401041, 61671072 and 61501046), and by Meteorological Information and Signal Processing Key Laboratory of Sichuan Higher Education (QXXCSYS201601).

## References

1. Niu, H., Li, C., Papathanassiou, A., Wu, G.: RAN architecture options and performance for 5G network evolution. In: IEEE Wireless Communications and Networking Conference Workshops, pp. 294–298. IEEE Press, New York (2014)
2. Pawar, S., Niu, H., Papathanassiou, A.: Front-haul compression using scheduling side information for cloud radio access networks. In: IEEE Global Communications Conference, pp. 1–6. IEEE Press, New York (2015)
3. Wang, K., Yang, K., Wang, X., Magurawalage, C.S.: Cost-effective resource allocation in C-RAN with mobile cloud. In: IEEE International Conference on Communication, pp. 1–6. IEEE Press, New York (2016)
4. Peng, M., Wang, C., Lau, V., Poor, H.V.: Fronthaul-constrained cloud radio access networks: insights and challenges. *IEEE Wirel. Commun.* **22**, 152–160 (2015)
5. Al-Dulaimi, A., Anpalagan, A., Bennis, M., Vasilakos, A.V.: 5G green communications: C-RAN provisioning of CoMP and Femtocells for power management. In: IEEE Ubiquitous Wireless Broadband, pp. 1–5. IEEE Press, New York (2015)

6. Song, G.H., Brady, J., Sayeed, A.M.: Beamspace MIMO transceivers for low-complexity and near-optimal communication at mmWave frequencies. In: *IEEE Acoustics, Speech and Signal Processing*, pp. 4394–4398. IEEE Press, New York (2013)
7. Zuo, J., Zhang, J., Yuen, C., Jiang, W., Luo, W.: Energy efficient user association for cloud radio access networks. *IEEE Access* **4**, 2429–2438 (2016)
8. Patil, P., Yu, W.: Hybrid compression and message-sharing strategy for the downlink cloud radio-access network. In: *IEEE Information Theory and Applications Workshop*, pp. 1–6. IEEE Press, New York (2014)
9. Chen, L., Jin, H., Li, H., Seo, J.B., Guo, Q., Leung, V.: An energy efficient implementation of C-RAN in HetNet. In: *IEEE Vehicular Technology Conference*, pp. 1–5. IEEE Press, New York (2014)
10. Tang, S., Li, X., Huang, X., Xiang, Y., Xu, L.: Achieving simple, secure and efficient hierarchical access control in cloud computing. *IEEE Trans. Comput.* **65**, 2325–2331 (2015)
11. Hu, F., Hao, Q., Bao, K.: A survey on software-defined network and OpenFlow: from concept to implementation. *IEEE Commun. Surv. Tutorials* **16**, 2181–2206 (2014)
12. John, W., Kern, A., Kind, M., Skoldstrom, P., Staessens, D., Woesner, H.: Split architecture: SDN for the carrier domain. *IEEE Commun. Mag.* **52**, 146–152 (2014)
13. Kreutz, H., Ramos, F., Verissimo, P., Rothenberg, C., Azodolmolky, S., Uhlig, S.: Software-defined networking: a comprehensive survey. *Proc. IEEE* **103**, 14–76 (2014)
14. Liang, C., Yu, F.R.: Wireless network virtualization: a survey, some research issues and challenges. *IEEE Commun. Surv. Tutorials* **17**, 358–380 (2014)
15. Agyapong, P., Iwamura, M., Staehle, D., Kiess, W., Benjebbour, A.: Design considerations for a 5G network architecture. *IEEE Commun. Mag.* **52**, 65–75 (2014)
16. Zhou, L., Ratnarajah, T., Xue, J., Khan, F.: Energy efficient cloud radio access network with a single RF antenna. In: *IEEE International Conference on Communications*, pp. 1–6. IEEE Press, New York (2016)
17. Wang, R., Hu, H., Yang, X.: Potentials and challenges of C-RAN supporting multi-RATs toward 5G mobile networks. *IEEE Access* **2**, 1187–1195 (2014)
18. Park, S.H., Simeone, O., Sahin, O., Shamai, S.: Joint decompression and decoding for cloud radio access networks. *IEEE Sig. Process. Lett.* **20**, 503–506 (2013)
19. Kenji, M., Shigeru, K., Jun, T., Akihiro, O.: Split-PHY processing architecture to realize base station coordination and transmission bandwidth reduction in mobile fronthaul. In: *IEEE Optical Fiber Communications Conference and Exhibition*, pp. 1–3. IEEE Press, New York (2015)
20. Kang, J., Simeone, O., Kang, J., Shamai, S.: Fronthaul compression and precoding design for C-RANs over ergodic fading channels. *IEEE Trans. Veh. Technol.* **65**, 5022–5032 (2016)
21. Liu, L., Zhang, R.: Optimized uplink transmission in multi-antenna CRAN with spatial compression and forward. *IEEE Trans. Sig. Process.* **63**, 5083–5095 (2015)
22. Shi, Y., Zhang, J., Letaief, K.B.: Statistical group sparse beamforming for green Cloud-RAN via large system analyses. In: *IEEE International Symposium on Information Theory*, pp. 870–874. IEEE Press, New York (2016)
23. Shi, Y., Zhang, J., Letaief, K.B.: CSI overhead reduction with stochastic beamforming for cloud radio access networks. In: *IEEE International Conference on Communication*, pp. 5154–5159. IEEE Press, New York (2014)
24. Larsson, E.G., Edfors, O., Tufvesson, F., Marzetta, T.L.: Massive MIMO for next generation wireless systems. *IEEE Commun. Mag.* **52**, 186–195 (2014)
25. Lee, W., Simeone, O., Kang, J., Shamai, S.: Multivariate fronthaul quantization for C-RAN downlink: channel-adaptive joint quantization in the Cloud. In: *IEEE International Conference on Communication*, pp. 1–5, IEEE Press, New York (2016)



26. Vu, T.X., Quek, T.Q.S., Nguyen, H.D.: Joint decoding and adaptive compression with QoS constraint for uplinks in cloud radio access networks. In: IEEE Global Telecommunication Conference. pp. 1–6, IEEE Press, New York (2015)
27. Liu, L., Zhang, R.: Downlink SINR balancing in C-RAN under limited fronthaul capacity. In: IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 3506–3510. IEEE Press, New York (2016)
28. Yoon, C., Cho, D.-H.: Energy efficient beamforming and power allocation in dynamic TDD based C-RAN system. *IEEE Commun. Lett.* **19**, 1806–1809 (2015)
29. Zhao, W., Wang, S.: Traffic density based RRH selection for power saving in C-RAN. *IEEE J. Sel. Areas Commun.* **99**, 1–11 (2016)
30. Peng, M., Li, Y., Jiang, J., Li, J., Wang, C.: Heterogeneous cloud radio access networks: a new perspective for enhancing spectral and energy efficiencies. *IEEE Wirel. Commun.* **21**, 126–135 (2014)
31. Yang, C., Chen, Z., Xia, B., Wang, J.: When ICN meets C-RAN for HetNets: an SDN approach. *IEEE Commun. Mag.* **53**, 118–125 (2015)
32. Rappaport, T.S., et al.: Millimeter wave mobile communications for 5G cellular: it will work. *IEEE Access* **1**, 335–349 (2015)
33. Heath, R.W., Gonzalez-Prelcic, N., Rangan, S., Roh, W., Sayeed, A.M.: An overview of signal processing techniques for millimeter wave MIMO systems. *IEEE J. Sel. Top. Sig. Process.* **10**, 436–453 (2016)
34. Han, S., Chih-Lin, I., Xu, Z., Rowell, C.: Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G. *IEEE Commun. Mag.* **53**, 186–194 (2015)
35. Chandra, K., Cao, Z., Bruintjes, T.M., et al.: mCRAN: a radio access network architecture for 5G indoor communications. In: IEEE International Conference on Communication Workshop, pp. 300–305, IEEE Press, New York (2015)
36. Mohammadi, M., Suraweera, H.A., Tellambura, C.: Uplink and downlink rate analysis of a full-duplex C-RAN with radio remote head association. In: European Signal Processing Conference, pp. 778–782. IEEE Press, New York (2016)
37. Zhang, X., Cheng, W., Zhang, H.: Full-duplex transmission in PHY and MAC layers for 5G mobile wireless networks. *IEEE Wirel. Commun. Mag.* **22**, 112–121 (2015)
38. Mohammadi, M., Suraweera, H.A., Tellambura, C.: Full-duplex Cloud-RAN with uplink/downlink remote radio head association. In: International Conference on Communications, pp. 1–6. IEEE Press, New York (2016)
39. Hu, Y.C., Patel, M., Sabella, D., Sprecher, N., Young, V.: Mobile edge computing—a key technology towards 5G. ETSI White Paper **11**, 11–15 (2015)
40. Peng, M., Yan, S., Zhang, K., Wang, C.: Fog computing based radio access networks: issues and challenges. *IEEE Netw.* **30**, 46–53 (2016)
41. Yu, Y., Zhang, J., Letaief, K.B.: Joint subcarrier and CPU time allocation for mobile edge computing. arXiv preprint, [arXiv:1608.06128](https://arxiv.org/abs/1608.06128) (2016)
42. Peng, M., Sun, Y., Li, X., Mao, Z., Wang, C.: Recent advances in cloud radio access networks: system architectures, key techniques, and open issues. *IEEE Commun. Surv. Tutorials* **18**, 2282–2308 (2016)
43. Zhou, S., Zhao, T., Niu, Z., Zhou, S.: Software-defined hyper-cellular architecture for green and elastic wireless access. *IEEE Commun. Mag.* **54**, 12–19 (2016)
44. De Domenico, A., Strinati, C.S., Capone, A.: Enabling green cellular networks: a survey and outlook. *Comput. Commun.* **37**, 5–24 (2014)
45. Alsheikh, M.A., et al.: Machine learning in wireless sensor networks: algorithms, strategies, and applications. *IEEE Commun. Surv. Tutorials* **16**, 1996–2018 (2014)
46. Chen, M., Mao, S., Liu, Y.: Big data: a survey. *Mob. Netw. Appl.* **19**, 171–209 (2014)

47. Beck, M.T., Werner, M., Feld, S., et al.: Mobile edge computing: A taxonomy. In: Sixth International Conference on Advances in Future Internet, pp. 48–54. IARIA, Wilmington (2014)
48. Yang, N., Wang, L., Geraci, G., Elkashlan, M., Yuan, J., Renzo, M.D.: Safeguarding 5G wireless communication networks using physical layer security. *IEEE Commun. Mag.* **53**, 20–27 (2015)
49. Mukherjee, A., Fakoorian, S.A.A., Huang, J., Swindlehurst, A.L.: Principles of physical layer security in multiuser wireless networks: a survey. *IEEE Commun. Surv. Tutorials* **16**, 1550–1573 (2014)
50. Lv, T., Gao, H., Yang, S.: Secrecy transmit beamforming for heterogeneous networks. *IEEE J. Sel. Areas Commun.* **33**, 1154–1170 (2015)
51. Gao, H., Lv, T., Wang, W., et al.: Energy-efficient and secure beamforming for self-sustainable relay-aided multicast networks. *IEEE Sig. Process. Lett.* **23**, 1509–1513 (2016)
52. Zhao, N., Yu, F.R., Jin, M., Yan, Q., Leung, V.C.M.: Interference alignment and its applications: a survey, research issues, and challenges. *IEEE Commun. Surv. Tutorials* **18**, 1779–1803 (2016)

### Author Biographies



**Meruyert Makhanbet** received the B. E. degree from Kazakh National Technical University named after K. I. Satpaev, Almaty, Kazakhstan, in 2015. She is currently pursuing the M. E. degree with the School of Information and Communication Engineering in Beijing University of Posts and Telecommunications (BUPT), Beijing, China. Her current research interest includes green communications.



**Zhang Xuwei** received the B. E. degree from Tianjin Polytechnic University, Tianjin, China, in 2015. She is currently pursuing the Ph.D. degree with the School of Information and Communication Engineering in Beijing University of Posts and Telecommunications (BUPT), Beijing, China. Her current research interests include resource allocation, wireless caching and multicast beamforming.



**Hui Gao** received the B. Eng. degree in information engineering and the Ph.D. degree in signal and information processing from Beijing University of Posts and Telecommunications (BUPT), Beijing, China, in July 2007 and July 2012, respectively. From May 2009 to June 2012, he also served as a Research Assistant for the Wireless and Mobile Communications Technology R&D Center, Tsinghua University, Beijing, China. From April 2012 to June 2012, he visited Singapore University of Technology and Design (SUTD), Singapore, as a Research Assistant. From July 2012 to February 2014, he was a Postdoc Researcher with SUTD. He is now with the School of Information and Communication Engineering, Beijing University of Posts and Telecommunications (BUPT), as an Assistant Professor. His research interests include massive MIMO systems, cooperative communications, ultra-wideband wireless communications.



**Himal A. Suraweera** received the B.Sc. Engineering (First Class Honors) degree from University of Peradeniya, Sri Lanka, in 2001, and the Ph.D. degree from Monash University, Australia, in 2007. Currently, he is a Senior Lecturer with the Department of Electrical and Electronic Engineering, University of Peradeniya. From January 2011 to May 2013, he was a Post-Doctoral Research Associate at the Singapore University of Technology and Design, Singapore. From July 2009 to January 2011, he was with the Department of Electrical and Computer Engineering, National University of Singapore, Singapore as a Research Fellow. From February 2007 to June 2009, he was a Research Fellow at Victoria University, Australia. His research interests include cooperative relay communications, full-duplex, energy harvesting communications, massive MIMO, cognitive radio, and wireless security.

# **Electrical and Electronic Engineering**

# Resonance and Impedance Matching for an Experimental Low-Frequency Wireless Power Transfer System

Raaju Hoolaus<sup>(✉)</sup> and Yasdeo Bissessur

Department of Electrical and Electronic Engineering, University of Mauritius,  
Réduit, Mauritius

raaju.hoolaus@gmail.com, yasdeob@uom.ac.mu

**Abstract.** Resonance allows a wireless power transfer (WPT) system to operate at greater distances compared to a non-resonant one. However, frequency splitting is observed which has the effect of limiting the maximum power transferred. This paper proposes a simple impedance matching algorithm to eliminate the frequency splitting phenomenon and maximise the output power which is cheaper to implement than other impedance matching methods. The WPT system is first converted into an equivalent circuit. Then the necessary conditions to satisfy the critically coupled condition using L matching circuits are formulated, from which the impedance matching algorithm is developed. The latter is verified from simulation and experimental results which are compared, analysed and improvements are also proposed.

**Keywords:** Wireless power transfer · Frequency splitting · Impedance matching

## 1 Introduction

Wireless Power Transfer (WPT) refers to the transmission of electrical energy from a transmitter to a receiver without any physical interconnection between them. In order to be able to extract a maximum amount of power from such a system, it is important to make use of magnetically coupled resonant transmitter and receiver coils [1]. Kurs *et al.* [1] recommended using capacitive loaded loops (external capacitors) rather than self-resonant coils since it confines most of the electric field inside the capacitor and can be operated at lower frequencies.

However, while pursuing high power transfer using resonant coils, an interesting phenomenon known as the frequency splitting phenomenon is observed [2]. It is dependent upon the number of turns in the transmitter and receiver coils, the resonant frequency and the distance between the coils.

A WPT system's working areas can be categorised into the over coupled, critically coupled and under coupled areas [3]. Frequency splitting only occurs in the over coupled mode i.e., when the transmitter and receiver coils are closed to each other, due to a high coupling coefficient resulting in the reflected impedance from the receiver to the transmitter,  $Z_{th_x}$  (a thevenin equivalent impedance referred to the transmitter) being higher than the initial impedance,  $R_{in}$  (internal impedance of the WPT system) of the

latter. This leads to a high transfer efficiency but low power transferred. In the critically coupled mode, frequency splitting is eliminated since  $Zth_{tx}$  is equal to  $R_{in}$  resulting in a maximum power transferred but efficiency amounts to only 50%. In the under coupled mode,  $Zth_{tx}$  is less than  $R_{in}$  due to the low coupling coefficient. Consequently, a reduction in both the transfer efficiency and power transferred is observed.

In order to maximise both the power transferred and efficiency, it is necessary to operate the WPT system in the critically coupled mode. Therefore, there is a need of impedance matching e.g., when operating in the over coupled mode,  $Zth_{tx}$  must be lowered to match  $R_{in}$  whereas when operating in the under coupled mode,  $Zth_{tx}$  must be increased to match  $R_{in}$ .

Impedance matching can be performed by adjusting the distance between the transmitter and receiver coils [1]. However, this method requires great precision in terms of matching distances and is practically difficult to realise. The operating frequency of the WPT system can easily be changed to perform impedance matching by tracking the frequency at which both the efficiency and power transferred are at maximum [3]. However, this method can only be used in the over coupled mode and within a narrow frequency band due to regulatory limits.

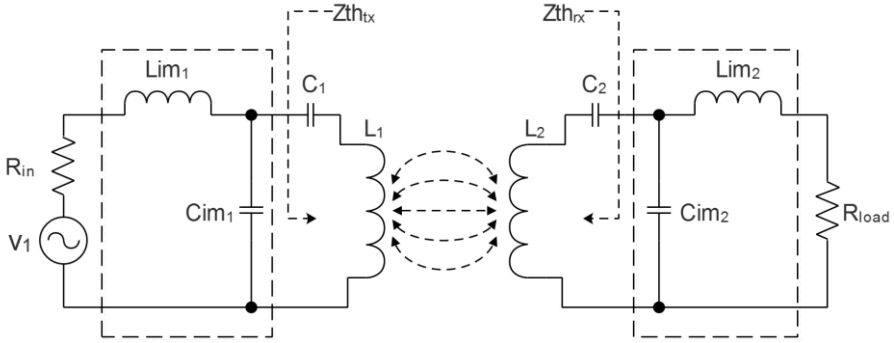
It is possible to operate the WPT system at a single frequency and perform impedance matching using networks consisting of reactive components in order to increase or decrease the reflected impedance from the receiver to the transmitter [4, 5]. Implementing such a system so that it automatically performs impedance matching as the distance between the transmitter and receiver coils is varied requires complex algorithms and expensive hardware e.g., vector network analysers to track the frequency splitting and switchable capacitor banks which makes the system bulky.

In the light of the above, this paper aims at solving the frequency splitting problem by making use of a network of reactive components to perform impedance matching on a static WPT system which uses external capacitors to implement the resonant transmitter and receiver coils. A simple and effective impedance matching algorithm is also proposed which is crucial to perform impedance matching on both the transmitter and receiver sides simultaneously and does not require expensive hardware compared to other algorithms as mentioned earlier.

## 2 The Impedance Matching Approach

### 2.1 General Considerations and Equivalent Circuit of the WPT System

The experimental WPT system consists of a square wave inverter which converts the DC voltage into an AC one at a fixed frequency,  $f_o$  at which the WPT system will resonate. It is represented by an AC voltage source,  $v_1$  and its internal resistance,  $R_{in}$ . Impedance matching is performed on both the transmitter and receiver sides simultaneously using L matching circuits since it provides the same results as  $\pi$  and T circuits but with less components. The impedance matching circuit on the transmitter side eliminates frequency splitting by satisfying the critically coupled condition while that on the receiver side provides a boosted output power. The transmitter and receiver



**Fig. 1.** Proposed WPT system with impedances matched on both sides with L Circuits where  $R_{load}$  represents the load resistor.

circuits both consists of identical coils and resonant capacitors,  $C_1$  and  $C_2$  respectively as shown in Fig. 1:

In order to calculate the values of the capacitors,  $C_{im1}$  and  $C_{im2}$  and inductors,  $L_{im1}$  and  $L_{im2}$  for the matching circuits, the WPT system is simplified into a single circuit by using a T-equivalent model which represents the coupled transmitter and receiver coils by their respective leakage reactance and mutual reactance.

The self-inductances,  $L_1$  and  $L_2$  of the transmitter and receiver coils respectively, can be calculated using the Lorenz formula [6] and the mutual inductance using the Neumann formula [7]. The leakage reactance of the transmitter and receiver coils is then found by subtracting the mutual inductance from the self-inductance of the respective coil.

Prior to introducing the impedance matching algorithm, impedance matching on the transmitter and receiver sides are considered individually as shown in the following section.

## 2.2 The Impedance Matching Algorithm

The WPT circuit can be further simplified by representing it as an equivalent thevenin impedance,  $Z_{th_{tx}}$  and  $Z_{th_{rx}}$  referred to the transmitter and receiver sides respectively as shown in Fig. 2.

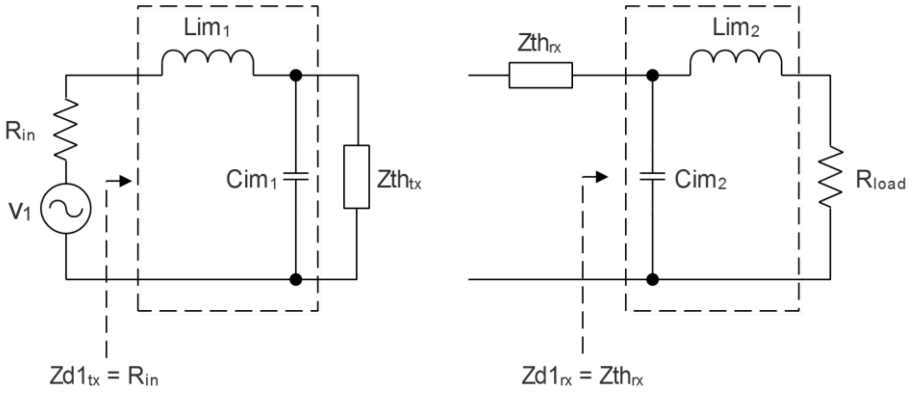
$Z_{th_{tx}}$  and  $Z_{th_{rx}}$  can easily be found by considering the WPT system as a cascade two port network and use the ABCD parameters to calculate the impedances.

Equating the real part of  $Zd1_{tx}$  to  $R_{in}$  and the imaginary part to zero, the matching capacitor,  $C_{im1}$  and inductor,  $L_{im1}$  are given by:

$$C_{im1} = \frac{Q_1}{\omega_o Z_{th_{tx}}} \quad (1)$$

$$L_{im1} = \frac{C_{im1} Z_{th_{rx}}^2}{\omega_o^2 C_{im1}^2 Z_{th_{tx}}^2 + 1} \quad (2)$$

Equating the real part of  $Zd1_{rx}$  to  $Z_{th_{rx}}$  and the imaginary part to zero, the matching inductor,  $L_{im2}$  and the capacitor,  $C_{im2}$  are given by:



**Fig. 2.** Simplified circuit of the WPT system referred to transmitter (*left*) and receiver (*right*) where  $Zd1_{tx}$  and  $Zd1_{rx}$  are the driving point impedance of the transmitter and receiver respectively.

$$Lim_2 = \frac{Q_2 R_{load}}{w_o} \tag{3}$$

$$Cim_2 = \frac{Lim_2}{w_o^2 Lim_2^2 + Z_{in}^2} \tag{4}$$

Where  $w_o$  represents the angular resonant frequency,  $Q_1$  and  $Q_2$  are the Q-factor of the simplified circuit referred to the transmitter and receiver respectively given by:

$$Q_1 = \sqrt{\frac{Zth_{tx}}{R_{in}} - 1} \tag{5}$$

$$Q_2 = \sqrt{\frac{Zth_{rx}}{R_{load}} - 1} \tag{6}$$

Equation (5) is valid if  $R_{in} < Zth_{tx}$  while Eq. (6) is valid if  $R_{load} < Zth_{rx}$ . These assumptions are generally true as the internal resistance of a system is much lower than its driving point impedance and the load resistance can be adjusted so that it is lower than  $Zth_{rx}$ . Equations (5) and (6) are used to calculate Q-factor required for the calculation of the matching capacitor in Eq. (1) and matching inductor in Eq. (3) respectively.

In order to match impedances on both the transmitter and receiver sides simultaneously,  $Zd1_{tx} = R_{in}$  at the transmitter and  $Zd1_{rx} = Zth_{rx}$  at the receiver must be achieved simultaneously. Therefore, the following impedance matching algorithm was devised:

1. Calculate  $Zth_{tx}$  assuming no matching circuit at the receiver
2. Calculate  $Lim_1$  and  $Cim_1$  from Eqs. (1) and (2) using  $Zth_{tx}$  as load resistance
3. Calculate  $Zth_{rx}$ , including the previously calculated matching circuit in step 2
4. Calculate  $Lim_2$  and  $Cim_2$  from Eqs. (3) and (4) using  $Zth_{rx}$  as source resistance.
5. Calculate the new value of  $Zth_{tx}$  which will be used in the next iteration, if any and including the previously calculated matching circuit in step 4.
6. Calculate  $Zd1_{tx}$  and the difference between  $Zd1_{tx}$  and  $R_{in}$



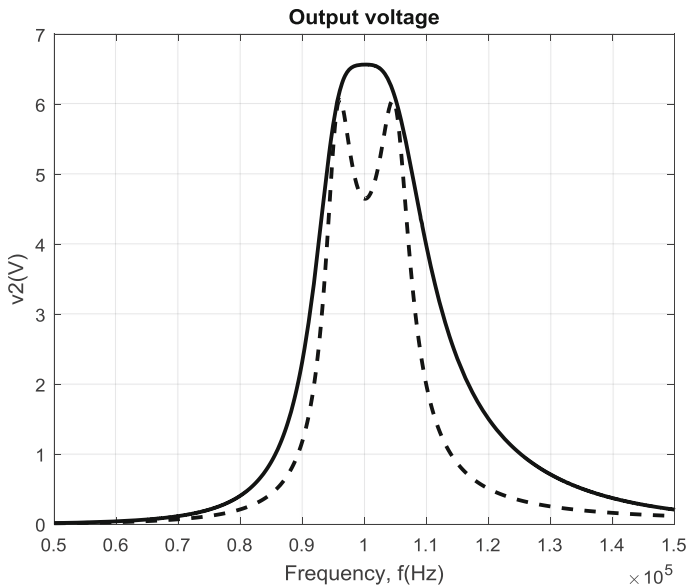
7. Calculate  $Zd1_{rx}$  and the difference between  $Zd1_{rx}$  and  $Zth_{rx}$
8. Obtain the maximum of the two differences calculated in steps 6 and 7 and compare it with the tolerance,  $\varepsilon = 0.01$ .

If the maximum difference is greater than  $\varepsilon$ , steps 2 to 8 are repeated, else this is the end of the matching algorithm.

### 3 Results and Discussions

The impedance matching algorithm was used to perform impedance matching on both the transmitter and receiver sides simultaneously for a WPT system with the radius of both the transmitter and receiver coils,  $r = 15$  cm, the number of the turns in the same,  $n = 15$  and separation distance,  $z = 15$  cm. The resonant frequency,  $f_o$  of the system was 0.1 MHz with an internal resistance,  $R_{in} = 2.5 \Omega$  and load resistance,  $R_{load} = 5 \Omega$ .

From Fig. 3, it can be observed that frequency splitting occurs at  $z = 15$  cm and is eliminated when impedance matching is performed. This results in the maximum output voltage of the WPT system over the entire range of distance, previously occurring at  $z = 25$  cm for the unmatched case due to the critically coupled condition satisfied naturally, to occur at  $z = 15$  cm. The impedance matching algorithm provides a solution to the impedance matching problem with 31 iterations and an accuracy of 1% for the WPT system with  $n = 15$  at  $z = 15$  cm. However, if the efficiency of the



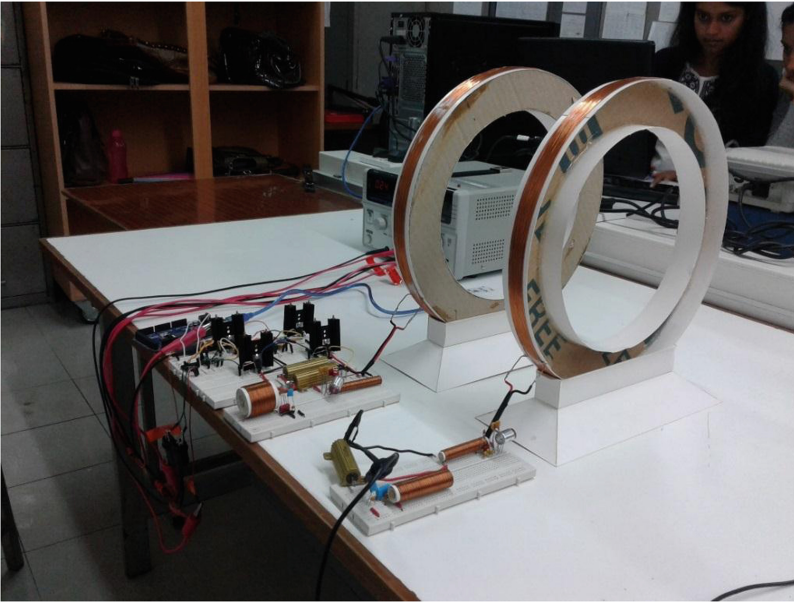
**Fig. 3.** Output voltage with impedances matched on both sides (*solid*) and unmatched case (*dotted*) for  $v_1 = 10$  V.

matched and unmatched WPT system is compared, a decrease from 60% (unmatched) to 37% (matched) is observed.

Figure 4 shows a practical implementation of the WPT system with impedance matching performed on both sides using air cored inductors and non-polarised capacitors to implement the L circuits. The square wave inverter was implemented using a full bridge configuration along with gate drive circuits triggered using a microcontroller. The input voltage,  $v_1$  was set to 10 V to limit the input current,  $i_1$  so that it is less than the maximum output current of the power supply.

Figure 5 and 6 show a comparison of the simulated and experimental results:

It can be observed that the output voltage of the WPT system,  $v_2 = 5.28$  V compared to a theoretical output voltage of 6.57 V. The efficiency,  $\eta = 27\%$  and output



**Fig. 4.** Experimental setup of WPT System for  $n = 15$  and impedance matching performed at  $z = 15$  cm

power,  $P_{out} = 3$  W compared to a theoretical efficiency of 36% and theoretical output power of 4.5 W respectively. The discrepancies between the theoretical and experimental results are due to practical tolerances in inductors and capacitors resulting in a mismatch between those used for the implementation of the WPT system and those calculated. This causes a shift in the resonant frequency and consequently a drop in  $v_2$  and  $\eta$ . Excessive heating of  $R_{in}$  due overshoots in  $v_1$  and power losses in the winding resistances of the coils also contribute to a decrease in  $v_2$  and  $\eta$ .

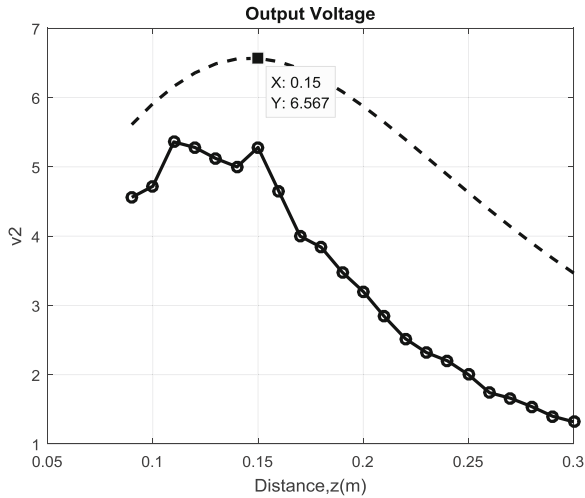


Fig. 5. Simulated (*dotted*) and experimental (*solid*) output voltage,  $v_2$  against distance,  $z$ .

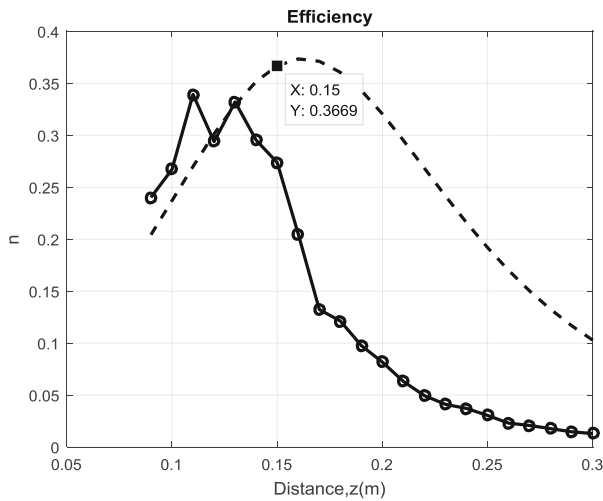


Fig. 6. Simulated (*dotted*) and experimental (*solid*) efficiency,  $\eta$  against distance,  $z$ .

## 4 Conclusions

From the results presented in the previous section, it was shown that frequency splitting can be eliminated by using the impedance matching algorithm to calculate the component values of the L matching circuits on both the transmitter and receiver sides. Also, the proposed system makes use of external capacitances to produce resonance instead of the self-capacitance of the coils which allows a much lower resonant frequency and makes the system less hazardous. The efficiency of the proposed system is

much less than that of other similar systems. This is due to the fact that the proposed algorithm aims at maximising the output power and not efficiency. It can therefore be concluded that the algorithm provides a simple and cheap solution to the frequency splitting problem for a static WPT system. This method relies heavily on accurate modelling of the transmitter and receiver coils. Frequency splitting is completely eliminated as shown in Fig. 3, however, due to the various reasons mentioned in the previous section, experimentally some discrepancies were observed. To improve the accuracy of the impedance matching algorithm, the transmitter and receiver coils can be modelled as a pair of coupled antennas, taking into consideration the non-uniform current distribution along the length of the coil and radiation effects, and then optimising the WPT system using the impedance matching algorithm.

## References

1. Kurs, A., Karalis, A., Moffatt, R., Joannopoulos, J., Fisher, P., Soljacic, M.: Wireless power transfer via strongly coupled magnetic resonances. *Science* **317**(5834), 83–86 (2007)
2. Wei, X., Wang, Z., Dai, H.: A critical review of wireless power transfer via strongly coupled magnetic resonances. *Energies* **7**(7), 4316–4341 (2014)
3. Sample, A., Meyer, D., Smith, J.: Analysis, experimental results, and range adaptation of magnetically coupled resonators for wireless power transfer. *IEEE Trans. Ind. Electron.* **58**(2), 544–554 (2011)
4. Beh, T., Kato, M., Imura, T., Oh, S., Hori, Y.: Automated impedance matching system for robust wireless power transfer via magnetic resonance coupling. *IEEE Trans. Ind. Electron.* **60**(9), 3689–3698 (2013)
5. Waters, B., Sample, A., Smith, J.: Adaptive impedance matching for magnetically coupled resonators. In: *PIERS Proceedings, Progress in Electromagnetics Research Symposium, Moscow, Russia*, pp. 694–701 (2012)
6. Rosa, E., Grover, F.: Formulas and tables for the calculation of mutual and self-inductance (Revised), Internet Archive (2016). <https://archive.org/details/formul812871912169169unse>. Accessed 15 July 2016
7. Thabet, T., John, W.: An approach to calculate the efficiency for an N-Receiver wireless power transfer system. *Int. J. Adv. Comput. Sci. Appl.* **6**(9), 94 (2015)

# Impact of Introducing Small Scale Distributed Generation on Technical Losses in a Secondary Distribution Network

Ismaël Adam Essackjee<sup>(✉)</sup> and Robert T.F. Ah King

Department of Electrical and Electronic Engineering, University of Mauritius,  
Réduit 80837, Mauritius

ismael.essackjee@umail.uom.ac.mu, r.ahking@uom.ac.mu

**Abstract.** System losses are important metrics for power utilities. The addition of distributed generation (DG) in the network can either improve or worsen system losses depending on the network configuration, placement of the DG or other factors. The impact that the integration of Small Scale Distributed Generation (SSDG) has on the losses in a low voltage distribution network in a highly residential has been presented in this paper. The model has been drawn in DIgSILENT Power Factory® using actual network parameters, load profile and photovoltaic (PV) generation data. The transformer losses and losses in the low voltage cables make up for the losses in the low voltage network and these components have been observed separately with increasing penetration of SSDG. The key contribution of this paper is the determination of that penetration of SSDG which results in the most optimum effect on network losses.

**Keywords:** Small Scale Distributed Generation (SSDG) · Network losses · Low-voltage secondary distribution networks · Maximum penetration level · Optimum integration

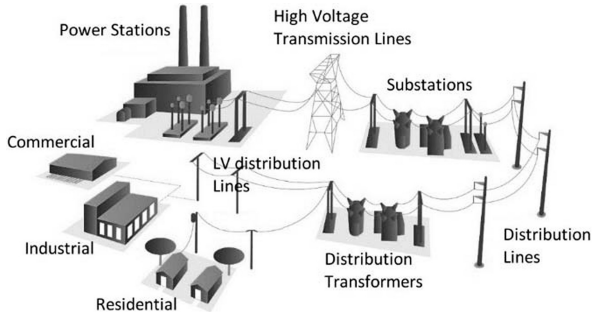
## 1 Introduction

There is increasing global awareness as to the pressing need to reduce carbon emission. This, along with various determining factors, has led to the serious consideration for the deliberate generation of power from sources of renewable energy. As such, Distributed Generation (DG) units are slowly integrating the Mauritian grid and contributing to the energy mix. In addition to the environment friendly aspect, DG does also favourably impact voltage regulation, system losses and helps in deferring huge capital investments for additional generation capacity and the associated network [1]. The Mauritian Government is encouraging the integration of Small Scale Distributed Generation (SSDG) on the low voltage network. In so doing, the traditional network, originally designed as a passive network, is transforming into an active network. With increasing penetration level of these SSDG and without careful assessment prior to their connection, the expected benefits can result in setbacks. This study assesses the impact that increasing penetration level of SSDG has on the line losses in a secondary distribution network of a typical highly residential area in Mauritius. Actual load and generation profiles have been used for this case study.

## 1.1 Background

### 1.1.1 Power Flow in a Traditional Power System

The power flow in a traditional power system starts from the large, centralized generation plants which are most of the time located some distance from the load centres. This bulk power is evacuated over transmission lines to load centres where they are distributed along feeders to residential, commercial or industrial consumers at a relatively lower voltage level as shown in Fig. 1.



**Fig. 1.** Traditional power flow (adapted from [2])

With DG units connected on the transmission or distribution sides of the power system, upstream power flow is reduced since power is usually consumed within the same region as it is being produced. This decreases the quantum of power loss in conductors and power transformers as compared to when the whole power was supplied from the large central power plants.

### 1.1.2 Transmission and Distribution Losses

Transmission and Distribution (T&D) losses are inherent to any power system and arise due to the dissipation of energy in current carrying components and equipment during the transmission and distribution of power. Power utilities strive to bring down this value which is a Key Performance Index in assessing the effectiveness of a utility. The losses can be decreased to an optimum level but cannot be eliminated [3]. According to [4], these losses can be decomposed into various components as shown in Table 1.

**Table 1.** Distribution of technical losses along the T&D network

System element	Power losses (%)	
	Minimum	Maximum
Step-up transformers and transmission system	2.0	4.0
Sub-transmission system and step down to distribution voltage level	2.0	4.5
Distribution Lines and Service Connections	3.0	7.0
<b>Total Losses</b>	<b>7.0</b>	<b>15.5</b>

It can be observed that the network which contributes mostly to the T&D losses is the Distribution Lines and Service Connections. Nonetheless, utilities seldom compute losses at secondary distribution level through modelling techniques but instead, estimate these losses using spreadsheet approximations [5].

### 1.1.3 Impact of DG on System Losses

Technical losses are inherent to any electrical network due mainly to  $I^2R$  losses in the current carrying elements and can be represented by [6]:

$$\text{Feeder Technical Losses} = \sum_{k=1}^{k=m} I_k^2 R_k \quad (1)$$

where  $k$  is the number of different subsections being considered (up to  $m$ ). The resistance,  $R$  will depend on the length of the different sections, conductor size, conductor type or conductor configuration.

It has been thought that the addition of DG would definitely decrease these losses since power produced would usually be consumed within a close range. A study by [7] has demonstrated that improvement in losses by 80% can be achieved by optimal placement of the DG. Indeed, it has been demonstrated that the location of the DG directly affects the resultant line losses [8]. However, [9] have demonstrated that while DG may unload lines and reduce losses, the reverse power flow from larger DG can give rise to excessive losses. Some research have even been devoted to the minimization of this component in system losses [6, 10–12]. This may also be true during off-peak period or days with low consumption in the area with DG connections. It was confirmed by [13] that the convenient location of DG units can significantly decrease T&D losses. However, they also noted that if customers are located a distance away from the DG sources, these losses can increase and affect the power distribution system.

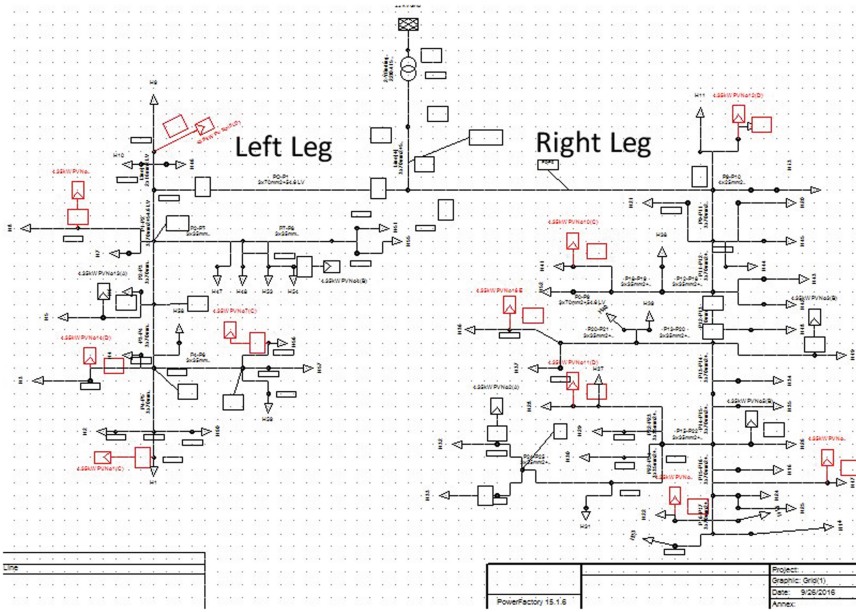
Similarly, [14] proved that the proper rating and placement of DG dictate the line loss in a network. The author described the importance of matching the DG rating to the load. In instances where the generated power cannot be consumed by the load, the resulting power flow resulted in higher line losses.

## 2 Methodology

A detailed modelling of a low voltage distribution network was necessary to have conclusive results to this study.

### 2.1 Modelling

DIgSILENT Power Factory® was used to model the low voltage distribution network. The parameters for the network components were obtained from their respective technical data sheets. This model is shown in Fig. 2 as having a distribution transformer with two downstream feeders supplying some 60 houses.

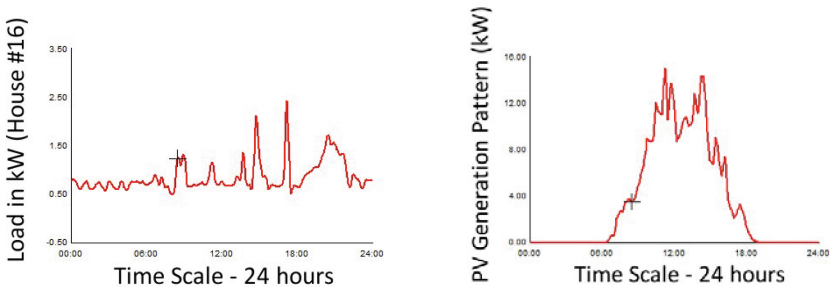


**Fig. 2.** DiGSILENT power factory® model of an actual residential area

## 2.2 Consumption and PV Generation Patterns

Out of the several methods used for the computation of losses by the industry, the Analysis of Hourly Load Level Scenarios generally provides more accurate results [5].

Metering data were sampled and recorded every 15 min using electronic meters and these data were fed to the model. A resulting typical profile is shown in Fig. 3a. Similarly, the generation pattern from a Photovoltaic (PV) system was sampled, recorded and fed to the DiGSILENT Power Factory® software. The same pattern was used for the different PV units connected to the network model during the simulation for increasing penetration levels, as shown in Fig. 3b. This is because the solar power incidence will be practically the same in this small geographical area of study.



**Fig. 3.** (a) Typical residential load profile (b) Typical PV generation profile



### 2.3 Penetration Level

The Penetration Level (PL) of DG is the amount of active power injected into the network ( $\sum P_{DG}$ ) to the network load capacity ( $\sum P_{LD}$ ) and can expressed as [15]:

$$PL = \frac{\sum P_{DG}}{\sum P_{LD}} \times 100 \quad (2)$$

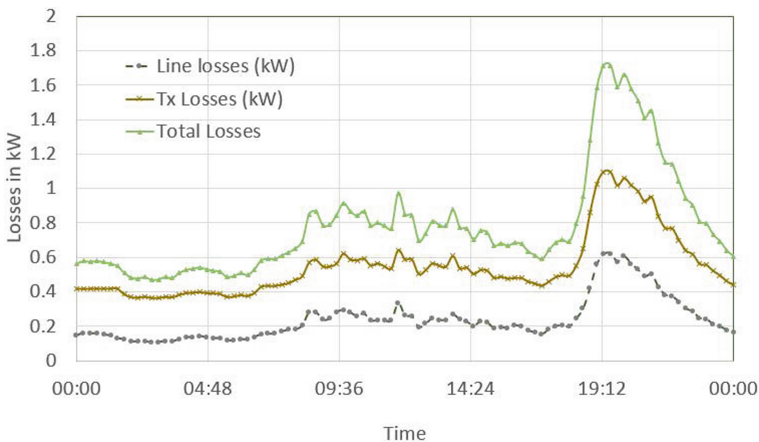
### 2.4 Computation of Losses

There are two main components contributing to the losses in the low voltage network: the transformer losses and the  $I^2R$  loss in the actual conductors of the network. The former is made up of the Core Loss which is practically constant and the Copper Loss which is dependent on the loading. System losses are highest during peak conditions. However, approximately 70% of the energy losses occur off peak [5].

## 3 Results and Discussion

### 3.1 Network Losses with no DG Connected

With no DG connected on the network, unbalanced load flows were simulated for every 15 min interval. The line and transformer losses were recorded and plotted over a 24-h period; along with the total losses given by the arithmetic sum of these two losses. Figure 4 depicts these losses.

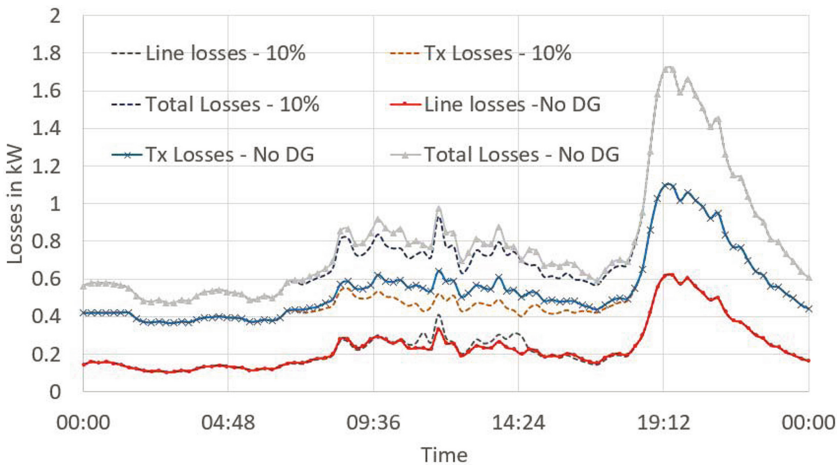


**Fig. 4.** Network losses with no DG connected

It is observed that the transformer losses predominate the line losses by nearly twice the value. As reported by [5], the maximum network losses actually occur during the (evening) peak. The energy loss has been computed as 5.76 kWh for the line and 13.29 kWh for the transformer. The system loss is 19.05 kWh.

**3.2 Network Losses with 10% of DG Penetration**

As per Eq. (2), a Penetration Level of 10% refers to a total connected SSDG capacity of 8.7 kW since the network load capacity has been found to be 87 kW. This was achieved using 2 units of 4.35 kW each, connected on the individual feeder. The values of the losses obtained during the load flows were plotted against the corresponding values with no DG connected. Figure 5 shows these different profiles.

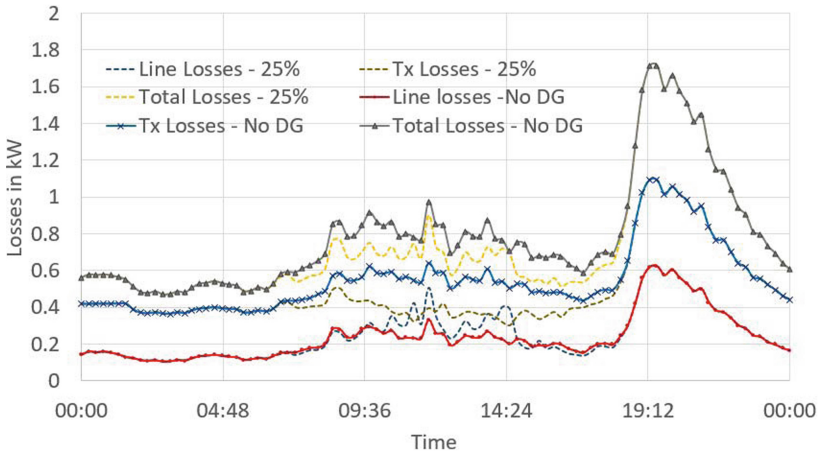


**Fig. 5.** Network losses with 10% of DG penetration

As soon as power is generated from the PV systems, the losses in the system change; but these changes are different for the different components. The line losses increase while the transformer losses decrease. However, the quantum of decrease in transformer losses is larger than the increase in line losses. The energy loss along the line is 5.87 kWh and that across the transformer is 12.57 kWh; resulting in a total energy loss of 18.44 kWh which is a 3% decrease in system losses.

**3.3 Network Losses with 25% of DG Penetration**

When the Penetration Level is increased to 25% achieved with 5 SSDG units of 4.35 kW each dispersed along the feeders and connected on different phases, the resulting effect on the components of the losses are as shown in Fig. 6.

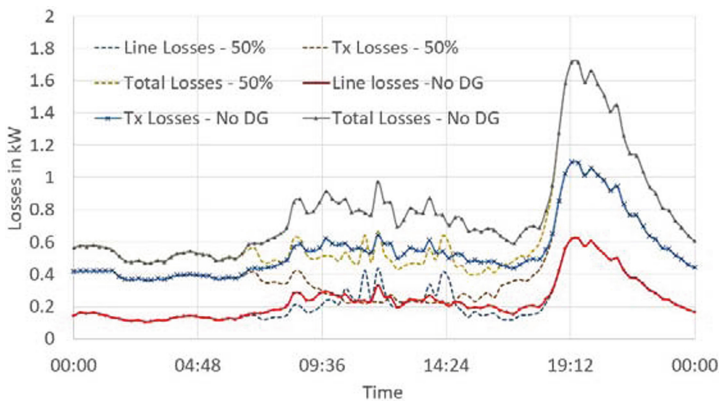


**Fig. 6.** Network losses with 25% of DG penetration

In line with the previous set of results, the line losses increase while the transformer losses decrease. During the peak generation phase of the PV systems at around mid-day, the resulting line losses grow larger than the transformer losses. Still, the overall system losses are lower than without connection of DG. In terms of kWh, the line losses have now increased to 6.07 kWh while the transformer related losses are 11.77 kWh. The resulting overall energy loss is now 17.84 kWh; which is a 6.3% decrease in system losses as compared to a system with no DG.

### 3.4 Network Losses with 50% of DG Penetration

If the number of connected SSDG units is doubled, a Penetration Level of 50% is reached. 10 SSDG units of 4.35 kW each were generating power along the feeders and the resulting losses are shown in Fig. 7.



**Fig. 7.** Network losses with 50% of DG penetration

With 50% of DG penetration, there is a decrease in the line losses at instances during the day; which was not the case for the previous penetration levels. At other instances, the line losses increased about the reference ‘no DG’ curve. There is noticeable decrease in the transformer losses, resulting in a similar fall in overall system loss. In terms of energy, line losses have this time decreased to 5.6 kWh and the transformer losses to 10.65 kWh. The overall system losses are 16.26 kWh which is a 14.7% decrease.

### 3.5 Network Losses with 75% of DG Penetration

With an additional 25% of SSDG units connected, it is observed (Fig. 8) that the line losses do have an initial decrease in the early morning but which drastically increase by mid-day. The transformer losses decrease consistently. However, for this scenario, the overall system losses do actually rise above the reference but for a very short duration. The line losses register an increase to 6.13 kWh and the transformer losses are 10.45 kWh. These correspond to a decrease in overall system losses by 13% (16.57 kWh).

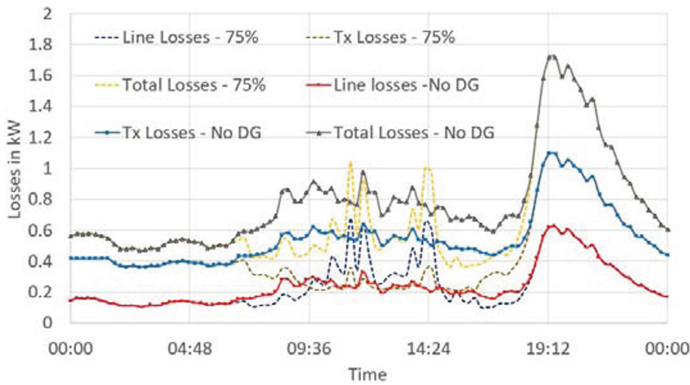
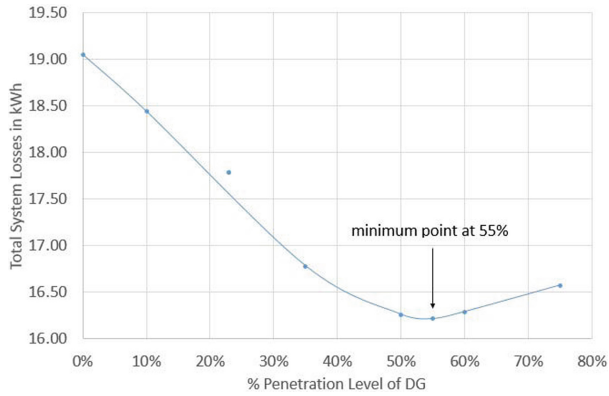


Fig. 8. Network losses with 75% of DG penetration

### 3.6 Optimum Penetration Level for Overall Minimum Losses

The results show that the different components contributing to losses in the low voltage network react differently to increasing penetration of DG. For penetration levels below 50%, the line losses increase slightly while for higher penetration levels, there is a decrease in losses along the line. The transformer losses decrease with increasing penetration levels. The transformer losses have very high influence on the overall system losses. Overall, it can be seen that losses in the low voltage network decrease in the presence of SSDG units. However, the fact the total energy losses at 50% penetration level is 16.21 kWh and that at 75% penetration level is 16.57 kWh conveys that there is an optimum penetration level that would result in the most optimum losses in the network. Figure 9 shows that system losses decrease with increase in penetration of DG and then rise giving an optimum value at about 55% of penetration.



**Fig. 9.** Variation of system losses in kWh with respect to % penetration level

## 4 Conclusion

A highly residential area was modelled in the DIgSILENT power factory software using actual network parameters, load profiles and PV generation pattern to assess the impact of the connection of SSDG units. There are two main components contributing to the losses in the low voltage network: transformer losses and line losses. Both however reacted differently to increasing SSDG units in the network. The transformer losses showed systematic decrease during the various penetration levels. The line losses, on the other hand, increased during low penetration of SSDG but afterwards decreased. At higher penetration level, the line losses increased again. It was demonstrated that the total system losses were highly influenced by the transformer losses. It can be established that the introduction of SSDG units in the low voltage network contributed to having a lower system losses in the network; whatever the penetration level. Yet, these reduction in losses varied with the penetration level of the SSDG. The study showed that the penetration level which resulted in minimum system losses for the network under study was 55%. These results may also deviate slightly with different placements of the SSDG units in the network but the general trend would be maintained. In the real situation, the overall T&D losses will still be further reduced since the decrease in transformer loading is synonymous to less current being drawn from the primary distribution and transmission networks; resulting in lowered  $I^2R$  loss in these components.

## References

1. Shi, Y., Yao, S., Wangl, Y.: Research on distribution system restoration considering distributed generation. In: The 4th China International Conference on Electricity Distribution (CICED), Nanjing, China, pp. 1–5 September 2010
2. CEB. Small Scale Distributed Generation (SSDG) Project. [http://ceb.intnet.mu/grid\\_code/project.asp](http://ceb.intnet.mu/grid_code/project.asp)

3. Bhalla, M.S.: Transmission and Distribution Losses (Power). The Energy and Resources Institute. <http://www.teriin.org/upfiles/pub/papers/ft33.pdf>
4. EPRI. Distribution System Losses Evaluation. Palo Alto, CA: 1016097 (2008)
5. Short, T., Swayne, T.: Assessment of Transmission and Distribution Losses in New York. PID071178 (NYSERDA 15464) Final Report, November 2012
6. Lakshmi, D.A., Subramanyam, B.: Optimal DG unit placement for loss reduction in radial distribution system – A case study. *APRN J. Eng. Appl. Sci.* **2**(6), 57–61 (2007)
7. Mithulananthan, N., Oo, T., Phu, L.V.: Distributed generator placement in power distribution system using genetic algorithm to reduce losses. *Thammasat Int. J. Sci. Technol.* **9**(3), 55–62 (2004)
8. Wei, B., Guoqing, H., Wenhui, S., Kaihui, F., Weiran, Z.: Distributed wind generation siting and sizing considering fluctuate wind resources. In: *IEEE PES Innovative Smart Grid Technologies*, pp. 1–7 (2012)
9. Ochoa, L.P., Padilha-Feltrin, A., Harrison, G.P.: Evaluating distributed generation impacts with a multiobjective index. *IEEE Trans. Power Deliv.* **21**(3), 1452–1458 (2006)
10. Lund, T.: Analysis of Distributed Systems with a High Penetration of Distributed Generation. Ph.D. Thesis. Technical University of Denmark (2007)
11. Gopiya, N.S., Khatod, D.K., Sharma, M.P.: Optimal allocation of distributed generation in distributed system for loss reduction. In: *IACSIT Coimbatore Conferences*, Singapore (2012)
12. Abbagana, M., Bakare, G.A., Mustapha, I., Musa, B.U.: Differential evolution based optimal placement and sizing of two distributed generators in a power distribution system. *J. Eng. Appl. Sci.* **4**, 61–67 (2012)
13. L'Abbate, A., Fulli, G., Starr, F., Peteves, S.D.: Distributed Power Generation in Europe: technical issues for Further Integration. JRC Scientific and Technical Reports
14. Chiradeja, P.: Benefit of Distributed generation: a line loss reduction analysis. In: *EEE/PES Transmission and Distribution Conference and Exhibition: Asia and Pacific Dalian, China*, pp. 1–5 (2005)
15. Farag, H.E., El-Saadany, E.F.: Voltage regulation in distribution feeders with high DG penetration: From traditional to smart. In: *Power and Energy Society General Meeting*, pp. 1–8 (2011)

# Noise Measurement and Analysis in a Power Line Communication Channel

B. Rajkumarsingh<sup>(✉)</sup> and B.N. Sokappadu

Department of Electrical and Electronic Engineering, University of Mauritius,  
Reduit, Mauritius

b.rajkumarsingh@uom.ac.mu,  
bhargava.sokappadu@umail.uom.ac.mu

**Abstract.** Due to the harsh environment of the power line, data transmission is greatly affected by noise, signal attenuation, interference, multiple reflections and varying impedance and loads. The noise channel at the University of Mauritius has been studied and the types of noise have been identified and modeled in Simulink. Results show that the noise with the highest impact on data transmission is impulsive noise. Narrowband noise has the least effect.

**Keywords:** Noise · Measurement · Power line communications · Coupling

## 1 Introduction

Power line communication (PLC) involves the transmission of data either in a domestic or industrial environment over existing mains power supply infrastructure which consists mainly of wires carrying AC supply at 230 V and 50 Hz. Since the setup of the PLC standards such as the HomePlug Power Alliance, nowadays, power line communication is being widely used in home automation and networking. It is also a popular alternative to wired communications since it can be implemented at a much reduced cost [1]. The mains supply provides a harsh environment and hence data transmission is affected by the following phenomena: noise, multipath propagation, signal attenuation, interference and frequency-dependent impedance of the conductors [2]. Noise in the power line is mainly produced by the various electrical appliances connected. From the findings of Lim et al. [3], it can be deduced that the primary sources of noise in the domestic environment are due to electrical loads from light dimmers, vacuum cleaners and corroded wiring junctions. Power line noise can be classified into two main types: Background noise and Impulsive noise [9, 10]. Different types of noises in the channel tend to distort the desired signal and therefore cause a change in the data received. From the analysis in [1], it is shown that the BER can reach up to 4 for an SNR of 0 dB when transmitting data over a channel with the appropriate modeling of the various forms of noise present. Noise analysis in power line communication has indeed been an expanding field of study ever since its introduction with the aim of improving data transmission quality. One of the foremost studies have as objective to design an appropriate coupling circuit in order to be able to filter out the AC component while allowing higher frequency signals to pass through so as to record

the noise signals as laid forward by Gassara et al. [4]. After the analysis of the noise spectrum, an approximate model of the channel can be developed. One of the prime modeling techniques concerning impulsive noise have been suggested in [5]. Owing to the design and successful implementation of Narrowband PLC systems, further research and analysis of high frequency noise spectrum in Broadband PLC has been carried out for ranges between 10 kHz and 100 MHz [6, 7]. Modeling techniques for Broadband Power line communications have been developed focusing mainly on Background and Appliance noise. The applications of PLC have ever since been increasing and can now even be put into use at the University of Mauritius. However, first an investigation on the quality of the power line channel on the campus is necessary as each channel will have its own model restrained by its own parameters and only after analysis, it can be determined whether the use of NB-PLC will be effective.

This research aims at the study of noise in the power line channel at the University of Mauritius and devising a mathematical model to represent it. The prime objectives of this research are the design and implementation of a coupling circuit (with the necessary safety components) for noise analysis, measurement and recording of the noise over a span of time, analysis of the recorded data to find out the types of noises present, modeling of these noises according to actual noise models available, modeling of the channel with the appropriate noises included. The outline of the paper is as follows. Section 2 describes the methodology employed. Key importance has been given to the filter design part following which details on how the noise data have been recorded are given. Section 3 describes the analysis of the different noises and the modeling parameters. In Sect. 4, the noise modeling in Matlab is explained. We end with conclusions in Sect. 5.

## 2 Methodology

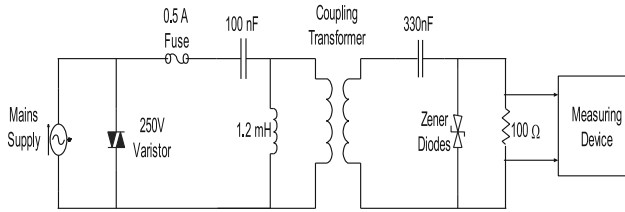
The main aim of the coupling circuit is to incorporate a filter so as to remove the 50 Hz AC signal component while allowing signals of other frequencies. The ideal implementation of this scenario would be to design a bandstop filter with a very narrow stopband near 50 Hz. However, this design is not feasible as this would result in a filter of a very high order hence increasing the complexity. Hence, the most practical solution would be to implement a low pass filter which would attenuate the low frequency components and most importantly, the 230 V AC signal should be attenuated sufficiently enough so that it does not appear as “noise” at the output. In this context, a third order highpass Butterworth filter was designed. The transfer function for the 3<sup>rd</sup> order highpass Butterworth filter is given below.

$$H(s) = \frac{s^3}{1 + 2s + 2s^2 + s^3} \quad (1)$$

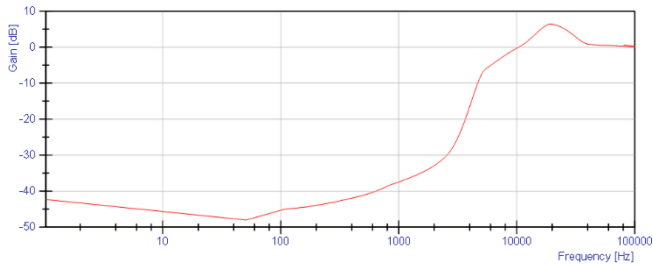
The final circuit together with its safety components is as shown in Fig. 1.

The actual magnitude response of the filter was measured using the softwares NI ELVISmx, LabVIEW and DIAdem. The frequency response is band limited to 100 kHz due to the limit in sampling rate of myDAQ. The magnitude response of the filter is shown in Fig. 2.





**Fig. 1.** Coupling circuit with high-pass filter



**Fig. 2.** Magnitude response of coupling circuit.

From the findings in [4], it can be deduced that the same approach as above can be used for the implementation of the receiver coupling circuit which is symmetric to the receiver side. Experiments conducted on the coupling circuit prove that the latter can be used for bidirectional transmission.

### 3 Analysis and Results

#### 3.1 Noise Analysis

From the whole set of data that has been recorded and compiled accordingly, several aspects of the power line channel have been analyzed, namely:

1. Noise Spectrum variation with time of the day.
2. Types of noise present.
3. Analyzing these noises and thereafter developing a statistical noise model.
4. Simulating the noise model using Simulink.

#### 3.2 Noise Spectrum

As mentioned in [8], the noise density spectrum (NDS) offers an easy way to interpret the various noises being generated with respect to frequency for a certain type of equipment. Unfortunately, noise is never a fixed factor but a time-varying one depending mainly on the electrical loads connected to the mains. Extrapolating from

this idea, a mean noise spectrum has been devised from the recorded data which gives a basic notion of the variation of the noise with frequency and time of the day at the University of Mauritius Microprocessors Lab. Figure 3 shows a plot of the mean noise power spectral density from 9 a.m. to 4 p.m.

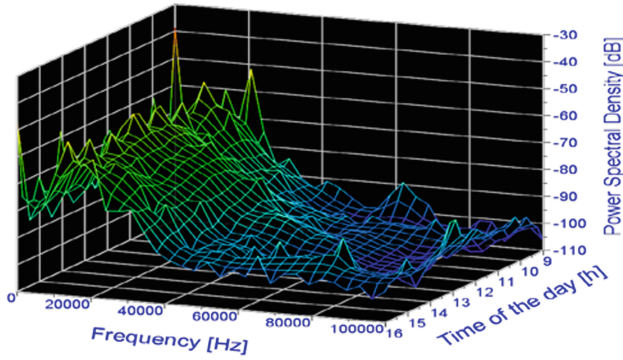


Fig. 3. Variation in power spectral density with time of the day

Figure 4 below shows the maximum and minimum noise power spectral density for the whole data.

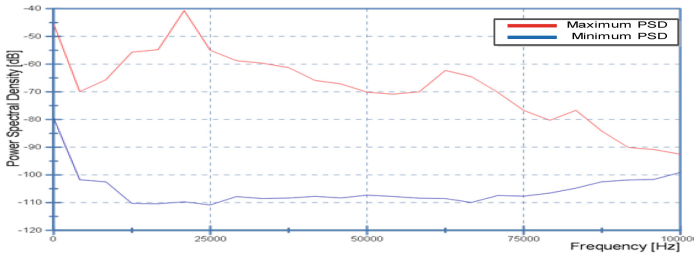
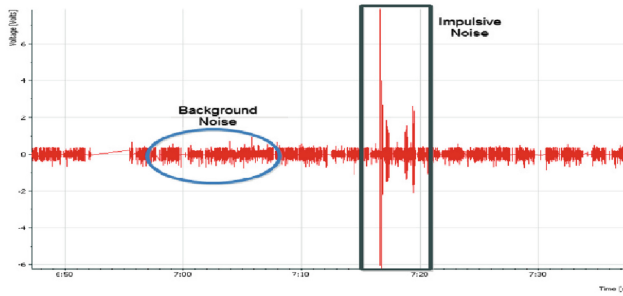


Fig. 4. Maximum and minimum power spectral density

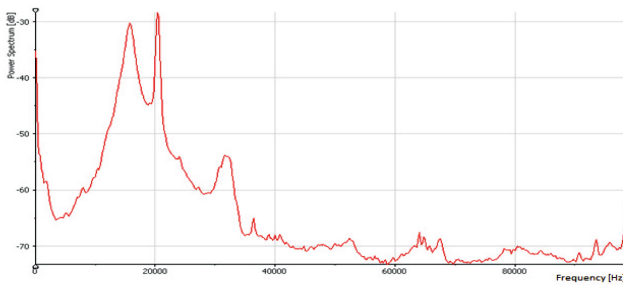
The noise spectrum only gives a trend of the noise according to time and frequency. In order to be able to understand further regarding the trend of the noise in the time domain, a statistical method needs to be employed [9]. From the noise waveform, the main noises present can be identified: namely the impulsive noise which is represented by a series of pulses of short duration and high in amplitude and the background noise which has a lower magnitude but is present throughout. Figure 5 shows a sample of background noise and impulsive noise.

The mean power spectrum for the whole set of data that has been recorded and is shown in Fig. 6.

From the graph of Fig. 14, the fundamental peaks are given in Tables 1 and 2.



**Fig. 5.** Sample of background and impulsive noise



**Fig. 6.** Mean power spectrum.

**Table 1.** Fundamental frequencies recorded from the mean power spectrum

Fundamental frequencies [Hz]	Magnitude [dB]
0	-35.16
15800	-30.24
20400	-28.34
31600	-53.83

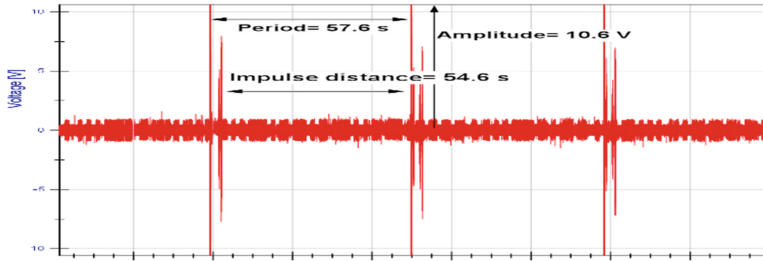
### 3.3 Measured Parameters for the Impulsive Noise

From Fig. 7, it can be clearly observed that there are impulses which appear as an impulse train which implies that it corresponds to periodic impulsive noise and those impulses that do not conform to the same periodicity are classified as aperiodic impulsive noise. In order to devise a proper model for the impulsive noise, a reliable set parameters namely the amplitude, impulse distance and time period, is required. From the data recorded, the mean values for periodic impulsive noise have been calculated and are shown in Fig. 7. This type of impulsive noise has been named as Type 1 impulsive noise.

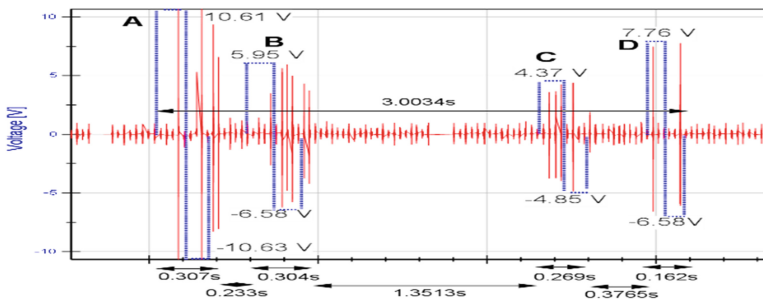
Figure 8 shows a magnified section of a single impulse from Fig. 7 with the magnitudes and time duration of the respective pulses found therein. The rectangular pulses can be used to model the envelope of the pulses and hence provide a far better approximation of the whole impulse.

**Table 2.** Portions of the power spectrum with lowest noise

Frequency range [kHz]	Magnitude range [dB]
55.8–59.8	–73.33 to –72.29
69.2–75.4	–73.27 to –72.76



**Fig. 7.** Mean values of the different parameters used for impulsive noise



**Fig. 8.** Parameters for modeling of an impulse using rectangular pulse

Most of the impulses that are present in the time-domain waveforms conform to the same pulse width with minor variations in amplitude. They vary mainly in periodicity. For this reason, the same model for a single impulse can be used and by varying the impulse distance, even aperiodic impulsive noise can be modelled.

The next type of impulsive noise that has been identified is as shown in Fig. 9 (named as Type 2 impulsive noise). The use of rectangular envelope has been deployed to model the impulse even though it is itself made up of pulses of shorter duration, because of limitations of myDAQ and the way LabVIEW records waveforms. The waveform is supposed to be a continuous time signal, however, LabVIEW records samples for a specific time interval of 5 ms and then takes another sample for recording after 0.029 s. Therefore, it is not possible to correctly identify the pulses within the rectangular pulse model that has been used and hence the whole repeating sequence has been modeled.

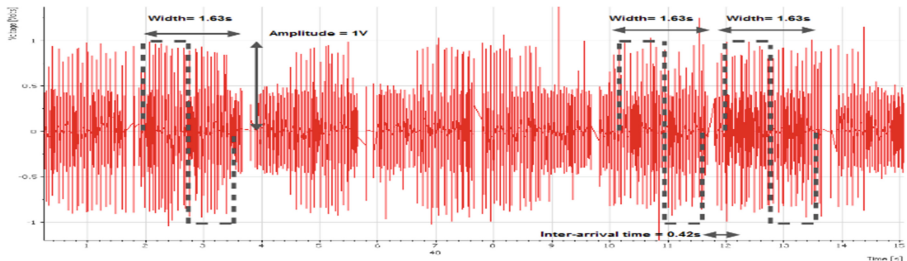


Fig. 9. Parameters for modeling of the second type of Impulsive noise

### 3.4 Measured Parameters of the Narrowband Noise

From the previous chapters, it can be deduced that narrowband noise appears as amplitude modulated sinusoids and occupy a small segment of the power spectrum. A highly magnified view of the noise waveforms results in Fig. 10.

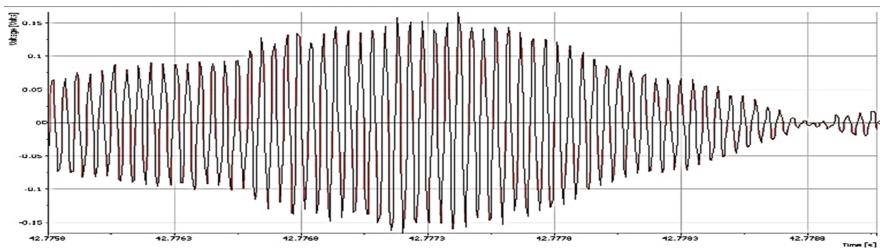


Fig. 10. Sample time-domain waveform of narrowband noise of frequency 19.8 kHz

Figure 10 consists of a sinusoidal signal having a period close to  $50.5 \mu\text{s}$  that has been amplitude modulated resulting to a frequency of approximately 19.8 kHz. From the frequency spectrum, it can be concluded that the peak at 20.4 kHz is due to narrowband noise similar to the above. Another portion (Fig. 11) of the noise waveform has been found to consist of a sinusoid quite similar as above but varying in time period and hence resulting in a frequency of 15.78 kHz. The latter corresponds to the first peak obtained at 15.8 kHz in the frequency spectrum. Therefore, it can be concluded that the two peaks at 15.8 kHz and 20.4 kHz constitute the narrowband noise in the channel. Another point that supports the idea of considering these two peaks as narrowband noise is that both of them occupy a narrow portion in the frequency spectrum as it is the case for signals causing narrowband interference. The signals at 15.8 kHz and 20.4 kHz occupy a bandwidth of 5.52 kHz and 1.41 kHz respectively.

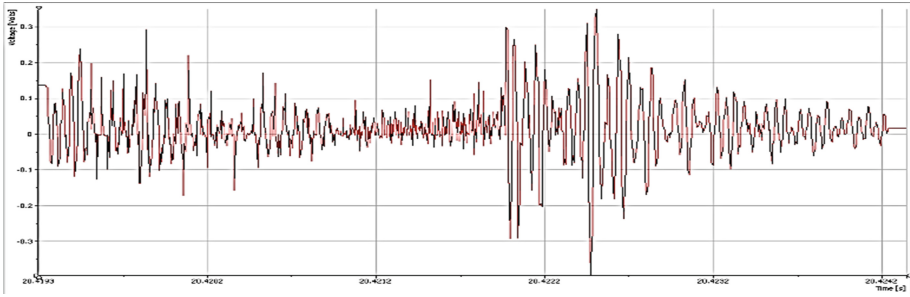


Fig. 11. Narrowband Noise identified at a frequency of 20.4 kHz

## 4 Noise Modeling in Simulink

### 4.1 Periodic Impulsive Noise

In relation to the different types of impulsive noises that have been analyzed in the section above, a model has to be developed for each type so that it can later on be used to model the channel as a whole. Regarding the noise modeling technique used by [5], a rectangular pulse envelope has been used to model each impulse and a set of rectangular pulse generators have been used to model the impulsive noises.

The Pulse Generator blocks in Simulink are used to generate one-sided rectangular pulse of specific pulse width, period, amplitude and phase. This means that to generate the negative part of a two-sided rectangular waveform with zero DC offset, another pulse generator is required. The model developed on Simulink, for the impulsive noise modeling is shown in Fig. 12.

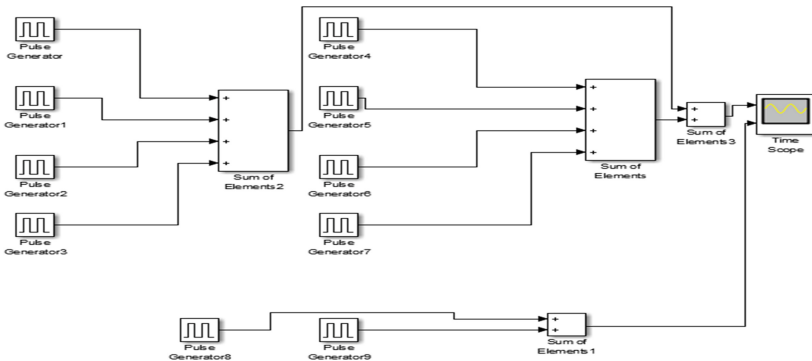
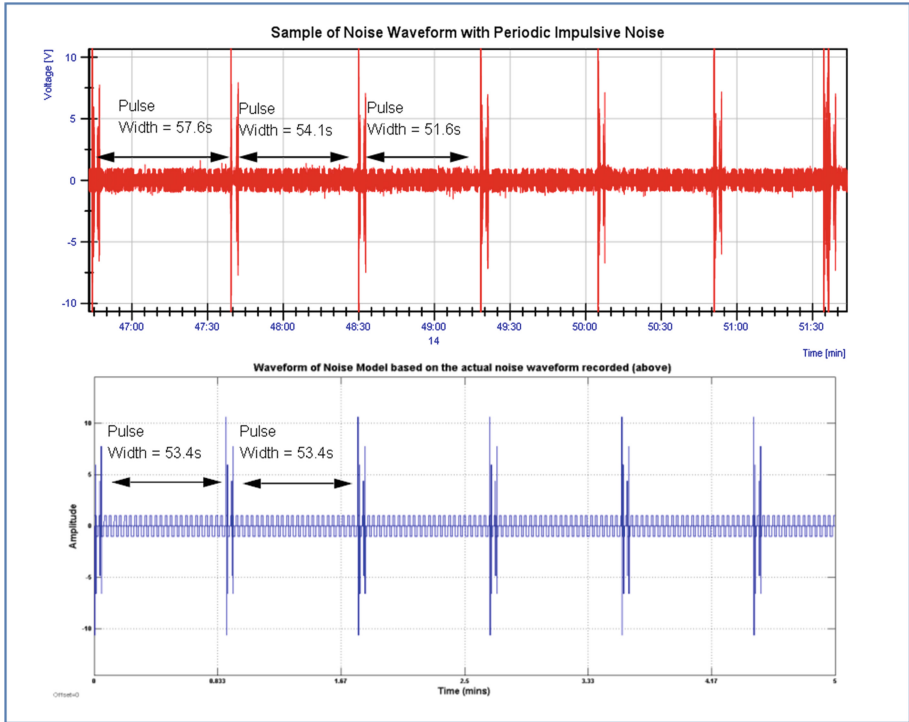


Fig. 12. Simulink model for impulsive noise generation

The output waveform from the generation of the impulse train has been obtained as in Fig. 13.



**Fig. 13.** Sample of recorded impulsive noise compared to modelled noise

The model generated for periodic impulsive noise may not entirely reflect the noise waveform recorded in actual because of the varying pulse width. In order to demonstrate the modeling of the recorded signal in Fig. 13, a pulse width of 53.4 s which conforms to the mean pulse width for the recorded noise waveform was used. However, for further modeling purposes, the mean pulse width of 57.6 s was used throughout for Type 1 impulsive noise.

## 4.2 Aperiodic Impulsive Noise

For the case of aperiodic impulsive noise, only the Type 1 impulsive noise has been considered since Type 2 conforms to periodic noise. The periodicity of the impulsive noise in aperiodic noise varies randomly as observed from Fig. 5. Therefore, to construct an appropriate model, the same impulse model as obtained from Fig. 8 was used, the worst case value for the inter-arrival time of the impulse was calculated to be 13.243 s while for the amplitude, the maximum amplitude of 10.63 V was used.

### 4.3 Background Noise

#### 4.3.1 Colored Noise

Background noise consists mainly of the predominant narrowband noise while the secondary type of noise is colored. Colored noise can be modeled by several existing models such as White (Gaussian Noise), Brown or Blue noise amongst others. However, the right model can be chosen only after analyzing the power spectral density and the probability density function of the noise. As the power spectral density has as main highlights, narrowband noise, the probability density function of the noise would provide an easier to approach to colored noise modeling. Figure 14 is the probability density function of the noises recorded. The y-axis corresponds to an arbitrary scale for the probability density as assigned by LabVIEW. The value  $x$  corresponds to the random variable which is the measured voltage of the received signal.

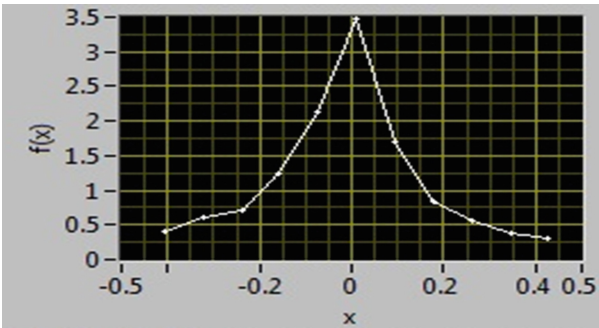


Fig. 14. Probability density function of the recorded noise

From Fig. 14, it can be deduced that the colored noise in the PLC channel profile follows a Gaussian distribution with a mean value,  $\mu$ , of 2.792 mV and a variance of  $0.01945 \text{ V}^2$ . Using Simulink, White noise bearing the above mentioned parameters can be simulated and the way the White noise affects data transmission is thereafter assessed.

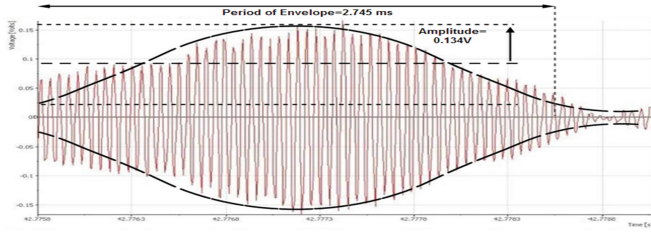
#### 4.3.2 Narrowband Noise

The research carried out in [5, 11] describe narrowband noise modeling as a series of amplitude modulated (AM) sinusoids. In conjunction with the narrowband noise that has been identified in the PLC profile, the same technique of simulating the narrowband noise at different frequencies using different sinusoids can be employed. Narrowband noise to some extent encloses some part of coloured noise as well. The envelopes of the narrowband signals are identified as shown in the following figures. Narrowband noise can be effectively removed using a bandpass filter which shall allow only the signal of desired frequency but a filter of high order will be required because of the high magnitude of the narrowband signal.

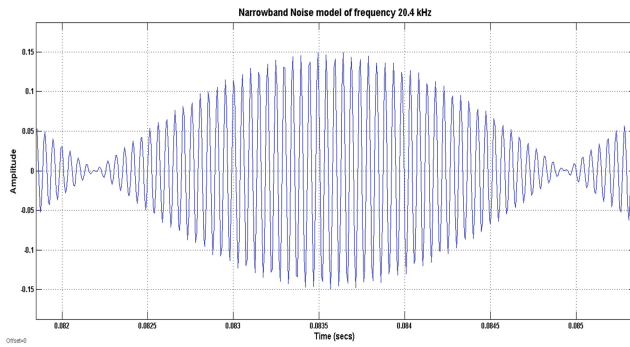


• **20.4 kHz Narrowband Signal**

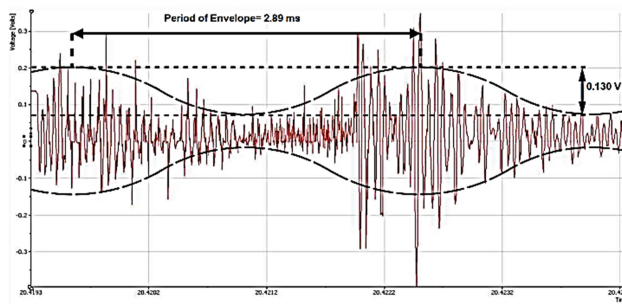
The model that can represent the signal in Fig. 15 would be an AM wave with carrier frequency of 20.4 kHz and amplitude 0.162 V and message signal (envelope) of frequency 364.3 Hz and amplitude 0.134 V. The corresponding model is shown in Fig. 16.



**Fig. 15.** Parameters used in modeling of narrowband noise of frequency 20.4 kHz



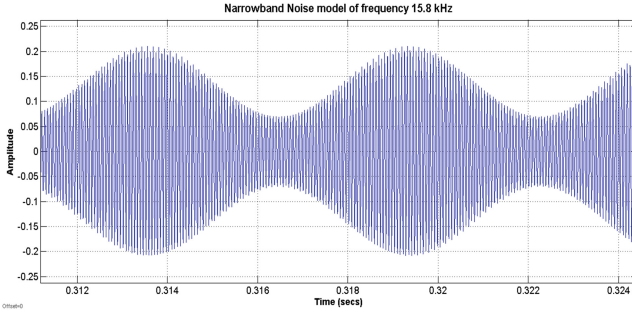
**Fig. 16.** Model of narrowband noise of frequency 20.4 kHz



**Fig. 17.** Parameters used in the modeling of narrowband signal at 15.8 kHz

- **15.8 kHz Narrowband Signal**

The same procedure as above can be repeated to obtain a model for the 15.8 kHz narrowband signal (Fig. 17) with the simulated signal shown in Fig. 18.



**Fig. 18.** Model for narrowband noise of frequency 15.8 kHz

## 5 Conclusion

Noise data was recorded for several time durations and the whole of the data was compiled to extract the necessary parameters of the noise. Noise modeling was performed for impulsive noises and for the background noise. For the case of aperiodic impulsive noise, the worst case scenario was considered whereby the impulse repeats itself every 13.2 s. It can be deduced that the noise with the highest impact on data transmission is Periodic Impulsive noise followed by aperiodic impulsive noise. It has also been shown that among the periodic impulsive noises, the Type 2 noise has greater significance with a minimum BER of 0.195 compared to 0.009 for the Type 1 noise. These values conclude that the performance of a channel is more affected by impulsive noises which are more frequent as is the case for Type 2 which has a period of 2.05 s compared to 57.6 s for the impulsive Type 1. Furthermore, another deduction can be that the frequency of occurrence and duration of an impulse has greater impact compared to the magnitudes of the impulse. Type 2 noise has a pulse width of 79.9% while Type 1 noise can be considered to have a pulse width of 5.21% when the impulse is considered as a whole even though Type 2 noise has a maximum magnitude of only 1 V when compared to 10.63 V for the Type 1 noise. In addition to the above, Narrowband noise has the least effect on data rate transmission followed by Gaussian noise.

## References

1. Mathew, S., Murukan, P.: Periodic impulsive noise reduction in OFDM based power line communication. *Int. J. Res. Eng. Technol.* **3**(5), 517–522 (2014)
2. Deshpande, S., Prasanna, I., Panda, S.K.: An efficient impedance matching technique for improving narrowband power line communication in residential smart grids. *Int. J. Adv. Res. Electr. Electron. Instrum. Eng.* **2**(7), 3206–3213 (2013)
3. Lim, C., Gunawan, E., So, P., Chen, S., Lie, T., Gua Y.N.: Development of a test bed for high-speed power line communications. In: *Proceedings of the International Conference on PowerCon 2000, Perth, WA* (2000)
4. Gassara, H., Bali, M., Duval, F., Rouissi, F., Ghazel, A.: Coupling interface circuit design for experimental characterization of the narrowband power line communication channel. In: *IEEE International Symposium on Electromagnetic Compatibility (EMC), Pittsburgh, PA* (2012)
5. Zimmerman, M., Dostert, K.: Analysis and modeling of impulsive noise in broadband powerline communications. *IEEE Trans. Electromagn. Compat.* **44**, 249–258 (2002)
6. Tanaka, M.: High frequency noise power spectrum, impedance and transmission loss of power line in Japan on intrabuilding power line communications. *IEEE Trans. Consum. Electron.* **34**, 321–326 (1988)
7. Hirayama, Y., Okada, H., Yamazato, T., Katayama, M.: Noise analysis on wide-band plc with high sampling rate and long observation time. In: *International Symposium on Power-Line communications and Its Application, Kyoto, Japan* (2003)
8. Voglgsang, A., Langguth, T., Komer, G., Steckenbiller, H., Knorr, R.: Measurement, characterization and simulation of noise on powerline channels. In: *International Symposium on Power Line Communications and Its Applications (ISPLC)* (2000)
9. Meng, H., Guan, Y., Chen, S.: Modeling and analysis of noise effects on broadband power-line communications. *Trans. Power Delivery* **20**(2), 630–637 (2005)
10. Cortes, J.A., Diez, L., Canete, J., Sanchez-Martinez, J.J.: Analysis of the infor broadband power-line noise scenario. *Trans. Electromagn. Compat.* **52**(4), 849–858 (2010)
11. Babic, M., Hagenau, M., Dostert, K., Bausch, J.: Theoretical postulation of PLC channel model. *IST Project Deliverable D4v2.0* (2005)

# Using the Development of the Information and Communications Technology Sector to Enhance Teaching in Electrical Engineering

Bhimsen Rajkumarsingh<sup>1</sup>(✉) and Sarjoosing Goolaup<sup>2</sup>

<sup>1</sup> Department of Electrical and Electronic Engineering, University of Mauritius, Reduit, Mauritius

b.rajkumarsingh@uom.ac.mu

<sup>2</sup> Nanyang Technological University, Singapore, Singapore  
SGoolaup@ntu.edu.sg

**Abstract.** Web labs provides the opportunity to work from a computer connected to the internet anywhere and at any convenient time. In this work, a web lab for the study of a specific earthing or grounding arrangement has been conceived and implemented. Students are able to investigate different scenarios and understand why proper sizing of the protective device is critical to avoid any electrical accident. The main aims of this research work were to formulate a model for electrical installations of the TT (Terra Terra) earthing configurations, design and develop a test bed using LabVIEW (Laboratory Virtual Instrument Engineering Workbench) front panel as a user interface for conducting the lab experiments remotely. To assess the effectiveness of the web controlled lab a pre-test and a post-test evaluation were carried out. It was found from the results that their understanding of the subject improved substantially.

**Keywords:** Engineering education · Improving classroom education · Pedagogical issues · LabVIEW

## 1 Introduction

Mauritius, an island of around 2040 km<sup>2</sup>, is located some 2000 km from mainland Africa. Mauritius has rapidly developed in the last three decades from a mono-agriculture economy based on sugar to a greater degree of industrialisation, bringing economic growth and a higher standard of living for its 1.2 million population. From the statistics of the ICT sector in Mauritius [1], the following fact figures regarding the development of the ICT sector in Mauritius are mentioned. The Information and Communications Technology (ICT) sector comprises ICT-related activities in Manufacturing, Telecommunications, Wholesale and Retail Trade, and Business Services such as call centers, software development and website development. Among the various activity segments, the IT Enabled Services together with the Business Process Outsourcing (ITES/BPO) has become a key industry contributing 45% to the total value added of the ICT sector. As at 2014 some 140 companies are operating in this industry [1]. The quality of internet access in the country has improved remarkably during the past few years. The incoming and outgoing capacity increased by 43.3%

from 11,421 Megabits per second (Mbps) to 17,077 Mbps and the international Bandwidth capacity (for both incoming and outgoing traffic) per inhabitant increased by 43.0% from 9,462.3 to 13,534.7 bits per second between 2013 and 2014. The percentage of households having access to computers has increased to 53.1% in 2014 from a level of 44.9% in 2012. The number of internet subscriptions per 100 inhabitants has increased to 58.3% in 2014 as compared to 54% in 2013. The distribution of individuals in 2014 by age group indicates that 46.5% of persons aged twelve years and above were internet users and around 80.7% in the age group of 12–19 years had a propensity to be more online. Due to an improvement in the access and use of ICT, Mauritius has made a significant progress towards becoming an information society. The ICT Development Index (IDI) of the country, which ranges from 0 (lowest ICT development) to 10 (highest ICT development), improved from 5.34 in 2013 to 5.67 in 2014. Latest provisional IDI figures compiled by the International Telecommunications Union (ITU) are available by country for 2013. According to these figures, Mauritius ranked 70th among 166 countries in 2013 in terms of ICT development while Denmark (IDI of 8.86) ranked first.

Today in Mauritius the internet is available in 735,000 households [1]. The latter opens a new avenue in the dissemination of education in the country. The broadband infrastructure is in place to enable speed up to 100 Mbps through the FTTH (Fibre to the Home) and LTE. We have a new paradigm ahead of us. The development in technology is happening faster than the development in education. The education environment has expanded and our students are now digital learners. We should reflect on the teaching tools so that we can optimize on the way education is more appropriate and fruitful in the new era of advanced IT access. The new penetration of ICT must be seized and used towards the improvement of teaching methods. Information flow is now not only from educators but through the web as well. We have a new tool in the landscape and we cannot neglect that opportunity to improve learning. It is believed that access to IT and internet helps in the endeavor to produce creative learners. We are embarking in a new teaching culture where we can complement the black/green/white board and paper textbooks. The exposure to multimedia tools through the broadband has now opened the way to even remote lab experiments. Teaching and learning is no longer confined within the walls of class or lab.

Computer simulations using virtual and remote Labs (Weblabs) have enhanced student learning and helped to develop the necessary skills so that the student could integrate the working environment more easily [2–10]. Students do not need to be physically present in the lab to perform the experiment. The real lab with real electrical systems and instrumentations can now be remotely operated through the internet. Weblabs offer to the students an opportunity to work from a computer connected to the internet anywhere and at any convenient time. It is a remote laboratory comprising of software and hardware tool that enables students to remotely access real equipment located in the university as if they were in a hands-on-lab session. In addition, the universities can share their equipment with institutions that are currently using Weblabs, like the MIT iLab project, the Labshare Sahara project, or WebLab-Deusto that is available only in certain universities. Another advantage of using Weblabs is that it permits to perform experiments remotely in electrical installations where the danger of electric shocks is predominant. The students will be able to look at and operate

electrical systems that they are likely to encounter in their professional life without being physically present in the labs. LabVIEW (Laboratory Virtual Instrument Engineering Workbench) is a graphical programming language that uses icons instead of lines of codes to create applications. LabVIEW has been reported as an excellent tool for simulating practical systems such as Electrical Motors and other systems in electronic, mechatronics and electrical engineering [16–19].

The total number of students enrolled in 2012/2013 reached 12254 at the University of Mauritius [15]. The University of Mauritius has continued to strive to increase access to Tertiary Education in line with the current Government's Policy. The percentage of students enrolled in the Faculty of Engineering has increased by 90% in 2012/2013 as compared to the level in 2002/2003. Hands-on learning can give students a better feel for the training material at hand in the context of engineering education.

Earthing is the intentional connection all of the components of an electrical system that could get charged, to the general mass of earth. In the TT system the transformer neutral is earthed and the consumer frame is earthed. In the TT (Terra Terra) system, an RCD is used as an additional device to protect persons against electric shock from direct and indirect contacts [11–14].

In this context, we have implemented a web based lab experiments in connection with the study of low voltage TT electrical installation using LabVIEW. The remote controlled TT system is designed to enable students to conduct experiments on a 240 V line without the risk of being electrocuted. The developed Web labs provide the opportunity to work from a computer connected to the internet anywhere and at any convenient time. The remote laboratory comprising of software and hardware tool enables students to remotely access real equipment located in the university as if they were in a hands-on-lab session.

The main objectives of the work were to design and implement web controlled experiments for the study of RCD tripping in a TT network, state the learning outcomes and to evaluate whether the learning outcomes have been achieved through the pre-test and the post-test exercises.

Section two states the learning outcomes. Section three gives the details about the hardware and software design employed respectively in the construction of the web lab. Section four describes the experiments developed. Section five gives an overview of the pre- and post-tests results. We end the paper with the conclusions.

## 2 Learning Outcomes

The TT system test bed has been developed in the Electrical and Electronic department of the University of Mauritius. This test bed aims at providing a better understanding of this type of earth connection. The TT system is widely used in the country, particularly in households. The learning outcomes are defined as below.

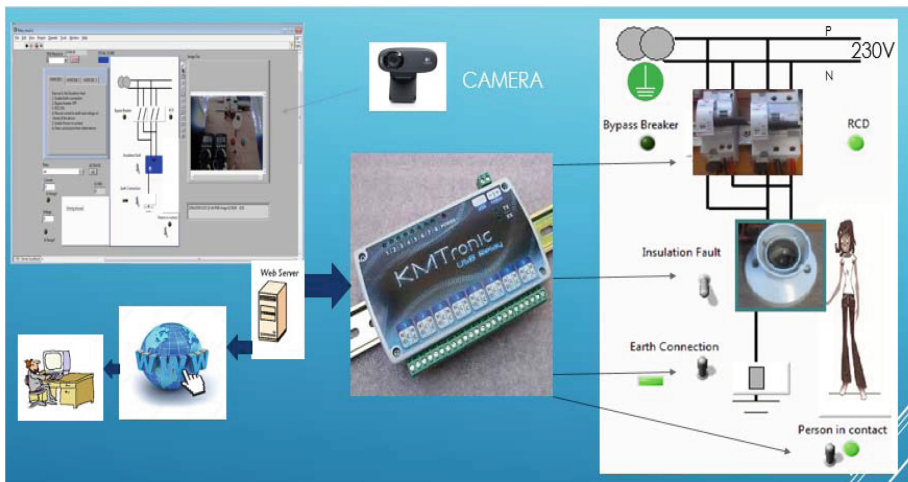
At the end of the experiments the student should be able:

- To understand the meaning of the TT system
- To state the requirements for electrical safety in a TT system
- To demonstrate the importance of an earth connection

- To understand the need of the RCD
- To explain the conditions for the tripping of the RCD
- To understand the protection against direct and indirect contacts
- To understand the protection against electric shock under normal conditions
- To understand the protection against electric shock under fault conditions

### 3 Description of the TT Remote Experiments

In this section, the hardware components used are described and the electrical circuitry is explained. In the next section, the software implemented in LabVIEW is explained. The setup for the proposed system is depicted in Fig. 1.



**Fig. 1.** System architecture

Real lab experiments have been developed and mounted with instruments and relays which can be operated and controlled remotely through data acquisition cards. The user will have to access a secure website where he can log in the system to perform or view the experiment. A webcam is used to allow the user to view the setup and he/she will be able to interact with the electrical system through a customized Graphical User Interface (GUI).

#### 3.1 Hardware Implementation

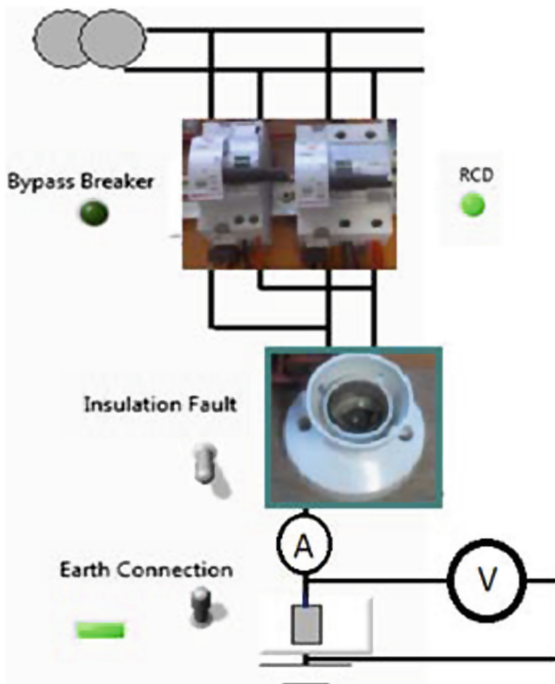
The electrical setup and the main components used are shown in Fig. 1. The aims of the experiments are to investigate the importance of the RCD and earthing on the safety of a TT system. In this context, a resistive load with an exposed conductive frame is used to investigate cases with and without insulation faults. The resistive load is connected

to a 230 V supply in a TT system. The installation is protected with an RCD. The user can create an insulation fault between the live wire and the conductive part of the frame by activating the ‘insulation fault’ button. Likewise, the contact of a person with the frame can be enabled/disabled through the ‘person in contact’ button. Also, the ‘earth connection’ button allows the user to connect/disconnect the frame from the ground. The resistance of the person is represented by a 1000 Ω resistor.

The RCD breaker rating is 30 mA and it trips when there is a difference exceeding that amount in the live and neutral line currents. The bypass circuit breaker is used to prevent the RCD from tripping while measurements are being carried out in the presence of an insulation fault. A USB-8 channel relay controller from KMTRONIC is used to control the switching of the motorized DX3 type circuit breakers and other electrical contacts.

### 3.2 Measurements

The voltage on the frame of the device and the current to earth are measured by two multimeters as shown in Fig. 2.



**Fig. 2.** Measurements of leakage current and frame voltage

The two multimeters together with the electrical board can be viewed from the webcam (Logitech C920). Hence, the user can see the tripping of the circuit breakers



through the web camera as well as record the values viewed on the multimeters as shown in Fig. 3. To program the web camera in LabVIEW, Vision acquisition is selected from the functions palette. Figure 3 shows the front panel of the camera program in LabVIEW. The web cam is interfaced to the computer through a USB cable.

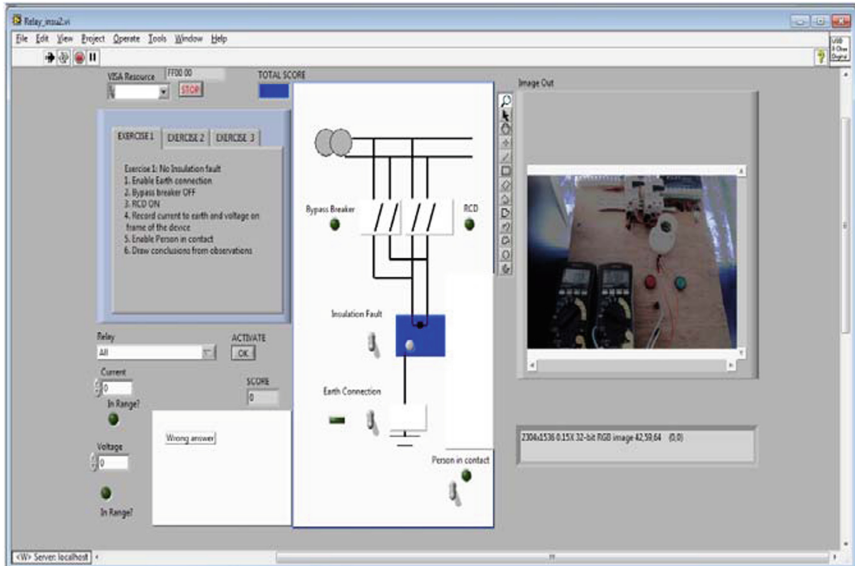


Fig. 3. Front panel.

### 3.3 Software Implementation

The program flow diagram is shown in Fig. 4. The user first selects one of the three exercises in order to undertake the experiments and has to follow the instructions in order to perform the experiment. In particular, he has to control the opening/closing of the circuit breakers and to enable/disable the connections regarding earth, insulation and a person in contact with the frame of the device. The bypass circuit breaker when activated is used to prevent the opening of the RCD when a fault is created so that the user can record the residual current and the corresponding voltage on the frame of the device.

Figure 3 gives a snapshot of the front panel through which the user interacts with the lab experiment. The user has to select the exercise that he intends to perform and he has to follow the instructions provided to undertake the experiment. The different components and connections are activated through the relay control. The current to earth and voltage across the frame of the device have to be read and inputted by the user. If the measurements have been incorrectly inputted the user will be prompted to reenter the correct values.

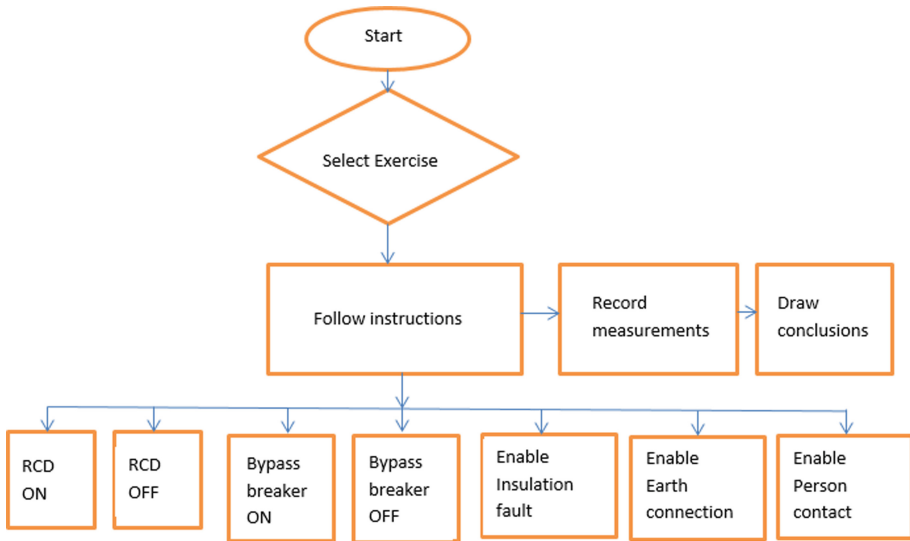


Fig. 4. Program flow diagram.

### 3.4 Learning and Teaching Activities Related to the Experiments

Each activity below provides students with prospects to develop their learning by applying concepts on earthing arrangement and practicing on virtual and real experiments. These activities also provide the lecturer feedback about the students' learning.

1. Introduction to LabVIEW environment and tools.
2. Lecture on earthing arrangement (6 h).
3. Students called to practice with the virtual experiments. This work has been described in a previous paper [16].
4. An evaluation is then conducted and used as pre-test.
5. The lecturer explains the main features of the remote lab experiments in a demo. The students are familiarized on the working environment.
6. An activity sheet is provided to the students describing the experiments and procedures.
7. An evaluation is conducted to know the usefulness and the value added by the remote labs as compared to the virtual lab (post-test).

## 4 The Experiments

The schematic displayed on the front panel is shown in Fig. 5.

There are three experiments to be performed by the user in order to understand the tripping of the RCD in the TT configuration. The student should have completed the virtual experiments on the TT configuration before undertaking the remote lab experiments [16]. To meet the learning outcomes mentioned in Sect. 2, three

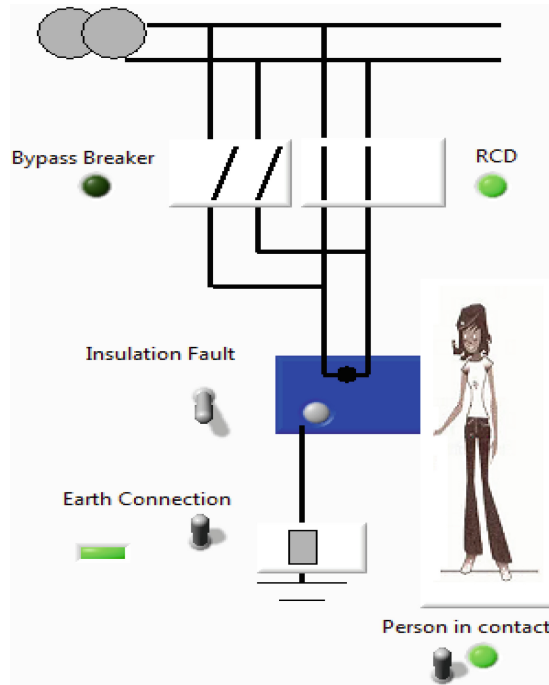


Fig. 5. Schematic of the user interface

experiments have been designed and the instructions on how to perform these experiments are explained in the following sections.

#### 4.1 Experiment 1: No Insulation Fault

At the end of this experiment the student should be able to understand the protection against electric shock under normal conditions. The student should understand the different types of basic protection required in an electrical installation in addition to the RCD. The instructions provided for undertaking this experiment are given below.

1. Enable Earth connection
2. Bypass breaker OFF
3. RCD ON
4. Record current to earth and voltage on frame of the device
5. Enable Person in contact
6. State how the protection against electric shock is ensured in these conditions through some basic methods in addition to the RCD
7. Draw conclusions from observations

#### 4.2 Experiment 2: Insulation Fault with EARTH CONNECTION

In this experiment the student investigates the importance of the earth connection in the protection of persons in the presence of an RCD and earthing. The student should be able to understand that a good connection to earth of the exposed conductive parts of the installation will lower the voltage on these parts to a safe level and that the RCD will trip once an insulation fault is detected when a leakage current flows to earth. The instructions for this experiment are given below.

1. Enable Earth connection
2. Bypass breaker ON
3. RCD OFF
4. Enable Insulation fault
5. 5. Record current to earth and voltage on frame of the device
6. Bypass breaker OFF
7. RCD ON
8. Draw conclusions from observations

#### 4.3 Experiment 3: Insulation Fault with NO EARTH CONNECTION

At the end of this experiment the student will understand that the inclusion of an RCD in the installation adds to safety in an electrical installation even when there is a breach of connection to earth. However, the RCD trips only when a person is in contact with the exposed conductive part of the installation. The leakage current required to trip the RCD passes through the body of the person to earth. If the RCD has been selected correctly the person will be safe. The instructions provided for this experiment are given below.

1. Disable Earth connection
2. Bypass breaker ON
3. RCD OFF
4. Enable insulation fault
5. Enable Person in contact
6. Record current to earth and voltage on frame of the device
7. 5. Bypass breaker OFF
8. RCD ON
9. Draw conclusions from observations

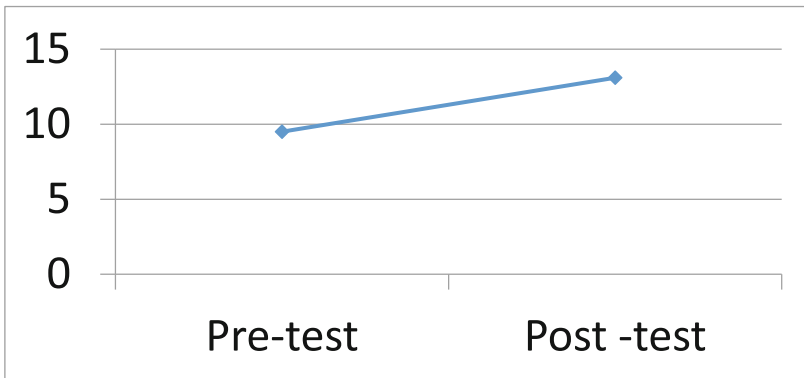
## 5 Results

Pre- and post-tests were conducted to gauge the student knowledge before and after the use of the web labs. Forty-eight participants were considered for the tests. All of the students were from year one and from the Electrical and Electronic engineering, Mechatronics, Civil and Mechanical Engineering programmes. The students were grouped into two per station. Each group was provided a worksheet with a hardcopy of

the three scenarios together with the procedure on the execution of each scenario. The pre-test was administered when the students had already some knowledge of the topic through the practice of virtual experiments. In particular, students had already conducted virtual investigations with the TT and TN earthing systems using LabVIEW [16, 18]. The post-test was targeted to measure the value added on learning after that the students have undertaken all the web labs. From the latter, information was also obtained on the appropriateness of the learning objectives and the improvements that can be brought to the instruction. It was found from the results that their understanding of the subject improved substantially as compared to the practice on virtual experiments [16, 18] as indicated by the results in Table 1 and Fig. 6.

**Table 1.** Pre- and post-tests results

	Pre-test	Post-test
Mean	9.5	13.14583
Variance	13.31915	8.510195
Observations	48	48
Hypothesized mean difference	0	
df	90	
t Stat	-5.40626	
P(T <= t) one-tail	2.62E-07	
t critical one-tail	1.661961	
P(T <= t) two-tail	5.24E-07	
t Critical two-tail	1.986675	



**Fig. 6.** Mean results of the pre and post tests

The paired t-test indicates that the mean difference between the pre-test and post-test is significantly different. It shows that when lectures and virtual experiments are supplemented by remote-laboratory work a complete learning experience is provided to the students. We have been able to provide the student with a learning

environment where he can carry out the experiments with earthing arrangements through internet and acquire the needed skills to develop his future jobs through the use of the remote Labs.

In general, the students responded positively to the use of the web-controlled lab. The main positive comments reported are as follows: The web-controlled lab save time in the sense that they do not have to be physically present in the lab. They do not need to setup the instruments and electrical connections and get them verified before starting the practical. The experiments are auto-didactic. The lessons can be learnt without the benefit of an instructor. The experiments do not need to be performed in a lab. They can practice at any time and at home. Many of them found that it was an easier and safer alternative. They can play around without the risk of being electrocuted. The experiments were a good aid in meeting the objectives mentioned in the curriculum. Even if they were not present in the lab they could still manipulate the instruments and get a feel of the real lab.

## 6 Conclusions

The availability of the internet to a wide number of households in Mauritius today has opened a new avenue in the dissemination of education in the country. Electrical installation is a major subject of Electrical Engineering which is taught during vacation training for year 1 students at the University of Mauritius. This module is followed by Electrical, Electronics/Communication, Civil, Mechatronics and Mechanical engineering students. Due to lack of equipment, space and staff, practical experimentation is limited. Additionally, taking into consideration the health and safety risk, students are not allowed to carry “hands-on” experimentation due to the danger of electric shocks. Besides safety risk to students, practical hands-on experimentation may also result in many unwanted device failures. In this context, we have implemented web based lab experiments in connection with the study of low voltage electrical installations using LabVIEW. In particular, a Web lab on the study of the TT earthing arrangement has been implemented and tested. The remote Web lab version has been implemented successfully. An assessment was conducted among forty-eight year one students selected among the BEng (Hons) Civil, Mechanical, Mechatronics, Electrical and Electronic Engineering program to evaluate the effectiveness of the web labs. The Web based labs were found to be effective as indicated from the pre-test and a post-test evaluation. It was found from the results that their understanding of the subject improved substantially. Since the execution of the experiments lasts shorter, students can book the station for use at a particular time. The Web based lab provides them with the opportunity to undertake the experiment at home rather than being physically present in the lab. Many of them found that it was an easier and safer alternative. They can play around without the risk of being electrocuted. The test bed can be further improved by including other types of earthing arrangement so that the students can get a broader picture of the different types of protections.

**Acknowledgment.** This work was supported and funded by the University of Mauritius.

## References

1. Central Statistics Office (CSO) Mauritius. <http://statsmauritius.govmu.org/English/StatsbySubj/Pages/ICT-Statistics-Year-2014.aspx>. Accessed June 2015
2. Buckman, B.A.: VI-based introductory electrical engineering laboratory course. *Int. J. Eng. Ed.* **16**(3), 212–217 (2000)
3. Rui, W., Rui, X., Hui, L.: Design and application of LabVIEW courseware in the signal processing teaching. In: *The 7th International Conference on Computer Science and Education*, Melbourne Australia, 14 July 2012
4. Kutlu, A., Taşdelen, K.: Remote electronic experiments using LabVIEW over controller area network. *Sci. Res. Essays* **5**(13), 1754 (2010)
5. Sancristobal, E., Martin, M.C.S., Tawfik, M., Pesquera, A., Gil, R., Diaz, G., Peire, J.: Remote labs as learning services in the education arena. In: *Global Engineering Education Conference (EDUCON 2011)*, p. 1189. IEEE (2011)
6. Khachadorian, S., et al.: A practical approach for applying online remote experiment: OnPReX. *Eur. J. Eng. Educ.* **36**(1), 21–34 (2011)
7. Orduna, P., Lopez-de-Ipina, J.G.-Z.D., Bailey, P.H., Hardison, J.L., DeLong, K., Harward, V.J.: Sharing laboratories across different remote laboratory systems. In: *12th International Conference on Advanced Learning Technologies*. IEEE (2012)
8. Sancristobal, E., Gil, S.M.R., Orduna, P., Tawfik, M., Pesquera, A., Diaz, G., Colmenar, A., Garcia-Zubia, J., Castro, M.: State of art, initiatives and new challenges for virtual and remote labs. In: *IEEE 12th International Conference on Advanced Learning Technologies* (2012)
9. Salzmann, C., Gillet, D., Huguenin, P.: Introduction to real-time control using LabVIEW with an application to distance learning. *Int. J. Eng. Educ., Special Issue: LabVIEW Application in Engineering Education* **16**(3), 255–272 (2000)
10. Oldenburg, J.T.: Using national instruments LabVIEW software in an introductory electronics course – MOSFET transistor parameter estimation and bias circuit design. In: *Proceedings of the 2004 American Society for Engineering Education Pacific Southwest Section Conference*. University of the Pacific, Stockton, 1–2 April 2004
11. Tricker, R.L.: *Wiring Regulations in Brief*, pp. 61–69. Elsevier Ltd., Oxford (2007)
12. Schneider Electric: Protection against electric shock. <http://www.electricalinstallation.org/enw/images/5/53/F-Protection-against-electric-shocks.pdf>. Accessed June 2015
13. Institution of Engineering and Technology: Requirements for Electrical Installations: BS 7671:2008 Incorporating Amendment No 1: 2011: IET Wiring Regulations, 17 edn
14. IEC TC/SC 23E: IEC 60898-1. Electrical accessories - Circuit-breakers for overcurrent protection for household and similar installations – Part 1: Circuit-breakers for a.c. operation. Ed. 1.2 b (2003)
15. University of Mauritius Annual Report. [http://www.uom.ac.mu/images/Files/AnnualReports/AnnualReport\\_2012.pdf](http://www.uom.ac.mu/images/Files/AnnualReports/AnnualReport_2012.pdf). Accessed June 2015
16. Rajkumarsingh, B., Goolaup, S., Galleegadoo, A.: Using LabVIEW software in an introductory residual current device course. In: *Proceedings of International Conference on Computer Education and Instructional Technology, ICCEI 2012*, pp. 1649–1654. WASET, Paris (2012)
17. Jabbar Khan, R.A., Mohammed, A., Junaid, M., Masood, M.A., Iftkhar, A.: LabVIEW based electrical machines laboratory for engineering education. *WSEAS Trans. Adv. Eng. Education.* **7**(5), 161–171 (2010)

18. Rajkumarsingh, B., Goolaup, S.: Modeling an earthing system using labview. In: Proceedings of IEEE Educon 2013, Berlin, Germany, pp. 763–768 (2013)
19. Poorna Chandra, B.R., Geevarghese, K.P., Gangadharan, K.V.: Design and implementation of Remote Mechatronics Laboratory for e-Learning using LabVIEW and Smartphone and Cross-Platform Communication Toolkit (SCCT). *Procedia Technol.* **14**, 108–115 (2014). 2nd International Conference on Innovations in Automation and Mechatronics Engineering, ICIAME 2014



# Quantifying the Pumping Energy Loss Associated with Different Types of Leak in a Piping System

L. Latchoomun<sup>1</sup>(✉), D. Mawooa<sup>1</sup>, Robert T.F. Ah King<sup>2</sup>,  
K. Busawon<sup>3</sup>, and R. Binns<sup>3</sup>

<sup>1</sup> Department of Electromechanical and Automation Engineering,  
Université des Mascareignes, Rose Hill, Mauritius  
{nlatchoomun, dmawooa}@udm.ac.mu

<sup>2</sup> Department of Electrical and Electronic Engineering,  
University of Mauritius, Réduit, Mauritius  
r.ahking@uom.ac.mu

<sup>3</sup> Department of Mathematics, Physics and Electrical Engineering,  
University of Northumbria, Newcastle upon Tyne, UK  
{krishna.busawon, richard.binns}@northumbria.ac.uk

**Abstract.** In this paper, the impact of different types of leaks on the excess pumping energy required in a distribution pipe is investigated. It is now well established that leakage is directly proportional to the pressure at the leak point in a pipe such that a pressure drop is inevitable. In order to compensate this drop in pressure, the pumping energy required to supply demand at a constant pressure has to be increased. How much more energy is required depends on the type and extent of the leak. For the 3 types of simulated leaks namely: orifice, circumferential and longitudinal slits with the same area of discharge, it is found that the excess pumping energy required for a circumferential crack in the pipe is the highest, followed by the orifice and then longitudinal slit as a result of the discharge to sustain customer demand and pressure head. The results of this research can be used to design a more robust control system or pump rescheduling strategy in order to save energy and water with regard to design, operation, and rehabilitation of old water distribution system.

**Keywords:** Polyvinyl chloride · Leakage exponent · Variable speed drive · Non-revenue water

## 1 Introduction

The requirement of pumping energy for airlifting water from boreholes and distribution to flat or upper lands in Mauritius is very significant since nearly fifty percent of resources are derived from ground water sources. Unfortunately, the present level of unaccounted-for water in the water distribution network which varies between 45 to 55%, causes much energy loss as can be imagined. In addition, the stress of an increasing population, the changing climate and the emergence of new industries are simply exacerbating the situation by putting much stress on the old existing

infrastructure of water distribution. The current water management practices are not robust enough to cope with the situation since the problem of leakage is persistent and the replacement of existing pipes is not only costly but a long term investment. The impact of leakage on pumping is also a matter of great concern in developed countries like the United States as pointed out by American Water Works Association (AWWA). It is estimated that 5 to 10 billion kWh of power generated in the US is spent on NRW annually (AWWA 2003). Besides the extra electrical energy required for pumping, losses in terms of the chlorination and pretreatment should also be accounted for. Colombo and Karney [1] found that there is a direct relationship between the pumping cost and leaks when the same service at demand points is maintained. They also developed an analytical approach to the energy efficiency of centrifugal pumps related to leak size and its location on a pipe segment. The dynamics of centrifugal pumps are very sensitive to demand pressure which in turn depends on the dynamics of demand, the head loss in the pipe due to age material wear and tear, and the topography of land among other things. In this research paper, a correlation is found between the excess pumping energy required and the three different types of leaks that are prevalent in a distribution pipe namely: a hole, a longitudinal crack and a circumferential crack.

## 1.1 Background

### 1.1.1 Pump Dynamics

With reference to Fig. 1, in most centrifugal pumps, as pressure decreases, the performance point moves out to the right of the curve while flow increases and as pressure increases, the flow decreases shifting the performance point to the left of the curve. This is what happens on the H-Q curve of the pump when a valve connected to its output is throttled. However, if this exercise is not carried out carefully, the risk of getting out of the best efficiency point zone (BEP) is quite real. The hydraulic power produced by the centrifugal pump at point 2 is given by the shaded area under the H-Q curve. On the other side, if a variable speed drive attached to the motor is used to shift from flow  $Q_1$  to  $Q_2$  on speed curve  $n_5$  of Fig. 2, less head  $H_2$  is imparted to the water thereby reducing the hydraulic power produced and most importantly, the pump performs in the BEP zone. Based on the affinity law (Eq. 1 on page 3), the governing equations regarding the flow, pressure head and speed can be used assuming that the pump efficiency stays constant at different speeds:

*Affinity Law for pumps*

$$\frac{Q_1}{Q_2} = \frac{n_1}{n_2} \quad \frac{H_1}{H_2} = \left(\frac{n_1}{n_2}\right)^2 \quad \frac{MP_1}{MP_2} = \left(\frac{n_1}{n_2}\right)^3 \quad (1)$$

where  $Q_1$ ,  $H_1$ ,  $MP_1$  and  $n_1$  are the flow, head, mechanical power and speed respectively at point 1 and  $Q_2$ ,  $H_2$ ,  $MP_2$  and  $n_2$  are the flow, head, mechanical power and speed respectively at point 2.

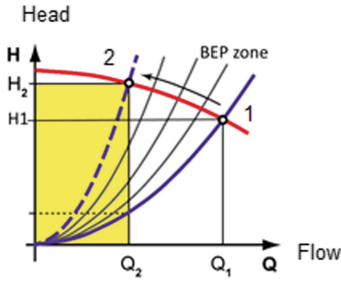


Fig. 1. Fixed speed H-Q curve

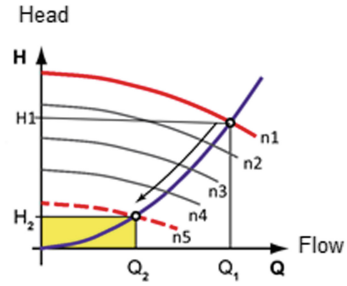


Fig. 2. Variable speed H-Q curve

### 1.1.2 Leak Dynamics

The rate of leakage  $Q_l$  is proportional to the square root of the static head  $P$  in a pipe according to Torricelli's theorem,

$$Q_l = C_1 A_1 \sqrt{2gP} \tag{2}$$

where,

$C_1$  = the discharge coefficient

$A_1$  = leak area

$G$  = acceleration due to gravity

$P$  = total pressure head at the point of leak.

This empirical formula was confirmed by several researchers namely Hikki [2], Greveinstein and Van Zyl [3], irrespective of the pipe material and size of hole. Nevertheless, leaks are not always of orifice type and therefore other shapes like circumferential and longitudinal cracks were investigated. Lambert [4] proposed a more general form of the leak equation:

$$Q_l = C \cdot P^N \tag{3}$$

where

$C$  is the leakage coefficient

$N$  is the leakage exponent.

Extensive research works have been performed with leaks of different shapes in pipes. For example Van Zyl and Clayton [5] investigated four factors that may explain the high values of  $N$  in pipes, namely leak hydraulics, soil hydraulics, pipe material behaviour and water demand. The third factor was considered the most important one suggesting that the leak area increases with increase in pressure. For flexible pipes like

PVC, the relationship is complicated by plastic deformation and hysteresis as investigated by Ferrante et al. [6] and Massari et al. [7]. Cassa and Van Zyl [8] used finite element analysis to show that all types of leak vary linearly under pressure through elastic deformation under the FAVAD concept:

$$A(h) = m \cdot P + A_0 \quad (4)$$

where,

$A$  = the new leak area and  $A_0$  = initial leak area

$P$  = pressure head at leak point

$m$  = slope of the head-area curve.

Rewriting the leak flow Eq. 4 according to FAVAD (Fixed and Variable Area Discharges):

$$Q_l = C_d \sqrt{2g} (A_0 P^{0.5} + m \cdot P^{1.5}) \quad (5)$$

where,

$A_0$  is the initial leak area

$m$  is the slope of the head-area curve.

According to Van Zyl et al. [9], for laminar flow  $N$  varies between 0.5 and 1 whereas for turbulent flow,  $N$  is less than 0.5. Greyvenstein [3] found that  $N$  may vary between 1.38 and 1.8 in uPVC pipes with longitudinal cracks and between 0.41 and 0.53 for circumferential cracks. On the other hand, M. Franchini and L. Lanza [10] showed that the leakage through pipes of different elastic materials can be represented by the classical Torricelli's equation by using a correction factor to cater for the variable leak area and hydraulic factors affecting the coefficient of discharge. He stated that the overwhelmingly high values of  $N$  obtained in laboratory can be attributed to the improper quantification of the leakage coefficient.

In this work, the maximum pressure of the pump has been limited to 1.5 bars in a uPVC pipe of diameter 32 mm and thickness 4 mm such that an expansion of the area of the openings is negligible as pointed out by Latchoomun et al. [11] for a pressure below 2 bars where the mean leakage exponent  $N$  for orifice, longitudinal slit and circumferential slit in uPVC was found experimentally to be about 0.51 (gradient of line) as shown in Fig. 3 below. Hence the value of  $m$  in Eq. 5 above may be neglected.

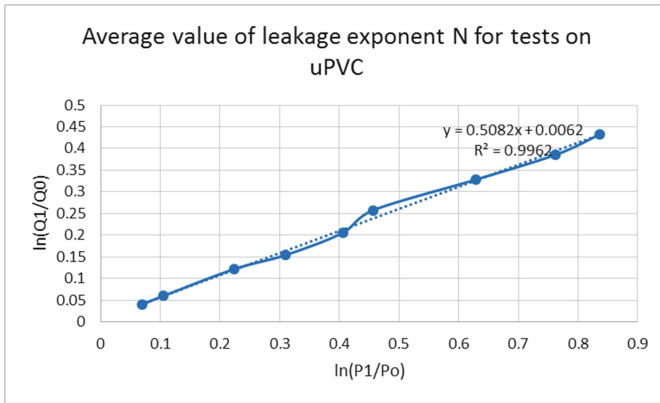


Fig. 3. Average leak exponent value N in uPVC

## 2 Methodology

First of all, care is taken to see that the flow in the channel is not laminar by calculating the Reynold’s number at different flows and pressures within the leaking pipe using the equation:

$$Re = \frac{VD\rho}{\mu} = \frac{VD}{\nu} \tag{6}$$

where

$Re$  = Reynolds number,  $V$  = average fluid velocity,  $D$  = pipe diameter,  $\rho$  = fluid density,  $\mu$  = absolute viscosity,  $\nu$  = kinematic viscosity.

The values of  $Re$  found in this experiment vary between 4761 and 21927 suggesting that the flow is turbulent and therefore Eq. 3 can be used safely. Now, if the laws of conservation of energy are applied to the fluid flow through a section AB of the smooth PVC pipe (Fig. 4) typically of a small length (3 m) with a small leak at the orifice, the head loss due to the roughness of the pipe and sensors may be neglected such that:

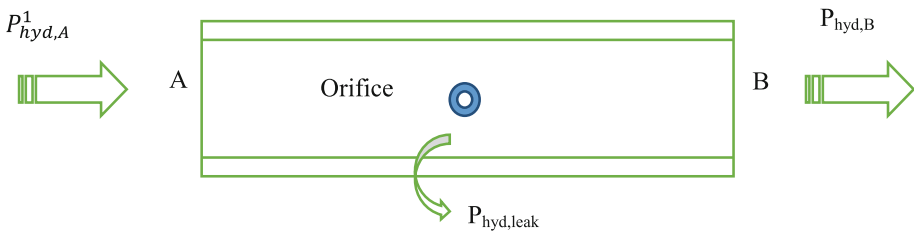


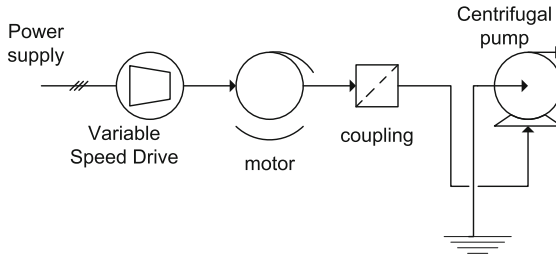
Fig. 4. Conservation of energy in a leaking pipe

$$P_{hyd,A}^1 = P_{hyd,leak} + FL_{leak} + P_{hyd,B} \tag{7}$$

Where  $P_{hyd}$  is the hydraulic power of the fluid at the point of interest and FL is the friction loss at leak point.

The only significant loss is due to friction at the emitter whose coefficient of discharge C in Eq. 3, is highly dependent on its shape and area. This loss can be quantified if the hydraulic power of the fluid at point A, B and the leak are computed.

Typically the output hydraulic power from a pump is given by:



**Fig. 5.** Pump drive train

$$P_{hyd} = \rho gQH \tag{8}$$

where  $\rho$  = density of water,  $g$  = gravitational acceleration,  $H$  = pressure head and  $Q$  is the flow.

If the pumping system is attached to a variable speed drive, an induction motor, the coupling and the centrifugal pump as shown above in Fig. 5, then the overall pumping system efficiency,  $\eta_s$  is given by:

$$\eta_s = \eta_v \cdot \eta_m \cdot \eta_c \cdot \eta_p \tag{9}$$

where

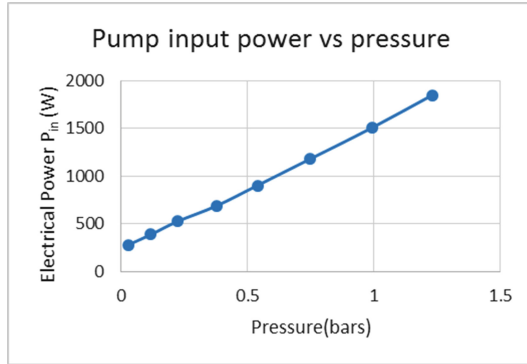
$\eta_v$  = efficiency of VSD,  $\eta_m$  = efficiency of motor,

$\eta_c$  = coupling efficiency,  $\eta_p$  = efficiency of pump

Therefore,

$$P_{hyd} = \rho gQH = \eta_s P_{in} \tag{10}$$

where  $P_{in}$  = input electrical power



**Fig. 6.** Pumping energy requirement at different pressures

Hence at point A which is the output of pump,

$$P_{in} = \frac{P_{hyd,A}}{\eta_s} \tag{11}$$

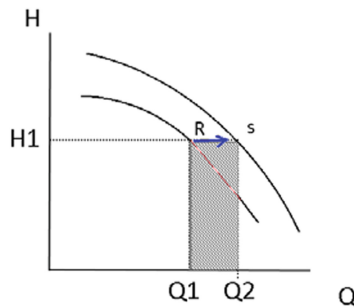
This relationship was verified experimentally by the graph of Fig. 6 below.

For the test first performed without leak, neglecting the power loss due to pipe roughness,

$$P_{hyd,A} = P_{hyd,B} \tag{12}$$

When leak is introduced maintaining the same demand flow and pressure at point B (same  $P_{hyd,B}$ ) through an increase in speed of the VSD as depicted in Fig. 7, there is a shift from point R to S on the Q-H curve. The governing power equation is now given by Eq. 7. Finally, the change in hydraulic power at A is:

$$\Delta P_{hyd,A} = P_{hyd,A}^1 - P_{hyd,A} = P_{hyd,leak} + FL_{leak} \tag{13}$$



**Fig. 7.** Speed variation using VSD on H-Q curve

Now, if it is assumed that the efficiency  $\eta_s$  of the system remains constant with change in speed, then from Eq. 11,

$$\Delta P_{in} \propto \Delta P_{hyd,A} \tag{14}$$

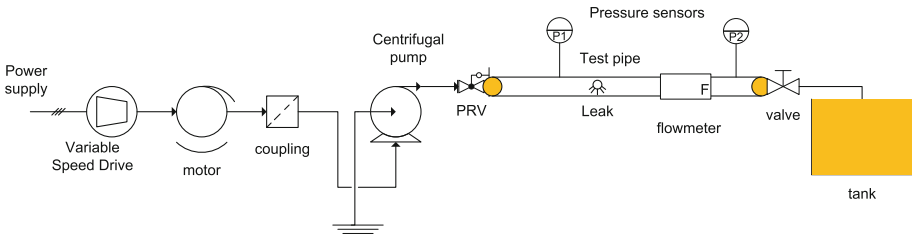
The shaded area under the H-Q curve of Fig. 7 is thus proportional to the electrical power loss in pumping i.e. the excess input pumping power is consumed as the hydraulic power of the leaking water and friction loss in the emitter such that:

$$\Delta P_{in} \propto (P_{hyd,leak} + FL_{leak}) \tag{15}$$

Equation 15 can thus be used to gauge the amount of additional pumping energy required for the different types of leak.

### 3 Laboratory Setup

As shown in Fig. 8 on next page, the pump drive train consists of a variable speed drive coupled motor, a reducer as coupling and the centrifugal pump. At the entrance, a pressure reducing valve (PRV) is connected to the sample test pipe which is a 3 m long uPVC pipe of 32 mm diameter. The leak is perforated midway and a pressure sensor  $p1$  is embedded in the pipe at a distance of about 0.5 m from it to avoid transients in the recorded data. The leak flow is measured by a separate setup whereas the flow and pressure at the demand side are recorded by  $F$  and  $p2$  respectively. A data recorder collects all the sensor measurements while an energy analyzer connected to the VSD logs the electrical power consumption of the pumping system. To be able to compare the performance of the 3 different types of leak, a fixed area of discharge of  $8 \text{ mm}^2$  is engineered in the test pipe such that an orifice of approximately 3 mm diameter, a longitudinal slit of  $2 \times 4 \text{ mm}^2$  and a circumferential slit of  $4 \times 2 \text{ mm}^2$  are obtained respectively.



**Fig. 8.** Laboratory setup of pumping system with VSD



## 4 Results of Experiment

As mentioned earlier, the first test is performed without leak for a demand  $Q_d$  of 6.5 L/min and a pressure of about 1 bar at point B. The input power is logged through the energy analyzer and then the test is carried out again under the same constraints of flow and pressure at the demand side with different types of leak in the pipe one at a time. The leak rate is also measured such that the hydraulic power associated with the leaking fluid can be calculated. Another interesting point is to calculate the pressure drop  $\Delta p$  caused by the leak as shown in Tables 1 and 2.

**Table 1.** Measured values for different types of leakage

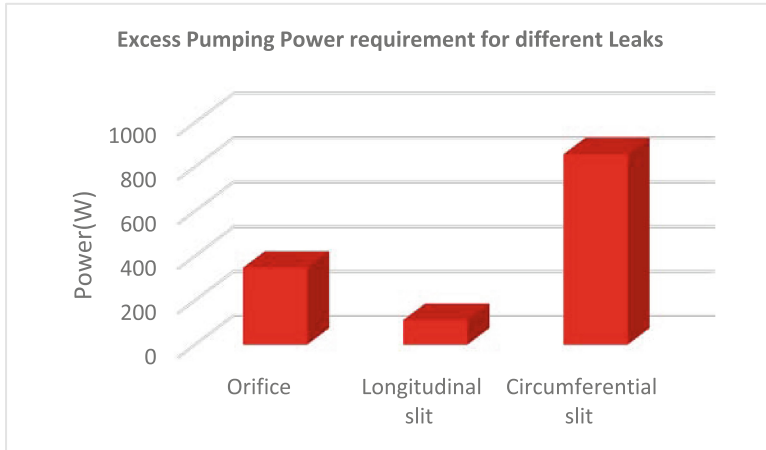
	p1 (bars)	p2 (bars)	$\Delta p$ (bars)	$Q_{leak}$ (L/min)	$P_{in}$ (leak)(W)	$P_{in}$ (no leak)(W)	$\Delta P_{in}$ (W)
Orifice (8 mm <sup>2</sup> )	1.259	1.219	0.040	3.67	1350.7	1004	346.7
Longitudinal slit (4 × 2 mm <sup>2</sup> )	1.118	1.002	0.116	5.29	1114.5	1004	110.5
Circumferential Slit (2 × 4 mm <sup>2</sup> )	1.255	1.097	0.158	3.87	1860	1004	856

**Table 2.** Computed energy values for different types of leakage

	$\Delta P$ (bars)	$P'_{hyd,A}$ (W)	$P_{hyd, B}$ (W)	$P_{hyd,Leak}$ (W)	$F \cdot L \cdot (W) = (P'_{hyd,A} - P_{hyd,B}) - P_{hyd,leak}$
Orifice (8 mm <sup>2</sup> )	0.04	21233.5	12948.5	7862.2	422.8
Longitudinal slit (4 × 2 mm <sup>2</sup> )	0.116	22357.7	11027.2	10053.9	1276.6
Circumferential slit (2 × 4 mm <sup>2</sup> )	0.158	22509.1	12465.5	8248.2	1795.4

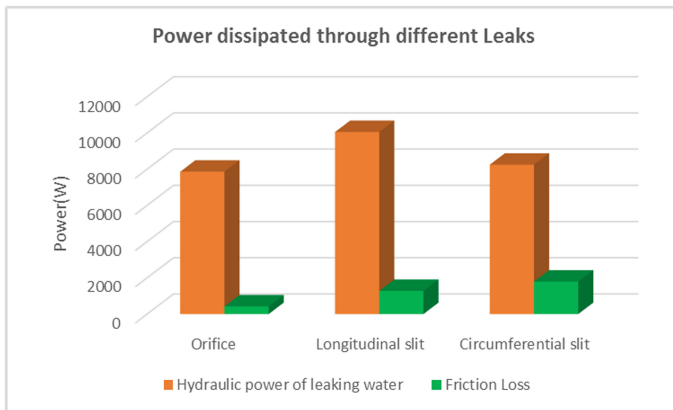
## 5 Discussions

For the same area of discharge (8 mm<sup>2</sup>), the circumferential slit in uPVC requires the highest excess pumping energy ( $\Delta P_{in} = 856$  W) in order to meet the demand constraints of flow and pressure whereas for the longitudinal slit, it is lowest (110.5 W) or about 7 folds smaller compared to the circumferential as depicted by Fig. 9. This can be explained by the high frictional losses associated with a high pressure drop ( $\Delta p$ ) of 0.158 bars for the vertical slit.



**Fig. 9.** Excess pumping power through leaks

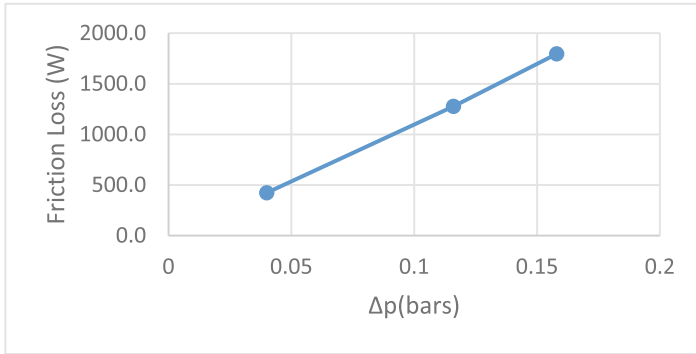
The latter can further be justified by the fact that only a surface of 2 mm is exposed to the direction of water flow in the longitudinal slit compared to 4 mm in the circumferential one, meaning that the stream of water meets much more resistance as it bursts out of the circumferential slit (Fig. 10).



**Fig. 10.** Hydraulic and frictional power loss in leaks

Another interesting fact is that the energy wasted in terms of hydraulic power,  $P_{hyd,leak}$  is biggest for the longitudinal slit due to its high leak flow rate. A longer width of 4 mm exposed in the direction of flow undoubtedly accounts for that.

Furthermore, there is a direct relationship between pressure drop and friction loss at the leak point as shown in Fig. 11. From this result one can predict that the leakage coefficient  $C$  (Eq. 3) for the circumferential slit would be the highest.



**Fig. 11.** Friction losses against pressure drop in the different leaks

Finally, although orifices tend to consume less pumping electrical energy compared to circumferential slits, they are more prevalent in old aged pipes, especially ductile iron.

## 6 Conclusion

It is thus clear that circumferential cracks within leaky pipes require the greatest amount of excess pumping energy compared to orifices and longitudinal cracks. This is mainly caused by a high friction loss at the point of leakage. On the other hand, it is found that for the same area of discharge, longitudinal cracks are subjected to the highest hydraulic energy waste due to a much bigger flow rate at the leak for a fixed demand at a constant pressure. Unfortunately, the relationship between the excess pumping energy and the change in hydraulic power of the fluid due to leakage is not straightforward and Eq. 15 could not be verified. Further investigation is required in order to quantify the wasted power due to leakage.

Quantifying how much energy is wasted on a leaky pipe is a very valuable information that can be used by management to fix its long term goals. Is it worth investing in a new piping system or rather put in place a strategy to control leakage and reduce the associated cost? As a further work, we can develop a Life Cycle Cost analysis to optimize resources. The results of this research work can help in this endeavour.

## References

1. Colombo, A.F., Karney, B.W.: Impacts of leaks on energy consumption in pumped systems with storage. *J. Water Resour. Plan. Manag.* **131**(2), 146–155 (2009)
2. Hikki, S.: Relationship between leakage and water. *J. Jpn. Waterworks Assoc.* **5**, 50 (1981)

3. Greyvenstein, B., Van Zyl, J.: An experimental investigation into the pressure-leakage relationship of some failed water pipes. *Aqua J. Water Supply Res. Technol.* **56**(2), 117–124 (2007)
4. Lambert, A.: What do we know about pressure: leakage relationships in distribution systems? In: *System Approach to Leakage Control and Water Distribution Systems Management*, IWA, Brno, Czech Republic (2000)
5. Van Zyl, J.E., Clayton, C.R.I.: The effect of pressure on leakage in water distribution systems. In: *Proceedings of CCWI2005 Water Management for the 21<sup>st</sup> Century*, University of Exeter, UK (2005)
6. Ferrante, M., Massari, C., Todini, E., Brunone, B., Meniconi, S.: Experimental investigation of leak hydraulics. *J. Hydroinformatics* **15**(3), 666–675 (2012)
7. Massari, C., Ferrante, M., Brunone, B., Meniconi, S.: Is the leak head–discharge relationship in polyethylene pipes a bijective function? *J. Hydraul. Res. IAHR* **50**(4), 409–417 (2012)
8. Cassa, A.M., Van Zyl, J.E.: Predicting the head-area slopes and leakage exponents of cracks in pipes. Exeter, Centre of Water Systems, University of Exeter, College of Engineering Mathematics and Physical Science, pp. 485–490 (2011)
9. Van Zyl J.E., Cassa, A.: Linking the power and FAVAD equation for modelling the effect of pressure on leakage. In: *Eleventh International Conference on Computing and Control for Water Industry (CCWI)*, University of Exeter, United Kingdom (2011)
10. Franchini, M., Lanza, L.: Use of Torricelli’s equation for describing leakages in pipes of different elastic materials, diameters and orifice shape and dimensions. In: *Proceedings of the 16<sup>th</sup> Conference on Water Distribution Systems Analysis, WADSA 2014* (2014)
11. Latchoomun, L., Ah King, R.T.F., Busawon, K., Mawooa, D., Kaully, R.G.: Laboratory investigation of the leakage characteristics of unburied HDPE pipes. In: *CCWI2015, Procedia Engineering*, vol. 119, pp. 91–100. doi:[10.1016/j.proeng.2015.08.858](https://doi.org/10.1016/j.proeng.2015.08.858)

# A Power Flow Control Scheme for a Photovoltaic to a Low Voltage Microgrid System

S.Z. Sayed Hassen<sup>(✉)</sup>, M.I. Jahmeerbacus,  
K. Sewraj, and M.S. Ruhomaun

Department of Electrical and Electronic Engineering,  
University of Mauritius, Reduit, Mauritius  
{z.sayedhassen, iqbal}@uom.ac.mu,  
keshav.sewraj@umail.uom.ac.mu, sameer070892@gmail.com

**Abstract.** The increasing penetration of renewable energy sources into the ac grid has brought several challenges for the power utilities, especially in the fields of power quality, stability, reliability, protection and control. In this paper, we present the design and implementation of a power flow control scheme for a photovoltaic to low voltage ac grid interface. The control algorithm is based on instantaneous tracking of the power reference signal by current regulation of the grid connected inverter. The proposed control and graphical user interface are implemented using LabVIEW. The ability of the developed system to control both the real and reactive powers injected into the grid is validated using both computer simulations and experimental tests.

**Keywords:** Photovoltaic · Microgrid · Power control

## 1 Introduction

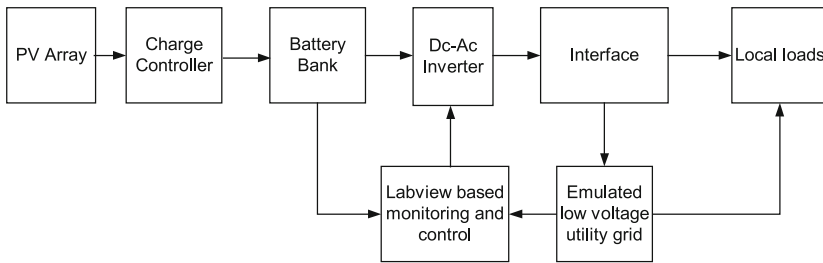
Distributed Generation (DG) typically includes photovoltaic panels, wind turbines and waste-to-energy sources, among others. Although renewable energy sources are environmental friendly compared to fossil fuels, they also come with a set of challenges for the power utilities, namely in the areas of power quality, harmonics, stability, protection and control. Moreover, massive deployment of distributed generation may affect key grid parameters such as line frequency and voltage, and calls for a coordinated approach for integrating DG in the network.

Microgrids can operate in both grid-connected and islanded mode [1], where generators and loads are usually interconnected at low voltage. On the utility side, a connected microgrid can be controlled as if it were one entity. Islanded microgrids however have characteristics which are quite different, compared to those of the conventional power system. In grids powered by synchronous generators, unbalances between the generated power and electrical power consumption are compensated by the inertia of the rotating system, resulting in a change of frequency. Non-mechanical energy sources, such as photovoltaic arrays, are however interfaced to the microgrid through power electronic converters [2, 3]. The control of both active and reactive powers by such converters enables better frequency and voltage regulation.

In this paper, we propose the design and implementation of a power flow control scheme between a photovoltaic array with battery energy storage and a low voltage microgrid. The control algorithms and graphical user interfaces are developed using LabVIEW to regulate power transfer from the dc side to the ac grid. The paper is organized as follows: Sect. 2 describes the main stages and modeling of the microgrid system, including the photovoltaic panel, and the power flow control from battery to grid through an inverter. Section 3 presents the main simulation and experimental results for power flow control between the dc side and the ac grid, with injection of real and reactive power at varying phase. A typical charging performance of the energy storage battery is also demonstrated. Section 4 finally provides some concluding remarks.

## 2 Modeling, Control and Design of Proposed System

A block diagram describing the proposed setup is shown in Fig. 1. Variations in solar irradiance affect both the current-voltage characteristics and the output power of the photovoltaic (PV) array. The battery bank is charged from the PV array with maximum power tracking operation and provides energy storage at a stable terminal voltage. The dc to ac inverter converts the battery output dc voltage to ac while matching the power and frequency requirements of the microgrid.



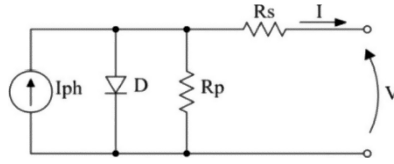
**Fig. 1.** Block diagram of proposed system

### 2.1 Modeling and Characteristics of PV Array

The equivalent circuit model of a PV cell is shown in Fig. 2, where  $I_{ph}$  and  $D$  are an ideal current source and diode, respectively. The non-ideal cell characteristics are represented by the insertion of resistors  $R_s$  and  $R_p$  [4]. The generated current ( $I_{ph}$ ) depends on the solar irradiance ( $S$ ) and the temperature  $T$  so that:

$$I_{ph} = (I_{sc} + k_i(T - T_r)) \left( \frac{S}{1000} \right) \quad (1)$$

where  $I_{sc}$  is a reference short-circuit current,  $T_r$  a reference temperature, and  $k_i$  the short-circuit temperature coefficient.



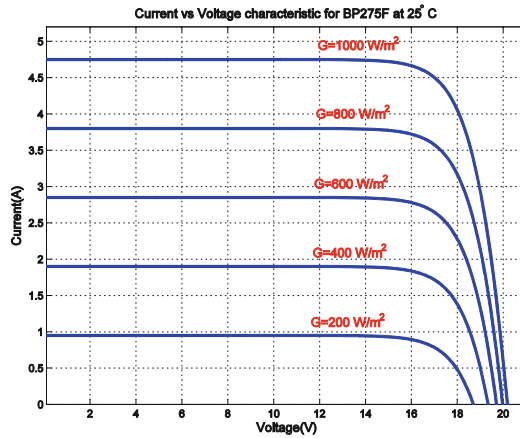
**Fig. 2.** Equivalent model of a PV cell [4]

The PV panel simulation model is represented by

$$I = I_{ph} - I_r \left[ e^{\frac{q(V+IR_s)}{\eta kT}} - 1 \right] - \frac{V + IR_s}{R_p} \tag{2}$$

where  $V$  is the output voltage of one PV panel,  $I_{ph}$  is the photocurrent,  $I_r$  is the saturation current,  $q$  is the electronic charge,  $\eta$  is the p-n junction quality factor,  $k$  is the Boltzmann’s constant and  $T$  is the panel temperature.

A 75 W BP275F PV module [5, 6] is used to implement the photovoltaic source. The current-voltage and the power-voltage characteristics of the PV module are generated using Eqs. (1) and (2), as shown in Figs. 3 and 4, respectively. The PV panel is specified at an operating temperature of 25 °C. However, due to panel exposure to direct sunlight and dusty environment for extended times, the actual power output is to be adjusted according to the ambient temperature and dust reduction factors [8, 9], respectively.



**Fig. 3.** Current-voltage characteristics of PV panel

The charging process of the lead-acid battery bank consists of a bulk stage at constant current and rising voltage, an absorption stage at constant voltage with falling current, and a float stage. As solar irradiance varies, the charge controller adjusts the instantaneous voltage and current to the battery, according to the charging mode.

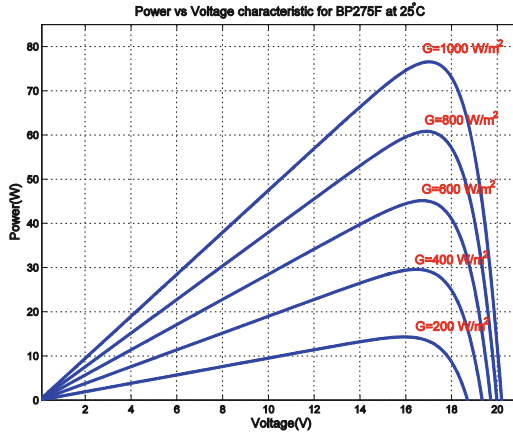


Fig. 4. Power-voltage characteristics with varying solar irradiance

### 2.2 Interface Between Dc Side and Low Voltage Ac Grid

The power conversion stage from the dc to low voltage grid interface is shown in Fig. 5. At a given time instant, power can be transferred from battery to the grid, or vice-versa. However, the cycle average real power transferred to the grid is positive. Transformer T provides electrical isolation between the single phase mains and the inverter output, and also steps down the mains voltage to  $v_g$ .

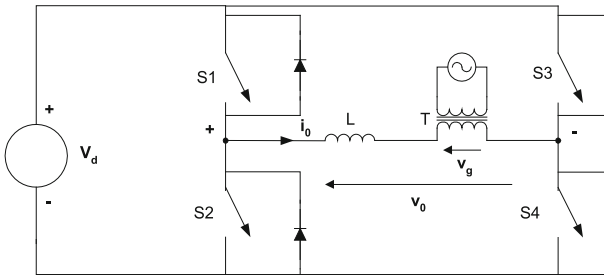


Fig. 5. Dc to low-voltage ac side interface through switch-mode inverter

The power semiconductor switches, S1 to S4, are controlled to track a sinusoidal reference current at a predefined phase angle, and to deliver the target real power to the grid. Since the inverter switches are operated in either the on or off states, the output voltage and current contain harmonics at the switching frequency and its multiples. The harmonic content in the output current therefore needs to be minimized to improve the power quality of the inverter to grid interface. Tracking of sinusoidal current is achieved by a hysteresis band current controller. The ac side inductor caters for the voltage difference between the inverter output and the grid voltage and also helps to filter out the high frequency harmonics in the injected current.



### 2.3 Power Control into Ac Side

It is required to control the active and reactive components of the power between the inverter and the ac grid. Hence the PV-inverter combination can be configured to operate at unity power factor, or as a reactive power source with leading or lagging power factors.

The instantaneous power due to the fundamental components of the injected current ( $i_1$ ) and utility voltage ( $v_g$ ) at grid frequency ( $f_1$ ) can be expressed as

$$p(\theta) = v_g(\theta)i_1(\theta) \tag{3}$$

where:

$$v_g(\theta) = V_m \sin \theta \tag{4}$$

$$i_1(\theta) = I_m \sin(\theta - \phi) \tag{5}$$

$$\theta = 2\pi f_1 t \tag{6}$$

$\phi$  is the phase lag of the injected current with respect to the grid voltage.

From Eq. (3), the average real power is given by

$$P = S \cos \phi \tag{7}$$

where  $S$  is the apparent power and is expressed as:

$$S = \frac{V_m I_m}{2} \tag{8}$$

The power flow control scheme is represented in Fig. 6. The signal  $e(k)$  represents the error between the reference and actual powers at the  $k^{th}$  sampling instant. The Proportional + Integral (PI) controller output,  $u(k)$ , represents the amplitude of the current to be injected into the grid. This control signal is multiplied by a sinusoidal template and applied to a hysteresis band controller to generate the switching signals of the dc to ac inverter.

#### 2.3.1 Reference and Actual Instantaneous Powers, and PI Power Controller

The power to be injected into grid can be described by the apparent power command ( $S^*$ ), and the phase angle ( $\phi^*$ ) between injected current and grid voltage. From Eqs. (3) to (6), the instantaneous power reference can be expressed as:

$$p^*(k) = S^* [\cos \phi^*(k) - \cos(4\pi f_1 k T_s - \phi^*(k))] \tag{9}$$

where  $T_s$  is the sampling time period.

To obtain the actual instantaneous power injected, voltage and current sensors are incorporated in the ac line. After appropriate signal conditioning and sampling of the measured signals  $v_g$  and  $i_o$ , the instantaneous power is computed as:

$$p(k) = v_g(k)i_0(k) \tag{10}$$

The PI controller outputs the required amplitude of the grid-injected current, so that:

$$u(k) = K_e e(k) + (K_i T_s - K_e) \cdot e(k - 1) + u(k - 1) \tag{11}$$

where  $K_e$  and  $K_i$  are the proportional and integral gains respectively, and

$$e(k) = p^*(k) - p(k) \tag{12}$$

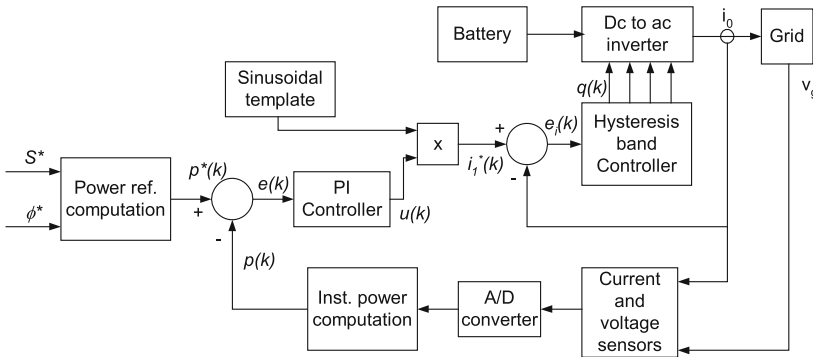


Fig. 6. Block diagram of proposed power controller

### 2.3.2 Hysteresis Band Current Controller

To adjust the real and apparent powers, the required phase shift of the inverter output current is described by Eq. (5). The reference current,  $i_1^*(k)$ , is obtained using the output signal from the PI controller. Hence:

$$i_1^*(k) = u(k) \sin[2\pi f_1 k T_s - \phi^*(k)] \tag{13}$$

where  $\phi^*$  is the phase shift command.

The hysteresis band controller generates switching signals  $q_1$  to  $q_4$  for the inverter power transistors S1 to S4, respectively, as shown in Fig. 7. Signal  $e_i(k)$  represents the error between the reference and actual line currents, and is given by:

$$e_i(k) = i^*(k) - i_0(k) \tag{14}$$

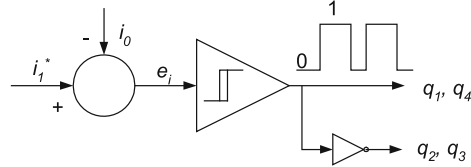
The hysteresis band controller output signals are given by:

$$q_1(k) = q_4(k) = \begin{cases} 1 & e_i(k) > h \\ 0 & e_i(k) < -h \\ q_1(k - 1) & \text{otherwise} \end{cases} \tag{15}$$

and

$$q_2(k) = q_3(k) = \bar{q}_1(k) \tag{16}$$

where  $h$  is the width of the hysteresis band.



**Fig. 7.** Generation of inverter switching signals

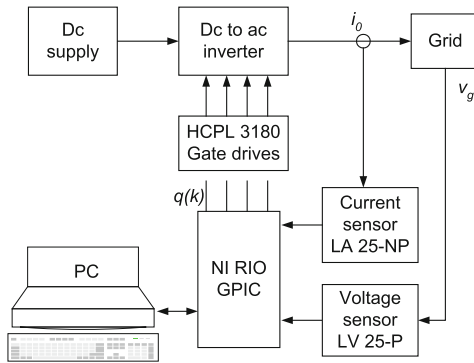
### 3 Simulation and Experimental Tests

To investigate the effectiveness of the proposed power control scheme, an experimental test bed with graphical user interface was implemented. The main parameters of the system under test are summarized in Table 1.

**Table 1.** Parameters of the battery to inverter to grid link

Parameter	Value
Injected power rating	20 VA
Power factor	0 (lagging/leading) to 1
Ac side voltage and frequency	7 V (rms), 50 Hz
Battery bank nominal voltage	24 V
Inverter output side inductance	5 mH

The main components of the test setup are shown in Fig. 8. A National Instruments NI RIO GPIC inverter controller card [6, 7] implements the control algorithm described in Sect. 2 and supplies the switching signals to the dc to ac inverter.



**Fig. 8.** Main components of the implemented system

The same board acquires the measured ac side signals through Hall-effect current and voltage sensors. The control software and graphical user interface are implemented with LabVIEW graphical programming for FPGA.

### 3.1 Battery Charging Profile

Figure 9 shows the charging profile of the battery over a test interval of 85 min. The bulk charging stage maintains a fairly constant current of about 2 A while the charging voltage rises gradually from 22 V to 26 V. In the absorption stage the charging current falls below 1.4 A, while the voltage remains between 25 V to 26 V.

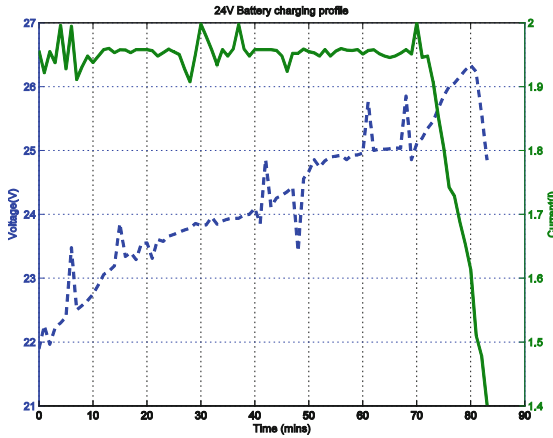


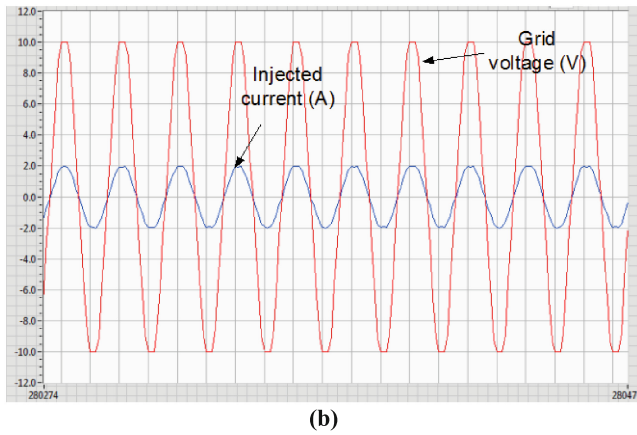
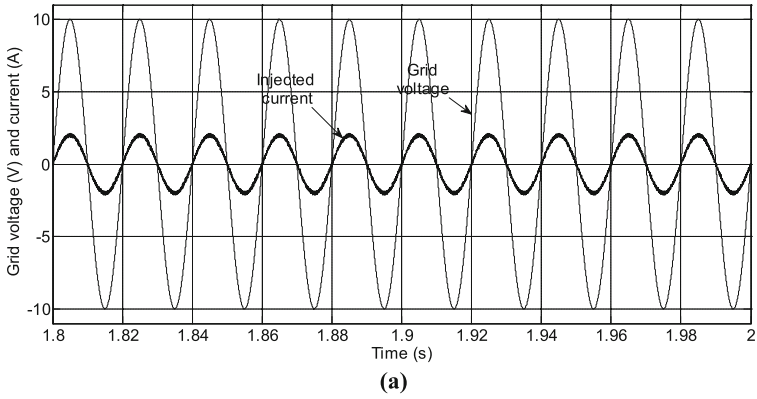
Fig. 9. Battery charging profile (solid line: current, dotted line: voltage)

### 3.2 Power Control from Inverter to Grid

The power control scheme described in Sect. 2 is validated through computer simulations and experimental results. The inverter and controller models are built using Simulink™ software, and tested for different power commands. The response plots are compared with the corresponding experimental results for power factors of unity and zero (leading), respectively.

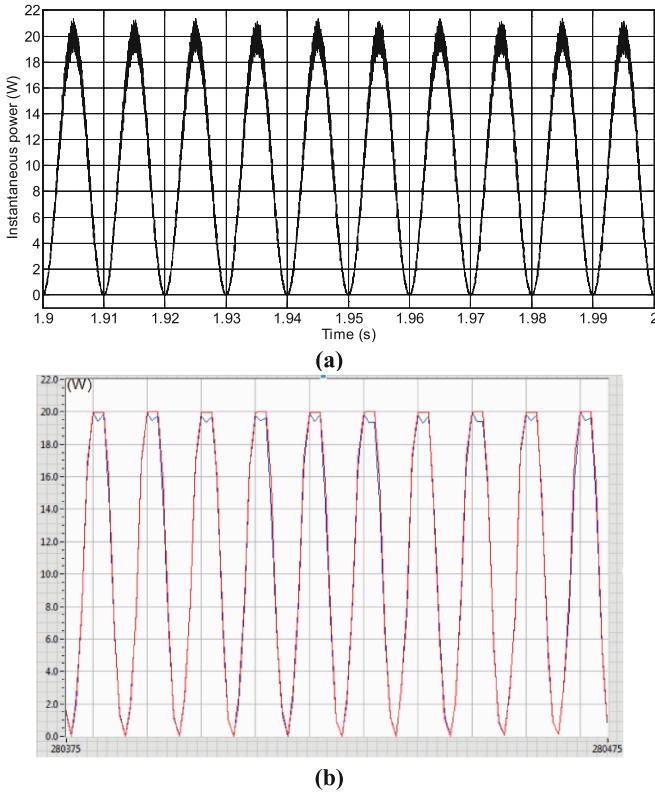
#### 3.2.1 Operation at Unity Power Factor

Figure 10 shows the response plots for the grid voltage and current when the injected power command is at 10 VA, unity power factor. The simulation and experimental results are in close agreement, showing the zero phase shift between injected current and grid voltage. The total harmonic distortion in the inverter output current is 4.2%.



**Fig. 10.** (a) Simulation and (b) Experimental results for  $i_0$  and  $v_g$  at  $S = 10 \text{ VA}$ ,  $\cos \phi = 1$

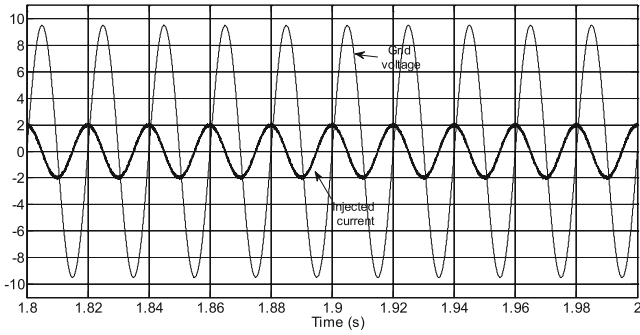
The instantaneous power variations between the dc and ac sides are shown in Figs. 11(a) and (b). Since the injected power has no reactive component, the power flow is always positive and has an average value of 10 W, corresponding to the command value.



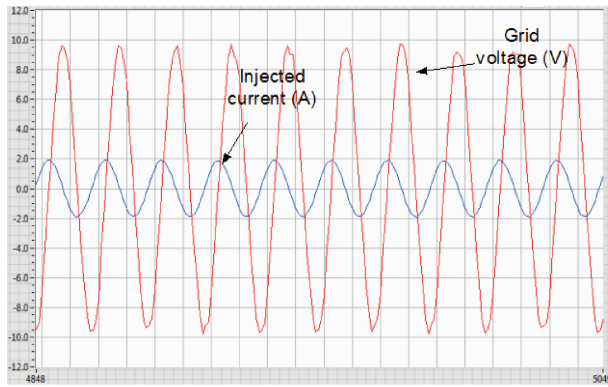
**Fig. 11.** (a) Simulation and (b) Experimental results for instantaneous power at  $\cos \phi = 1$

**3.2.2 Operation at Zero Power Factor, Leading**

The power command is next adjusted to inject purely reactive power (leading) into the ac side. The corresponding simulation and experimental results are shown in Figs. 12 and 13. The inverter output current is seen to lead the grid voltage by  $90^\circ$ . As a result, the instantaneous power response has zero average value and the reactive power is 9.5 VAR.

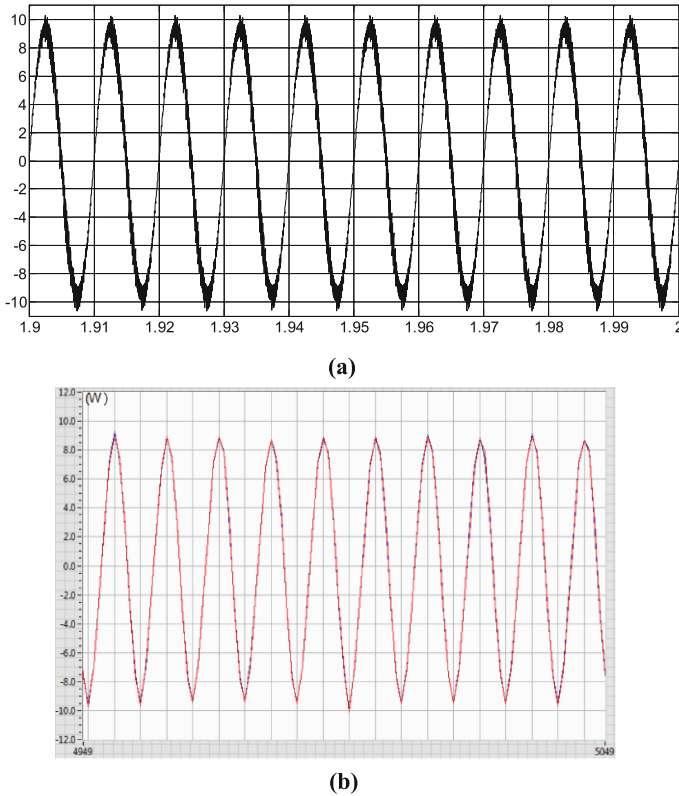


(a)



(b)

**Fig. 12.** (a) Simulation results for  $i_0$  and  $v_g$  at  $S = 9.5$  VA,  $\cos \phi = 0$  (lead) and (b) Experimental results for  $i_0$  and  $v_g$  at  $S = 9.5$  VA,  $\cos \phi = 0$  (lead)



**Fig. 13.** (a) Simulation and (b) Experimental results for instantaneous power,  $\cos \phi = 0$  (lead)

## 4 Conclusion

The modeling and design of a power flow control scheme from a photovoltaic array to low voltage ac grid have been presented. The proposed method is based on generating the output current command for the switch-mode dc to ac inverter, based on the error between the reference and actual instantaneous powers. A current tracking scheme using hysteresis band control ensures that the inverter output current has the required amplitude and phase shift with respect to the grid voltage. The implemented system has been tested under various power commands, including purely real and purely reactive power injection. In all cases, the instantaneous power responses closely follow the reference profiles, and the injected currents into the ac side are nearly sinusoidal with low harmonic distortion.

**Acknowledgments.** The authors wish to thank the University of Mauritius for funding this research work (No. R182), and for all the facilities provided.



## References

1. Vandoorn, T.L., et al.: Microgrids – hierarchical control and an overview of the control and reserve management strategies. *IEEE Ind. Electron. Mag.* **7**(4), 42–55 (2013)
2. Serban, I., Teodorescu, R., Marinescu, C.: Analysis and optimization of the battery energy storage systems for frequency control in autonomous microgrids, by means of hardware-in-the-loop simulations. In: 3rd IEEE International Symposium on Power Electronics for Distributed Generation Systems (PEDG), pp. 374–379 (2012)
3. Loh, P.C., Li, D., Blaabjerg, F.: Autonomous control of interlinking converters in hybrid AC-DC microgrids with energy storages. In: Energy Conversion Congress and Exposition (ECCE), pp. 652–658 (2011)
4. Gomes De Brito, M.A., et al.: Evaluation of the main MPPT techniques for photovoltaic applications. *IEEE Trans. Ind. Electron.* **60**(3), 1156–1167 (2013)
5. Ruhomaun, M.S.: Solar charge controller with maximum power point tracking. B.Eng. (Hons) dissertation, University of Mauritius (2015)
6. Sayed Hassen, S.Z., Jahmeerbacus, M.I., Sewraj, K., Ruhomaun, M.S.: Development and control of an experimental test bed for a microgrid using LabVIEW. Internally funded project, University of Mauritius (2015)
7. Sewraj, K.: Control of microgrid-connected dc to ac inverter with LabView interface., B.Eng. (Hons) dissertation, University of Mauritius (2015)
8. Zorilla, J., et al.: Analysis of dust losses in photovoltaic modules. In: Proceedings of World Energy Congress, pp. 2985–2992, May 2011
9. Rashid, M. (ed.): *Power Electronics Handbook*, 3rd edn. Elsevier Inc., Amsterdam (2011)

# Performance Analysis of a Hybrid Algorithm for Power Loss Reduction by Distribution Network Reconfiguration

Sarah Marappa Naiken and Robert T.F. Ah King<sup>(✉)</sup>

Department of Electrical and Electronic Engineering,  
University of Mauritius, Reduit 80837, Mauritius  
sarahmnaiken@gmail.com, r.ahking@uom.ac.mu

**Abstract.** Power loss reduction by Distribution Network Reconfiguration (DNR) is the process of finding the network topology offering the least losses. In this paper, an efficient hybrid algorithm is presented for the DNR problem. Through the addition of 3 deterministic refinements (a 'warm start', elitism and the hill climbing strategy), in any heuristic algorithm, both time efficiency and reliability of convergence can be achieved. Testing was done on 6 standard test systems: 16, 33, 70, 118, 135 and 880 bus systems. The results obtained revealed better network configurations that provided lesser power loss than other reported results.

**Keywords:** Distribution network reconfiguration · Hybrid flower pollination algorithm · Hybrid genetic algorithm · Power distribution systems · Power loss reduction · Radial networks

## 1 Introduction

Over the last decades, power loss reduction has been the driving force that has fuelled most contemporary research. Indeed, producing more power with the same amount of energy will bring us one step closer to global sustainability. Therefore, in this paper, an attempt at loss reduction in radial distribution networks is made. Surely, in between energy generation and consumption, 70% of the power losses occurring have been traced to distribution networks [1], thereby making the optimization of distribution networks mandatory.

Various techniques, such as reconfiguration processes, injection of distributed generations in the grid and the installation of capacitors, have been developed for this purpose [2]. But, the focus of this research is Distribution Network Reconfiguration (DNR). Requiring no additional resources than the network itself, DNR involves the opening and closing of switches to find a network configuration that produces minimal power loss [3].

Nevertheless the complexity of the DNR problem far outweighs its cost effectiveness. Indeed, the non-linear nature of distribution networks requires iterative power flows, which make deterministic approaches computationally burdening [4–7]. DNR has thus been linked to heuristic algorithms that offers fast and easy convergence to a

solution, without guarantee of true global optimality. Therefore, many studies have focused on the use of other new heuristic algorithms that will ensure such a feat. Algorithms such as the Ant Colony Search algorithm [8], Gravitational Search algorithm [9], Simulated Annealing algorithm [10], Artificial Immune System Optimization algorithm [11] have hence been conceptualized.

While some of these new algorithms indeed produce better results, will it not be better to focus on the refinement of already existing ones? In this research therefore, a new hybrid algorithm is thus developed (Sect. 3) by combining the reliability of deterministic algorithm with the fast convergence of heuristic algorithms, thereby taking the desired characteristics from already good concepts.

This paper hence demonstrates:

- (i) The effectiveness, both in terms of time and reliability, of DNR combined with the hybrid algorithm in finding the network configuration that produces the least power loss (Sect. 8.1).
- (ii) The extension of the proposed hybrid refinements to any other heuristic algorithm (Sect. 8.2).
- (iii) The ability of the hybrid algorithm to work with the same level of efficiency for networks of increasing sizes, that is, an increasing level of computational difficulty (Sect. 8.3).

## 2 Problem Definition

The main aim of this paper is the power loss reduction in radial distribution networks, by DNR, using the developed hybrid algorithm, while ensuring that the analysis is close to real life conditions.

Therefore, the objective function can be presented as follows.

$$\text{Minimise } f = \text{power loss}$$

subject to the following constraints:

### 1. Radial Network

Only the radial network configurations are considered while the non-radial one are penalized.

$$\varphi(x) = 0$$

### 2. Kirchhoff's Current Law

$$g_i(I, k) = 0$$

### 3. Kirchhoff's Voltage Law

$$g_i(V, k) = 0$$

### 3 The Hybrid Algorithm

Prior to the improvement of an algorithm, its weakness must first be studied. In this section, the purely heuristic algorithm is thus analyzed, after which ways of eliminating its limitations are proposed.

#### 3.1 Hybrid Genetic Algorithm (HGA)

The first algorithm chosen for the purpose of this research is Real-coded Genetic Algorithm [12].

**Purely Heuristic GA.** The purely heuristic GA consists of the following steps [13].

1. Creation of an initial random population
2. Loop:
  - 2.1 Roulette Wheel Selection
  - 2.2 Crossover/Mating
  - 2.3 Mutation
3. Termination

**Refinement of GA.** The 3 refinements proposed for GA are as follows.

1. Inclusion of the initial configuration in the initial population

Since the initial configuration of any distribution network is already known, it can be included in GA as an initial solution. Thus, instead of letting GA start with a totally random initial population, a sub-optimal solution is provided to give a ‘warm start’. This hence allows faster convergence to the required solution [14].

2. Elitism

Usually, the crossover and mutation process, despite their aim of finding better individuals, end up losing some of the already good solutions present. To prevent this, the concept of elitism can be implemented. After each stage of the genetic algorithm, the elite individuals are stored and included back in the population for the next generation. Hence if they are among the fittest, they will thus be kept in the population up to the last generation, thereby eliminating the risks of losing the global minimum.

3. Hill Climbing Strategy

To ensure that GA does not get stuck to a local optimum, it is mandatory to check if there is no solution better than the one produced. In [14] it is said that the sub-optimal solution is most of the time close to its global solution. For the case of DNR, the change of only a few tie switches can thus lead to a better solution, if any. Therefore, after the operations of GA have ended for a generation, the fittest individual is selected. Local transformations are then made on that individual to produce a better version that has even greater fitness. This is done by opening ‘n’ closed switches and simultaneously closing ‘n’ opened switches (n is an integer). This deterministic approach is

stopped as soon as a better version is produced. This newly engineered individual is also included in the next population. However, in case no better solutions are found, the search stops after a failed number of attempts. GA however continues its heuristic search. This hill climbing strategy thus brings a deterministic touch to the algorithm thereby increasing its chances of finding the true global solution.

### 3.2 Hybrid Flower Pollination Algorithm (HFPA)

The same refinements that have been applied for GA can be applied to FPA as well. The flower pollination algorithm conceptualized in [15] is used here. The Hybrid FPA consists of the following steps.

1. Creation of the initial population of flowers
2. Inclusion of initial configuration to the initial population
3. Loop
  - 3.1 Fitness evaluation of each flower
  - 3.2 Selection of the best flower
  - 3.3 Local Pollination
  - 3.4 Global Pollination
  - 3.5 Fitness evaluation of new population of flower
  - 3.6 Retention of higher fitness flowers (elitism)
  - 3.7 Implementation of deterministic search (hill climbing strategy)
4. Termination

## 4 Load Flow Development

With power loss as the objective function, a means of calculating this loss is required. In [16], a fast and efficient load flow analysis for radial distribution networks has been developed. The same approach is used here. Through the construction of a bus-injection to branch-current (BIBC) and a branch-current to bus-voltage (BCBV) a DLF matrix is formed, which can then be used to solve Eqs. (1–3) iteratively.

$$I_i^k = \left( \frac{P_i + jQ_i}{V_i^k} \right)^* \quad (1)$$

$$[\Delta V^{k+1}] = [DLF][I^k] \quad (2)$$

$$[V^{k+1}] = [V^0] + [\Delta V^{k+1}] \quad (3)$$

where  $V_i^k$  and  $I_i^k$  are the bus voltage and equivalent current injection of the bus  $i$  at the  $k^{th}$  iteration.  $P_i$  and  $Q_i$  represent the real and reactive power of the bus respectively.

From there, the power loss was calculated from Eqs. (4–5).

$$[B] = [BIBC][I] \quad (4)$$

$$\text{power loss} = [B] * R \quad (5)$$

where  $R$  represents the branch resistance.

## 5 Mathematical Modelling of DNR

For any distribution network, to ensure radiality, there is a predefined number of sectionalized switches and tie switches. In this paper, a particular network configuration has been represented by the tie switch numbers in an array. The tie switch can be any branch number of the system.

### 5.1 The Repair Algorithm

For any configuration, repetition in tie switch number is not allowed. As soon as such a configuration was detected, a repair algorithm was immediately applied. A unique copy of the switch numbers were retained and new switch numbers were generated to fill in the array.

## 6 DNR Using HGA

DNR using HGA basically applies the steps described in Sect. 3.1 to the DNR problem. This section gives a detailed analysis of the implementation. It must be noted that for a given distribution system, the parameters are defined according to the type and size of the network (Table 1).

**Table 1.** Details of system data

Test system	16 [17]	33 [17]	70 [18]	118 [19]	135 [20]	880 [21]
No. of branches	16	37	79	133	156	900
No. of nodes	16	33	70	123	223	880
No. of generator nodes	3	1	2	1	1	7
Generator node	1–3	1	1, 70	1	0	1–6, 880
No. of sectionalising branch	11	32	68	118	135	873
No. of tie branch	3	5	11	15	21	27
Base power, $S_{\text{base}}$ /MVA	10	10	10	10	100	50.68
Base voltage, $V_{\text{base}}$ /kV	10	12.66	11	11	13.8	130.8
Search space of system	364	$4.35 \times 10^5$	$7.78 \times 10^{12}$	$2.43 \times 10^{19}$	$5.44 \times 10^{25}$	$3.00 \times 10^{51}$

### 6.1 Creation of Random Initial Population

At the start of the algorithm, an individual was created by randomly creating an array of randomly generated numbers from 1 to the number of branches in the system. In case repetition of switches occurred, the repair algorithm was applied. Depending on the number of individuals required, the process was repeated.

### 6.2 Fitness Evaluation

In line with the aim of power loss reduction, the smaller the power loss the greater the fitness. Hence the fitness function can be defined as:

$$f'(x) = \frac{1}{f(x) + \varphi(x)} \quad (6)$$

where

- $x$  : Individual under test
- $f'(x)$  : Fitness function
- $f(x)$  : Objective function (power loss)
- $\varphi(x)$  : Constraint

### 6.3 Roulette Wheel Selection

Each individual having its own fitness level, is given a portion of the wheel. Thus, the greater its fitness, the larger share of the wheel it occupies. A random number was then generated to denote the amount by which the wheel had spun. The point where the wheel had stopped was the point of selection.

$$portion = \frac{\text{fitness of individual}}{\text{sum of all fitnesses}} \quad (7)$$

### 6.4 Crossover

The selected individuals are assigned a random number between 0 and 1. Depending on that number, the individual undergoes one of the two processes below.

**Formation of Mating Pool.** For individuals having been given a number greater than the probability of crossover  $P_c$ , they are carried forward to the ‘offspring population’ unchanged.

**Generation of Offspring.** Individuals having the number smaller than  $P_c$  are grouped into pairs (representing parents) and are made to undergo single point crossover. The repair algorithm was again applied in case duplication of switch numbers occurred.

## 6.5 Mutation

Here as well the individuals are assigned a random number. For those having a number less than the mutation rate, one of their switches is replaced by another randomly generated number. Switch duplication was handled by the repair algorithm (Sect. 5.1).

## 7 DNR Using HFPA

The steps in Sect. 3.2 was applied for DNR using HFPA. Just as in HGA procedures, a random number (rand) was assigned to each flower in the population. Depending on its value with respect to a switch probability  $p$ , the flower was made to undergo one of the two types of pollination below.

### 7.1 Local Pollination

This occurs when rand is greater than  $p$ . The new flower is obtained using the following Eq. 8.

$$x^{t+1}_i = x^t_i + \epsilon(x^t_j - x^t_k) \quad (8)$$

where

$x^{t+1}_i$  : Next flower

$x^t_i$  : Current flower

$x^t_j$  and  $x^t_k$  : Random flowers in the plant

$\epsilon$  : Random number taken from a uniform distribution [0, 1]

### 7.2 Global Pollination

This occurs when rand is less than  $p$ . The new flower is obtained using the Eq. 9.

$$x^{t+1}_i = x^t_i + L(g_* - x^t_i) \quad (9)$$

where

$L$  is the random number taken from the Lévy distribution.

## 8 Results

For this research, all programming was done using the MATLAB software. The version used was MATLAB R2013a. The simulations were performed using an Intel core i5 1.80 GHz processor with 4.00 GB of RAM. Also, the random number used were defined using the function `rng (seed)` so that replication of the results obtained could be done.



Testing was done on systems of increasing sizes to test the efficiency of the developed algorithm with greater computational difficulty. The details of the test systems are given in Table 1.

### 8.1 Effect of the Deterministic Refinements on the Heuristic Algorithms

In this section, the deterministic refinements proposed in Sect. 3.1 were tested independently then combined on GA. The results are shown in Table 2 and demonstrate the following:

- (i) The use of the initial configuration alone brings no evident improvement. But when combined with elitism and hill climbing, the large test systems become able to escape the non-radial configurations.
- (ii) Elitism alone is enough to ensure convergence in small test systems but for large networks the deterministic search is necessary.
- (iii) The hill climbing strategy cannot be used by itself due to the overly large search space area of the network.
- (iv) The 3 refinements combined together produces the best results. This can be seen more blatantly for the 118 bus system, where only the full hybrid algorithm was able to converge to the true optimum solution.
- (v) A peculiarity can be seen for the 70 bus system, which can be explained by the fact that for such relatively small test system the ‘initial configuration’ does not yet have much effect.

**Table 2.** Analysis of deterministic refinements on GA

Test system	No. of times global optimum reached in 100 runs			
	16	33	70	118
Purely heuristic	41	2	0	0
Elitism	100	65	0	0
Initial configuration (warm start)	33	2	0	0
Hill climbing	99	88	0	0
Elitism and initial configuration	100	62	0	0
Initial configuration and hill climbing	98	89	0	0
Elitism and hill climbing	100	100	2	0
Full hybrid	100	100	1	80

## 8.2 Analysis of the Hybrid Efficiency Using Different Heuristic Algorithms

This section is used to demonstrate whether the hybrid algorithm operates with the same level of efficiency when applied to different heuristic algorithms.

**Table 3.** Analysis of the efficiency of HGA and HFGA

Test system	Method	No. of times global optimum reached in 100 runs	No. of iterations to convergence	Time for 1 run (s)
16	HGA	100	2	20.060
	HFGA	100	2	3.901
33	HGA	100	10	82.018
	HFGA	100	10	11.544
70	HGA	1	109	289.124
	HFGA	98	9	118.176
118	HGA	80	412	519.784
	HFGA	86	75	217.326
135	HGA	80	463	73.542
	HFGA	99	102	319.430
880	HGA	0	915	1637.541
	HFGA	93	239	792.481

From Table 3 it can be seen that despite the same refinements being applied to both GA and FGA, HFGA demonstrate a better efficiency than HGA (both in terms of time to convergence and reliability). This thus reinforces the notion of the ‘No Free Lunch’ Theorem (NFL) [22]. Hence, other heuristic algorithms can be tested and achieve even better efficiency since the efficiency of the optimization algorithm and the problem being solved is independent.

## 8.3 Comparison of the Hybrid Algorithm with Related Works

This section shows the power loss in MW obtained using different algorithms for the test systems used.

From Table 4 it can be seen that the hybrid refinements indeed lead to a better search of the true global solution. This is noted for the 70, 118 and 880 bus systems, where new configurations, offering lesser power loss (in bold), were obtained.

The 0.1396 MW and 0.1395 MW power loss that was produced by HGA, HFGA and RGA in [4] corresponded to the same tie switch configuration. The slight discrepancy was accounted by the fact that a different load flow determination program was used. Approximations in the different methods thus produced a slight difference of 0.0001 MW which is negligible.

**Table 4.** Comparison of hybrid solutions with related works

Test system	16	33	70	118	135	880
Initial configuration	0.5114	0.2027	0.2268	1.2967	0.3204	1.4964
HGA	0.4661	0.1396	<b>0.1173</b>	<b>0.8536</b>	0.2802	0.4587
HFPA	0.4661	0.1396	<b>0.1173</b>	<b>0.8536</b>	0.2802	<b>0.4581</b>
RGA [4]	0.4661	0.1395	–	–	–	–
GA [17]	0.4661	0.1396	–	–	–	–
Branch exchange [23]	–	–	–	0.8856	–	–
TSA [24]	–	–	–	0.8842	–	–
ITSA [19]	–	–	–	0.8659	–	–
Method in [20]	–	–	–	–	0.2855	–
QP [25]	–	–	–	–	0.2808	0.4593
MIQCP [26]	–	0.1395	0.2014	0.8697	0.2802	–
MST [14]	–	0.1407	0.2087	0.8943	0.2894	–
LS [14]	–	0.1399	0.2036	0.8835	0.2864	–
DISTOP [27]	–	0.1395	0.2039	0.8919	0.2959	–

## 9 Conclusion

The problem of power loss reduction through the optimization of distribution networks by DNR can no doubt be enhanced using the hybrid algorithm developed. In this paper, the superiority of any hybrid algorithm, compared to its purely heuristic version, is observed. The hybrid algorithm guarantees convergence to the true global solution due to the hill climbing strategy, which constantly attempts to find a better solution. While ensuring reliable convergence, the hybrid algorithm also provides fast convergence to the solution, thereby providing time efficiency. Also, the use of other heuristic algorithms can lead to even better or worse efficiency than HGA and HFPA (NFL Theorem). Further scope for this research thus involves an analysis of other heuristic algorithms when hybridized. Also, more deterministic refinements can be included to the process.

## References

1. Suresh, N., Gowri Manohar, T.: Optimal citing of custom power controllers in distribution systems for loss reduction. In: Second National Level Conference on Arena of Intellectual Innovations, GSSS Institute of Engineering & Technology for Women, Mysore (2009)
2. Ramesh, L., Chowdhury, S.P., Chowdhury, S., Natarajan, A.A., Gaunt, C.T.: Minimization of power loss in distribution networks by different techniques. *Int. J. Electr. Comput. Energ. Electron. Commun. Eng.* **3**(4), 661–667 (2009)
3. Franco, J., Rider, M., Lavorato, M., Romero, R.: A mixed-integer LP model for the reconfiguration of radial electric distribution systems considering distributed generation. *Electr. Power Syst. Res.* **97**, 51–60 (2013)

4. Radha, B., Ah King, R.T.F., Rughooputh, H.C.S.: A modified genetic algorithm for optimal electrical distribution network reconfiguration. In: Congress on Evolutionary Computation, CEC 2003, vol. 2, pp. 1472–1479 (2003)
5. Srinivasa, R., Narasimham, S.: A new algorithm for the network reconfiguration of distribution feeders for loss minimization. *IEEE Trans. Power Deliv.* **7**(3), 1484–1491 (1992)
6. Singh, U.: Radial distribution system reconfiguration for loss minimization using exhaustive search techniques. M. Eng. Thesis, Thapar University (2014)
7. Morton, A., Mareels, I.: An efficient brute-force solution to the network reconfiguration problem. *IEEE Trans. Power Deliv.* **15**(3), 996–1000 (2000)
8. Su, C., Chang, C., Chiou, J.: Distribution network reconfiguration for loss reduction by ant colony search algorithm. *Electr. Power Syst. Res.* **75**(2–3), 190–199 (2005)
9. Kottal, L., Priyadarshini, R., Prakash, R.: Network reconfiguration for loss reduction of a radial distribution system. *IJARCCCE* **4**(5), 686–690 (2015)
10. Santander, L.G., Chacra, F.A., Opazo, H., Lopez, E.: Minimal loss reconfiguration based on simulated annealing meta-heuristic. In: IEEE Conference on Electronics, Communications and Computers, pp. 95–99 (2005)
11. Alonso, F., Oliveira, D., Zambroni De Souza, A.: Artificial immune systems optimization approach for multiobjective distribution system reconfiguration. *IEEE Trans. Power Syst.* **30**(2), 840–847 (2015)
12. Ravibabu, P., Venkatesh, K., Kumar, C.S.: Implementation of genetic algorithm for optimal network reconfiguration in distribution systems for load balancing. In: IEEE Conference on Computational Technologies in Electrical and Electronics Engineering, pp. 124–128 (2008)
13. Stender, J., Hillebrand, E., Kingdon, J.: Genetic Algorithms in Optimisation, Simulation, and Modelling. IOS Press, Amsterdam (1994)
14. Ahmadi, H., Martí, J.: Minimum-Loss Network Reconfiguration: A Minimum Spanning Tree Problem. *Sustain. Energy Grids Netw.* **1**, 1–9 (2015)
15. Yang, X., Karamanoglu, M., He, X.: Multi-objective flower algorithm for optimization. *Procedia Comput. Sci.* **18**, 861–868 (2013)
16. Teng, J.-H.: A direct approach for distribution system load flow solutions. *IEEE Trans. Power Deliv.* **18**(3), 882–887 (2003)
17. Zhu, J.: Optimal reconfiguration of electrical distribution network using the refined genetic algorithm. *Electr. Power Syst. Res.* **62**(1), 37–42 (2002)
18. Das, D.: A fuzzy multiobjective approach for network reconfiguration of distribution systems. *IEEE Trans. Power Deliv.* **21**(1), 202–209 (2006)
19. Zhang, D., Fu, Z., Zhang, L.: An improved TS algorithm for loss-minimum reconfiguration in large-scale distribution systems. *Electr. Power Syst. Res.* **77**(5–6), 685–694 (2007)
20. Mantovani, J., Casari, F., Romero, R.: Reconfiguration of radial distribution systems using the voltage drop criteria. *SBA Control Autom.* **11**(3), 150–159 (2000)
21. Kavasseri, R., Ababei, C.: REDS: REpository of Distribution Systems (2016). <http://www.dejazzer.com/reds.html>. Accessed 26 Feb 2016
22. Wolpert, D., Macready, W.: No free lunch theorems for optimization. *IEEE Trans. Evol. Comput.* **1**(1), 67–82 (1997)
23. Baran, M., Wu, F.: Network reconfiguration in distribution systems for loss reduction and load balancing. *IEEE Trans. Power Deliv.* **4**(2), 1401–1407 (1989)
24. Mishima, Y., Nara, K., Satoh, T., Ito, T., Kaneda, H.: Method for minimum-loss reconfiguration of distribution system by tabu search. *Elect. Eng. Jpn.* **152**(2), 18–25 (2005)

25. Taylor, J., Hover, F.: Convex models of distribution system reconfiguration. *IEEE Trans. Power Syst.* **27**(3), 1407–1413 (2012)
26. Ahmadi, H., Marti, J.: Distribution system optimization based on a linear power-flow formulation. *IEEE Trans. Power Deliv.* **30**(1), 25–33 (2015)
27. Shirmohammadi, D., Hong, H.: Reconfiguration of electric distribution networks for resistive line losses reduction. *IEEE Trans. Power Deliv.* **4**(2), 1492–1498 (1989)

# Voltage Stability Maximization by Distribution Network Reconfiguration Using a Hybrid Algorithm

Robert T.F. Ah King<sup>(✉)</sup> and Sarah Marappa Naiken

Department of Electrical and Electronic Engineering,  
University of Mauritius, Reduit 80837, Mauritius  
r.ahking@uom.ac.mu, sarahmnaiken@gmail.com

**Abstract.** Voltage stability maximization by Distribution Network Reconfiguration (DNR) is the process of finding a network configuration offering the least voltage deviation at the buses. Through the use of an Improved Voltage Deviation Index (IVDI) the stability of the entire distribution network has been studied. Using a hybrid algorithm, the DNR problem was tested on 6 standard test systems: 16, 33, 70, 118, 135 and 880 bus systems. The results showed that DNR using the IVDI can indeed lead to the improvement of the entire system stability.

**Keywords:** Distribution network reconfiguration · Hybrid flower pollination algorithm · Hybrid genetic algorithm · Power distribution systems · Voltage stability maximization · Improved voltage deviation index · Radial networks

## 1 Introduction

Nowadays, the need to save our finite energy resources has led to a greater emphasis on energy efficiency and hence power loss reduction. But, this should not be at the expense of the reliability of our power systems. Indeed, voltage instability often translates into voltage collapse, which in turn leads to blackout or abnormally low voltage levels at the buses of our distribution networks [1]. Thus, the maximization of voltage stability is mandatory.

Throughout the years, various techniques have been formulated to determine the stability of the buses: the use of P-V and Q-V curves, L-index, FVSI, LQP, VCPI, modal analysis, amongst others [2]. In this way, the determination of the weakest bus was possible and from that appropriate measures, such as the installation of capacitors, tap-changing transformers and other related switching equipment could be done. However, through these techniques, only the stability of the weakest bus was improved and also additional costly resources were required.

This research therefore proposes a cost-free approach to the problem. Through the use of Distribution Network Reconfiguration (DNR), the mere opening of sectionalizing switches and the closing of tie switches finds a radial network configuration with improved stability [3]. But the complexity of DNR lies with its associated algorithm. In [4] the efficiency of a hybrid algorithm is shown through increasingly difficult problems

using DNR for power loss reduction. Hence, this paper uses the same hybrid algorithm for the maximization of voltage stability (Sect. 5). Moreover, an improved voltage deviation index (IVDI) is presented which enables the analysis of the entire network's stability, rather than just the weakest bus of the system (Sect. 3). This paper hence demonstrates:

- (i) The ability of DNR to improve the stability of the entire distribution network through the use of the IVDI (Sect. 7.1).
- (ii) The effectiveness of the IVDI in the analysis of the network stability (Sect. 7.2).
- (iii) The efficiency of the hybrid algorithm irrespective of the optimization case (Sect. 7.3).

## 2 Problem Definition

The main aim of this paper is the voltage stability maximization in radial distribution networks, by DNR using the hybrid algorithm presented in [4], while ensuring that the analysis is close to real life conditions. Therefore, the objective function can be presented as follows.

Minimise  $f = \text{Total Voltage Deviation}$

Subject to the following constraints:

### 1. Radial Network

Only the radial network configurations are considered while the non-radial one are penalized.

$$\varphi(x) = 0$$

### 2. Kirchhoff's Current Law

$$g_i(I, k) = 0$$

### 3. Kirchhoff's Voltage Law

$$g_i(V, k) = 0$$

## 3 The Improved Voltage Deviation Index (IVDI)

For the analysis of the stability of a network, this section shows an improvement to the Voltage Deviation Index (VDI) in [5]. The principle of the VDI basically implies that the closer the voltage level of a particular load bus is to the generator bus, the lesser the voltage deviation of the load bus and therefore the greater its stability. While being able to provide a direct and easy method to stability analysis, the only limitation of the method is that the voltage stability of a particular network configuration was determined by the stability of the weakest bus only. Thus, an extension is made to the method to enable the analysis of the entire system's stability.

The stability analysis begins with a load flow implementation, which determines the voltage level at each bus. In [6], a fast and efficient technique is proposed. Through the construction of a bus-injection to branch-current (BIBC) and a branch-current to bus-voltage (BCBV) a DLF matrix is formed, which can then be used to solve Eqs. (1–3) iteratively.

$$I_i^k = \left( \frac{P_i + jQ_i}{V_i^k} \right)^* \quad (1)$$

$$[\Delta V^{k+1}] = [DLF][I^k] \quad (2)$$

$$[V^{k+1}] = [V^0] + [\Delta V^{k+1}] \quad (3)$$

where  $V_i^k$  and  $I_i^k$  are the bus voltage and equivalent current injection of the bus  $i$  at the  $k^{th}$  iteration.  $P_i$  and  $Q_i$  represent the real and reactive power of the bus respectively.

Using the bus voltage levels obtained, the absolute values are taken and converted into per unit. Substituting these values in Eq. (4), the voltage deviation at each bus is calculated [5].

$$\Delta V_{D_i} = \frac{V_G - V_i}{V_G} \quad (4)$$

where

$V_G$  : Voltage at the Generator node (1 pu)

$V_i$  : Absolute Voltage at the Load Bus

$\Delta V_{D_i}$  : Voltage deviation Index at a bus

The total stability of a particular network configuration was finally found through the addition of the total voltage deviation of all the buses in the system using Eq. (5).

$$V_D = \sum_i^N \Delta V_{D_i} \quad (5)$$

where

$V_D$  : Total voltage deviation of the system

$N$  : Total number of load buses

Therefore, the smaller the value of  $V_D$  the greater the voltage stability of the distribution network.



## 4 Mathematical Modelling of DNR

In this paper, the maximization of the voltage stability of radial distribution networks is being considered. Thus, to ensure radiality, there is a predefined number of sectionalized switches and tie switches for any distribution network. In this paper, tie switch numbers in an array are used to represent a particular network configuration. The tie switch can be any branch number of the system.

### 4.1 The Repair Algorithm

Repetition in tie switch number is not allowed for any configuration. A repair algorithm [4] was immediately applied if such a configuration was detected such that a unique copy of the switch numbers were retained and new switch numbers were generated to fill in the array.

## 5 The Hybrid Algorithm

In [4], a hybrid algorithm is presented which combines deterministic refinements to any heuristic algorithm. This paper considers the hybridization of 2 heuristic algorithms: Genetic Algorithm (GA) and Flower Pollination Algorithm (FPA). GA consists of a selection, crossover and mutation parameter while FPA makes use of the concept of local and global pollination. However due to this greedy search, obtaining the true global solution becomes difficult. Hence, the following refinements are included in the algorithms.

1. Use of a 'warm start'

This is achieved through the inclusion of the initial configuration in the initial search.

2. Elitism

This involves the retention of good solutions after each computation process of the algorithm to prevent loss of solutions.

3. Hill Climbing Strategy

This is a local search which enables the finding of better solutions by opening 'n' sectionalized switches and simultaneously closing 'n' tie switches.

## 6 DNR Using HGA and HFPA

The same technique as in [4] is used for the application of the hybrid algorithm to the DNR problem, with the only difference in the objective function. Figures 1 and 2 give a description of the steps for the implementation of the hybrid GA (HGA) and the hybrid FPA (HFPA), respectively.

Regardless of the algorithm, the fitness of any individual can be evaluated using Eq. 6.

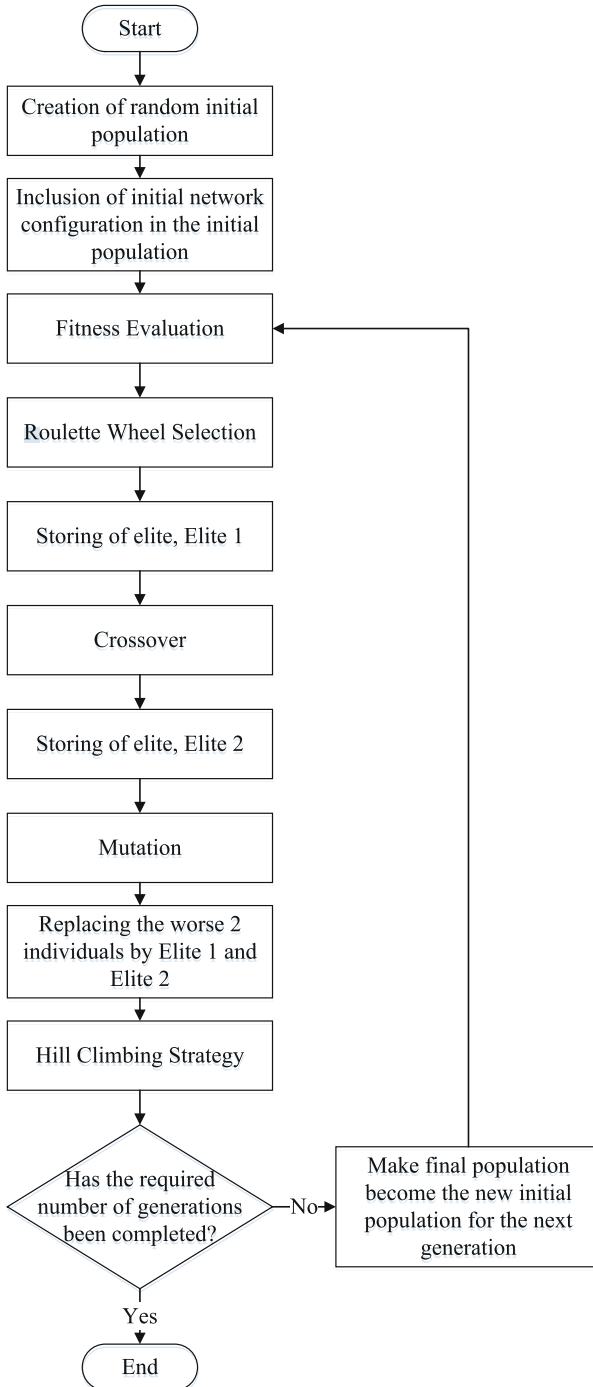
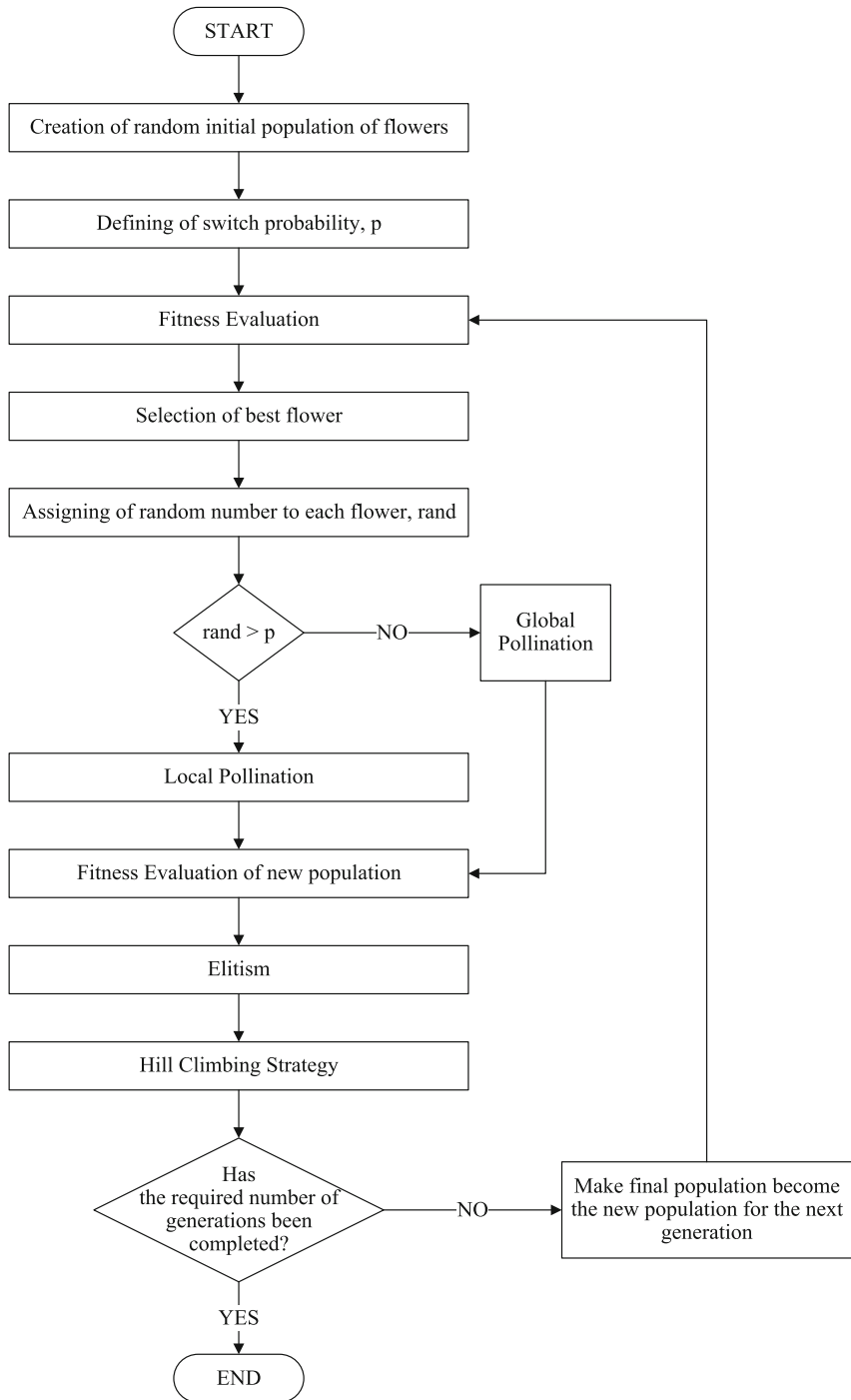


Fig. 1. Flowchart of HGA

**Fig. 2.** Flowchart of HFPA

$$f'(x) = \frac{1}{f(x) + \varphi(x)} \quad (6)$$

where

- $x$  : Individual under test  
 $f'(x)$  : Fitness function  
 $f(x)$  : Objective function (Total Voltage Deviation)  
 $\varphi(x)$  : Constraint

## 7 Results

For this research, all programming was done using the MATLAB R2013a software. The simulations were performed using an Intel core i5 1.80 GHz processor with 4.00 GB of RAM. Also, the random number generator was initialized using the function `rng(seed)` so that results could be reproduced.

To test the efficiency of the developed algorithms, simulations were performed on systems of increasing sizes with greater computational difficulty as shown in Table 1.

**Table 1.** Details of system data

Test system	16 [7]	33 [7]	70 [8]	118 [9]	135 [10]	880 [11]
No. of branches	16	37	79	133	156	900
No. of nodes	16	33	70	123	223	880
No. of generator nodes	3	1	2	1	1	7
Generator node	1–3	1	1, 70	1	0	1–6, 880
No. of sectionalising branch	11	32	68	118	135	873
No. of tie branch	3	5	11	15	21	27
Base power, $S_{\text{base}}$ /MVA	10	10	10	10	100	50.68
Base voltage, $V_{\text{base}}$ /kV	10	12.66	11	11	13.8	130.8
Search space of system	364	$4.35 \times 10^5$	$7.78 \times 10^{12}$	$2.43 \times 10^{19}$	$5.44 \times 10^{25}$	$3.00 \times 10^{51}$

### 7.1 Effect of DNR on the Voltage Stability of Distribution Networks

This section demonstrates the effect of DNR on the voltage stability of the network. From Table 2, it can be seen that the optimization of the distribution networks by DNR indeed leads to a decrease in the total voltage deviation of the system and thus causes an improvement in stability.

### 7.2 Voltage Profile Analysis of the Distribution Networks

The effect of DNR on the stability of the weakest and strongest buses of the examined systems is analysed here. As shown in Table 3, DNR for minimisation of total voltage

**Table 2.** Analysis of total voltage deviation after DNR

Test system	Total voltage deviation (pu)		
	Initial configuration	HGA	HFPA
16	0.211045	0.184465	0.184465
33	1.701121	1.051281	1.051281
70	3.415624	3.158206	3.158206
118	5.240284	3.755860	3.755860
135	3.407807	3.015136	3.015136
880	11.389184	3.971487	3.971080

deviation did not always improve the voltage profile at the weakest and strongest bus. Rather it aimed at improving the stability of the entire system.

In Table 4, the voltage at the weakest bus was seen to approach close to that of the generator bus after DNR, except for the 135 bus system. Even though there was no improvement in the stability of its weakest or strongest bus, the stability of the other buses were in turn improved, thereby leading to an overall improvement in system stability.

Moreover, it was seen for the 880 bus system that a decrease of 0.0002 pu occurred at the strongest bus of the network after DNR. That did not imply a decrease in the stability of the system. Rather, this small decrease in stability at the strongest bus led to the enhancement of the voltage of the other buses of the system (as seen for the weakest bus).

### 7.3 Efficiency of the Hybrid Algorithm

In [4], the superiority in the converging ability of the hybrid algorithms for power loss reduction by DNR was analyzed. This section is thus used to demonstrate whether the hybrid algorithm operates with the same level of efficiency when applied to another optimization case (the minimization of the total voltage deviation).

From Table 5, it can be seen that:

- (i) Application of the hybrid algorithm for the minimization of the total voltage deviation by DNR can indeed be done.
- (ii) HFPA was able to find the true global optimum solution for all test systems while HGA was unable to do the same, thereby implying greater reliability of HFPA.
- (iii) In terms of the time and number of iterations required to convergence, HFPA was again better than HGA.

**Table 3.** Identification of weakest and strongest bus

Test system	Weakest bus number			Strongest bus number		
	Initial conf.	HGA	HFGA	Initial conf.	HGA	HFGA
16	12	12	12	14	13	13
33	18	33	33	2	2	2
70	67	29	29	51	51	51
118	77	77	77	100	100	100
135	116	106	106	63	63	63
880	769	396	396	825	604	604

**Table 4.** Voltage deviation at weakest and strongest bus

Test system	Voltage at weakest bus			Voltage at strongest bus		
	Initial conf.	HGA	HFGA	Initial conf.	HGA	HFGA
16	0.9693	0.9716	0.9716	0.9948	0.9923	0.9923
33	0.9131	0.9356	0.9356	0.9970	0.9971	0.9971
70	0.9059	0.9247	0.9247	0.9931	0.9946	0.9946
118	0.8688	0.8688	0.8688	0.9963	0.9963	0.9963
135	0.9307	0.9596	0.9596	0.9999	0.9999	0.9999
880	0.9561	0.9916	0.9916	0.9999	0.9997	0.9997

**Table 5.** Performance analysis of HGA and HFGA

Test system	Method	No. of times global solution reached in 100 runs	No. of iterations to convergence	Time for 1 run (s)
16	HGA	100	2	83.660
	HFGA	100	2	4.498
33	HGA	100	5	122.745
	HFGA	100	5	29.868
70	HGA	99	103	235.468
	HFGA	100	15	98.734
118	HGA	94	195	418.723
	HFGA	98	46	176.410
135	HGA	86	279	627.119
	HFGA	100	96	243.971
880	HGA	0	817	1294.317
	HFGA	94	249	628.349

## 8 Conclusion

Distribution Network Reconfiguration can no doubt be used as an optimization technique for the maximization of the stability of a given network. In this paper, the proposed Improved Voltage Deviation Index (IVDI) enables the fast and easy analysis of the entire system stability. Through consideration of the voltage deviation of the

entire system, rather than just the weakest bus, greater improvement in system stability was achieved. Thus, using DNR as an optimization tool, network stability was achieved without requiring any additional cost. Also, the hybrid algorithm in [4] works with the same level of efficiency, irrespective of the optimization case, thereby confirming its effectiveness.

## References

1. Kundur, P., Paserba, J., Ajarapu, V., Andersson, G., Bose, A., Canizares, C.A., Hatziargyriou, N., Hill, D., Stankovic, A., Taylor, C., Van Cutsem, T., Vittal, V.: Definition and classification of power system stability. *IEEE Trans. Power Syst.* **19**(2), 1387–1401 (2004)
2. Reis, C., Barbosa, F.: A comparison of voltage stability indices. In: *IEEE Mediterranean Electrotechnical Conference, MELECON 2006*, pp. 1007–1010 (2006)
3. Barker, P., De Mello, R.: Determining the impact of distributed generation on power systems, i. radial distribution systems. In: *IEEE Power Engineering Society Summer Meeting 3*, pp. 1645–1656 (2000)
4. Naiken, S.M., Ah King, R.T.F.: Performance analysis of a hybrid algorithm for power loss reduction by distribution network reconfiguration. In: *1st International Conference on Emerging Trends in Electrical, Electronic and Communications Engineering (ELECOM 2016), Lecture Notes in Electrical Engineering* (2016)
5. Nguyen, T., Truong, A.: Distribution network reconfiguration for power loss minimization and voltage profile improvement using cuckoo search algorithm. *Int. J. Electr. Power Energy Syst.* **68**, 233–242 (2015)
6. Teng, J.-H.: A direct approach for distribution system load flow solutions. *IEEE Trans. Power Deliv.* **18**(3), 882–887 (2003)
7. Zhu, J.: Optimal reconfiguration of electrical distribution network using the refined genetic algorithm. *Electr. Power Syst. Res.* **62**(1), 37–42 (2002)
8. Das, D.: A fuzzy multiobjective approach for network reconfiguration of distribution systems. *IEEE Trans. Power Deliv.* **21**(1), 202–209 (2006)
9. Zhang, D., Fu, Z., Zhang, L.: An improved TS algorithm for loss-minimum reconfiguration in large-scale distribution systems. *Electr. Power Syst. Res.* **77**(5–6), 685–694 (2007)
10. Mantovani, J., Casari, F., Romero, R.: Reconfiguration of radial distribution systems using the voltage drop criteria. *SBA Control Autom.* **11**(3), 150–159 (2000)
11. Kavasseri, R., Ababei, C.: REDS: REpository of Distribution Systems (2016) <http://www.dejazzer.com/reds.html>. Accessed 26 Feb 2016

# Design and Implementation of a Smart Dual Axis Solar Tracker with an Anti-theft Alarm Mechanism

Anshu Prakash Murdan<sup>(✉)</sup>, Rameshwar Jugurnauth,  
and Ravishwara Rakesh Nirsimloo

Department of Electrical and Electronic Engineering,  
University of Mauritius, Reduit, Mauritius  
{a.murdan, r.jugurnauth}@uom.ac.mu,  
ravishwara.nirsimloo@umail.uom.ac.mu

**Abstract.** Photovoltaic (PV) systems have played a key role over the last decade in the evolution of the electricity sector. Solar trackers are commonly used to increase the energy output from PV panels. However less focus have been laid on the security of the panels. It has been reported that the past few years has seen a significant increase in solar panel robbery. To address this issue, a microcontroller based, anti-theft, dual axis solar tracking prototype was designed and implemented. A comparative analysis was then performed between the dual-axis solar tracking system, and a fixed photo module. Results show that, on average the dual axis solar tracker produced 25.1% more energy than the stationary photo module, for the tropical island of Mauritius. An inexpensive microcontroller was used which enabled Short Message Service (SMS) text messages to and from the tracker, for control and signaling of tampering and/or theft of PV panels.

**Keywords:** Microcontroller · Solar tracker · PV panel · Dual-axis

## 1 Introduction

Electricity is the backbone of modern society, but it is at the cost of depleting our major natural resources. When fossil fuels are burned to generate electricity, they release harmful greenhouse gases into the atmosphere. Renewable energy sources, like solar and wind provide potential solutions to the environmental problems being caused by fossil fuels. The various forms of solar energy – solar thermal, and solar photovoltaic (PV) offer a clean, climate-friendly, very abundant and in-exhaustive energy resource to mankind. PV systems capture the sun's higher frequency radiation (visible and ultra violet) in an array of semiconductor, photovoltaic cells which convert the radiant energy directly into electricity. Photovoltaic systems - standalone and grid-connected systems have been in use for more than twenty years [1]. PV systems power output is dependent on direct sunlight, so about 10–25% is lost if a tracking system is not used,



since the cell will not be directly facing the sun at all times [2]. Consequently, solar tracking is increasingly being applied as a sustainable power generating solution. Solar tracking system is a device for orienting a solar panel or concentrating a solar reflector or lens towards the sun. Concentrators, especially in solar cell applications, require a high degree of accuracy to ensure that the concentrated sunlight is directed precisely to the powered device. Precise tracking of the sun is achieved through systems with single or dual axis tracking. In Mauritius, from January to June 2015, only 8% of energy generated (excluding the use of bagasse in thermal power station) were from renewable resources, and that include the 10 GWh solar energy generated from solar farms [3]. Over the years, the electricity consumption by Mauritians has increased considerably, owing to the economic and technological growth of the country. The Central Electricity Board (CEB) is the electricity utility in Mauritius, and its greatest source for power generation is coal and oil, which are non-renewable resources. In 2013, 55.74% of the CEB electricity supplied to consumers came from independent power producers [4]. Since Mauritius is a tropical island with sunny weather all year round, PV systems works at their best, and many individuals and corporations have now started generating their own electricity and selling the surplus. With schemes and grants provided by the government e.g. Maurice Ile Durable (MID) fund, green loans, subsidies, Small Scale Distributed Generation (SSDG), solar energy has become more accessible and popular among Mauritians. However, current rooftop PV systems can be further improved to harness more solar energy. With current fixed PV systems, the solar panels are not used to their maximum. The sun moves through different paths and angles throughout the year; thus making it very difficult to extract the maximum power output from a fixed PV panel. Thus, a system/technique needs to be implemented in order to make sure the solar panels are always oriented towards the sun.

### **Related Works**

Solar tracking is by far the most efficient way of making sure that the PV panel is perpendicular to the incident sun rays. Passive solar trackers are slow to respond to changes in the sun's position, and they are vulnerable to wind gusts compared to active solar trackers which use sensors and motors to minimize the pointing error in order to maximize energy production. Although it is not essential to implement a solar tracking system, it has been proven to increase the intensity of sunlight received by 10–100%, regardless of the time of the year or location. However, in driving the system, it will use 2–3% more electricity [5]. According to Barsoum et al., solar tracker will increase energy output by approximately 20% or even more, and a dual axis solar tracker will double this, giving an increase in energy output by approximately 40% when compared to a PV system on a horizontal surface [6]. The tracking system's brain is its controller which runs an algorithm, allowing it to send signals for the motors' movement. Advancements in solar tracking algorithms has allowed the development of solar tracking systems, each with different accuracies and abilities.

Most of the reviewed works cited above demonstrate the efficiency gains with solar trackers. Security of PV panels and solar trackers have not been the focus of research despite the fact that theft of PV panels and components have been reported for more than a decade. According to the California State Sherriff's Association, the Napa Valley winery industry has been especially hard hit by PV thieves; more than 400 solar panels worth over \$400,000 were stolen from wineries between June 2008 and September 2009. In April 2010, a San Francisco-area high school reported a holiday week theft of 108 newly installed panels, with the loss calculated at over \$40,000 [7]. While several PV manufacturers inscribe a serial number on their panels, there are no consistent standards for indelible marking or identification protocols that would lend themselves to a formalized registry for PV theft deterrence. In 2016, a crime group arrested in Germany was suspected of at least 22 PV thefts at the solar facilities in Brandenburg, Bavaria and Hesse [8].

In this project, a simple, flexible, and cheap wire-break alarm was implemented to inform the user in case the solar panel is being removed from its frame, or is being tampered with. It does not require any complex calibrations or excavations, and can even be made with plastic fiber, which is a material that does not age or affected by harsh weather conditions. The system controller is made up of an Arduino microcontroller and a GSM module which communicates with the user. The microcontroller has a dual purpose, sun tracking and providing security, by sending a text message to the person in charge, to alert him in case the PV panel is being tampered with. In case there are high winds, the user can also park the panel in the position of least resistance to the wind, that is  $180^0$  on the horizontal axis for minimal surface area exposed to the wind, by sending a text message to the controller. Another text message may be sent to resume tracking. At night, the system will automatically position the PV panel to a preconfigured position in order to wait for sunrise to resume solar tracking. A working prototype was built, comprising of two motor drivers for the mount movement in three dimensions. Photodiodes are used to detect the panel orientation relative to the sun's position. The data processing mechanism is done by the microcontroller.

## 2 System Flowchart

The system flowchart, for the Arduino microcontroller consists of the following functions: INITIALISE, NIGHT, PARK, WIREBREAK, RESUME and TRACKING. As shown in the Fig. 1, the controller will collect the light sensors' values (INITIALISE function), for each clock cycles. Then based on the values recorded, it will determine whether it is day or night. If it is night the tracker will move the panel to a preconfigured position (NIGHT function) so as to wait for sunrise; else it will check the GSM network to see if the park instruction – 'P' has been received. The 'PARK' mode is the position of the panel which offers least resistance to the wind. The tracker will remain in the 'PARK' mode until the 'RESUME' instruction 'R' has been received. The controller will then check the wire break system's status to see if it has been

activated. The WIREBREAK function causes an alarm to be activated when a thin wire attached to the rear of the solar panel breaks; this triggers the controller to send a ‘miscall’ and an alert SMS message to the user. Whenever the tracker is going to the ‘PARK’ or ‘RESUME’ modes, confirmation SMS messages will be sent to the user to inform him. Also, to reduce the power consumption by the motors, the system will enable the motors for ten seconds and then disable them for three minutes before being enabled again. Finally the TRACKING function is executed, whereby the tracker moves the panel so as to capture maximum sunlight.

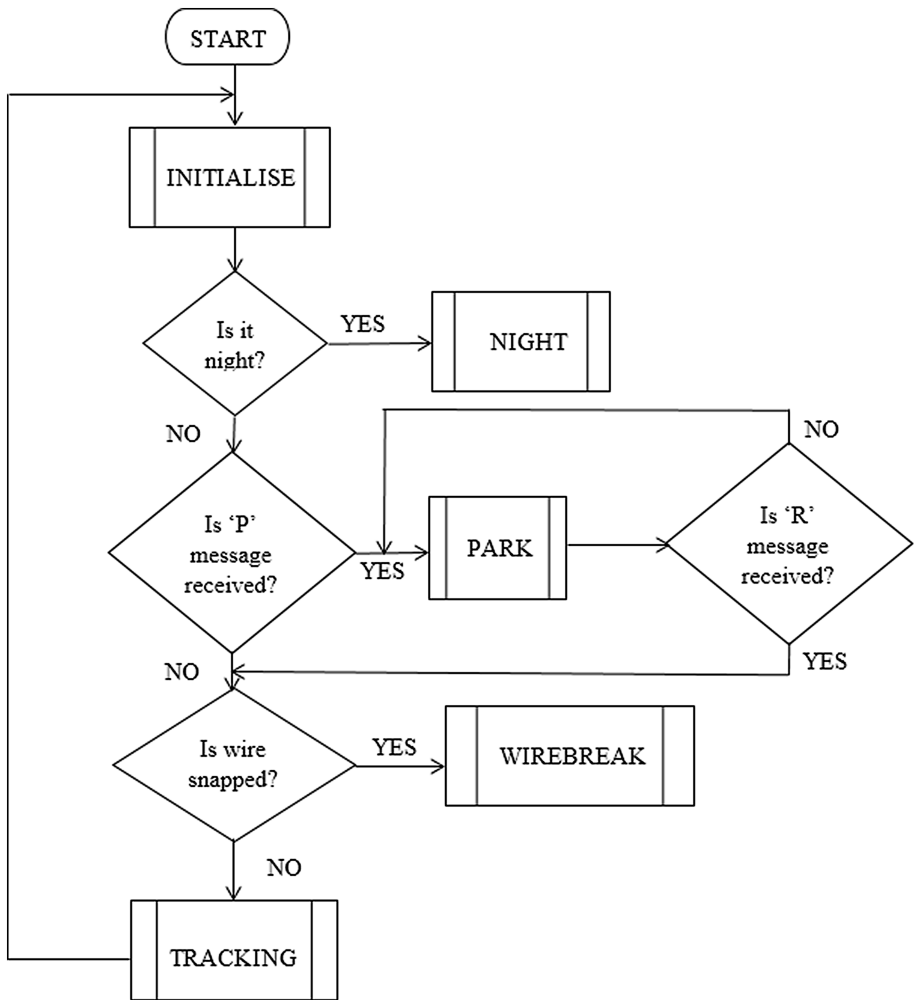


Fig. 1. System flowchart

### 3 Proposed Design

The tracker is a closed loop active system with light sensors (Fig. 2).

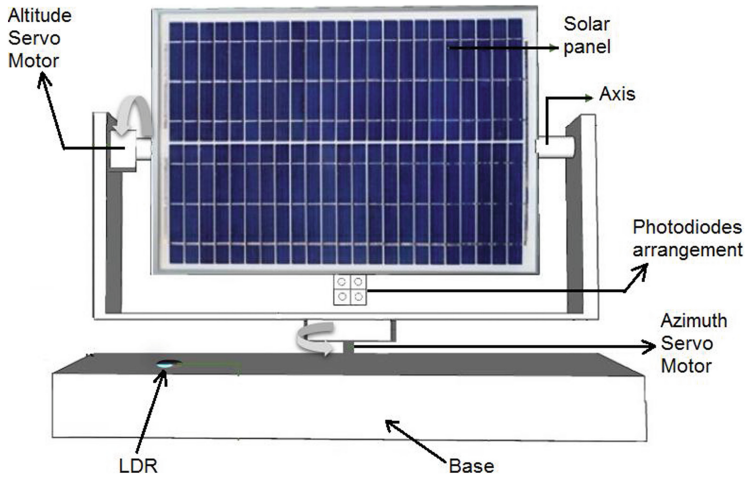


Fig. 2. Proposed design

#### 3.1 Solar Panel

The panel selected for the prototype is single diode, epoxy, waterproof and capable of generating 5 W of power. The specification of the PV panel is shown in Table 1.

Table 1. Solar panel specifications.

Related power	5 W
Voc	18.2 V
Vop	17.3 V
Short circuit current	0.29 A
Working current (Iop)	0.27 A
Output tolerance	±3%
Temperature range	-40 °C to +80 °C
SLA battery voltage	12 V
Dimensions (L×W×H) mm	200 * 210 * 3 mm
NET weight (KG)	0.25

#### 3.2 Motor

A servomotor is used for the prototype. Since the system is closed loop, the controller will be requiring exact feedback position of the motor to determine when it should stop moving, thus a servo motor fits our requirements.

### 3.3 Microcontroller

A microcontroller is just like a computer, but in a smaller scale, with lesser processing power. It is used to control processes that are repetitive in nature. A microcontroller is used to interface all sensor inputs to the output devices and carry out the processing of data. For our prototype, we will be using the 'ARDUINO Mega 2560' microcontroller which allows for easy and fast prototyping, and even has an in-built analog to digital converter to sample the light sensors' analog inputs.

### 3.4 GSM Module

For the project, a cheap and reliable system is required to communicate with the prototype; thus a GSM module fits our requirements as it does not require any complex configurations or installation. A GSM module is a device used to send/receive data from a remote location through a GSM network. There exists numerous GSM modules on the market, but the SIM900 module was selected for this project, since Mauritian mobile operators operate on the 900/1800 MHz band. The SIM900 is a Quad-band GSM/GPRS module which delivers GSM/GPRS performance. It works on frequencies 850/900/1800/1900 MHz, and along with internet access, it allows the user to make voice calls as it has ports to connect a microphone and loud speaker. It also incorporates TCP/IP protocol and supports AT commands.

### 3.5 Light Sensor

Any device that exhibits a change in its electrical behaviour when exposed to changing light intensity can be used for solar tracking purposes. One light sensor is the photodiode, which is more complex than the Light Dependent Resistor (LDR) but its functioning is different as it produces current depending on the light intensity reflecting on its surface [9]. Mainly, it is a diode which is reverse biased and turned off when a certain light intensity threshold is exceeded [10]. Compared to LDRs, photodiodes react faster and can sense even the smallest change in light intensity when exposed to very bright light. Thus we chose photodiodes to sense the sunlight position, and an LDR to sense whether it is day or night. As depicted in the prototype design, for the light sensing part of the system, we used four similar photodiode sensors each connected in series with a resistor, creating a Wheatstone bridge circuit. They were positioned in the shadow counter system as shown below, in order to detect the intensity of the light in all four orientations (up, down, left and right) (Fig. 3).

The structure helps in isolating all four photodiodes, thus, depending on the position of light source, one of the four photodiodes will receive higher insolation. Illustrated in Figs. 4 and 5, when one part of the quadrant is receiving sunlight (sensor1), the opposite part is being shadowed (sensor2). Thus that photodiode will result in

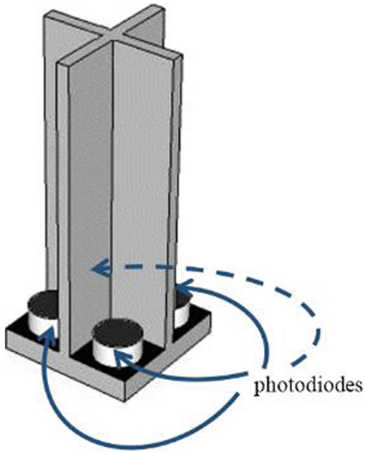


Fig. 3. Photodiodes

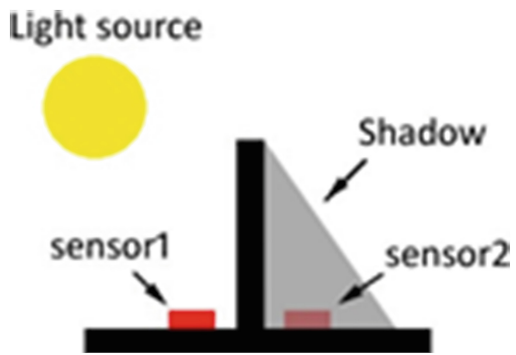


Fig. 4. Shadow counter-type illustration

a lesser voltage drop allowing us to know where to orient the PV panel for highest irradiance. Since the photodiodes are attached to the solar panel, the values recorded by them will keep varying while the panel is moving, until a tolerance value for each of them is sensed. The Fig. 5 shows the final prototype.

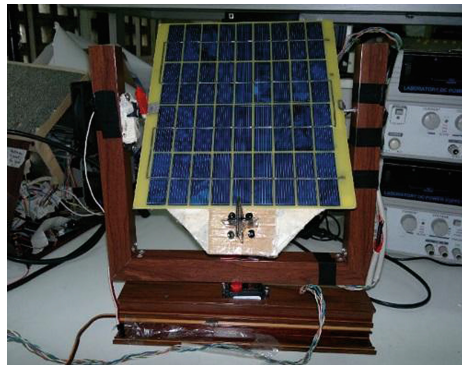


Fig. 5. Final prototype

### 3.6 Experimental Setup

In order to measure the performance of the solar tracker, two PV systems were set up. One panel was used for the tracking system and the other panel of the same dimension and capacity was used as a fixed (flat lying panel) system. The tests were

performed at the GPS location ( $-20.222215, 57.542560$ ), in Mauritius. The solar tracking prototype as well as the fixed panel, were installed on the roof of a building, which was free from shades. They were placed side by side, and the testing was performed between the 24<sup>th</sup>–29<sup>th</sup> December 2015. The microprocessor enabled both functions – tracking and GSM communication with the user, depending on certain conditions. The microcontroller features were tested first, and the response of the panels were recorded.

## 4 Results and Discussion

Random text messages and miscalls were sent to the solar tracker to test if it would enter the ‘PARK’ mode. It was observed that these did not affect the tracker in any way. When the panel was suddenly removed from the frame; a miscall and text message was automatically sent to the user to inform him that the panel has been tampered with. Also, when in the ‘PARK’ mode, the LDR was blocked to make the system ‘think’ that it is night. It was observed that the tracker stayed in the same position until the message “R” was sent by the user. The power supply to the servo motors was cut for about five hours, and when the power was ‘ON’ again, the tracker could determine the new position of the sun, and moved accordingly to face the sun directly. When the panel was touched or shaken, it did not trigger the alarm system; the alarm was only activated when the panel was pulled from the frame. Another test was done on the wirebreak to see how it will deal with the wire breaking and reconnected again (a situation may arise whereby a thief will try to reconnect the wirebreak system to stop the tracker from sending the miscall and alert message to the user) and we observed that the tracking stopped, and the miscall and alert message was still sent to the user (Table 2).

### Output Power Measurements

On the 28<sup>th</sup> December 2015, (sunny day), the microprocessor was initialized for tracking and the voltage as well as the current measurements were recorded every 30 min, on both panels, from 09:00–16:00. The output power was calculated, and a comparative graph of output power against time, with and without the tracking system was plotted, as shown in Fig. 6. The experimental results indicate that the proposed dual-axis tracker generates a higher power output as compared to a fixed panel. The energy generated is calculated from the area under the corresponding curve. The results show that the output of the dual-axis solar tracker is 25.1% higher than that of the flat lying fixed panel. Also it is noticed that at around noon, when the sun is directly above the ground, both systems (fixed and tracking) generates almost the same amount of power. At all other time of the day, the tracking system generates a significantly higher amount of energy.

**Table 2.** Testing microcontroller features

Testing of functions	User action	Response of PV panels
PARK	SMS sent by user over GSM network	Panels moved to 'PARK' position which offers minimum wind loading. Confirmation SMS was sent back to user by the microcontroller
RESUME	SMS sent by user over GSM network	Panels resumed tracking. Confirmation SMS was sent back to user by the microcontroller
NIGHT	The panels were left in the 'TRACKING' mode	At around 18:35, when light intensity diminished considerably, panels automatically moved to a preconfigured position to wait for sunrise. An SMS was sent to the user to signal its position. Panels stayed in the same position during the whole night. Next morning, at around 6:30, the LDR detected sufficient light, and panels started moving, and tracking started automatically. SMS confirmation was sent to user, informing that 'TRACKING' had started
TRACKING	No action taken by user	Depending on the shadowing on the photodiodes (which are separated in four quadrants), the panels moved in such a direction so as to favor maximum sunlight intensity on all four photodiodes. For three minutes, the panel remained stationary, then started moving again searching for the maximum sunlight intensity
WIREBREAK	The Panel were moved from its frame at night	A misscall was sent to the user, and after a few seconds, a text message was also sent to the user
	The Panel were moved from its frame during the day	Tracking stopped, a misscall was sent to the user, and after a few seconds, a text message was also sent to the user



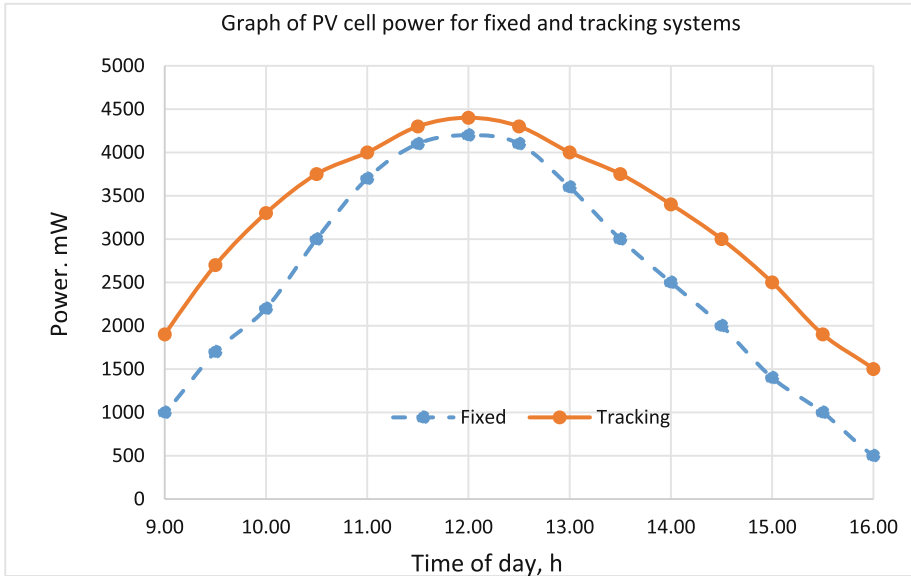


Fig. 6. Power output for fixed panel and the dual-axis solar tracker

## 5 Conclusion

Tests performed revealed that the energy outputs of the panel with dual-axis tracking system were 25.1% higher than the fixed panel. The security features were also tested, and the panels behaved as expected. The wire-break alarm circuit used almost negligible amount of energy and was a very cheap solution compared to Closed Circuit Television (CCTV) systems. It was noted that the experimental results were close to those recorded by many other research works [11, 12]; however, the actual values would be slightly lesser as the tracking system also consumed energy in driving the motors and powering its controller. The prototype can be further improved by uploading a solar positioning algorithm available off the shelf to the microcontroller to allow it to position the motors more accurately. Further works can be implemented to improve the sensors arrangements separated by quadrants; a system can be devised to prevent water from reaching the photodiodes. A higher rating solar panel can be used to power the controller and servo motors.

## References

1. Bazilian, M., Onyeji, I., Liebreich, M., MacGill, I., Chase, J., Shah, J., Gielen, D., Arent, D., Landfear, D., Zhengrong, S.: Re-considering the economics of photovoltaic power. *Renew. Energy* **53**, 329–338 (2013)
2. Bushong, S.: Advantages and disadvantages of a solar tracker system. <http://www.solarpowerworldonline.com/2016/05/advantages-disadvantages-solar-tracker-system.htm>. Accessed 20 Aug 2016
3. Mauritius, A.R.: Renewable energy agency bill – Electricité: seuls 12% produits de sources renouvelables. *Le Défi*, p. 5, 31 Aug 2015
4. Central electricity board: let there be light, annual report 2013, Central Electricity Board, Mauritius (2013)
5. Mousazadeh, H., Keyhani, A., Javadi, A., Mobli, H., Abrinia, K., Sharifi, A.: A review of principle and sun-tracking methods for maximizing solar systems output. *Renew. Sustain. Energy Rev.* **13**(8), 1800–1818 (2009)
6. Barsoum, N., Vasant, P.: Simplified solar tracking prototype. *Global J. Technol. Optim. (GJTO)* **1**, 38–45 (2010)
7. Putting a trace on solar PV panels. <http://www.semi.org/en/anti-counterfeiting-task-force-tackle-putting-trace-solar-pv>. Accessed 15 Aug 2016
8. A group of PV thieves arrested in Germany. [https://www.pv-magazine.com/2016/03/23/a-group-of-pv-thieves-arrested-in-germany\\_100023864/#ixzz4J7Ych7sQhttp://www.pv-magazine.com/news/details/beitrag/a-group-of-pv-thieves-arrested-in-germany\\_100023864/#ixzz4J7Xi1UMW](https://www.pv-magazine.com/2016/03/23/a-group-of-pv-thieves-arrested-in-germany_100023864/#ixzz4J7Ych7sQhttp://www.pv-magazine.com/news/details/beitrag/a-group-of-pv-thieves-arrested-in-germany_100023864/#ixzz4J7Xi1UMW). Accessed 11 Aug 2016
9. Making sense of light sensors EE times. [http://www.eetimes.com/document.asp?doc\\_id=1272314](http://www.eetimes.com/document.asp?doc_id=1272314). Accessed 22 Aug 2016
10. Agarwal, T.: Basics of phototransistors - types, features and applications. In: *Electronics, ElProCus - Electronic Projects for Engineering Students* (2013)
11. Chong, K.K., Wong, C.W.: General formula for one axis sun- tracking system and its application in improving tracking accuracy of a solar collector. *Solar Energy* **83**, 298–305 (2009)
12. Lewis, G.: Optimum tilt of solar collectors. *Solar WingTech.* **4**, 407–410 (1987)

# A Low-Cost Autonomous Cleaning System for Photovoltaic Arrays

Mohammad Fardeen Islam, Vishwamitra Oree<sup>(✉)</sup>,  
and Anshu Prakash Murdan

Electrical and Electronic Engineering Department,  
University of Mauritius, Reduit, Mauritius  
{v.oree, a.murdan}@uom.ac.mu

**Abstract.** In recent years, installation of photovoltaic arrays has increased significantly worldwide driven by attractive incentives and reduced prices. However, the efficiency of the photovoltaic panels is negatively impacted by deposits of dust and dirt on their surface. Manual cleaning is highly labor-intensive, costly and uses excessive amounts of water. To address these shortcomings, an autonomous system is proposed to clean photovoltaic panels. It is based on two interconnected motorized structures that move along the surface of the PV panels to ensure that all sections are cleaned. A sensor network provides feedback on various aspects of the cleaning process. Tests showed that the proposed system can result in a gain of about 22% in daily average power output after three weeks of operation compared to a normal photovoltaic system.

**Keywords:** Microcontroller · Feedback · Motorized · Photovoltaic

## 1 Introduction

Solar photovoltaic (PV) installations worldwide have been soaring during the last ten years, with the cumulative installed capacity growing at a mean yearly rate of about 49% [1]. A combination of technological developments, drops in prices and government subsidies has driven this rapid growth. While the investment on a solar PV array can be significant, its payback period will depend essentially on its energy output over time. Several factors affect the power generated by the array, including the efficiency of the modules, the orientation of the panels and the ambient temperature. The accumulation of dust, dirt and debris also adversely influences the performance of the array, particularly in arid environments [2]. It has been reported that the accumulation of dust on PV arrays reduced the power output by up to 50% [3]. It is difficult to characterize dust deposition and its associated impact on the system performance given that it is a complex phenomenon affected by diverse site-specific environmental and meteorological parameters [4].

It is therefore important to monitor the cleanness of the PV panel surfaces and eliminate buildup of dust. Manual cleaning of PV arrays is a tedious task, is not cost-effective in the long run and uses large amounts of water. Several automatic alternatives exist in the market nowadays. Their designs focus on performing the

cleaning of the panels as fast as possible. One innovative cleaning technique, known as electrostatic cleaning technology, applies an electrostatic charge material on a transparent plastic sheet on the surface of the solar panels [5]. When sensors detect that the accumulation of dust has exceeded a pre-set level, the material is energized. An electrically charged wave cascades through the covering material on the surface and lifts the dust away from the surface. Nevertheless, this technology requires special material along with special strong and flexible structures to withstand the vibrations caused by the waves [5].

Another widespread cleaning methodology uses sprinkler systems. These mechanisms simulate rainfall by spraying fine jets of water mixed with soap solution on the surface of the PV panels. The system requires a pumped water supply through pipes mounted on the top of the PV panels, a water filtration system and a soap dispensing system [6]. While the system is simple to implement, it has some limitations. It uses a large amount of water and does not have a feedback system to determine whether cleaning was done effectively. Furthermore, it requires regular refilling of soap solution. Recently, a large number of robotic systems have been developed to minimize the need for human intervention in PV array cleaning. One of the most popular large-scale PV array cleaning robot is the Ecoppia E4, an autonomous and water-free system that uses powerful microfiber brushes to remove dust [7]. Mounted on a metallic frame, the brushes spin downwards through gravity, generating an intense flow of air that blows the dust off from the surface. The main issue with robotic systems is that they are expensive.

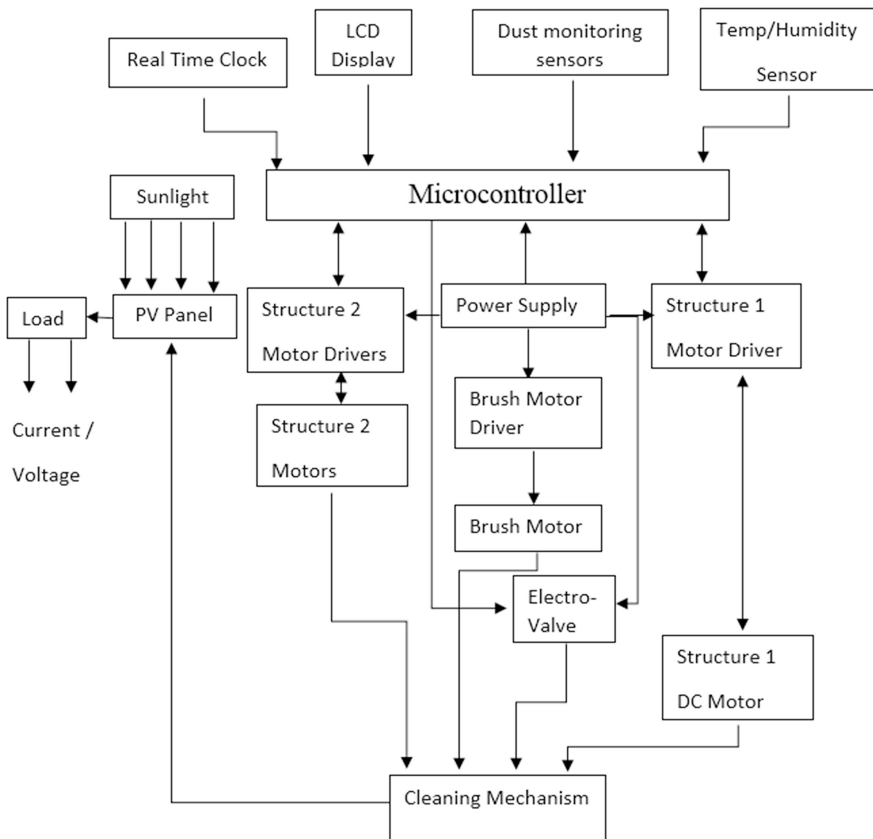
This paper proposes a PV array cleaning system that addresses the shortcomings identified in the designs described above. In addition to being low-cost and fully autonomous, it integrates a feedback system that indicates whether all sections of the PV array surface have been cleaned properly. Moreover, the mechanism can operate on the surface of a PV array at different angles. Finally, its low cost implies that it does not require a large-scale array to be cost-effective like most commercially available robotic cleaning systems.

## 2 Methodology

The purpose of installing a cleaning mechanism in PV arrays is to ensure that the performance and efficiency of the modules do not decline as dust, dirt and debris accumulate on their surface. The device should therefore be cost-effective so that the investment on it is recovered through gains in the panel output power within a reasonable period of time. Other than the cost implications, the main guiding principles for the design include operation with minimum human intervention, optimized cleaning, robustness to harsh environment such as strong winds and efficient use of water. The use of water for cleaning PV arrays is also beneficial to the efficiency as lower temperatures resulting from the cooling effect of water generally enhances the current-voltage characteristics of solar cells. A study reported that the conversion efficiency of amorphous silicon and crystalline silicon solar cells decrease by about 0.2% and 0.4% per degree rise in temperature respectively [8].

### 2.1 System Block Diagram

Figure 1 shows the block diagram of the proposed design. At the heart of the system is a microcontroller, which reads and processes information from various sensors to monitor the operation of the device. The Arduino Due R3, a single-board microcontroller is used here. Arduino is an open-source prototyping platform that provides its users with many facilities, in particular, easy-to-use hardware and software [9]. It can be easily interfaced with a host of sensors and allows users to control different outputs and components. Its simplicity, low-cost and user-friendly platform make the Arduino ideal for this application. Infra-Red sensors are placed at strategic locations on the mechanism in order to monitor the extent of dust accumulation on the surface of the PV panel. Based on the values read by the sensors, different motors are actuated by the microcontroller to implement the cleaning function.



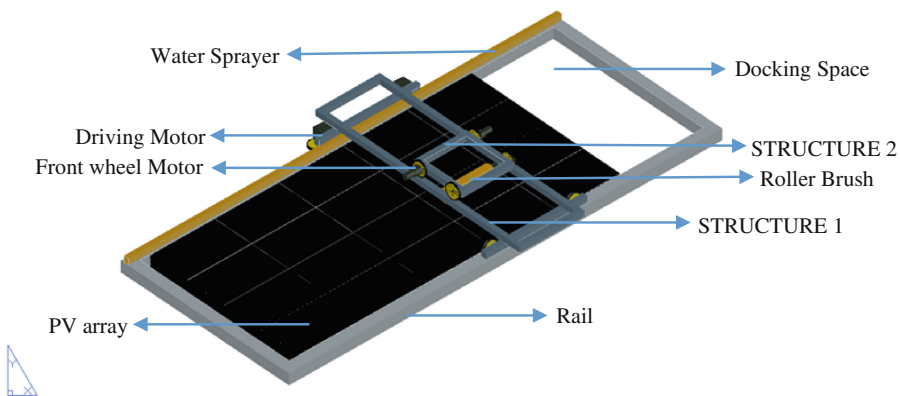
**Fig. 1.** Block diagram of the autonomous PV panel cleaner.

The Arduino activates electro-valves to control the flow of water on the panel and facilitate the cleaning process. A rain sensor is used to keep track of the conditions of the surrounding environment. Detection of rain may preclude activation of the

electro-valves, thus avoiding water wastage. Limit switches are used for the detection of extremities of the PV array so as to prevent the mechanism from moving past the side edges. A Real-Time Clock enables the microcontroller to keep track of the current time for the system to start its operation. Moreover, a DHT11 sensor is fixed on the surface of the PV array to record its temperature. If the temperature is higher than  $35^{\circ}\text{C}$  during the day, jets of water are sprayed intermittently by activating the electro-valves at regular intervals for short periods. Cooling of the solar panels is performed to curtail degradation of the efficiency. Recorded values of important system parameters are displayed on an LCD display to provide information about the status of operation. Finally, a dedicated power supply with adequate autonomy is required for the system so as to ensure autonomous operation.

## 2.2 Structure of Cleaning Mechanism

The proposed design consists of two different structures as illustrated in Fig. 2. The rectangular and larger Structure 1 moves horizontally across the panel on guide rails and wheels installed at the top and bottom of the PV array. Structure 2 is mounted on Structure 1 and moves vertically up and down the latter through two motors located on the upper wheels so that the cleaning brush fixed on Structure 2 covers the whole PV panel. Both structures are driven by stepper motors that move them accurately as required by the microcontroller. Structure 2 also accommodates a roller brush that spins with the help of a DC motor. Coupling the motion of the two structures in a sequenced manner enables the brush to evenly reach every section of the PV panel surface. The cleaning process is undertaken at night only so as not to affect the efficiency of the PV panel which is optimal during the day. A pipe through which water flows is placed along the upper guiding rail. Small jets of water can be sprayed onto the panel by controlling the current through electro-valves placed at regular intervals in the pipe. The designs of the mechanical drives, structures and their movement through rails and wheels ensure reliability and durability of the system. Only the brush is in contact



**Fig. 2.** Cleaning system design with two main structures.

with the surface of the panels. Besides, the brush used is made of microfibers so that the protective surface of the PV panels is not scratched during cleaning. A docking space is provided for the mechanism so that the latter does not affect the output of the PV panels by casting a shadow on them when it is not in operation.

### 2.3 Detection of Dust

Most existing mechanisms used to clean PV panels are programmed to start at a specific time of the day without determining whether the panel actually needs cleaning or not. Such a mode of operation is wasteful in terms of water use and power required to operate the system. Thus, it may defeat the whole purpose of the cleaning system which is to enhance the output power of the PV panel. To address this issue, IR sensors are placed at carefully chosen locations on Structure 1. The sensor emitter sends an IR beam on the panel and depending on the cleanliness of the panel surface, different amounts of light can be detected by the sensor receiver. In contrast to clean surfaces which reflect light unidirectionally, the presence of dirt directs light in different directions resulting in less light reaching the receiver. Figure 3 illustrates this phenomenon for the PV array. The microcontroller reads the inputs of the IR sensor each time before the structure moves horizontally across the PV array. The microcontroller is programmed to start spinning the brush only if any of the IR sensors detects a voltage less than a threshold value determined through calibration experiments. This ensures that the mechanism cleans only locations on the panel that are dirty. For example, bird droppings may affect only limited areas of the panel and may not require brushing the remaining areas if they are already clean.

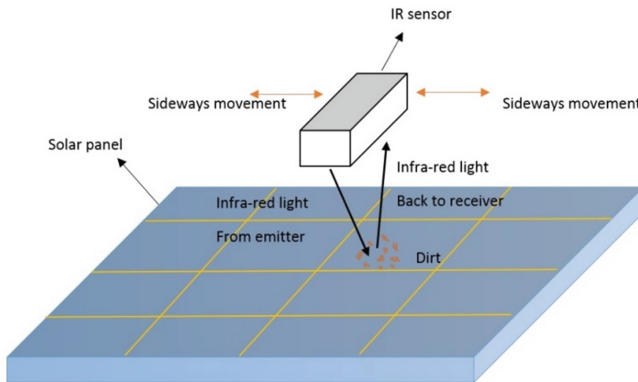


Fig. 3. Operation of IR sensor for providing cleanliness feedback.

### 2.4 System Operation

When the system is started, the Structure 1 moves one step from right to left of the PV array. During the displacement, the IR sensors scan the surface to detect dust accumulation above the pre-set threshold. If no significant dust layer is detected, Structure 1

moves one more step to the left. Else, the water sprayer is turned on allowing water to flow down the surface of the array at that position and the brush spins to start the cleaning process. Structure 2 then moves downwards to clean the whole surface area enclosed within Structure 1. This process is repeated until the leftmost edge of the PV array is detected when the system goes back to the docking area where it waits for the next activation signal. The latter is generated when the RTC time is equal to the set time for cleaning. The detailed operational flowchart of the cleaning process is shown in Figs. 4 and 5.

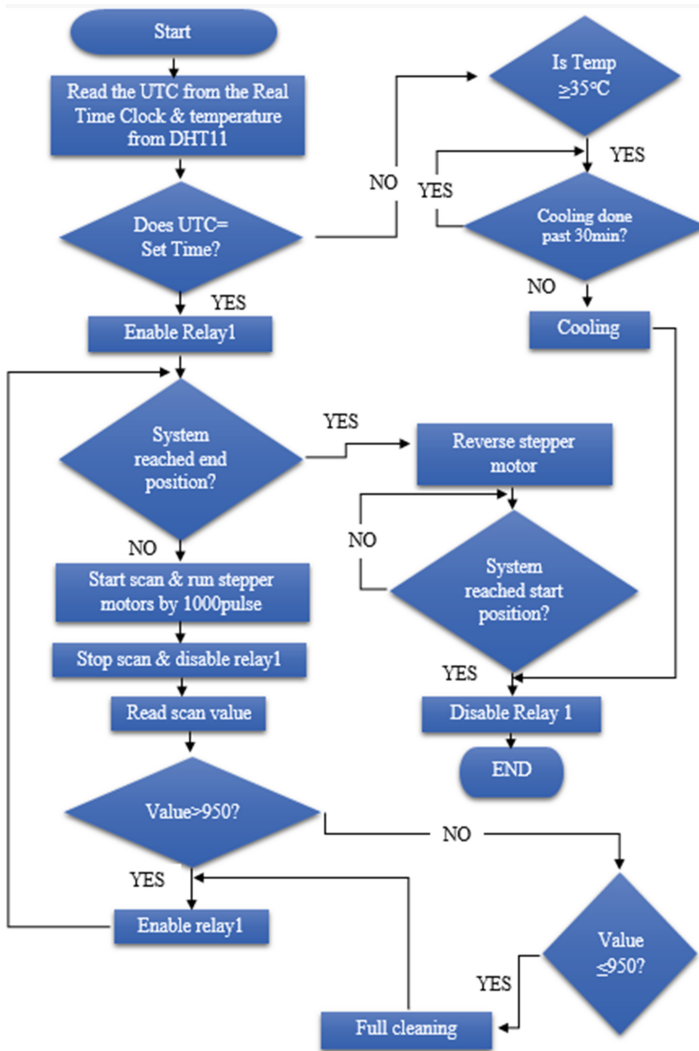


Fig. 4. Cleaning system operational flowchart – Part 1.



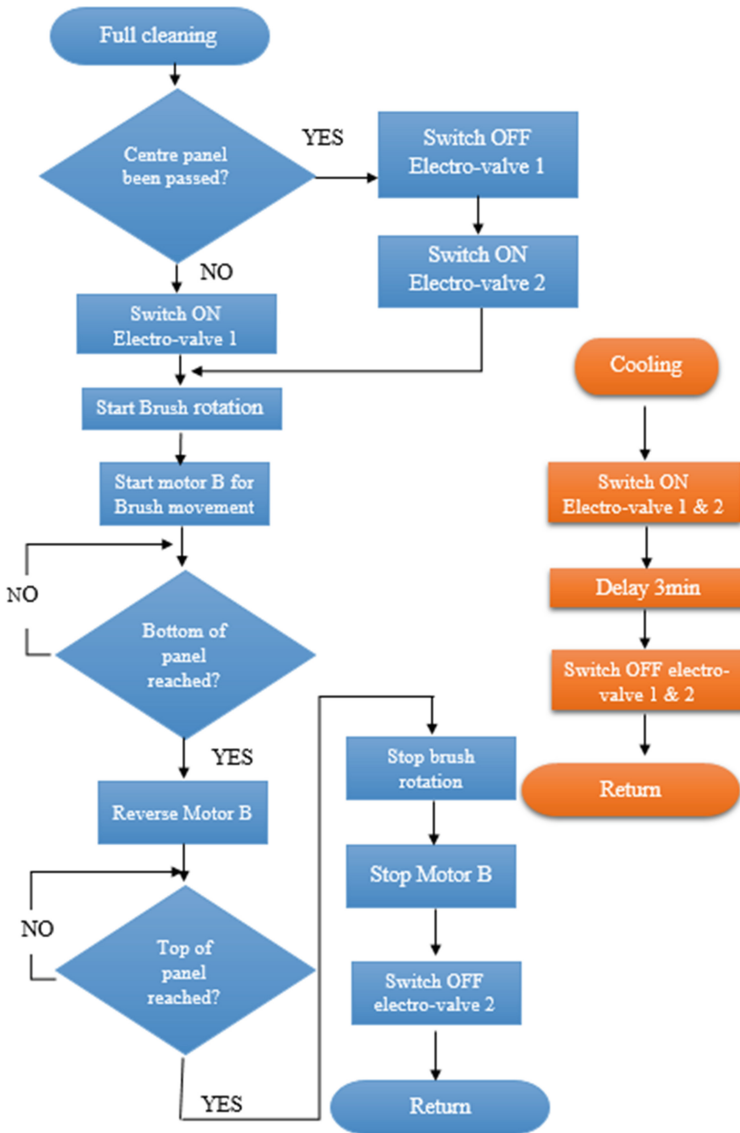


Fig. 5. Cleaning system operational flowchart – Part 2.

The prototype of the low-cost autonomous PV array cleaning system is shown in Fig. 6. The microcontroller together with all other signal conditioning and interfacing electronic circuitry is housed in a metallic box with the LCD screen fitted on the cover of the box.



**Fig. 6.** Prototype of the low-cost autonomous PV array cleaner.

### 3 Results and Discussion

The prototype was tested on a 200 kW panel exposed to the environment for two weeks. The cleaning system was exceptionally programmed to operate during the day so as to be able to visualize the difference between the cleaned section and the uncleaned one. Figure 7 shows a photograph taken during the cleaning process after a single pass of Structure 2. The cleaning effectiveness can be clearly distinguished.

The improvement in the energy output of the PV system with the prototype was also evaluated against a baseline PV panel placed near to it, while ensuring the same inclination. Both systems were tested during a period of three weeks from the 13<sup>th</sup> March to 3<sup>rd</sup> April 2016. The location was carefully chosen to ensure that no shade or other disturbances affected the PV panels. A load in the form of a heating element was connected to the panels and the output voltage and current were recorded at intervals of 15 min between 11.00 and 14.00 h during four days every week. The results are given in Table 1. During Week 1, cleaning of the test panel was performed only on Friday night but no cooling operation was performed. No cleaning and cooling operation was performed during Week 2. However, it can be observed that the average power recorded by the end of the week was stable at around 10 to 12% better than that for the baseline case. Week 3 was characterized by windy conditions resulting in an increase in dust accumulation on the panel. Two cleaning operations were performed on Sunday and Friday nights while cooling operations were effected twice on Saturday as temperature prevailing on that particular day was beyond the threshold value. By the end of the testing period, it was observed that average daily power output from the test panel exceeded that from the baseline by about 22%.



**Fig. 7.** Operation of the prototype showing the cleaned right section of the PV panel after one pass of Structure 2 while the left section remains dirty.

**Table 1.** Average daily values of temperature, power and solar radiation during test period.

	Average temperature/ $^{\circ}\text{C}$	Average power (W) prototype	Average power (W) baseline	Solar radiation/ $(\text{W}/\text{m}^2)$	% increase in power output
Monday Wk 1	32	0.437	0.434	1011	0.293
Thursday Wk 1	33	0.459	0.448	1118	2.52
Saturday Wk 1	35	0.473	0.441	1129	7.24
Sunday Wk 1	33	0.460	0.428	1096	7.53
Monday Wk 2	28	0.371	0.338	854	10.0
Thursday Wk 2	35	0.478	0.433	1134	10.4
Saturday Wk 2	32	0.446	0.397	1073	12.3
Sunday Wk 2	33	0.456	0.428	1069	11.4
Monday Wk 3	35	0.491	0.424	1130	15.9
Thursday Wk 3	35	0.481	0.413	1101	16.4
Saturday Wk 3	36	0.484	0.402	1108	20.4
Sunday Wk 3	34	0.474	0.388	1074	22.1

The total cost of the system was 350 USD. For large-scale PV arrays, additional costs will be incurred on longer guide rails at a cost of about USD 8 for each additional panel.

## 4 Conclusions

The accumulation of dust on PV panels in dry and arid regions calls for an efficient cleaning system to enhance the power yield. A low-cost autonomous PV cleaning system is proposed in this paper that can be easily extended to large-scale arrays at minimal costs. The system consists of a vertically moving motorized structure accommodating a microfiber brush, mounted on another horizontally moving motorized structure. Jets of water are released by electro-valves located in a pumped piping system to facilitate the cleaning process and to cool the panels. IR sensors are fitted on the structures to provide feedback of the cleaning efficiency. Tests carried during three weeks revealed that the proposed system provided a gain of 22% in daily average power output compared to that of a baseline PV panel with no cleaning.

The proposed system can be further improved by making it lighter so that its power consumption is reduced. In addition, it can be designed in a scalable manner to enable easy assembly for arrays of different sizes. Further improvements will consist of integrating remote monitoring and control as well as self-diagnostic features.

## References

1. International Energy Agency. World Energy Outlook 2013 (2013)
2. Elminir, H.K., Ghitas, A.E., Hamid, R.H., El-Hussainy, F., Beheary, M.M., Abdel-Moneim, K.M.: Effect of dust on the transparent cover of solar collectors. *Energy Convers. Manag.* **47**, 3192–3203 (2006)
3. Adinoyi, M.J., Said, S.A.: Effect of dust accumulation on the power outputs of solar photovoltaic modules. *Renew. Energy* **60**, 633–636 (2013)
4. Mani, M., Pillai, R.: Impact of dust on solar photovoltaic (PV) performance: research status, challenges and recommendations. *Renew. Sustain. Energy Rev.* **14**, 3124–3131 (2010)
5. Jalbuena, K.R.: Mars-inspired technology makes PV panels self-cleaning (2010). <http://www.ecoseed.org/technology/13801-mars-inspired-technology-makes-pv-panels-self-cleaning>
6. Heliotex: Automatic solar panel systems (2016). <http://www.solarpanelcleaningsystems>
7. Ecoppia Empowering Solar Website: Ecoppia E4 (2015). <http://www.ecoppia.com/ecoppia-e4>
8. Luque, A., Hegedus, S.: Handbook of Photovoltaic Science and Engineering, 2nd edn. Wiley, Hoboken (2011)
9. Arduino Website. Arduino: Introduction (2016). <https://www.arduino.cc/en/Guide/Introduction>

# Voltage Control of a Power System with Exciter and Generator Saturation

R. Ramjug-Ballgobin<sup>(✉)</sup> and S.G. Calchand

Faculty of Engineering, University of Mauritius, Reduit, Mauritius  
r. ramjug@uom.ac.mu

**Abstract.** This paper presents the control of voltage of a saturated power system using an Automatic Voltage Regulator (AVR) model. An unsaturated model of the system is first considered and then, this initial model is altered to include the effects of exciter and generator saturations. Since the amplifier has saturation limits in practice, these are also considered in the saturated system. The individual effect of the two types of saturations are analysed followed by an analysis of their combined effect. In order to obtain satisfactory responses in terms of settling time, percentage overshoot and steady state error, three types of compensating schemes were implemented, namely, Proportional Integral Derivative (PID), Fuzzy controller and PI-Fuzzy controller.

**Keywords:** Voltage control · AVR · Saturation · PID · Fuzzy

## 1 Introduction

For smooth operation of electrical appliances, their input voltage should remain within some prescribed values. This is achieved in power distribution network through the use of voltage controllers. The main causes of malfunction of the appliances is due to factors like overheating of generators and motors, reduction in transmission losses and to maintain the ability of the system to withstand and prevent voltage collapse. Voltage fluctuation is a major problem that is encountered at the consumers' end. This is due to the load variation on the supply power system. When load increases on the supply side on the supply system, voltage decreases at the consumers' terminal as an increase in voltage drop is experienced in the power system components and vice versa when the load is decreased [1]. Electrical equipment are designed to operate within a specific range of voltage and frequency. Hence the performance of the appliances is adversely affected and this may damage the latter if there is a prolonged operation of the equipment outside the permissible range of voltage [2]. Therefore the terminal voltage of all equipment needs to be controlled and maintained within a specified range [1]. The principal method of generator reactive power control is through generator excitation control using AVR [3]. Automatic Voltage Regulator (AVR) is an important equipment in an electric power system, whose purpose is to maintain the magnitude of the terminal voltage of a synchronous generator within the required limit by its field excitation control [4]. The main function of AVR is to control the voltage magnitude of the synchronous generator at a specific limit [3]. A basic AVR model consists of an amplifier, exciter, generator, sensor and a comparator as shown in Fig. 1 [2]. When a

step response is applied to the system without any control, this produces some oscillations and consequently the system response is reduced. In the past, voltages were controlled manually but this actually takes a lot of time which is enough to damage a particular equipment [5]. The introduction of controllers allow the mentioned problems to be solved. Proportional Integral Derivative (PID) is a type of controller which is used the most in AVR system [6]. Fuzzy Logic Controller are the other types of controllers that can be used in order to enhance the performance of the AVR system in terms of a decrease in settling time, percentage overshoot and steady state error. Using Fuzzy Logic Controller voltage error problems are effectively and quickly reduced. Therefore the control of the power system operation is easy to handle and this does not cause any specific damage to the equipment [5]. The main objective was to design controllers as this would enhance the system performance and response by reducing the settling time, percentage overshoot and the steady state error.

### 1.1 PID Controller

A Proportional Integral Derivative (PID) controller is among the most widely used of controllers in control systems [7]. The main reason for its wide use is that it performs for a wide range of process. Moreover for a wide variations of operating conditions, a robust performances can be obtained and it is easy to use [4]. The main role of the PID controller is to adjust the input to the systems so that the system response is a desired one. Thus we can tune it to make the system faster, reduce oscillations or eliminate steady state errors. This is done by PID as follows, it calculates the error value which is the difference between the measured process variable and a set point. It also improves the response and eliminate the steady state error [8]. This controller has three parameters namely the proportional (P), Integral (I) and Derivative (D) [8]. The purpose of the derivative is that it adds a finite zero to the open-loop transfer function and improves the transient response. The integral adds a pole at the origin and increases the type of the system. Hence to a step input, the steady state error will be eliminated [9]. The three parameters are then added together and these values are used to adjust the process via a control element [8]. The equation of the PID in the Laplace domain is given by

$$U_c(s) = \left[ k_p + k_i \frac{1}{s} + k_{DS} \right] E(s) \quad (1)$$

### 1.2 Fuzzy Logic Controller

Fuzzy logic systems are appropriate for approximate reasoning. Their response is much better, faster and smoother compared to conventional systems [10].

There are three processes that are involved in the design of a fuzzy logic system. They are the fuzzification interface, fuzzy inference system and the defuzzification [10].

1. Fuzzification: It changes the crisp logic to fuzzy logic. The fuzzy logic is split into 5 or 7 parts [5]. The fuzzification method used is the Mamdani one.
2. Fuzzy Inference system

In this section the fuzzy inputs are mapped on the corresponding fuzzy outputs related to their membership. The membership function used were triangular-shaped one. The Model with 5 MF with 25 rules were used (Table 1).

**Table 1.** Rule-based output Fuzzy rules for 5 MF of the FLC

	LN	SN	Z	SP	LP
LN	LN	LN	LN	SN	Z
SN	LN	LN	SN	Z	SP
Z	LN	SN	Z	SP	LP
SP	SN	Z	SP	LP	LP
LP	Z	SP	LN	LP	LP

Where

- LN: Large Negative
- SN: Small Negative
- Z: Zero
- SP: Small Positive
- LP: Large Positive

### 3. Defuzzification

In order to obtain the crisp output a defuzzification must be used. The defuzzification used in this paper is the center of gravity, commonly known as the centroid formula which is given the equation below

$$u(t) = \frac{\int \mu_1(x) \cdot x \, dx}{\int \mu_1(x) \, dx} \tag{2}$$

where,

- $\mu_1(u)$  is the aggregated membership function
- $x$  is the output variable

### 1.3 PI-Fuzzy Controller

The PI-Fuzzy controller uses a fuzzy and a PI controller together to improve the control action. The Fuzzy Logic Controller is used to tune the integral gain,  $k$  of the AVR system. An AVR system with PI control scheme was implemented [11].

## 2 Method

An AVR system model was considered for the voltage control. The unsaturated model, model with exciter and generator saturations individually and both were analysed in order to observe their initial response. Different types of controllers such as PID, Fuzzy and PI-Fuzzy were designed so that a comparative analysis later on can be performed. The systems were implemented in block diagrams in the s-domain and then simulation was performed. The controllers designed were tuned using the control system toolbox to obtain the required response.

### 2.1 Systems

The system response was analysed in terms of its settling time, percentage overshoot and steady state errors. The AVR models which was used to analyse the system response were, unsaturated AVR model system, AVR model with exciter saturation, AVR model with generator saturation and controllers added to the mentioned AVR model. The main objective was to see which controller enhanced the system response.

#### 2.1.1 Unsaturated Models

The AVR system proposed by [2] was used. A unit step input was applied at the input. Figure 1 shows a basic model of an AVR. The block diagrams were represented in the form of transfer function when simulated. The transfer function and the system parameters used were simulated using the values as show in Table 2.

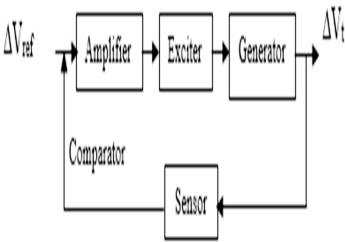


Fig. 1. Basic AVR model

Table 2. Parameters values

	Transfer function	Gain constant	Time constant
Amplifier	$\frac{K_A}{1 + sT_A}$	$K_A:10$	$T_A:0.1$
Exciter	$\frac{K_E}{1 + sT_E}$	$K_E:1$	$T_E:0.4$
Generator	$\frac{K_G}{1 + sT_G}$	$K_G:1$	$T_G: 1$
Sensor	$\frac{K_S}{1 + sT_S}$	$K_S:1$	$T_S:0.01$

#### 2.1.2 AVR with Exciter Saturation

In this section an exciter saturation was added to the system. This can be added be implemented to the system by means of an exciter saturation block negatively fed back to the exciter block diagram and the exciter saturation is represented by the equation below. The derivation is shown in [12]. The exciter saturation is represented by

$$S_{E_{max}} = (8.6455 \times 10^{-4})e^{1.5046E_{fd}} \tag{3}$$

where  $E_{fd}$  is the exciter output voltage.



### 2.1.3 AVR with Generator Saturation

Generator saturation,  $S_G$  is represented by the block diagram having the following equation

$$S_G = 0.0478e^{6.037(v_t - 0.8)} \quad (4)$$

where,  $v_t$  is the terminal voltage of the AVR.

A block diagram representing the equation of the generator saturation as shown above in Eq. (2) was negatively fed to the generator field block. The derivation is shown in [12].

## 2.2 PID Controller

A PID controller was first added to the AVR system. Matlab/Simulink inbuilt PID Controller was used. The following PID parameters were used in order to get the desired response. Initially the following gains were used,  $K_p = 1.00$ ,  $K_I = 0.28$  and  $K_D = 0.25$ , [3]. Using the automatic tuning feature whereby a percentage overshoot of less than 5% and a settling time of less than 3 s was set, the new gains obtained were  $K_p = 0.123$ ,  $K_I = 0.107$  and  $K_D = 0.0103$ .

## 2.3 Fuzzy Controller

A fuzzy controller was then designed. In this case as well since the inbuilt block for fuzzy was available in Matlab/Simulink, the same block was used. "Fuzzy" was typed in the command window which open the FIS editor where FIS was created. The latter defines all of the parameters of the fuzzy block. For the Fuzzy controller, the gains depend within a range where the control variables exist at any instant. The value of the gains are manually tuned until the fuzzy control appears satisfactory.

## 2.4 PI-Fuzzy Controller

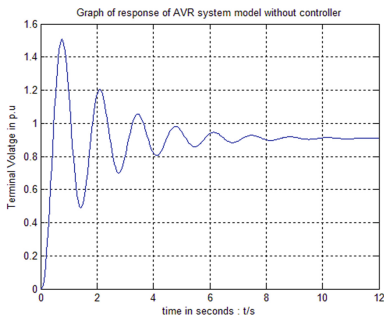
The same procedure was repeated as that of fuzzy controller used to design the fuzzy rules and the fuzzy membership functions. The PI controller was connected to the input of the Fuzzy controller and hence the response of the AVR was analysed. The value of the gains are adjusted manually until a desired response is obtained.

## 3 Result

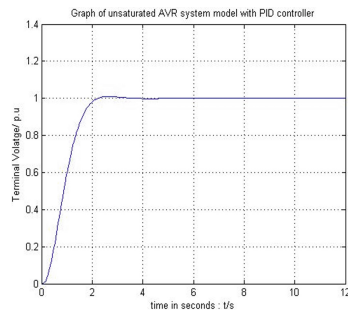
The systems were analysed taking the reference input voltage set at unit step response that is at 1pu. Initially the system without any controller was analysed. The system's response with different kinds of controllers were then studied and analysed (Figs. 2, 3, 4, and 5).

**Table 3.** Settling time, Percentage overshoot and Steady state error of the uncompensated AVR model and with PID Controller

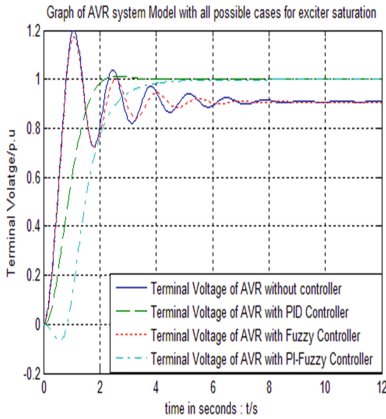
	Case	Model of AVR without any controller	Model of AVR with PID controller	Model of AVR with Fuzzy controller	Model of AVR with PI-Fuzzy controller
Settling time/s	Unsaturated	6.9874	2.6833	4.771	3.741
	Exciter saturation	6.017	2.007	4.918	3.811
	Generator saturation	4.553	2.165	4.154	3.334
	Both saturation	5.226	2.191	4.175	3.404
Percentage overshoot/%	Unsaturated	65.72	0	42.1	0
	Exciter saturation	31.93	1.0257	28.9	0
	Generator saturation	30.27	0	27.5	0
	Both saturation	29.70	0	27.1	0
Steady state error	Unsaturated	0.0909	0	0.0927	0
	Exciter saturation	0.0912	0	0.0931	0
	Generator saturation	0.0934	0	0.0957	0
	Both saturation	0.0933	0	0.0961	0



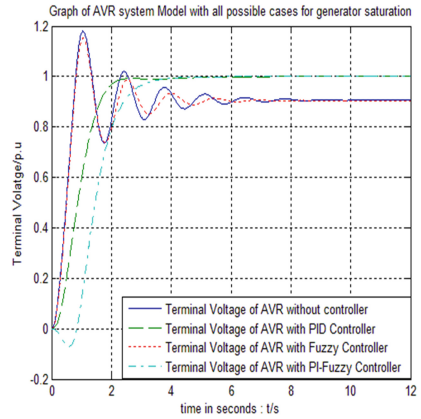
**Fig. 2.** Response of unsaturated AVR system without controller



**Fig. 3.** Response of unsaturated AVR system with PID Controller



**Fig. 4.** Response of AVR system with exciter saturation for all possible cases



**Fig. 5.** Response of AVR system with generator saturation for all possible cases

## 4 Discussion

### 4.1 Unsaturated and Uncompensated AVR System Model

It can be observed that from Table 3, for the unsaturated system, the settling time is 6.874 s and the percentage overshoot is 65.72%. Upon addition of the exciter and generator saturation function, it could be noted that this damped the system response. This reduced both the settling time and percentage overshoot. When the saturation was treated individually, exciter saturation had a greater damping effect than generator saturation. Since the value of settling time, percentage overshoot and steady state error were quite significant, these were needed to be reduced. One method of overcoming this problem is to design controllers and study the effect of the system response to see whether the response is a desired one or not.

### 4.2 Model of AVR with PID Controller

When a PID controller is added to the system, the settling time is decreased to 2.6833 s whereas the percentage overshoot and the steady state error were completely eliminated for the saturated case. Similar observations were observed with the saturated system, only for exciter saturation a percentage overshoot of 1.03% is noted. Hence it can be deduced that with the addition of the PID controller, this has indeed enhanced the system response.

### 4.3 Model of AVR with FUZZY Controller

When a fuzzy controller is added to the AVR system model, the settling time decreased to 4.771 s while comparing it with the unsaturated model. However the percentage

overshoot was quite significant and the steady state error also was present. Hence, this response of the system was not the desired one.

#### 4.4 Model of AVR with PI-Fuzzy Controller

When the PI-Fuzzy was implemented in the system, both the settling time as well as the percentage overshoot had reduced. Steady state error as well was completely eliminated with this controller. However, PID had the smaller settling time compared to PI-Fuzzy. The purpose of the integral is the PI-controller was to decrease the steady state response. Since the error in the steady state was nearly zero, the controller output equals the output of the integral of the PI-controller. This is what explained for the negative part in the terminal AVR voltage.

#### 4.5 Effect of Exciter Saturation of the System Response

The exciter saturation function  $S_E$  has an exponential function and is related to the exciter voltage. This block is taken as a negative feedback path and hence, as exciter voltage increases it is the saturation effects which determines the outcome of the output voltage.

#### 4.6 Effect of Generator Saturation of the System Response

The saturation function  $S_G$  which was added to the generator block as well has an exponential function as the latter is related to generator terminal voltage. This block is taken as a negative feedback path and hence, as generator terminal voltage increases it is the saturation effects which determines the outcome of the output voltage. The generator saturation had a greater damping effect compared to exciter saturation.

## 5 Conclusion

From the simulation results obtained, it could be observed that the desired response was obtained with the PID controller. The steady state error was completely eliminated and a percentage overshoot of less than 2% was obtained. The settling time was smaller with PID controller compared to the other types of controllers that were used in this paper. Hence this was the most appropriate response obtained among all which were tested. However, it could also be seen that it is not necessary for a particular controller to give the best control action in all situations. According to different systems and different situations, controllers can give different kinds of performances. Overall, it was observed that the PID controller had enhanced the system performance compared to the AVR model without controllers and fuzzy controller.

## References

1. Sivanagaraju, S., Shreenivasan, G.: Power System operation and controls, p. 421. Pearson Education, New Delhi (2010)
2. Prasad, L.P., Gupta, H.O., Tyagi, B.: Application of policy iteration technique based adaptive optimal control design for automatic voltage regulator of power system. *Electr. Power Energy Syst.* **63**, 940–949 (2014)
3. Saadat, H.: Power System Analysis, pp. 555–565. McGraw-Hill Education, Singapore (2004)
4. Dastranj, M.R., Rezaei, M., Rad, P.S.: Design FOPID control of an automatic voltage regulator (AVR) system imperialist competitive algorithm. *Int. J. Softw. Eng. Appl.* **8**, 143–152 (2014)
5. Srinivasa Rao, T.C., Ponnala, R., Subramanyam, T.C., Srinivas, N.: Frequency error and voltage control by using pi and fuzzy logic controllers for multi area inter connected power system. *Int. J. Comput. Appl.* **77**, 15–23 (2013)
6. Mukherjee, V., Ghoshal, S.P.: Intelligent particle swarm optimized fuzzy PID controller for AVR system. *Electr. Power Syst. Res.* **77**, 1689–1698 (2007)
7. Stephanopoulos, G.: Chemical Process Control: An Introduction to Theory and Practice, pp. 70–75. PTR Prentice Hall, Upper Saddle River (1984). 07632
8. Sreeraj, P.V.: Design and Implementation of PID Controller with Lead Compensator for Thermal Process. *Int. J. Comput. Appl.* **67**, 26–31 (2013)
9. Vasanthi, S., Gopila, M., Gnanambal, I.: Fuzzy and PID excitation control system with AVR in power system stability analysis. *Int. J. Eng. Adv. Technol.* **1**, 95–99 (2012)
10. Ugale, C.P., Dhumale, R.B., Dixit, V.V.: DC-DC converter using fuzzy logic controller. *Int. Res. J. Eng. Technol.* **2**, 593–596 (2015)
11. Loukianov, A.G., Sanchez, E., Lizarde, C.: Force tracking neural block control for an electro-hydraulic actuator via second-order sliding mode. *Int. J. Robust Nonlinear Control* **18**, 319–332 (2007)
12. Anderson, P.M., Fouad, A.A.: Power System Control and Stability, pp. 425–432. The Iowa State University Press, Ames Iowa (1997)

# Estimation of Solar Photovoltaic Parameters Using Pattern Search Algorithm

M. Derick<sup>1</sup>, C. Rani<sup>1</sup>(✉), M. Rajesh<sup>1</sup>, K. Busawon<sup>2</sup>, and R. Binns<sup>2</sup>

<sup>1</sup> School of Electrical Engineering, VIT University, Vellore, India  
derick.mathew2016@vitstudent.ac.in,  
{crani,mrajeshkumar}@vit.ac.in

<sup>2</sup> Faculty of Engineering and Environment, Northumbria University, Tyne, UK  
{krishna.busawon,richard.binns}@northumbria.ac.uk

**Abstract.** The interest towards solar Photovoltaic (PV) based power generation has increased worldwide due to climate change and depletion of fossil fuels. This has led to the need for accurate solar PV modelling under different environmental conditions. Since solar PV shows nonlinear characteristics, optimization technique is the best tool for modelling and estimation of PV parameters. Thus, in this paper, a Pattern Search (PS) algorithm is proposed to optimize the parameters of solar photovoltaic panels. The objective is to identify the parameters of single diode model based PV, in such a way that the difference between PV experimental current and simulated current is minimal. The effectiveness of the algorithms is investigated through simulation in MATLAB/Simulink environment at different solar irradiance and temperature and it is compared with the experimental data of solar module Kyocera – KC200GT 215. Results clearly reveal that the proposed technique shows better results in terms of its accuracy, convergence and CPU execution time.

**Keywords:** Modelling · Pattern search · Single diode PV module

## 1 Introduction

Today, due to climate change and the continuous decline of conventional fuel sources, the usage of renewable and inexhaustible energy sources is gradually increasing. Among them solar energy seems to be the best option. In India 140% of growth is expected in utility scale, among them 80% in states of Tamil Nadu, Telangana and Karnataka. In addition to that, Government of India decided to revise the National Solar Mission target to increase the grid connected solar power project from 20,000 MW to 100,000 MW by the year 2021–22 [1]. So, large scale PV plants are used for power generation that is fed to the grid. However, due to its high initial capital cost and poor panel efficiency it is not popular among the middle class population. With regards to practical implementation, the solar photovoltaic system should be optimized before its installation. This can be assured by precise modelling and simulation of PV module. The modelling of solar PV is generally done using equivalent diode models to describe the current-voltage (I-V) relations over a wide range of temperatures and solar irradiances. In the literature, one can find two models to describe the I-V characteristics of

solar PVs: single diode model and double diode model [2, 3]. On one hand, the single diode model is quite simple in its modelling and several researchers have verified its accuracy by analyzing and testing its five unknown parameters. On the other hand, the double diode model provides more accuracy but is quite complicated, with increased number of parameters.

In this work, we focus on the single diode model of the solar PV due to its simplicity. The parameters of the solar PV vary with temperature and irradiance. Hence, precise estimation of the parameters is required to model the solar cell accurately. Popular approaches or techniques employed for parameter estimation are categorized as analytical techniques [4, 5], numerical extraction [6, 7] and artificial intelligent techniques [8, 10].

In the analytical technique, mathematical equations are used to find the parameters. Most of the values in the equation are not provided in manufacturer datasheet. As a result, this method is not deemed accurate [8]. Numerical extraction technique is considered to be an accurate method compared to the analytical method since it uses the measured data for curve fitting. However, the application of curve fitting to the nonlinear equations of a diode is difficult, and as a result numerical extraction approach is not so popular.

On the other hand, artificial intelligence techniques are considered as excellent in dealing with nonlinear equations. In the recent years, different optimization techniques have been introduced to estimate the parameters of solar PV; namely, the Genetic Algorithm (GA) [8], Pattern Search (PS) optimization [9], Artificial Immune System (AIS) [10], Bacterial Foraging Algorithm (BFA) [12], Simulated Annealing (SA) [11], and Differential Evaluation (DE) [13].

In this work we focused on the PS optimization technique to estimate the solar PV parameters. An outline of the paper is as follows: In Sect. 2, the mathematical modelling of solar PV is presented. Section 3 gives the details of the proposed methodology for solar PV parameter estimation. The discussion of simulation and results of PV module using Matlab-Simulink is presented in Sect. 4. Finally, conclusions are given in Sect. 5.

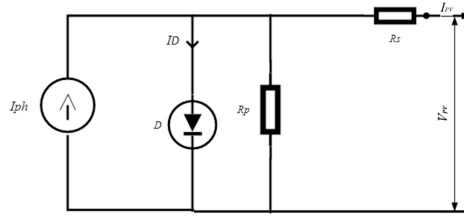
## 2 Mathematical Modelling

Many model have been proposed and developed by several researchers to estimate the solar PV parameters accurately. Among them, the most popular models are single diode and double diode model. In our work, we used single diode model to represent the behavior of solar PV due to its simplicity and accuracy. The equivalent circuit of solar PV is shown in Fig. 1.

By using Kirchhoff's current law, solar output current  $I_{PV}$  can be write as given below:

$$I_{PV} = I_{ph} - I_D - I_P \quad (1)$$

Here  $I_{ph}$  is the photon current generated by the incident light,  $I_D$  is the diode current and  $I_P$  is the shunt current flowing through  $R_P$ .



**Fig. 1.** Single – diode solar PV equivalent circuit

$$I_{PV} = I_{ph} - I_o \left( \exp \left( \frac{V_{PV} + I_{PV} R_s}{N_s V_t} \right) - 1 \right) - \left( \frac{V_{PV} + (I_{PV} R_s)}{R_p} \right) \quad (2)$$

Here  $I_o$  is the reverse saturation current of diode and  $N_s$  is number of series cell per module,  $R_p$  is the parallel resistance which exist mainly due to the leakage current of p-n junction and depends on the fabrication methods of the solar PV cell.  $R_s$  is the series resistance, resembles contact resistance of metal with P and N layers.  $V_t$  is thermal voltage of diode which depends on junction temperature and is given by:

$$V_t = a \frac{KT}{q} \quad (3)$$

Here  $a$  denotes diode ideality factor,  $T$  expresses junction temperature in Kelvin (K),  $q$  is electron charge ( $1.6021765 \times 10^{-19}$ C), and  $K$  is the Boltzmann constant ( $1.38065 \times 10^{-23}$  J/K). The cell temperature  $T$  and irradiance  $G$  have great influence on photon current, so  $I_{ph}$  can be found using the following equation.

$$I_{ph} = (I_{ph-s} + K_I(T - T_S)) \frac{G}{G_S} \quad (4)$$

Here  $I_{ph-s}$  is the photon current at standard test conditions ( $T_S$  is the temperature at standard test condition  $T_S = 25$  °C, and  $G_S$  is solar radiation at standard test condition  $G_S = 1000$  W/m<sup>2</sup>), and  $K_I$  is the current temperature coefficient.

The photon current at standard test conditions is given by:

$$I_{ph-s} = I_{SC-s} \frac{R_p + R_s}{R_p} \quad (5)$$

Here  $I_{SC-s}$  is the short circuit current at standard test conditions. The reverse saturation current  $I_o$  can be written as follows [14];

$$I_o = \frac{I_{SC-s} + K_I(T - T_S)}{\exp \left( \frac{V_{OC-s} + K_V(T - T_S)}{N_s V_t} \right) - 1} \quad (6)$$



Here  $V_{OC-S}$  is the open circuit voltage at standard test condition,  $K_V$  is the voltage temperature coefficient.

The objective function that is given in Eq. (2) does not have direct solution because it is an implicit function of  $a, R_s, R_p, I_{PV}$  and  $I_o$ . So, single diode model requires the computation of above said five parameters. There are two prospects in finding the above parameters: (i) the first option is to find all these parameters using optimization technique, (ii) the second option is finding  $a, R_s, R_p$  using optimization technique and calculate  $I_o, I_{ph}$  using the above Eqs. (4) and (6). In our work, the second approach is adopted to find optimized values of ideality factor, series and parallel resistances using Pattern Search optimization technique, in order to minimize computation complexity. The objective function is to optimize the average absolute error between the measured value and the calculated value. The average absolute error ( $E_a$ ) is described as:

$$E_a = \frac{1}{n} \sum_{i=1}^n abs(I_{PV(i)(m)} - I_{PV(i)(c)}) \quad (7)$$

Here  $n$  is the number of different points taken into account,  $I_{PV(i)(m)}$  is the measured solar PV current and  $I_{PV(i)(c)}$  is calculated solar PV current. Absolute error at  $i^{\text{th}}$  point E is given by:

$$E = f_i(V_{PV(m)}, I_{PV(m)}, X) \quad (8)$$

Here X is the vector representing the model parameters. The optimization function is chosen as:

$$f_i(V_{PV(m)}, I_{PV(m)}, X) = abs(I_{PV(m)} - (I_{ph} - I_D - I_p)) \quad (9)$$

To model the solar PV, the proposed optimization technique is tested using the values such as open circuit voltage  $V_{OC}$ , short circuit current  $I_{SC}$ , voltage temperature coefficient  $K_V$  and current temperature coefficient  $K_I$ .

### 3 Methodology

Pattern Search algorithm is used to find the optimized parameters of solar PV. The algorithm is organized to examine a set of points which is near to the current position intended to realize a point that gives a smaller objective function value than at the current position. It finds the solution for non-continuous, not differentiable and multi model, multiple local optima. It shows an adaptable and balanced operation to upgrade the global and local search. The flowchart of pattern search is represented in Fig. 2.

The main steps of pattern search algorithm are Initialization, Starting and Pattern movement and are explained as follows.

**Initialization.** In this process all the constants provided in the manufacturer datasheet, expansion factor  $\alpha$ , contraction factor  $\beta$  are employed with proper selection of starting point  $X_0$ .

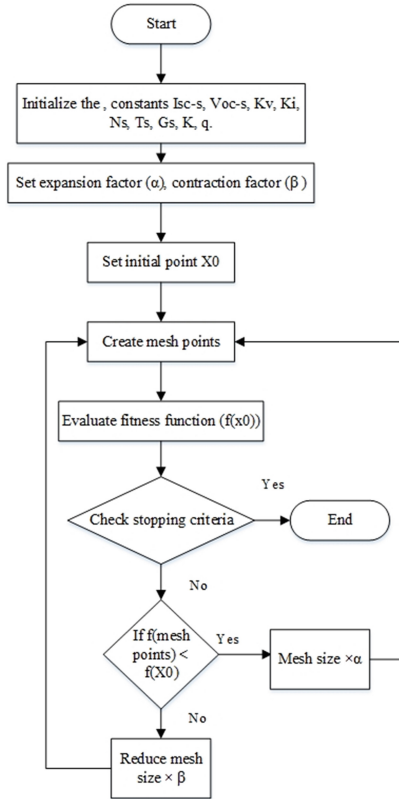


Fig. 2. Flowchart of Pattern search

**Starting.** This process is begun by creating a set of points around the starting point  $X_0$  called mesh points. They are added with  $X_0$  to form a pattern. The mesh points are a set of vectors and mesh size is a scalar. If the mesh size = 1, then the mesh points are given by (1 0), (0 1), (-1 0), (0 -1) and the corresponding patterns are given by  $X_0 + [1\ 0]$ ,  $X_0 + [0\ 1]$ ,  $X_0 + [-1\ 0]$ ,  $X_0 + [0\ -1]$ . The respective fitness function value of each mesh point is computed. If a point with fitness value less than fitness value of  $X_0$  is found, then it is considered as a successful poll and the new mesh point is set as  $X_1$ .

**Pattern Movement.** In the second iteration, the current mesh size is multiplied by the expansion factor  $\alpha$  then the corresponding patterns now becomes  $\alpha * [1\ 0] + X_1$ ,  $\alpha * [0\ 1] + X_1$ ,  $\alpha * [-1\ 0] + X_1$ ,  $\alpha * [0\ -1] + X_1$ . For an unsuccessful poll, the current point will remain the same; that is  $X_1 = X_0$ . The mesh size is multiplied by contraction factor  $\beta$ , so mesh size will be smaller. This process is continued until an optimal solution of objective function is found within the set tolerance.

### 4 Simulation and Result Analysis

Matlab-simulink model is developed to test the proposed optimization techniques. The proposed optimization technique is available in the global optimization toolbox of MATLAB has been used for minimizing the objective function. The data given in [15] which is experimentally measured values of multicrystal PV module Kyocera – KC200GT 215 using Gaobo GSMT – H- 3A 100 solar module is used for determining the error. According to the variations in the solar PV, the lower and upper constraints used in optimization technique will also vary. The ideality factor  $a$  has a value between 1 and 2. The series resistance  $R_s$  generally has less value varies from  $0.015 \Omega$  to  $0.5 \Omega$ , whereas the parallel resistance  $R_p$  has higher values between  $600 \Omega$  to  $1000 \Omega$ .

The expansion factor, Mesh contraction factor and number of iterations are set as given in Table 1. After using the optimization toolbox of PS, simulation of the PV module is done using MATLAB-Simulink to obtain all five optimized parameters. Figure 3 shows solar PV simulation model of a PV equivalent circuit, implemented using single diode model. The Matlab optimization toolbox is used for several trials for the best convergence of the objective function.

**Table 1.** Values setting in PS

Expansion factor	2
Mesh contraction factor	0.5
The maximum number of iterations	300

Table 2 indicates the values of  $a$ ,  $R_s$  and  $R_p$  for different number of iterations which gives the best fitness value for the objective function.

Normally absolute error at constant current region will be less, but error in constant voltage region is comparatively more. Therefore, the accuracy and efficiency of the method will be decided by evaluating the error at the voltage source region. Solar PV module under different environmental conditions are simulated and it is compared with experimentally measured data. Figure 4 shows the error at different irradiance and different temperature for the proposed method. From this figure it is concluded that PS shows more accurate, fast and consistent solution.

**Table 2.** Parameter values obtained by simulation.

No of iterations	$a$	$R_s(\Omega)$	$R_p(\Omega)$
300	1.7	0.0339	624.3828
500	1.546	0.1002	683.6521
800	1.620	0.0976	702.8256

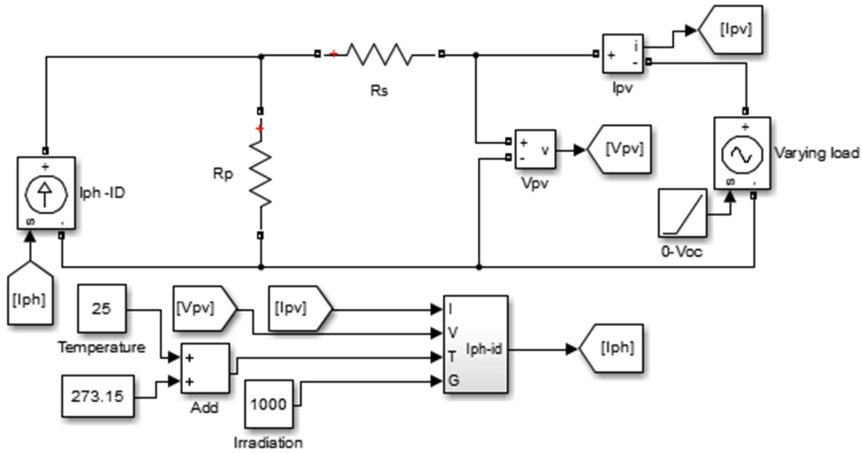


Fig. 3. The PV system simulated by Matlab/Simulink

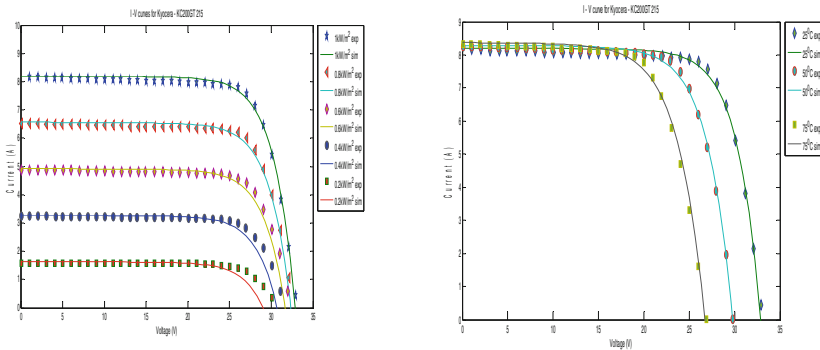


Fig. 4. I-V curve at different irradiance and temperature

### 5 Conclusion

In this paper, Pattern Search optimization technique is used for estimating PV module parameters. The solar PV module was simulated in Matlab-Simulink environment by using the parameters obtained through pattern search optimization technique. Finally, the obtained set of parameters are used to evaluate the absolute error between simulated and measured data. The results obtained, clearly reveals that pattern search technique gives better results in terms of its accuracy and convergence speed.

## References

1. Ministry of New and Renewable Energy. Scaling up of grid connected solar power projects from 20000 MW by the year 2021–22, to 100000 MW by the year 2021–22 under National solar mission. No. 30/80/2014-15/NSM, July 2015
2. Farivar, G., Asaei, B.: A new approach for solar module temperature estimation using the simple diode model. *IEEE Trans. Energy Convers.* **26**, 1118–1126 (2011)
3. Barth, N., et al.: PV panel single and double diode models: optimization of the parameters and temperature dependence. *Sol. Energy Mater. Sol. Cells* **148**, 87–98 (2016)
4. Chan, D.S.H., Phang, J.C.H.: Analytical methods for the extraction of solar – cell single- and double – diode model parameter from I-V characteristics. *IEEE Trans. Electron. Devices* **34**, 286–293 (1987)
5. De Soto, W., Klein, S.A., Bechman, W.A.: Improvement and validation of a model for photovoltaic array performance. *Sol. Energy* **80**(1), 78–88 (2006)
6. Wolf, P., Benda, V.: Identification of PV solar cells and modules parameters by combining stational and analytical methods. *Sol. Energy* **93**, 151–157 (2013)
7. Ishaque, K., Salam, Z., Taheri, H., Shamsudin, A.: A critical evaluation of EA computation methods for photovoltaic all parameter extraction based on two diode model. *Sol. Energy* **85**(9), 1768–1779 (2011)
8. Ismail, M.S., Moghavvemi, M., Mahila, T.M.I.: Characterization of PV panel and global optimization of its model parameter using genetic algorithm. *Energy Convers. Manag.* **73**, 10–25 (2013)
9. Ye, M., Wang, X., Yousheng, X.: Parameter extraction of solar cells using particle swarm optimization. *J. Appl. Phys.* **105**(9), 094502 (2009)
10. Rajasekar, N., et al.: Application of modified particle swarm optimization for maximum power point tracking under partial shading condition. *Energy Procedia* **61**, 2633–2639 (2014)
11. El-Naggar, K.M., Al Rashidi, M.R., Al Hajri, M.F., Al-Othman, A.K.: Simulated Annealing algorithm for photovoltaic parameter identification. *Sol. Energy* **86**(1), 266–274 (2012)
12. Rajasekar, N., Kumar, N.K., Venugopalan, R.: Bacterial foraging algorithm based PV parameter estimation. *Sol. Energy* **97**, 255–265 (2013)
13. Da Costa, W.T., et al.: Identification of photovoltaic model parameters by differential evolution. 2010 IEEE International Conference on Industrial Technology (ICIT). IEEE (2010)
14. Villalva, M.G., Gazoli, J.R., Ruppert Fiho, E.: Comprehensive approach of modelling and simulation of photovoltaic array. *IEEE Trans. Power Electron.* **24**, 1198–1208 (2013)
15. Pauls, C.: Optimization approaches for parameter estimation and maximum power point tracking (MPPT) of photovoltaic systems. Dissertation University of Liverpool (2014)

# An Improved Reconstruction Method for Compressively Sampled Magnetic Resonance Images Using Adaptive Gaussian Denoising

Henry Kiragu<sup>(✉)</sup>, George Kamucha, and Elijah Mwangi

School of Engineering, University of Nairobi,  
P.O. BOX 30197-00100, Nairobi, Kenya  
hkiragu@mmu.ac.ke,  
{gkamucha, elijah.mwangi}@uonbi.ac.ke

**Abstract.** In this paper, an improved Compressively Sampled Magnetic Resonance Imaging (CS-MRI) method that suppresses reconstruction noisy artifacts is proposed. The algorithm involves random undersampling of the k-space data of an MR image followed by reconstruction of the k-space data coefficients in a wavelet sparsifying domain. The high frequency noise in the reconstructed coefficients is suppressed in the Fourier transform domain by an adaptive Gaussian low pass filter. The reconstructed MR image is finally obtained by Inverse Discrete Fourier Transformation (IDFT) of the denoised k-space data. Experimental results demonstrate the robustness of the proposed method to sub-Nyquist sampling associated artifacts in terms of terms of Structural SIMilarity (SSIM) index and Peak Signal to Noise Ratio (PSNR) assessments.

**Keywords:** Gaussian denoising · Compressive sampling · MRI · SSIM · PSNR

## 1 Introduction

Compressive Sampling (CS) is a signal acquisition and reconstruction paradigm that is used to undersample a signal and then reconstruct its full-length version in a suitable transform domain. The procedure relies on the signal sparsity as well as the incoherence between measurement and representation orthonormal bases [1, 3].

Magnetic Resonance Imaging (MRI) is a non-invasive medical imaging method that results in better medical images than other imaging modalities in terms of soft tissue contrast. The MRI technique is non-invasive and also uses non-ionizing radiation unlike Computed Tomography (CT). The MRI technique can also be used to image some conditions such as brain oxygen saturation changes due to neuronal activity, measurement of blood flow velocity and measurement of concentration of metabolites [2]. The data acquisition stage in conventional MRI takes a long time which makes it difficult for weak patients and children to remain still. The patient movements result in noisy artifacts in the acquired MR images [1, 4, 5]. Magnetic Resonance (MR) images are acquired in a coded form and also are sparse in some transform domains such as

Fourier or wavelet transform domains. These characteristics make MRI a suitable candidate for CS techniques [6, 8].

A CS-MRI simulation that makes use of wavelet domain k-space data is proposed by Liu et al. [9]. Their method guarantees good reconstruction quality but requires major modifications of the existing MRI equipment to allow acquisition of wavelet encoded k-space data. Zangen et al. [5] have proposed a compressive sampling based MRI reconstruction method. Simulation results demonstrate that the method reconstructs the edges and also reduces undersampling-related artifacts. However, their algorithm requires long computational time and is therefore unsuitable for real time imaging. Akanksha [10] proposes an Orthogonal Matching Pursuit (OMP) based CS image reconstruction method that uses multi-wavelet transformation. Experimental results show good visual quality in the reconstructed image. The image acquisition process is however quite long and also requires major modifications in the design of MRI equipment. Tesfamicael and Barzideh [11] have proposed an algorithm that involves clustered CS. The method produces better results than its non-clustered versions but the reconstructed images are noisy since it requires a sampling ratio of more than 0.7 to reconstruct an image that has a PSNR of about 20 dB. In [12], a CS method that utilizes sparsity and energy distribution of the MR images in spectral space is proposed. The method yields better performance than random sampling Fourier transform-based CS methods but results in speckle noise when a large number of approximation coefficients are discarded.

In this paper, a CS-based MRI method that exploits the sparsity as well as the profile of Fourier coefficients is proposed. The method is robust to reconstruction noise due to incorporation of Gaussian filtering. It is faster than other CS-MRI methods due to the low sampling ratios required for a given image quality. For example, results show that the method requires a sampling ratio of 0.3 to reconstruct an image with a PSNR of about 20 dB compared to a sampling ratio of 0.7 required to produce the same PSNR using the method proposed in [11]. Also, the method does not require modifications in the design of the MRI equipment since image acquisition is performed in the Fourier transform domain as in conventional MRI.

The rest of this paper is organized as follows: Sect. 2 gives the requisite theoretical background. Section 3 presents the proposed MRI algorithm while Sect. 4 presents some of the results obtained using the proposed algorithm as well as other methods. Conclusion and future work are given in Sect. 5.

## 2 Background Theory

The theory on CS, MRI, and image quality metrics is outlined in this section.

### 2.1 Compressive Sampling

Compressive Sampling (CS) is a technique that obtains a full-length sparse signal from a measurement vector whose length is much smaller than the signal length. To compressively sample an  $N$ -length signal  $f$ ,  $M$  linear measurements are taken. The acquisition process can be modeled as:

$$y = \Phi f \quad (1)$$

where  $y$  is a vector of length  $M \ll N$  and  $\Phi$  is an  $M \times N$  measurement matrix [1, 12].

In an orthonormal sparsifying domain, the  $N$ -length signal  $f$  can be written as:

$$f = \sum_{i=1}^N x_i \psi_i = \Psi x, \quad (2)$$

where  $\Psi$  is an  $N \times N$  representation matrix and  $x$  is an  $N$ -length vector of coefficients of the signal in the  $\Psi$  domain [1]. The measurement process can be expressed as:

$$y = Ax, \quad (3)$$

where  $A = \Phi\Psi$  is an  $M \times N$  sensing matrix that obeys the Restricted Isometry Property (RIP) [1]. The matrices  $\Phi$  and  $\Psi$  should be incoherent [13, 14].

Optimization and greedy methods are used to recover the reconstructed image tractably. The Optimization methods used are such as the Basis Pursuit (BP) and the Least Absolute Shrinkage and Selection Operator (LASSO). The greedy methods include the Orthogonal Matching Pursuit (OMP) and the Stage-wise OMP (StOMP). These methods are faster than the optimization ones for highly sparse signals [1, 7].

## 2.2 Magnetic Resonance Imaging

Clinical Magnetic Resonance Imaging (MRI) uses a strong longitudinal magnetic field and Radio Frequency pulses to generate high quality images. Other than yielding better soft tissue contrast than Computed Tomography [CT], MRI is non-ionizing, non-invasive as well as non-allergic to the patient. The MRI equipment integrates the applied time-dependent gradient transverse magnetization over the spatial volume to yield the MRI signal  $f(k)$  given by:

$$f(k) = \int |m(z)| e^{-j2\pi(k \cdot z)} dz, \quad (4)$$

where  $z \in \mathbb{R}^3$ ,  $m(z)$  is the transverse magnetization,  $f(k)$  is the Fourier transform of the spatially dependent magnetization  $m(z)$  that is sampled on the frequency curve  $k$ . This Discrete Fourier Transform (DFT) of the MRI image is its  $k$ -space data from which the MR image is obtained by inverse Fourier transformation [1, 2, 7].

## 2.3 Image Quality Measures

In this section, the PSNR, Mean Squared Error (MSE) and SSIM image quality metrics are discussed. The PSNR and MSE of an image of size  $m \times n$  pixels are respectively defined as follows:



$$\text{PSNR} = 10 \log_{10} \left( \frac{mnL^2}{\sum_{x=1}^m \sum_{y=1}^n [f - g]^2} \right), \quad (5)$$

$$\text{MSE} = \frac{\sum_{x=1}^m \sum_{y=1}^n [f - g]^2}{mn}, \quad (6)$$

Where  $f$  is the ground-truth image,  $g$  is the reconstructed image and  $L$  is the maximum pixel intensity value of the ground-truth image. The MSE and PSNR are simple and convenient to compute but they do not match well with the characteristics of the human visual system (HVS) [15]. The Structural SIMilarity (SSIM) index which is consistent with the HVS measures the visual quality of the reconstructed image by extracting the luminance, contrast, and the structural components of the original and the reconstructed images and then comparing the components. With each the three components assigned the same relative importance, the SSIM index of the reconstructed image  $g$  is defined as:

$$\text{SSIM}(f, g) = \frac{(2\mu_f\mu_g + K_1)(2\sigma_{fg} + K_2)}{(\mu_f^2 + \mu_g^2 + K_1)(\sigma_f^2 + \sigma_g^2 + K_2)}, \quad (7)$$

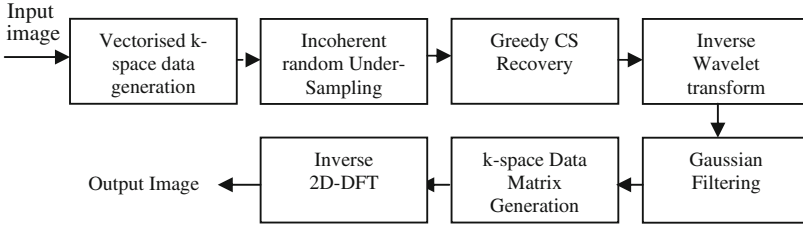
where  $f$  and  $g$  are the ground-truth and the reconstructed images respectively. Their means are  $\mu_g$  and  $\mu_f$  while  $\sigma_g$  and  $\sigma_f$  are their standard deviations. The correlation between  $f$  and  $g$  is  $\sigma_{fg}$ . The constants  $K_1$  and  $K_2$  are included to avoid instability when the values of the means and variances of the images are close to zero [15, 16].

### 3 Proposed Method

In this section, a proposed CS-MRI method is presented. The algorithm constitutes of the stages shown if Fig. 1. The input MR image is first converted into its k-space data by obtaining its two-dimensional Discrete Fourier transform (2D-DFT) followed by rearranging the 2D-DFT coefficients of the MR image by moving the zero-frequency component to the center of the array. The zero-frequency centered DFT matrix is converted into a vector and then randomly undersampled. Next, the MR image coefficients are reconstructed in the wavelet transform domain using the OMP algorithm. The wavelet coefficients are then converted back into vectorized k-space data through Inverse wavelet transformation before being denoised by an adaptive Gaussian low-pass filter. The denoised data is converted into a matrix which is ultimately converted into the output MR image through inverse DFT.

The zero-frequency centered vectorized k-space data coefficients of an N-length ground-truth MR image has a Gaussian-like profile centered about the centre coefficient index. The CS recovery of the data using the OMP method results in high frequency noisy artifacts whose level decreases as the sampling ratio increases.

The adaptive Gaussian filter  $w(i)$  included in the proposed method has a bandwidth that adaptively increases with increase in the sampling ratio ( $M/N$ ) as follows:



**Fig. 1.** Block diagram of the proposed procedure.

$$w(i) = e^{-k(i-\frac{N}{2})^2} \quad (8)$$

where  $i$  is the DWT coefficients index and  $k$  is a parameter whose value was experimentally approximated by:

$$k(M/N) = a(M/N - b)^3 \quad (9)$$

where for  $N = 2048$ ,  $a = 1.5 \times 10^{-5}$ ,  $b = 1$ , and  $M/N$  is the sampling ratio.

## 4 Simulation Results

In this section, performance of the proposed method in comparison with both the OMP and StOMP is presented using MATLAB simulation results. Twenty MR test images were used in the experiments with the results of five of them included in this section. The five images include three (Spine, Knee and hand)  $64 \times 32$  pixels images and two (heart and brain)  $32 \times 32$  pixels images. Figure 2 shows the effect of the adaptive Gaussian low pass filtering on the k-space data of the reconstructed spine image at a sampling ratio of 0.3. Part (a) shows the profile of the k-space data of the ground-truth spine MR image. The artifacts resulting from OMP recovery of the image that has been compressively sampled at a sampling ratio of 0.3 are shown in part (b). The adaptive Gaussian filter function is shown part (c). The denoising effect of the Gaussian filter is shown in part (d) which resulted in quality improvements of 1.7 dB and 0.0356 in terms of PSNR and SSIM respectively.

Figure 3 shows the reconstruction results of the proposed method for three different images at different sampling ratios. Column (a) presents the ground-truth images, column (b) shows reconstruction results at  $M/N = 0.2$ , column (c) shows reconstruction results at  $M/N = 0.4$  while the results at  $M/N = 0.6$  are presented in column (d). The visual quality of the reconstructed image increases with increase in the sampling ratio.

The performances of the proposed method, OMP and StOMP methods at  $M/N = 0.4$  are shown in Fig. 4 for three different images.

Column (a) shows the ground-truth images, column (b) shows reconstruction results of the OMP method, column (c) shows reconstruction results of the StOMP

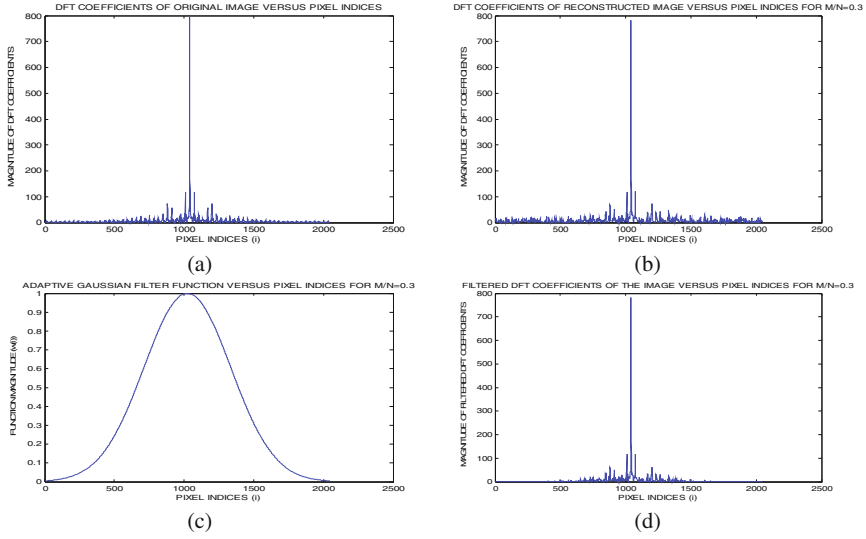


Fig. 2. Effect of Gaussian filtering denoising

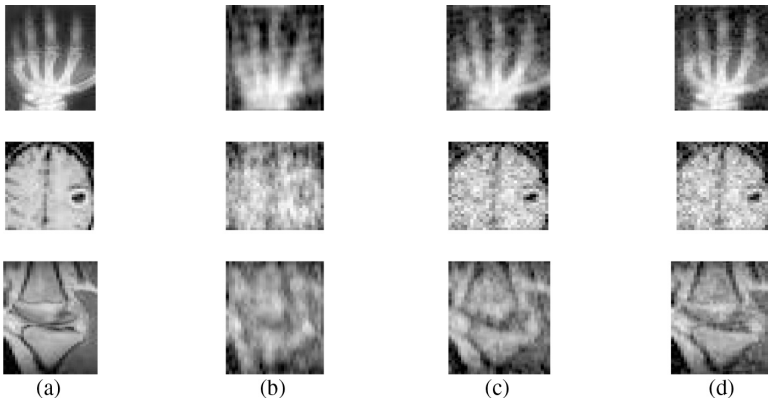


Fig. 3. Performance of the proposed method

method while the results of the proposed method are presented in column (d). The proposed method yielded less noisy reconstructed images than the other two methods.

The PSNR and SSIM quality results for five different MR images reconstructed using the proposed method, the OMP and StOMP methods are presented in Table 1. The proposed method yielded better reconstruction quality than the other two methods.

Figure 5 summarizes the performance of the proposed method. Row (a) from left to right shows the improvement in the quality of the reconstructed images in terms of PSNR and SSIM Row (b) from left to right shows that the proposed method yields images with better quality levels than both the OMP and StOMP methods.

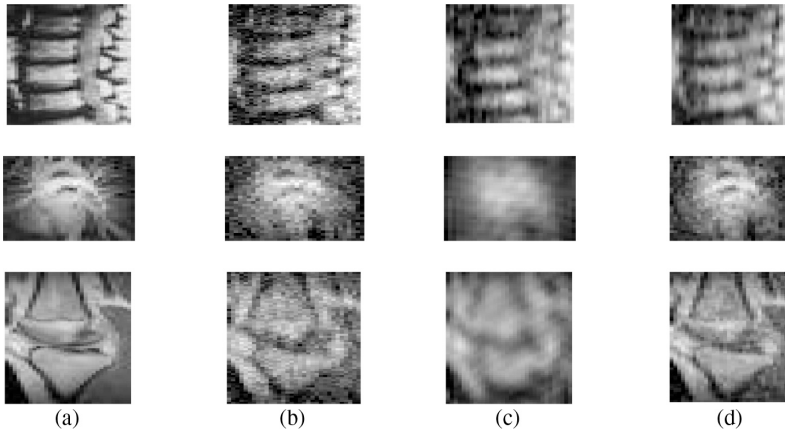







Fig. 4. Comparison of the proposed method with OMP and StOMP methods

Table 1. Reconstruction quality comparison

MR image	M/N	OMP		StOMP		Proposed method	
		PSNR (dB)	SSIM	PSNR (dB)	SSIM	PSNR (dB)	SSIM
	0.1	13.53	0.6303	14.34	0.6223	15.07	0.6927
	0.3	16.28	0.8314	16.73	0.8024	17.96	0.8670
	0.5	20.03	0.9267	19.20	0.8982	21.37	0.9424
	0.7	23.36	0.9677	20.06	0.9461	24.25	0.9714
	0.1	15.44	0.5635	15.05	0.5407	17.20	0.6497
	0.3	19.60	0.8554	19.11	0.7842	21.54	0.8994
	0.5	22.81	0.9292	21.26	0.9012	24.20	0.9466
	0.7	26.25	0.9675	24.05	0.9416	27.33	0.9741
	0.1	16.01	0.7011	15.07	0.6877	17.57	0.7693
	0.3	19.85	0.8891	20.06	0.8811	20.88	0.9091
	0.5	21.37	0.9228	21.78	0.9211	22.56	0.9403
	0.7	23.71	0.9550	23.31	0.9510	24.22	0.9604
	0.1	10.62	0.2885	11.53	0.2116	12.13	0.3514
	0.3	15.00	0.7369	16.17	0.7550	16.09	0.7775
	0.5	17.42	0.8479	17.47	0.8149	17.97	0.8614
	0.7	20.84	0.9340	19.03	0.8924	20.92	0.9349
	0.1	15.49	0.8084	15.50	0.7294	17.43	0.8664
	0.3	21.30	0.9482	20.47	0.9316	22.79	0.9620
	0.5	24.58	0.9754	23.81	0.9698	25.64	0.9803
	0.7	27.24	0.9871	26.52	0.9804	27.89	0.9886

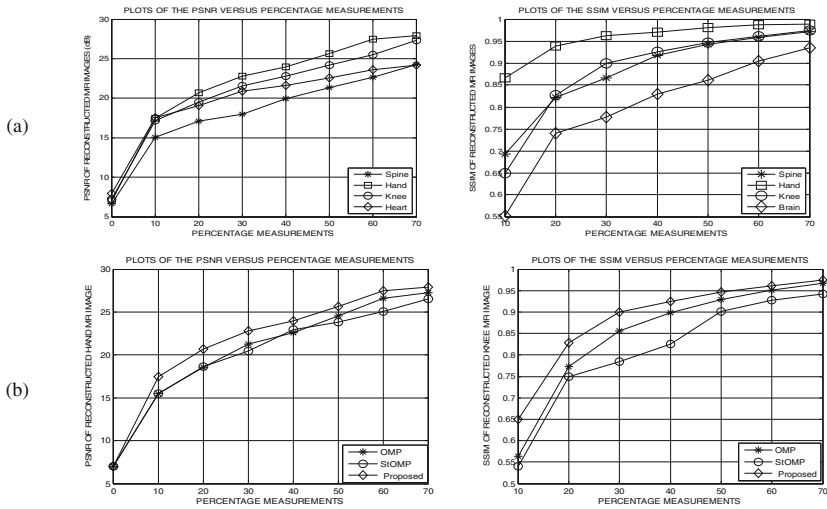


Fig. 5. Proposed method performance summary.

## 5 Conclusion

An adaptive Gaussian denoising CS-MRI algorithm has been presented in this paper. The algorithm is faster than conventional CS-MRI methods and also robust to reconstruction noise. The results included in this paper show an improvement in the quality of the reconstructed images in comparison to other CS-MRI algorithms. Future research work will involve an improvement of the denoising function as well investigation of the denoising in wavelet, shearlet, DCT and curvelet domains with an aim of employing super-resolution techniques to improve the reconstruction quality.

## References

1. Kiragu, H., Kamucha, G., Mwangi, E.: A robust magnetic resonance imaging method based on compressive sampling and clustering of sparsifying coefficients. *IEEE MELECON 2016 Conference, Limassol, Cyprus*, pp. 1–6, April (2016)
2. Lustig, M.: *Sparse MRI*, Ph.D. dissertation. Department of Electrical Engineering, Stanford University (2008)
3. Donoho, D.L.: Compressed sensing. *IEEE Trans. Inf. Theory* **52**(4), 1289–1306 (2006)
4. Chandarana, H., et al.: Free-breathing contrast-enhanced T1-weighted gradient-echo imaging with radial k-space sampling for paediatric abdominopelvic MRI. *European Soc. Radiol.* (2013)
5. Zangen, Z., Khan, W., Paul, B., Ran, Y.: CS-based MRI reconstruction using complex double-density dual-tree DWT. *Int. J. Biomed. Imaging* (2013)
6. Lustig, M., Donoho, D.L., Santos, J.M., Pauly, J.M.: Compressed sensing MRI: a look at how CS can improve on current imaging techniques. *IEEE Signal Process. Mag.* (2008)

7. Foucart, S., Rauhut, H.: *A Mathematical Introduction to Compressive Sensing*, 1st edn. Springer Science and Business Media, New York (2013)
8. Hashemi, R.R.H., Bradley, W.G., Lisanti, C.J.: *MRI the Basics*, 3rd edn. Lippincott Williams & Wilkins, Philadelphia (2010)
9. Liu, Z., Nutter, B., Mitra, S.: Fast MRI simulation with sparsely encoded wavelet domain data in a compressive sensing framework. *J. Electron. Imaging* (2013)
10. Akanksha, T.: Image reconstruct using compressive sensing. *Int. J. Tech. Res. Appl.* **4**(2) (2016)
11. Tesfamicael, S.A., Barzideh, F.: Clustered compressive sensing: application on medical imaging. *Int. J. Inf. Electron. Eng.* **5**(1) (2015)
12. Tina, J.M., Jayashree, M.J.: Compressed sensing MRI reconstruction using wavelets for high performance. *IJSRES* **3**(4) (2016)
13. Candes, E.J., Wakin, M.B.: Introduction to compressive sampling: a sensing/sampling paradigm that goes against the common knowledge in data acquisition. *IEEE Signal Proc. Mag.*, 21–30 (2008)
14. Kiragu, H., Kamucha, G., Mwangi, E.: A fast procedure for acquisition and reconstruction of magnetic resonance images using compressive sampling. *IEEE AFRICON 2015 Conference, Addis Ababa, Ethiopia* (2015)
15. Ziad, M.A., Kadhim, K.K.: Assessment of the quality of medical images (CT & MRI) by using wavelet transformation (WT). *Int. J. Emerg. Res. Manag. Technol.* **4**(7), 2278–9359 (2015)
16. Wang, Z., Bovik, A.C.: Image quality assessment: from error measurement to structural similarity. *IEEE Trans. Image Process.* **13**(1) (2004)

# A Novel Implementation of FPGA Based Finite Difference Time Domain (FDTD) Technique for Two Dimensional Objects

Srinivasa Rao Gandham<sup>(✉)</sup>, Kartheek Bodireddy,  
Boya Pradeep Kumar, and Chandra Sekhar Paidimarry

Department of ECE, University College of Engineering,  
Osmania University, Hyderabad, Tealngana, India  
gsraokmm@gmail.com, kartheekboddireddy@gmail.com,  
pradeep.boge@gmail.com, sekharpaidimarry@gmail.com

**Abstract.** Finite-difference time-domain (FDTD) is a versatile modeling technique which employs finite differences as approximations to both spatial and temporal derivatives. It is a numerical analysis technique used for modeling electrodynamics. The FDTD modeling is applied in wide range of applications such as biomedical imaging and biomedical treatment. In investigating the propagation, scattering and radiation of electromagnetic waves, the FDTD is used as a powerful tool. In this paper, we implemented 2D FDTD algorithm in Xilinx ISE platform by using Absorbing Boundary Conditions and Sinusoidal input as an exciting source. We have developed Synthesizable 2D-FDTD algorithm using fixed-point arithmetic. Synthesis and simulation are carried out on Virtex-5 FPGA and Analog Devices -BLACKFIN-533 target devices. It is experimentally found that proposed FPGA realization is 55.55% faster than that of DSP implementation.

**Keywords:** 2D-FDTD · FPGA · DSP implementation · Electromagnetic analysis · Verilog HDL · Absorbing Boundary Condition

## 1 Introduction

The finite difference time domain (FDTD) technique is widely used in electromagnetic problems to analyze radio propagation. The FDTD method implements finite differences and replace all the derivatives with finite differences. It also discretizes space and time and consequently the electric (E) and magnetic (H) fields are distributed in both space and time. Update equations are obtained by solving the resulting difference equations [1]. A computational grid is constructed to apply FDTD algorithm. E and H fields are evaluated only in the defined computational grid. The wave propagation phenomenon is explained by using Update equations. In FDTD, computational cell chosen is further divided into small cells. The input is applied at one point and corresponding field effects are calculated as field progresses in spatial domain with respect to time. The conventional FDTD technique involves huge computations and it requires large amount of time at the implementation stage. Most of the researchers are trying to develop FDTD algorithms in Field Programmable Gate Array (FPGA) devices to

reduce the computation complexity and increase the performance [2]. It is found from the simulation results of [3] that Kawaguchi and Matsuoka have developed simulation based FDTD algorithm in C-language and VHDL platforms. For real time verification the simulation based FDTD algorithms need to be synthesizable for realization in FPGA and DSP target devices.

In this paper the synthesizable 2D-FDTD algorithm is presented and realized in Virtex-5 FPGA and Blackfin-533 DSP target devices. The main contribution in this paper is development of an optimized 2D-FDTD algorithm to reduce the computational complexity and provide better performance. The update equations are developed in synthesizable form for RTL coding to realize the algorithm in FPGA target device. The input source is generated and accurately applied to the update equations of FDTD algorithm. The performance of algorithm is compared with conventional algorithms and reported.

The use of FPGA technology for one-dimensional FDTD cell is described in [7]. This paper analyzes the implementation of 1D FDTD cell in hardware and simulates the 1D FDTD successfully using a pipelined bit-serial arithmetic architecture. A 10 cell 1D FDTD design is implemented and runs at 37.7 MHz. However, the boundary conditions are not considered in this design. In addition, an FPGA implementation of the 2D FDTD algorithm is described in [8]. This paper analyzes the benefits of the use of FPGA technology to speed up the pseudo-2D FDTD algorithm with both the fixed-point C code and the floating-point FORTRAN code.

We present a novel implementation of FPGA based 2D FDTD algorithm in Xilinx ISE platform using Absorbing Boundary Conditions. In future, the proposed 2D FDTD algorithm can be used for bio-medical applications to detect the Breast Cancer.

The rest of the paper is organized as follows: the two dimensional FDTD algorithm is discussed in Sect. 2. Implementation of synthesizable 2D-FDTD algorithm is presented in Sect. 3. Implementation results and comparison results are reported in Sect. 4. Finally, Sect. 5 has concluding remarks.

## 2 Updating 2D-FDTD Equations

The computational grid Chosen is divided into small cells. The electric and magnetic field components are expressed in terms of past, future and neighbouring cells as shown in Fig. 1. The magnetic field components are distributed in the plane of propagation (X and Y). The electric field component is evaluated from these magnetic components. The Faraday's law and Ampere's law as given in [4, 5] are as follows,

$$\frac{\partial \mathbf{H}}{\partial t} = \nabla \times \mathbf{E}. \quad (1)$$

and

$$\varepsilon \frac{\partial \mathbf{E}}{\partial t} = \nabla \times \mathbf{H}. \quad (2)$$

Since there are no variations of either the fields or the excitations in one of the directions, the derivatives with respect to 'z' drops out from the two curl equations. The Maxwell's equations can be written as:



$$\frac{\partial H_x}{\partial t} = -\frac{1}{\mu} \frac{\partial E_z}{\partial y}. \quad (3)$$

$$\frac{\partial H_y}{\partial t} = \frac{1}{\mu} \frac{\partial E_z}{\partial x}. \quad (4)$$

$$\frac{\partial H_z}{\partial t} = \frac{1}{\mu} \left( \frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} \right). \quad (5)$$

$$\frac{\partial E_x}{\partial t} = \frac{1}{\epsilon} \frac{\partial H_z}{\partial y}. \quad (6)$$

$$\frac{\partial E_y}{\partial t} = -\frac{1}{\epsilon} \frac{\partial H_z}{\partial x}. \quad (7)$$

$$\frac{\partial E_z}{\partial t} = \frac{1}{\epsilon} \left( \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} \right). \quad (8)$$

Equations (3) to (8) are grouped to make two modes of wave propagation equations (TE and TM modes). The first set has  $E_x$ ,  $E_y$  and  $H_z$  and the other set has  $H_x$ ,  $H_y$  and  $E_z$ . First set is called TE mode and second set is called TM mode. These two modes are completely independent from one another. The TE and TM mode equations are given below:

TM Mode Equations:

$$\frac{\partial H_x}{\partial t} = -\frac{1}{\mu} \frac{\partial E_z}{\partial y}. \quad (9)$$

$$\frac{\partial H_y}{\partial t} = \frac{1}{\mu} \frac{\partial E_z}{\partial x}. \quad (10)$$

$$\frac{\partial E_z}{\partial t} = \frac{1}{\epsilon} \left( \frac{\partial H_y}{\partial x} - \frac{\partial H_x}{\partial y} \right). \quad (11)$$

and TE Mode Equations:

$$\frac{\partial E_x}{\partial t} = \frac{1}{\epsilon} \frac{\partial H_z}{\partial y}. \quad (12)$$

$$\frac{\partial E_y}{\partial t} = -\frac{1}{\epsilon} \frac{\partial H_z}{\partial x}. \quad (13)$$

$$\frac{\partial H_z}{\partial t} = \frac{1}{\mu} \left( \frac{\partial E_x}{\partial y} - \frac{\partial E_y}{\partial x} \right). \quad (14)$$

In this work, TM mode propagation is chosen for testing the field components in Absorbing body. None of the field components varies in z-direction in the chosen FDTD grid. Hence, there is certainly no propagation in z-direction. The spatially discrete equations for the TM mode are written in an analogous fashion by considering that the discrete ‘E’ variables are located at the cell edges and including an appropriate half cell shift as shown in Fig. 1.

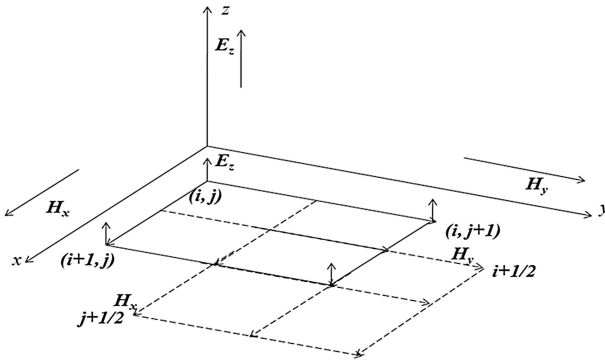


Fig. 1. An FDTD computational cell in TM mode

The final update equations for  $H_x$  and  $H_y$  are:

$$H_x^{q+\frac{1}{2}}\left[m, n+\frac{1}{2}\right] = \frac{1-\gamma_m}{1+\gamma_m} H_x^{q-\frac{1}{2}}\left[m, n+\frac{1}{2}\right] - \frac{1}{1+\frac{1}{\gamma_m}\mu\Delta_y} \frac{\Delta_t}{\epsilon} (E_z^q[m, n+1] - E_z^q[m, n]). \tag{15}$$

$$H_y^{q+\frac{1}{2}}\left[m+\frac{1}{2}, n\right] = \frac{1-\gamma_m}{1+\gamma_m} H_y^{q-\frac{1}{2}}\left[m+\frac{1}{2}, n\right] + \frac{1}{1+\frac{1}{\gamma_m}\mu\Delta_x} \frac{\Delta_t}{\epsilon} (E_z^q[m+1, n] - E_z^q[m, n]). \tag{16}$$

where  $\gamma_m = \frac{\sigma_m\Delta_t}{2\mu}$  and the electric field nodes are updated with the equation

$$E_z^{q+1} = \frac{1-\gamma}{1+\gamma} E_z^q[m, n] + \frac{1}{1+\gamma} \left( \frac{\Delta_t}{\epsilon\Delta_x} \left\{ H_y^{q+0.5}\left[m+\frac{1}{2}, n\right] - H_y^{q+0.5}\left[m-\frac{1}{2}, n\right] \right\} - \frac{\Delta_t}{\epsilon\Delta_y} \left\{ H_x^{q+0.5}\left[m, n+\frac{1}{2}\right] - H_x^{q+0.5}\left[m, n-\frac{1}{2}\right] \right\} \right). \tag{17}$$

where  $\gamma = \frac{\sigma\Delta_t}{2\epsilon}$ ,  $\Delta_x$  = Spatial offset and  $\Delta_t$  = Temporal offset.

### 3 Implementation of Synthesizable 2D-FDTD Algorithm

In this work, the selected computational grid is divided into  $45 \times 45$  small cells. FDTD algorithm sets up the model space and describes the properties of the source excitation and material chosen. The electric and magnetic fields are updated and depend on each other at every time step in all cells of the computational grid. The 2D-FDTD algorithm, whose flow diagram is shown in Fig. 2, is based on the Eqs. (15), (16) and (17).

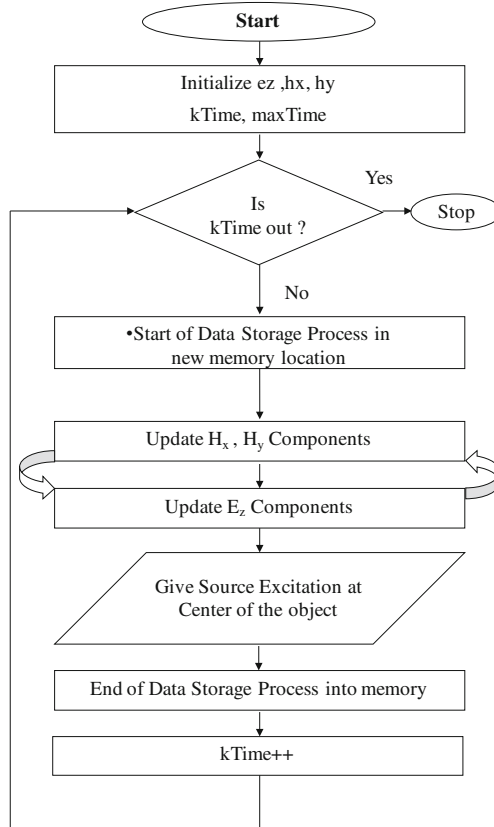


Fig. 2. Flow-graph of 2D-FDTD

In 2D-FDTD algorithm the future time steps of E-fields depend upon the present E-fields and future H-fields [6]. Similarly, future time steps of H-fields depend upon the present H-fields and future E-fields. In this work, the 2D-FDTD algorithm updates both the E-fields and H-fields in interlaced manner completely within the defined computational grid. Standard update equations cannot be used at node  $E_z[0, 0]$  since physical space does not have previous nodes  $H_x[0, -0.5]$  and  $H_y[-0.5, 0]$  and also there is no  $E_z[45, 45]$  node for  $H_x[45.5, 45]$  and  $H_y[45, 45.5]$ .

Instead of using Hardwired source or Additive source, in this work a sinusoidal source is used for excitation. The excitation is applied at the center of the computational grid i.e., at cell (20, 20) and the result is observed at the cell (30, 30). It is understood that the energy propagation takes two time steps for a wave front to cross the cell within the acceptable boundary condition. It is observed that the sinusoidal signal which is originated at the center propagates outward and is absorbed without reflecting back into the problem space. While implementing the FDTD equations in Field Programmable Gate Arrays, the sinusoidal values are stored in SRAM cells as source for exciting (giving input) the 2D-FDTD module. To generate binary data of  $\sin(\omega t)$  MATLAB tool is used, where  $\omega$  is frequency of source input and 't' is time step.

The Sample electric and magnetic field nodes are arranged in space and time. The methodology was adopted to achieve future fields through FDTD algorithm from the past fields. Here  $\Delta x$  denotes the step size in x-direction and  $\Delta t$  is the step size in time. The future value of magnetic field  $H_y$  depends on its previous value and the surrounding electric fields. Thus the future and past values are advanced by a half time step in the magnetic field.

### 4 Simulation Results

The 2D-FDTD algorithm is simulated in MATLAB and Xilinx ISE platforms and then realized in DSP and FPGA target devices. The sinusoidal wave is considered as an input source and goes in the forward direction. If it enters into free space it does not reflect to the source medium. The RTL based sinusoidal input source is generated by using of LUTs and corresponding output values are stored in SRAM as depicted in Fig. 3.

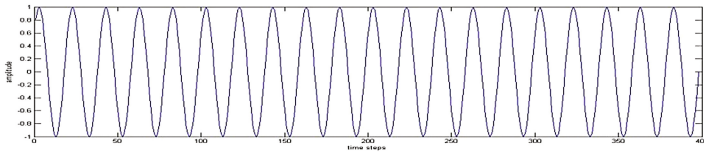


Fig. 3. Input source generated from LUTs

The 2D-FDTD algorithm is implemented in MATLAB platform and corresponding results are shown in Fig. 4.

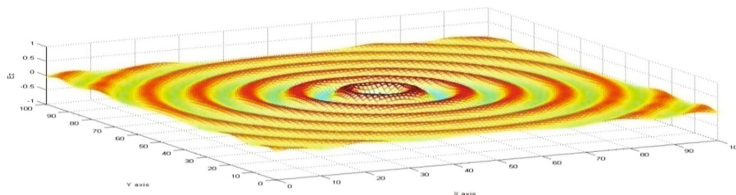


Fig. 4. EM energy distribution in presence of PEC and PMC observed in MATLAB

The MATLAB 2D FDTD algorithm realized in Blackfin-533 DSP target device and timing analysis is reported in Fig. 5.

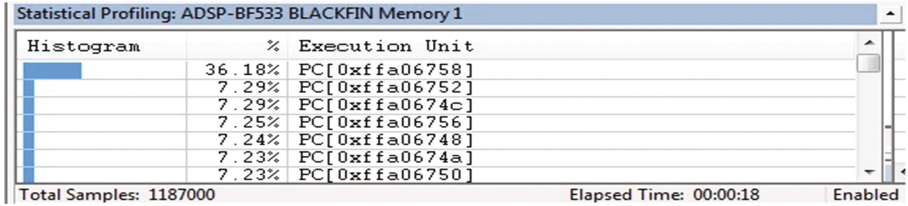


Fig. 5. Analog Device Blackfin-533 DSP target device timing profile

MATLAB implemented 2D FDTD algorithm is designed in Verilog HDL by using Xilinx ISE platform and simulated in Xilinx Isim simulator as shown in Fig. 6.

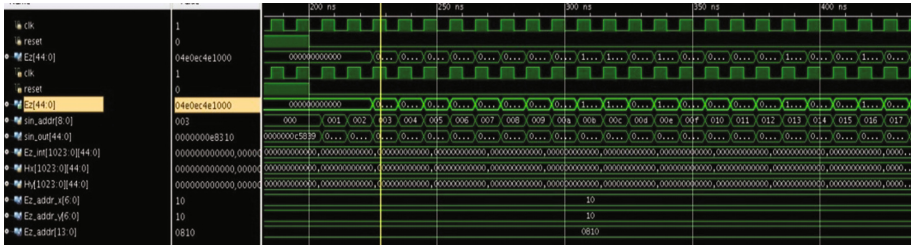


Fig. 6. Xilinx Isim simulator results implemented in Verilog HDL.

To simulate the algorithm, Xilinx ISE 13.2 platform is used and realized in Virtex-5 FPGA. In simulation 400 iterations are considered to analyze the object in MATLAB and Xilinx platforms. In each iteration result of  $E_z$  field is stored in memory. This is expressed by using Table 1, which shows the  $E_z$  values at 1000 ns (Table 2).

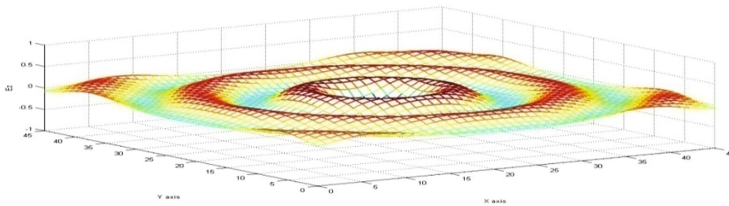
Table 1. Observations of  $E_z$  at 1000 ns generated by RTL

x	0	587785	587785	587785
0	809017	809017	809017	809017
0	951056	951056	951056	951056
0	100000	100000	100000	100000
0	951056	951056	951056	951056
0	809017	809017	809017	809017
0	587785	587785	587785	587785
0	309017	309017	309017	309017
0	0	0	0	0

**Table 2.** Comparison table of target devices

S.no	Target device	Computational time	Power
1	Virtex-5 FPGA	8 s	1.9 W
2	Blackfin-533 DSP	18 s	2.17 W

The Corresponding Electromagnetic (EM) energy wave is shown in Fig. 7, which shows the output of Virtex-5 FPGA target device generated data files and plots the output with help of MATLAB.

**Fig. 7.** Verilog HDL data files output in MATLAB.

It has been observed that virtex-5 FPGA requires 8 s to implement the 2D-FDTD algorithm, while Analog Device Blackfin-533 DSP target device requires 18 s.

## 5 Conclusion

In this work, 2D-FDTD algorithm is implemented using Verilog HDL and realized in FPGA target device with accurate input source and compared with DSP target device. This implementation will help in effective utilization of available digital resources to analyze the complex structures in less time as compared with traditional resources. The same algorithm is implemented in MATLAB to validate the generated data files by Verilog HDL. The numerical results indicate that FPGA hardware implementation achieves significant accelerated speed compared to DSP target device.

## References

1. Schneider, J.B.: Understanding the Finite-Difference Time-Domain Method, Handbook (2012)
2. Chen, W., Leeser, M.: Finite difference time domain: a case study using FPGAs. In: Reconfigurable Computing the Theory and Practice of FPGA-Based Computation, chap. 32. Department of Electrical and Computer Engineering Northeastern University (2008)
3. Kawaguchi, H., Matsuoka, S.: Design study of 3D FDTD/FIT dedicated computer on FPGA. In: Progress in Electromagnetic Research Symposium 2004, Pisa, Italy, 28–31 March 2004

4. Yee, K.: Numerical solution of initial boundary value problems involving Maxwell's equations in isotropic media. *IEEE Trans. Antennas Propag.* **14**(3), 302–307 (1966)
5. Taflov, A.: Application of the finite-difference time-domain method to sinusoidal steady state electromagnetic penetration problems. *IEEE Trans. Electromagn. Compat.* **22**(3), 191–202 (1980). doi:[10.1109/TEMC.1980.303879](https://doi.org/10.1109/TEMC.1980.303879)
6. Kawaguchi, H.: Improved architecture of FDTD dataflow machine for higher performance electromagnetic wave simulation. *IEEE Trans. Magn.* **52**(3), 1–4 (2016)
7. Schneider, R.N., Turner, L.E., Okoniewski, M.M.: Application of FPGA technology to accelerate the finite-difference time-domain (FDTD) method. In: *Proceedings of the 2002 ACM/SIGDA Tenth International Symposium on Field-Programmable Gate Arrays*, pp. 97–105 (2002)
8. Chen, W., Kosmas, P., Leeser, M., Rappaport, C.: An FPGA implementation of the two-dimensional finite-difference timedomain (FDTD) algorithm. In: *FPGA 2004 Proceedings of the 2004 ACM/SIGDA 12th International Symposium on Field Programmable Gate Arrays*, pp. 213–222. ACM, New York (2004). ISBN: 1-58113-829-6

# An Iterative Back-Projection Technique for Single Image Super Resolution with Natural Texture Preservation

Boniface M. Ngocho<sup>(✉)</sup> and Elijah Mwangi

School of Engineering, University of Nairobi,  
P. O. Box 30197, Nairobi 00100, Kenya  
bmngocho@students.uonbi.ac.ke,  
elijah.mwangi@uonbi.ac.ke

**Abstract.** High resolution images obtained from single image super resolution often suffer from excessive smoothness which blurs edges and gives smooth areas an un-natural texture. This paper proposes a super resolution method that combines wavelet super resolution and an edge enhancing back-projection filter to produce a super resolution image with sharp edges and natural texture. The low resolution image is first up-sampled using the discrete wavelet transform. High frequency sub-bands are derived from DWT of a bicubic interpolated low resolution image. The up-sampled image is then subjected to iterative back projection using a back-projection filter based on the Laplacian of Gaussian. The proposed method is tested on RGB colour images. It is found to produce better results than bicubic interpolation and selected edge directed interpolation methods in terms of PSNR and SSIM. It is also found to approximate the edge sharpness of edge directed interpolation while giving a more natural texture in smooth areas. The proposed method also eliminates spurious colours that have been observed in most edge-directed methods.

**Keywords:** Super resolution · Wavelets · Back-projection · Log filter

## 1 Introduction

Single image super resolution is a signal processing method that seeks to create a high resolution image from a single low resolution image. This is necessary when the spatial resolution of available image does not meet the requirements for human or machine interpretation. A shortfall in the achieved resolution may be as a result of limitations in the imaging or storage system, environmental factors or the need to use old images with current systems. Common methods of image resolution enhancement are mathematical interpolation, statistical estimation, wavelet-based interpolation and learning methods which make use of a library of example images.

Interpolation involves convolution of the low resolution image with an interpolation kernel such as Keys' bicubic function, B-splines and Lanczos functions [1].

In wavelet-based super resolution, the available image is treated as the low resolution sub-band of a wavelet multiresolution. The high frequency sub-bands are estimated from the high frequency results of the wavelet decomposition of the low



resolution image. Various methods have been applied for estimating the high frequency sub-bands. Many studies combine mathematical interpolation and wavelet super resolution. The high frequency sub-bands are obtained by interpolating those obtained by a DWT decomposition of the low resolution image [2] or from decomposition of a bicubic interpolated low resolution image [3, 4]. It can be shown [5] that the use of sub-bands from the bicubic interpolated low resolution image outperforms other methods that involve interpolation of sub-bands in terms of both Peak Signal to Noise Ratio (PSNR) and Structural Similarity Index Measure (SSIM).

In addition to estimation of the missing pixels, additional processing is usually carried out to improve the quality of the super resolution image. Iterative Back-Projection (IBP) [6, 7] is a process that intends to constrain the super resolution image to the available low resolution image. In IBP, a low resolution image is derived from the initial super resolution image. The error between the original and derived low resolution images is taken through a super resolution process and then added to the original super resolution image to create a new super resolution image. The process is repeated until it converges. The process is summarised in (1) and (2).

$$e_n = LR - (SR_n * h) \downarrow . \quad (1)$$

$$SR_{n+1} = SR_n + (e_n \uparrow) * p \quad (2)$$

In (1), a high resolution image,  $SR_n$  is down-sampled and subtracted from the original low resolution,  $LR$ , to give the error image,  $e_n$ . In (2), the error image,  $e_n$ , undergoes convolution with a back-projection kernel  $p$ . The result is added to the  $n^{th}$  super resolution image. The result is a new super resolution image,  $SR_{n+1}$ . The relationship between successful error images,  $e_n$  and  $e_{n+1}$  is as represented in (3).

$$\begin{aligned} e_{n+1} &= LR - (SR_n * h) \downarrow \\ &= LR - ((SR_{n-1} + (e_n \uparrow) * p) * h) \downarrow \\ &= LR - ((SR_{n-1} + ((LR - (h * SR_{n-1}) \downarrow) \uparrow) * p) * h) \downarrow \\ &= LR - SR_{n-1} * h \downarrow - (LR - SR_{n-1} * h \downarrow) \uparrow * p * h \downarrow \\ &= e_n - e_n \uparrow * p * h \downarrow \\ e_{n+1} &= e_n * (\delta - p * h \downarrow) \end{aligned} \quad (3)$$

In (3),  $h$  is the (usually unknown) point spread function of the imaging system,  $p$  is the back-projection filter and  $\delta$  is the delta function in matrix form [6, 7].

In order for the IBP iteration to converge, the  $(n + 1)^{th}$  error image must have a smaller magnitude than the  $n^{th}$  error image. From (3), we can conclude that this will happen if the condition in (4) is met.

$$\|(\delta - p * h \downarrow)\|_1 < 1 \quad (4)$$

Where  $\|f\|_1$  is the  $l_1$ -norm of the enclosed function,  $f$ , which, for a digital filter in matrix form is the sum of all the coefficients. The operators and variables  $\delta$ ,  $p$  and  $h$  can all be represented as digital filter kernels.

The super resolution process is not able to estimate all the high frequency information missing from the available low resolution image. The missing information usually results in blur or chessboard effects on edges. Isotropic application of IBP (in which case,  $p$  in (2) is set to 1) will not improve edge information. Dai et al. [7] use bilateral filter on the error image, with the initial super resolution image providing the pixel value guidance in the bilateral filter. Other approaches have used isotropic IBP but added edge sharpening techniques on the super resolution image. Wang et al. [8] propose isotropic back-projection followed by Laplacian of Gaussian (LoG) and Gaussian filtering of the resultant super resolution image. Makwana and Mehta [9] have similar approach to Wang but use a Gabor filter or Canny filter. Bareja [10] proposes use of Infinite Symmetric Exponential Filter (ISEF) for edge enhancement of the super resolution image.

The approach proposed here will apply back-projection in accordance with (2) with an edge-enhancing back-projection filter. In order to achieve natural texture in the final image, the initial super resolution image will be produced through wavelet super resolution. The high frequency sub-bands will be estimated by wavelet multiresolution of a bicubic interpolated input image. The process will involve design of a back-projection filter with characteristics to match a wide range of images.

The rest of the paper is arranged as follows. Section 2 discusses the development of the proposed algorithm, including the design of the back-projection filter. Section 3 presents the computer simulation results. The conclusion is given in Sect. 4.

## 2 Proposed Method

### 2.1 Resolution Enhancement

As mentioned in Sect. 1, the up-sampling method to be used will apply wavelet transform. The process used is the one described in [5]. The wavelet transform and inverse transform can be done using any biorthogonal wavelets. For this investigation, the Haar wavelet has been used.

### 2.2 Back-Projection Filter

Both wavelet super resolution and bicubic interpolation are implemented separately on rows and columns using separable kernels. Information on diagonal edges is therefore not effectively taken into account in the super resolution process. The proposed design of the super resolution filter will attempt to compensate for this by using a major axis rotated by  $\pi/4$  radians from the horizontal. The Laplacian of Gaussian (LoG) filter is a linear filter which combines the noise reducing function of a Gaussian filter and the edge enhancing function of a Laplacian function [11, 12].

The continuous LoG function is achieved by getting the second derivative of the Gaussian function, which results in (5) for the 2 dimensional function

$$LoG(x, y, \sigma) = \frac{1}{\sigma^3 \sqrt{2\pi}} \left( \frac{x^2 + y^2}{\sigma^2} - 1 \right) e^{-\frac{(x^2 + y^2)}{2\sigma^2}} \quad (5)$$

The LoG filter is a function of the Euclidian distance from the reference point  $(x,y)$  and the standard deviation,  $\sigma$ , of the Gaussian distribution. Methods of approximating the LoG filter in discrete form have been suggested for application in feature detection [13, 14]. One condition for a LoG filter kernel is the relationship between the kernel size,  $n$ , and the standard deviation of the Gaussian function,  $\sigma$ , which is presented in (6) [12].

$$n = 2 * ceil(3\sigma) + 1 \quad (6)$$

In (6),  $n \times n$  is the size of the kernel and  $\sigma$  is the standard deviation of the Gaussian function. The *ceil* function rounds the arguments to the next whole number. For this investigation, a  $5 \times 5$  kernel has been used, which gives a standard deviation,  $\sigma$  of  $2/3$ . The filter kernel is segmented to achieve a structure as presented in Fig. 1.

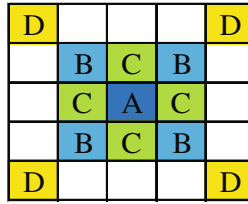


Fig. 1. Structure of  $5 \times 5$  filter kernel with axis rotated  $\pi/4$  radians

The structure in Fig. 1 has four zones, A, B, C and D. The filter coefficients in each zone are equal, giving a symmetrical linear phase filter. Let the coefficients in zones A, B, C and D be  $a$ ,  $b$ ,  $c$  and  $d$  respectively. The coefficients in the cells outside these zones are zero. In order to achieve unity gain in smooth areas, the coefficients are required to satisfy (7).

$$a + 4b + 4c + 4d = 1 \quad (7)$$

This requires the total sum of the coefficients in the filter kernel to be 1. The filter coefficients are determined by performance in iterative back-projection process. This starts with a high resolution ground truth image,  $GT$ . The image is down-sampled to produce a low resolution image,  $LR$ . The  $LR$  image is up-sampled using the method present in [5] to produce a high resolution image,  $SR$ . The initial super resolution image is subjected to back-projection using the proposed filter. Different estimates of filter coefficients are used to identify the coefficient sets that achieves the highest PSNR in super resolution image.

To ensure that the convergence condition in (4) is met, it is proposed to multiply the filter coefficients with a scaling factor,  $\lambda$ . The back-projection in (2) is therefore

modified to be as in (8). The scaling factor is experimentally adjusted to achieve convergence of the iteration.

$$SR_{n+1} = SR_n + (e_n \uparrow) * (p \times \lambda) \tag{8}$$

### 2.3 Proposed Algorithm

A flow chart of the proposed method is given in Fig. 2. The super resolution process is:

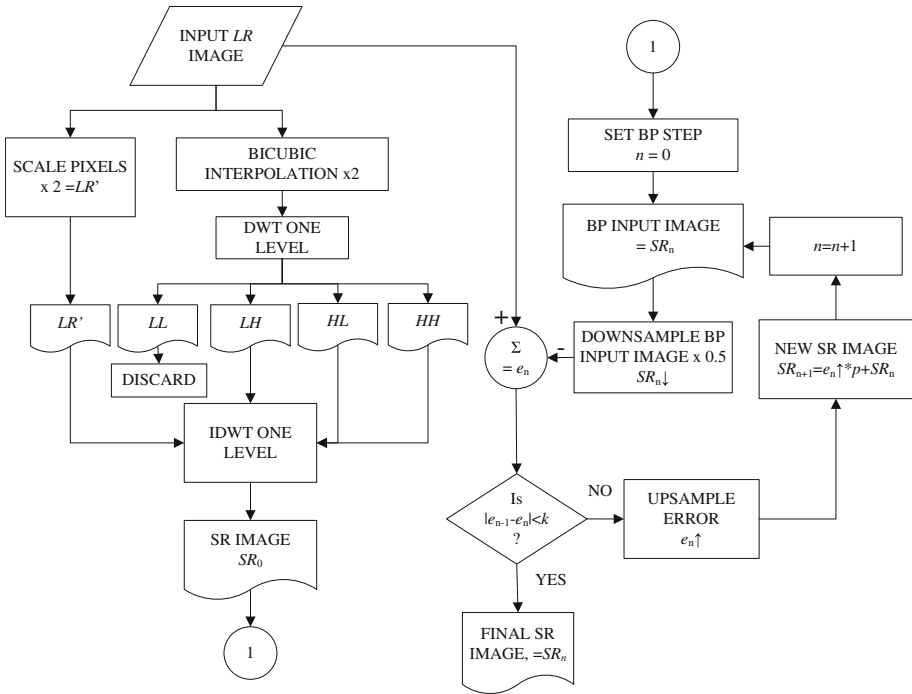


Fig. 2. Flow chart of super resolution process

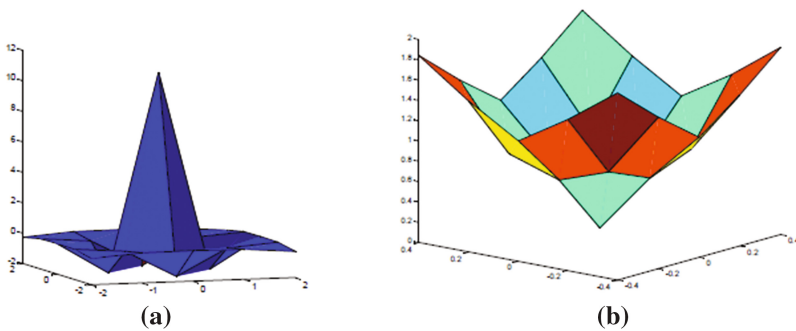
- i. Low resolution image is up-sampled by a factor of 2 using wavelet super resolution to obtain a high resolution image  $SR_0$ .
- ii. High resolution image is down-sampled by a factor of 2.
- iii. Down-sampled image is subtracted from original low resolution image to get error image,  $e_{n+1}$ .
- iv. Error image is added to the latest high resolution image to get a new high resolution image,  $SR_{n+1}$ .
- v. Steps (ii) to (iv) are repeated until the difference between successful error images falls below a certain minimum,  $k$ .

For this investigation, the iteration was stopped when the difference in PSNR between successful high resolution images was below 0.01 dB. This measure is used as representing the situation where difference between successful error images is approximating zero.

### 3 Results & Discussions

#### 3.1 Back-Projection Filter and Scaling Factor

The final selected values for the back-projection filter are:  $a = 11$ ,  $b = -0.25$ ,  $c = -0.2$  and  $d = -0.25$ . The resultant filter kernel is as shown in Fig. 3 in the time and frequency domains. It can be observed that in the spatial domain, the filter approximates the ‘Mexican hat’ shape of the LoG function. In the frequency domain, the filter has a low pass characteristic, a situation that results from approximation of the band-pass characteristic of the LoG function using a small kernel. A scaling factor for the filter was determined and it was found that the highest value for which IBP converged was  $\lambda = 0.07$ .



**Fig. 3.** Three dimensional view of proposed back-projection filter kernel (a) Gain against spatial displacement: (b) Gain against frequency

#### 3.2 Super-Resolution:- Test Images

The proposed algorithm was tested on the 24 images in the Kodak Lossless True Colour Images database [15]. In order to test performance with images of a different nature, tests were also carried out on 12 aerial images from the University of Southern California School of Engineering Image database [16]. Results for six selected images are presented here. The selected Kodak images are *kodim03*, *kodim04*, *kodim05*, and *Kodim20*. These are 512 by 768 pixels, 8 bit RGB format colour. The selected aerial images are *Aerial2103* and *Aerial2105*. They are 512 by 512 pixels, 8 bit RGB format colour.

In order to measure the performance of the process, the test images are treated as the ground truth. The ground truth image is down sampled by a factor of 2 using

bilinear interpolation to produce the input low resolution image. The low resolution image is subjected to the super resolution process to produce a super resolution image. The super resolution image is compared with the ground truth image to measure the performance. Performance is measured qualitatively in terms of PSNR and SSIM.

### 3.3 Quantitative Results

Results are presented in Tables 1 and 2 in terms of PSNR and SSIM for the proposed method and comparative methods (isotropic back projection, bicubic interpolation and two edge directed methods, [17, 18]).

**Table 1.** PSNR results

Image		K03	K04	K05	K20	A03	A05
Bicubic	R	32.83	32.51	25.83	29.74	27.77	29.29
	G	32.70	32.38	25.72	29.59	32.15	28.62
	B	32.63	32.36	25.85	29.81	38.91	32.51
Isotropic IBP	R	33.00	32.95	26.21	30.10	28.07	29.89
	G	32.87	32.73	26.16	29.94	32.34	29.04
	B	32.69	32.71	26.19	30.13	38.97	32.80
Edge Guided 1 [17]	R	31.34	30.39	24.23	27.93	26.27	26.49
	G	31.37	30.47	24.07	27.85	30.96	26.26
	B	31.72	30.41	24.40	28.16	38.10	30.55
Proposed	R	<b>33.02</b>	<b>32.97</b>	<b>26.23</b>	<b>30.11</b>	<b>28.03</b>	<b>29.91</b>
	G	<b>32.89</b>	<b>32.74</b>	<b>26.17</b>	<b>29.95</b>	<b>32.35</b>	<b>29.06</b>
	B	<b>32.72</b>	<b>32.73</b>	<b>26.21</b>	<b>30.14</b>	<b>38.98</b>	<b>32.83</b>
Edge Guided 2 [18]	R	25.58	24.65	22.27	20.33	24.49	21.43
	G	25.73	25.35	22.26	20.43	21.70	21.53
	B	26.00	25.28	22.52	21.11	19.93	20.11

It is observed from the results in Table 1 that the proposed method has better results in terms of PSNR than all the other methods tested. It is observed from Table 2 that in terms of SSIM, the proposed method performs better than the edge directed methods and bicubic interpolation. It however performs worse than isotropic back-projection for some of the test images. This is especially evident for images with a lot of high frequency content like *Kodim05*.

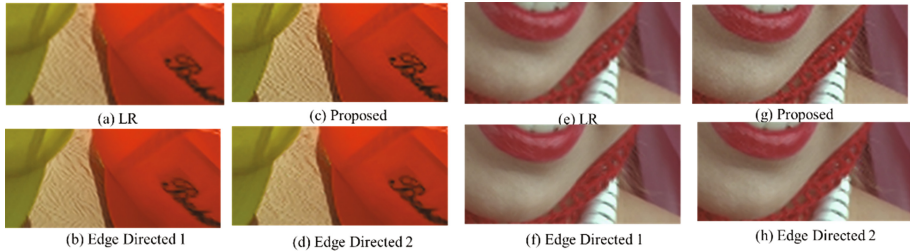
The other images in the Kodak and University of Southern California databases gave similar results [19].

**Table 2.** SSIM results

Image		K03	K04	K05	K20	A03	A05
Bicubic	R	0.918	0.892	0.84	0.915	0.776	0.866
	G	0.921	0.887	0.848	0.904	0.872	0.844
	B	0.91	0.887	0.829	0.864	0.926	0.86
Isotropic IBP	R	0.926	0.904	<b>0.863</b>	0.924	0.785	0.879
	G	0.928	0.898	<b>0.87</b>	0.913	0.877	0.858
	B	0.919	0.898	<b>0.852</b>	0.873	0.927	0.871
Edge Guided 1 [17]	R	0.903	0.853	0.782	0.896	0.744	0.81
	G	0.906	0.854	0.79	0.885	0.851	0.789
	B	0.893	0.853	0.769	0.843	0.917	0.817
Proposed	R	<b>0.926</b>	<b>0.904</b>	0.863	<b>0.924</b>	<b>0.785</b>	<b>0.879</b>
	G	<b>0.928</b>	<b>0.898</b>	0.87	<b>0.913</b>	<b>0.878</b>	<b>0.858</b>
	B	<b>0.919</b>	<b>0.898</b>	0.852	<b>0.873</b>	<b>0.927</b>	<b>0.871</b>
Edge Guided 2 [18]	R	0.893	0.848	0.778	0.886	0.741	0.804
	G	0.897	0.848	0.787	0.875	0.838	0.78
	B	0.882	0.847	0.762	0.833	0.9	0.804

### 3.4 Observation of Super Resolution Images

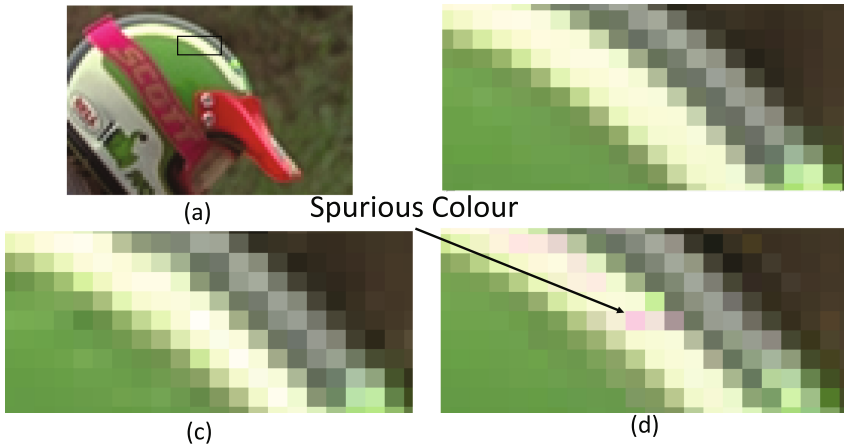
Figure 4 presents zoomed-in extracts for the image *Kodim03* and *Kodim04* selected from the low resolution image and the super resolution image. Extracts from selected edge enhancing super resolution methods are also included for comparison.



**Fig. 4.** Zoomed-in extracts from the images *Kodim03*. (a) LR image; (b) SR using edge directed 1 (c) SR using proposed method (d) SR using edge directed method 2 and. *Kodim04*. (e) LR image; (f) SR using edge directed 1 (g) SR using proposed method (h) SR using edge directed method 2.

It is observed that the images produced by the edge directed methods have sharper edges but the textured areas are not very well defined when compared to the image produced by the proposed method. In the *Kodim03* image, it can be noted that the writing on the hat is clearer in the image from the proposed method than from the edge directed methods. Similar observations can be made in the extracts from the image *Kodim04*. The woman's chin and lips are very smooth in the super resolution image using edge directed methods but have a natural texture in the image using the proposed method.

Figure 5 shows the image *Kodim05* with a selected part zoomed about 3000 times so that individual pixels are visible. It can be observed that the super resolution from edge directed method produces spurious colours in some areas. This is not observed for the proposed method.



**Fig. 5.** Zoomed-in extracts from the image *Kodim05*. (a) Zoom area; (b) Ground truth (c) SR from proposed method (d) SR using edge directed 1.

## 4 Conclusion

This work has demonstrated that the use of a suitable back-projection filter can improve the quality of super resolution images while preserving the natural texture in the super resolution image. This is achieved without the need for post processing of the super resolution image.

Further improvements of the back-projection filter could improve the sharpness of edges. For example, use of a larger back-projection filter kernel could allow for a filter with increased attenuation at high frequencies. This would reduce noise in the back-projected error. Such investigation would constitute future work not covered in this study.

## References

1. Thevenaz, P., Blu, T., Unser, M.: Image interpolation and resampling. In: Bankman, I. (ed.) *Handbook of Medical Imaging*, 2nd edn, pp. 465–494. Elsevier/Academic Press, San Diego (2009)
2. Bagawade, R.P., Bhagawat, K.S., Patil, P.M.: Wavelet transform techniques for image resolution enhancement: a study. *Int. J. Emerg. Technol. Adv. Eng.* **2**(4), 167–172 (2012)
3. Tai, S.-C., Kuo, T.-M., Iao, C.-H., Liao, T.-W.: A fast algorithm for single image super resolution in both wavelet and spatial domain. In: *International Symposium on Computer, Consumer and Control*, Taichung, Taiwan (2012)



4. Garg, A., Naidu, S.V., Yahia, H., Singh, D.: Wavelet based resolution enhancement for low resolution satellite images. In: 9th International Conference on Industrial and Information Systems (ICIIS), Gwalior, India (2014)
5. Ngocho, B.M., Mwangi, E.: Single image super resolution with improved wavelet interpolation and iterative back-projection. *IOSR J. VLSI Signal Process.* **5**(6), 16–24 (2015)
6. Irani, M., Peleg, S.: Motion analysis for image enhancement: resolution, occlusion and transparency. *J. Vis. Commun. Image Represent.* **4**(4), 324–335 (1993)
7. Dai, S., Han, M., Wu, Y., Gong, Y.: Bilateral back-projection for single image super resolution. In: *IEEE International Conference on Multimedia and Expo*, Beijing, July 2007
8. Wang, J.-T., Liang, K.-W., Chang, S.-F., Chang, P.-C.: Super-resolution image with estimated high frequency compensated algorithm. In: *9th International Symposium on Communications and Information*, Icheon, South Korea, 28–30 September 2009
9. Makwana, R.R., Mehta, N.D.: Single image super-resolution VIA iterative back projection based Canny edge detection and a Gabor filter prior. *Int. J. Soft Comput. Eng.* **3**(1), 2231–2307 (2013)
10. Bareja, M.N., Modi, C.K.: An improved iterative back projection based single image super resolution approach. *Int. J. Image Graph.* **14**(4), 1450015 (2014)
11. Gonzalez, R.C., Woods, R.E.: *Digital Image Processing*, 3rd edn. Prentice-Hall, Upper Saddle River (2008)
12. Gunn, S.R.: On the discrete representation of the Laplacian of Gaussian. *Pattern Recogn.* **32**(8), 1463–1472 (1999)
13. Agrawal, M., Konolige, K., Blas, M.R.: CenSurE: center surround extremas for realtime feature detection and matching. In: Forsyth, D., Torr, P., Zisserman, A. (eds.) *ECCV 2008*. LNCS, vol. 5305, pp. 102–115. Springer, Heidelberg (2008). doi:[10.1007/978-3-540-88693-8\\_8](https://doi.org/10.1007/978-3-540-88693-8_8)
14. Laguna, R., Barrientos, R., Blazquez, L.F., Miguel, L.J.: Traffic sign recognition application based on image processing techniques. In: *19th World Congress: The International Federation of Automatic Control*, Cape Town, South Africa (2014)
15. Franzen, R.: Kodak lossless true color image suite. Eastman Kodak, 27 January 2013. <http://r0k.us/graphics/kodak/>
16. University of Southern California, Signal and Image Processing Institute. The USC-SIPI Image Database: Version 5, April 2006. <http://sipi.usc.edu/database/database.php?volume=aerial>
17. Li, X., Orchard, M.T.: New edge directed interpolation. *IEEE Trans. Image Process.* **10**(10), 1521–1527 (2001)
18. Zhang, L., Wu, X.: An edge-guided image interpolation algorithm via directional filtering and data fusion. *IEEE Trans. Image Process.* **15**(3), 2226–2238 (2006)
19. Ngocho, B.M.: Single image super resolution using the wavelet transform and an iterative back-projection technique. Unpublished: M.S. thesis. School of Engineering, University of Nairobi (2016)

# Development of Active Acoustic Noise Cancellation Using the Professional Audio Development Kit (PADK) Featuring the TMS320C6727 DSP

Sajaad Boodoo<sup>1</sup>(✉), Yasdeo Bissessur<sup>1</sup>, and Roshun Paurobally<sup>2</sup>

<sup>1</sup> Department of Electrical and Electronic Engineering,  
University of Mauritius, Reduit, Mauritius

sboodoo@govmu.org, yasdeob@uom.ac.mu

<sup>2</sup> Mechanical and Industrial Engineering Department,  
Qatar University, PO Box 2713, Doha, Qatar  
mpaurobally@qu.edu.qa

**Abstract.** The PADK is a highly versatile DSP platform for implementing real-time audio applications. It features the Texas Instruments TMS320C6727, a powerful floating-point DSP processor and can process audio from eight input channels simultaneously. This makes it an ideal platform for developing multi-channel active acoustic noise control algorithms. This paper demonstrates results of system identification and feedback control using this kit. Using the Least Mean Square (LMS) algorithm, the filter coefficients are updated in real time to obtain an accurate impulse response. Potential applications of this particular platform are in the implementations of active acoustic noise control (ANC) systems in real time.

**Keywords:** PADK · C6727 · Least Mean Square · System identification · Feedback control

## 1 Introduction

The problem of acoustic noise is becoming increasingly important nowadays with the significant use of consumer electronic appliances and industrial and medical equipment. Acoustic noise is a public nuisance and causes hearing loss, disturbs speech communication and lowers performance in mental and physical tasks. Traditional control methods involve passive systems, but these methods are very ineffective at low frequencies as they will involve heavy and bulky components. Active noise control (ANC) systems have been developed and are efficient at low frequencies. The ANC system is based on an electro acoustical technique, which uses the principle of superposition. An anti-noise is generated which has the same amplitude but opposite in phase, to cancel the unwanted noise acoustically, thus resulting in reduced residual noise. The ANC technique using a loudspeaker to generate an anti-noise was first proposed in an 1936 patent by Lueg [1]. In practice, the characteristics of the noise source and the acoustic environment, which include the frequency, amplitude and

phase, are varying. The control performance of ANC systems is largely dependent on the accuracy of the amplitude and phase of the anti-noise generated. To deal with the changing parameters, ANC systems make use of adaptive filters. In order to have the best performance, it is necessary for these adaptive filters to have increased speed of convergence and robustness to interference. Nowadays, the development of powerful low cost digital signal processors makes use of the latest adaptive algorithms to have improved system performance for practical ANC applications. Most importantly, these adaptive processes need to be run in real time so as to cater for the changing parameters. The configurations of ANC systems can be classified into either feedforward control or feedback control. The feedforward configuration consists of a reference sensor, a secondary source and an error sensor. A reference signal is available from the reference sensor for the adaptive filter, together with an error signal which serves as an update for the coefficients of the adaptive filter. In the feedback structure, there is an error sensor and a secondary source but no reference signal. It has also been shown by researchers that multiple channel systems, if correctly positioned, can provide better noise cancellation than single channel systems [2]. Recent work on wireless active noise control [3], shows that for such a system to be practical a fast and powerful DSP platform is required. Hence, in the development of better ANC systems, it is necessary to have a platform which offers both increase in processing speed as well as enabling multiple channel systems such as the PADK. During the literature review, it has been found that very little work has been carried out on the PADK. One of the few papers relating to the development of the PADK platform is found in [4], but it is in german language and not accessible to many readers. Moreover, the work carried out in [4], is different with the work presented in this paper. The paper is organized as follows. An overview of the PADK is presented in Sect. 2. Sections 3 and 4 give a description of how the experiments are set up for system identification and regenerative feedback control respectively. In Sect. 5, the results are presented and discussed. Concluding remarks are given in Sect. 6.

## 2 Overview of PADK

The PADK is a multichannel DSP kit dedicated for audio applications and can accommodate up to eight analog input and eight analog output sources [5]. The available sampling rates range from 32 kHz to 192 kHz. It includes Analog to Digital Converters, Digital to Analog Converters, digital audio inputs and outputs. The PADK has an important feature and powerful tool known as the dMAX (Dual Data Movement Accelerator) [6]. The dMAX handles user-programmed data transfers between the internal data memory controller and the device peripherals on the C6727 DSP. It also allows movement of data to/from any addressable memory space, including memory, peripherals and external memory. The kit has also the capability of ping-pong data buffering. The latter is a technique that allows the CPU activity to be detached from the dMAX controller activity hence allowing for two sets of data buffers for all incoming and outgoing data streams. This allows DSP algorithms to be run continuously while new samples are being accumulated in the dMAX buffers.

The Code Composer Studio development environment is used to build and debug embedded real time software applications with the PADK. An emulator which acts an interface and controls the information to and from the target is required between the PC and the PADK. Texas Instruments use several common interface connections between the target and the emulation. For this current application, the Blackhawk Emulator USB 560v2 is used. Figure 1 shows the configuration of the set up of the PADK, the emulator and the PC.



**Fig. 1.** Professional Audio Development Kit with the Blackhawk Emulator USB 560v2

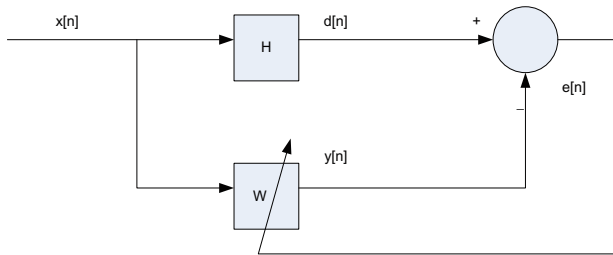
Several researchers have developed active ANC algorithms with other platforms featuring other DSPs. But till date, there is no low-cost platform that provides the same flexibility and computational power as the PADK. For example, the DSK 6713 has also been used [7, 8], but it has shortcomings. It has only one signal channel control input and output compared to the PADK. Also, the dMAX controller in the PADK offers increase in speed of processing. Table 1 below shows some of the differences in performance of the TMS 320 C6713 and the TMS 320 C6727 [9, 10].

**Table 1.** Performance characteristics

	TMS 320 C6713	TMS 320 C6727
Max MFLOPS	1350	1800
Max MIPS	1800	2400
Peak MMACS	450	600

Another difference between the two DSP kits is that the PADK process signals in block unlike the DSK C6713. This feature has to be taken into account in the programming of the adaptive algorithm. A basic operation that is required in any active ANC is the process of system identification. This consists of exciting the system with a reference signal and observing its response. From these signals, it is possible to extract the impulse response using a technique such as the LMS algorithm. The LMS algorithm is used as it is easy to implement in real time and due to its robustness in the

presence of noise. Other techniques such as Recursive Least-Squares (RLS), Kalman filtering etc..., can be used but they require far more computations than the LMS. The LMS algorithm was developed by Widrow and Hoff in the late 1950s. It is based on the method of steepest descent gradient whereby its objective is to minimize a cost function. Its derivation will not be given here as it is readily available in various references [11]. It searches for a minimum (if it exists), by taking steps in the negative gradient direction, by adjusting the filter coefficients to minimize the instantaneous squared error [12]. The LMS algorithm is widely used in many applications. Figure 2 shows the block diagram of a system identification unit using the LMS algorithm. In Fig. 2, the filter  $W$  will try to model the unknown system  $H$ , by changing its coefficient values.  $d[n]$  is the desired signal, which the filter output  $y[n]$  tries to match as close as possible.  $e[n]$  is computed and is used in the adaptive algorithm to update the filter coefficients.



**Fig. 2.** Block diagram representation of system identification using LMS algorithm.

The following equations can be used to implement the LMS algorithm in real time.

$$e[n] = d[n] - y[n] \quad (1)$$

$$w[n + 1] = w[n] + \mu e[n]x[n] \quad (2)$$

Equation (2) is the coefficient update equation, where  $\mu$  is the step size and determines the convergence speed. The filter  $W$  will alter its coefficients  $w[n]$  in order to reduce the error  $e[n]$ . A small value of  $\mu$  usually implies a long time to converge, whereas a large value of  $\mu$  may never allow the system to converge. This is one disadvantage of the LMS in that it has limitations on the value of the step size and a tradeoff between the minimum error and convergence time exists [13].

### 3 System Identification

System identification is an important procedure in various applications areas, where an unknown model is specified in terms of input-output desired response models. It can be found in various application areas such in control and communication networks. In this section, the procedure and configuration of the PADK for the system identification of

an electrical sub-system and an electro-acoustic path will be outlined. Consequently, the LMS algorithm can also be verified by the following experiment.

### 3.1 System Identification of the Series Combination of the ADC and the DAC of the PADK

Figure 3 shows the block diagram of the PADK connection configuration used for the system identification experiments. In this experiment, the impulse response of an electrical sub-system consisting of the DAC and ADC of the PADK, can be determined. From the impulse response, the delay associated with that electrical sub-system can be found. In Fig. 3, one of the outputs of the DAC is connected to one of the inputs of the ADC via a cable. The sampling frequency of the PADK was set at 32 kHz in the example. A block of white noise signals is generated internally by the PADK. The input reference signal (white noise) is fed to the LMS algorithm block for updating the weighing coefficients of the digital filter. The LMS algorithm will minimize the mean square error signal which is the difference between the outputs of the adaptive filter and that from the unknown system. The filter coefficients that will be generated represent the impulse response of the unknown system which is the series combination of the DAC and ADC of the PADK. The filter length was set at 50. The results are presented in Sect. 5.1.

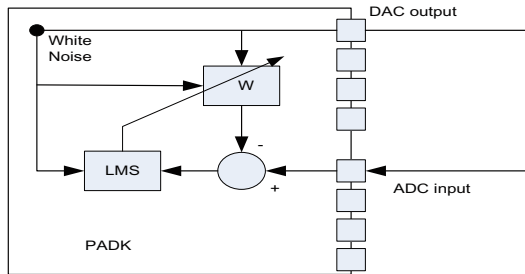


Fig. 3. System identification setup with the PADK

### 3.2 System Identification of the Secondary Path in a Single Channel System for Acoustic Active Noise Control

In this experiment, a single channel system has been implemented consisting of a power amplifier (QD-4960 inter-M), a loudspeaker (Pioneer TS-WX 303) and an electret microphone (with preamplifier Maxim MAX4466). Figure 4 shows a picture of the set up. The objective is to find the impulse response of the secondary path of an active acoustic noise control system, which consists of the DAC, the power amplifier, the loudspeaker, the microphone, the ADC and the acoustics of the room environment. The block of white noise signals generated by the PADK is fed to the loudspeaker through the power amplifier. The acoustic signal is then captured by the microphone and fed again to the ADC input of the PADK. The configuration is the same as that in

Fig. 3 except that the system also includes an electro-acoustic path, between the DAC output and the ADC input, which consists of the power amplifier, loudspeaker and the microphone. Several experiments have been carried out with this set up using different number of filter coefficients. It has been found that to obtain a more accurate impulse response of this electro-acoustic path, the filter length shall be set to the maximum available frame size of the PADK. The results are shown in Sect. 5.2.



Fig. 4. System identification setup in acoustic domain.

## 4 Regenerative Feedback Control

A regenerative feedback control system synthesizes its own reference signal based on the filter output and the error signal. The main idea of such a configuration is to estimate the primary noise and use it as a reference signal to the ANC filter [14]. This technique is used when the desired reference signal is not available. Also, applications of such technique can be found in global control of manufacturing plant noise and in active headsets for noise protection [15]. A regenerative feedback control system has been implemented in the electrical domain using the block diagram shown in Fig. 5. In this experiment, the source of noise is simulated as a sinusoidal wave of frequency 100 Hz.  $P(z)$  represents the primary path which results in the disturbance  $d(n)$  to be controlled.

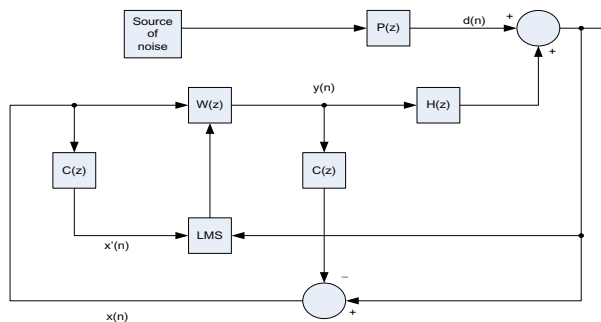
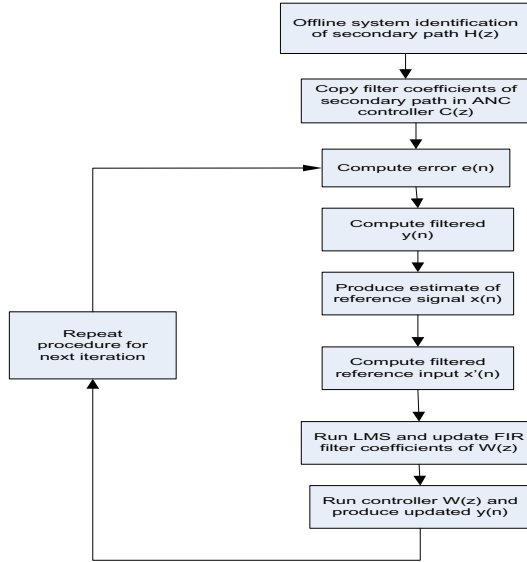


Fig. 5. Block diagram representation of regenerative feedback control in electrical domain [10]



**Fig. 6.** Flowchart of the steps involved in regenerative feedback control

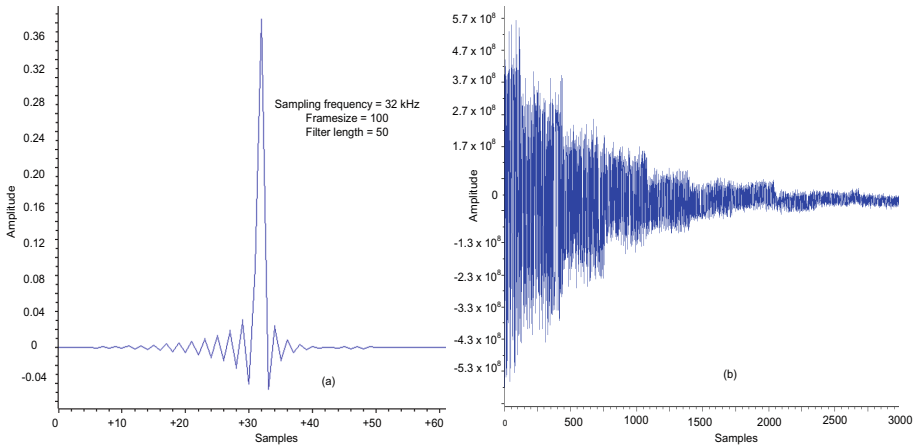
One channel of the DAC is used as the output of the controller whereas one channel of the DAC is used as the source of the electrical generated noise.  $H(z)$  is the secondary path which has been modeled offline by  $C(z)$ . Figure 6 shows a flowchart of the steps involved in the regenerative feedback control. The results are given in Sect. 5.3.

## 5 Results and Interpretation

### 5.1 System Identification of the Series Combination of the DAC and ADC

In this experiment, the value of the normalized step size to be used has been found by running the system many times. A large value of the step size is used first to make the system diverge. A fraction of this value is then used for the system identification, such that there are not many variations in the coefficients generated. Figure 7(a) below shows the filter coefficients for a normalized step size value of 0.03, after the system has been run for a few seconds. The coefficients generated represent the impulse response of the series combination of the ADC and DAC of the PADK. From experiments, it has been found that the effective delay of this electrical sub-system is 70 samples. In Fig. 7(a), it is seen that the impulse response shows a peak at sample 32 after a delay of 38 samples has been introduced in the program. In Fig. 7(b), it is seen that the error signal values clearly decrease indicating that the system converges (after approximately 62.5 ms (2000 samples)).





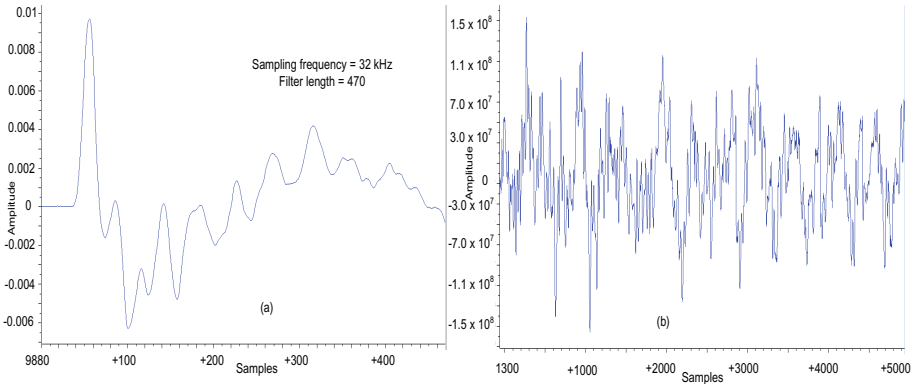
**Fig. 7.** (a) Impulse response of the DAC and ADC of the PADK (b) Error signal for system identification of electrical sub-system

## 5.2 System Identification of the Secondary Path of the Acoustic Active Noise Control System

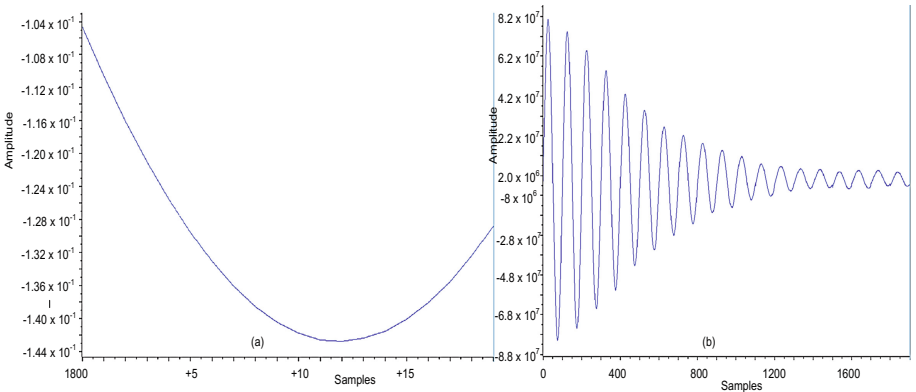
Figure 8 shows the results generated for the secondary path of a single channel electro-acoustic system, which consists of the DAC, power amplifier, loudspeaker, microphone and the ADC. From the impulse response (Fig. 8(a)), it is seen that there is a peak followed by oscillations. From the results of the error signal in Fig. 8(b), it is found that the system does not show a reasonable convergence. Also, the values of the filter coefficients generated are significantly small. Several experiments have been carried out using different values of the step size with no better results even if the system has been run for a few minutes. To better model this electro-acoustic system, it is needed to use a lower sampling rate which will consequently need fewer number of filter coefficients. However, the lowest available sampling frequency of the PADK is 32 kHz. Also, the waveform of the impulse response can also be due to a reverberant acoustic environment.

## 5.3 Regenerative Feedback Control in Electrical Domain

Figure 9 (a) shows the filter coefficients for the feedback controller in which a filter length of 20 coefficients has been used. The value of the step size is chosen using the same technique as described in Sect. 5.1. In this experiment, a normalized step size value of 0.1 has been used for the feedback controller. The values of the filter coefficients generated are sufficiently high for convergence to occur. As shown from the flowchart in Fig. 6, an offline system identification is carried out first after which the feedback controller starts to tune to minimize the error. From the results in Fig. 9(b), the system shows acceptable convergence after approximately 50 ms (1600 samples).



**Fig. 8.** Secondary path of an electro-acoustic single channel system (a) Impulse response (b) Error signal



**Fig. 9.** Regenerative feedback control (a) Filter coefficients of controller (b) Error signal

To carry out the regenerative feedback control in the acoustic domain, a proper model of the electro acoustic path of the system depicted in Sect. 5.2 needs to be obtained. It is intended as future work to externally enable the PADK to have a lower sampling rate than 32 kHz. For example, a sampling frequency of between 2–5 kHz will be sufficient to control low frequency noise below 300 Hz effectively in practice. This will allow fewer filter coefficients for the modeling and controller, thus making the system efficient and practical for the multi-channel case.

## 6 Conclusions and Future Work

In this work, the results of real time processing with the PADK in system identification and regenerative feedback control have been shown. Experiments have been carried out using the LMS adaptive algorithm. Filter coefficients have been generated showing

convergence of the system. In the acoustic domain, the impulse response of the secondary path has been obtained. However, in practice this needs further development as it is necessary to have a lower sampling rate than the available 32 kHz of the PADK. Future work will involve the implementation of multiple channel active acoustic noise control systems using the feedback and feedforward configurations with this kit.

**Acknowledgments.** One of the authors, Sajaad Boodoo, acknowledges the postgraduate bursary received from the Tertiary Education Commission, for carrying out this work.

## References

1. Lueg, P.: Process of silencing sound oscillations. U.S. Patent 2043416, 9 June 1936
2. Kajikawa, Y., Gan, W., Sen, M.K.: Recent advances on active noise control: open issues and innovative applications. *APSIPA Trans. Sig. Inf. Process.* **1**, e3 (2012). doi:[10.1017/ATSIP.2012.4](https://doi.org/10.1017/ATSIP.2012.4)
3. Paurobally, M.R., Sassi, S., Moses, K., Khadembashi, D.: Wireless active noise control results. In: 23rd International Congress on Sound and Vibration, ICSV 23, Athens, Greece, July 2016
4. Holters, M., Zolzer, U.: *Aktive Lärminderung in Wohn- und Schlafräumen*, May 2008. <http://www.dbu.de/OPAC/ab/DBU-Abschlussbericht-AZ-24483.pdf>
5. Lyrtech Signal Processing: Professional Audio Development Kit. Technical Reference Guide, September 2005
6. Texas Instruments: Reference Guide. TMS 320C672x DSP Dual Data Movement Accelerator (dMAX), November 2005, Revised 2007
7. Gaikwad, P., Jain, M., Pranita, Y., Bhandari, S.U.: Active noise control using field programmable gate arrays. *Int. J. Adv. Res. Electr. Electron. Instrum. Eng.* **2**(2), 793–799 (2013)
8. Chassaing, R., Reay, D.: *Digital Signal Processing and Applications with the TMS 320 C6713 and TMS 320 C6416 DSK*, 2nd edn. Wiley, Hoboken (2008)
9. Texas Instruments: TMS 320 C6713 Floating Point Digital Signal Processor. SPRS186L, December 2001, Revised November 2005
10. Texas Instruments: TMS 320C6727, TMS 320C6726, TMS 320C6722, Floating Point Digital Signal Processor. SPRS268E, May 2005, Revised January 2007
11. Haykin, S.: *Adaptive Filter Theory*, 3rd edn. Prentice Hall, Upper Saddle River (1996)
12. Dornean, I., Topa, M., Kirel, B.S., Szopos, E.: System identification with the least mean square adaptive algorithm. In: *Interdisciplinary in Engineering Scientific International Conference*, TG Mures, Romania, 15–16 November 2007
13. Elliot, S.J.: *Signal Processing for Active Control*. Academic Press, London (2001)
14. Kuo, M.S., Panahi, I., Chung, K.M., Horner, Y., Nadeski, M., Chyan, J.: *Design of Active Noise Control Systems with the TMS320 Family*. SPRA042, June 1996, Texas Instruments Incorporated
15. Kuo, S.M., Mitra, S., Gan, W.S.: Active noise control system for headphone applications. *IEEE Trans. Control Syst. Technol.* **14**(2), 331–335 (2006)

# Cooperative Spectrum Sensing in the DSA: Simulation of Spectrum Sensing Time Consumptions by Cognitive Radio Secondary Users

Masiala Mavungu<sup>(✉)</sup> and A.L. Nel

Mechanical Engineering and Engineering Management,  
University of Johannesburg (UJ), (Auckland Park Kingsway Campus) (APK):  
Corner Kingsway and University Rd,  
Rossmore, Johannesburg 2001, Gauteng, South Africa  
masialamavungu567@gmail.com, andren@uj.ac.za

**Abstract.** Wireless Communication and Signal Processing have become areas of flourishing research and innovation. Mathematical Modeling and Computational Simulations emerge as powerful ways to describe, analyze and control situations, solve problems and then interpret the results. The aim of this paper is to analyze the optimal sensing time made by cognitive radio secondary users during cooperative spectrum sensing in the Distributed Spectrum Utilization. The optimal time made by each cognitive radio user participating in cooperative spectrum sensing is modelled as a stochastic differential equation. The combination of all these sensing of the spectrum leads to a system of stochastic differential equations which is solved using a Runge-Kutta method. Computational simulations are provided to show that the total spectrum sensing time made by cognitive radios in the cooperative spectrum sensing is less than the total time they made in the spectrum sensing without cooperation.

**Keywords:** Wireless Communication · Dynamic Spectrum Utilization · Spectrum sensing · Digital signal processing

## 1 Introduction

Dynamic Spectrum Utilization (DSU) is a new technology to ameliorate the utilization of spectrum. Before utilizing an opening in the spectrum the cognitive radio (CR) users must perform spectrum sensing to obtain necessary information about the real-time of the spectrum to be accessed.

Spectrum sensing is a key function required in CR to prevent harmful interference between licensed users and to identify the available spectrum for opportunistic spectrum use improving the spectrum utilization. However, detection performance in practice is often compromised by multipath fading, shadowing and receiver uncertainty issues. To mitigate the impact of these issues, cooperative spectrum sensing has been shown to be an effective method to improve the detection performance by exploiting spatial diversity. CSS can reduce the detection time and improve the spectrum sensing

and utilization [4]. The main idea behind CSS is improvement of the sensing performance by exploiting the spatial diversity in the observations of spatially located CR users. By cooperation, CR users can share their sensing information making a combined decision more accurate than the individual decisions [5].

Spectrum sensing is time consuming so that modelling the time consumed by the CR secondary users during the sensing operation is a relevant concern. The performance of the spectrum sensing depends on the sensing time and the fusion scheme that is used when cooperative spectrum sensing is applied [2]. The longer spectrum sensing is performed, the more accurate the results are about the spectrum status.

There exist in the literature works discussing CSS time. Reference [2], proposes a method to maximize the throughput of the secondary users where the sensing time is one of the parameters with the main result of significant achievement in the improvement and the throughput of the secondary users when the parameters for the combination of the sensing results and the sensing time are jointly optimized. Reference [3], considers a sensing-throughput tradeoff problem for optimizing the sensing time that maximizes the secondary users' throughput while reducing or completely preventing interference to the primary user. Reference [4], reports the state of the art in cooperative sensing around the issues of first analyzing the cooperation method with the fundamental components of cooperative sensing and then presenting the impacting factors of achievable cooperative gain and incurred cooperation overhead. In addition, they identify open research challenges related to each issue in cooperative sensing along with the discussion. Reference [6], presents a sensing time optimization algorithm which has the purpose of maximizing the spectral efficiency of an Ultra Wide-band (UWB) based CR system by finding the optimal tradeoff between the length of the sensing window and the detection probability in the low-SNR regime. Under the constraint of probability of detection and that of instantaneous interference, reference [7] proposes an efficient algorithm finding optimal sensing time and an optimal power allocation scheme.

In this paper, we aim at mathematically modelling the CSS and simulating the optimal sensing time made by CR secondary users during such a CSS process.

The rest of the paper is organized as follows: Sect. 2 models every CR secondary user's time consumption in the CSS as a stochastic differential equation and combining the differential equations into a system of such equations. Section 3 solves the system of stochastic differential equations to obtain the time consumption function for each cognitive radio secondary user. Section 4 concludes the paper by deliberating on the results achieved.

## 2 Mathematical Model

For every CR secondary user  $i$ , define  $T_i(t)$  to be the optimal time to sense the spectrum status on time period  $[0, t]$ . When a selected CR secondary user  $i$  is the only one involved in sensing spectrum then there is no specific cooperative interaction. In general,  $T_i(t)$  satisfies the following of differential equation:

$$\frac{dT_i}{dt} = T_i(a_i - b_{ii}T_i). \tag{1}$$

which is a pure logistic equation [1].

$\forall i, a_i$  and  $b_{ii}$  are coefficients of proportionality. By considering the presence of an additional and  $N - 1$  CR secondary users in the same group, the above equation can be rewritten as

$$\frac{dT_i}{dt} = T_i\left(a_i - b_{ii}T_i - \sum_{j=1, j \neq i}^N b_{ij}T_j\right). \tag{2}$$

Which can be rewritten as

$$\frac{dT_i}{dt} = a_iT_i - b_{ii}T_i^2 - \sum_{j=1, j \neq i}^N b_{ij}T_iT_j. \tag{3}$$

Where  $\sum_{j=1, j \neq i}^N b_{ij}T_iT_j$  is the sum of interactions between CR secondary user  $i$ 's time consumption with the other CR secondary users  $j$ 's time consumption. It can be clearly seen from the  $-\sum_{j=1, j \neq i}^N b_{ij}T_iT_j$  term above the presence of each CR secondary term above the presence of each CR secondary user reduces the sensing times available to other CR users.

The shared spectrum environment is subject to significant and considerable randomness. Such a randomness is involved in every sensing duration of every CR secondary user. The above equation can be rewritten as:

$$\frac{dT_i}{dt} = T_i\left(a_i - b_{ii}T_i - \sum_{j=1, j \neq i}^N b_{ij}T_j\right) + \xi_i. \tag{4}$$

Where  $\xi_i$  is a colored noise such that  $(\xi_i(t))_{t \geq 0}$  is Ornstein-Uhlenbeck stochastic process and is defined by:

$$d\xi_i = \alpha_i(\beta_i - \xi_i)dt + \gamma_i dW_i \quad i = 1, \dots, N \tag{5}$$

where  $W_i$  is a white noise such that  $(W_i(t))_{t \geq 0}$  is a Wiener stochastic process and  $a_i, b_{ii}, b_{ij}, \alpha_i, \beta_i$  and  $\gamma_i$  are positive constant parameters of proportionality, and is the number of CR secondary users.

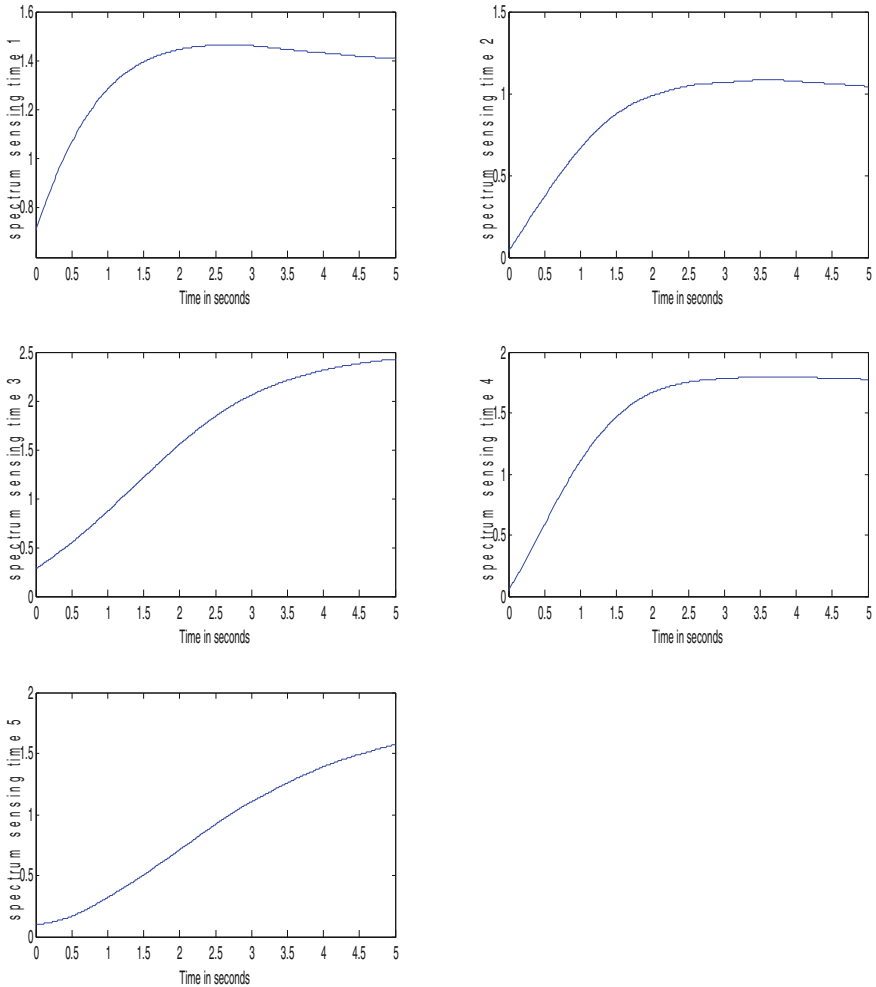
By letting  $\theta_i = \frac{dT_i}{dt}$  and by considering the specific case for  $N = 5$  CR secondary users, we obtain the interactive dynamics of all the 5 CR secondary users by a system of 10 stochastic differential equations as follows:

$$\frac{dT_i}{dt} = T_i\left(a_i - b_{ii}T_i - \sum_{j=1, j \neq i}^N b_{ij}T_j\right) + \xi_i \tag{6}$$

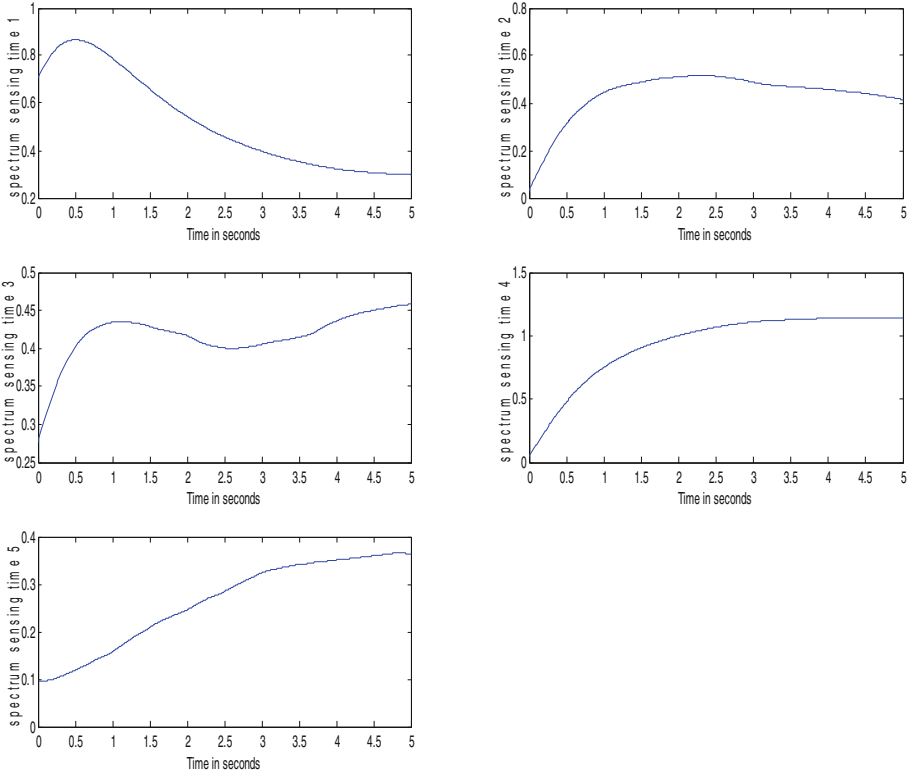
$$\frac{d\xi_i}{dt} = \alpha_i(\beta_i - \xi_i) + \gamma_i\theta_i \quad i = 1, \dots, 5. \tag{7}$$

### 3 Computational Simulations and Results

By using a fourth-order Runge-Kutta method to solve the above system of 10 stochastic differential equations we obtain the spectrum sensing for every cognitive radio secondary user when there is no cooperation and when there is cooperation. For every cognitive radio  $i$ , the time spent when there is cooperation is smaller then the time spent when there is no cooperation (Figs. 1, 2 and 3).

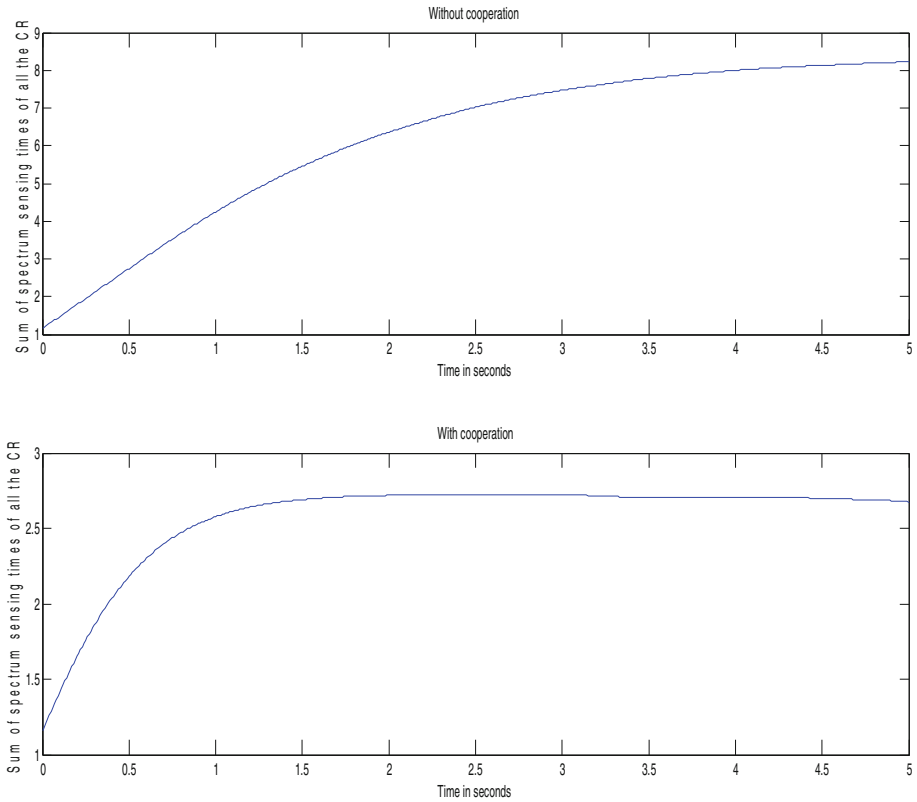


**Fig. 1.** Plot of spectrum sensing times of all the cognitive radio secondary users without cooperation.



**Fig. 2.** Plot of spectrum sensing times of all the cognitive radio secondary users in cooperation.





**Fig. 3.** Comparison of the sum of spectrum sensing times of all the cognitive radio secondary users without cooperation and in cooperation.

## 4 Conclusion

The aim of this paper was to mathematically model the optimal spectrum sensing time made by the cognitive radios secondary users during cooperative spectrum sensing before deciding on spectrum utilization and to numerically and computationally simulate them. To solve the problem, we used the following approach:

Firstly, we model every cognitive radio secondary user's time consumption in the cooperative spectrum sensing as a stochastic differential equation and the combination of all the differential equations as a system of stochastic differential equations.

Secondly, we coded the obtained system of stochastic differential equations as well as the fourth-order Runge-Kutta numerical method into a Matlab function files and then we designed a Matlab script function file to call the Runge-Kutta function to solve the system of stochastic differential equations. We plot each cognitive radio secondary user's optimal sensing time consumption when there is no cooperation and when there is cooperation with the others. Such simulations show that the total spectrum sensing

time made by the all the CRs for primary user detection in the cooperative spectrum sensing is less than the total spectrum sensing time made in the spectrum sensing without cooperation.

## References

1. Morris, W.H., Stephen, S., Robert, L.D.: *Differential Equations, Dynamical Systems, An Introduction to Chaos*, vol. 60, 2nd edn. Elsevier Academic Press, Cambridge (2004)
2. Edward, C.Y., Ying-Chang, L., Yong, L.G., Yonghong, Z.: Optimization of cooperative sensing in cognitive radio networks: a sensing-throughput tradeoff view. *IEEE Trans. Veh. Technol.* **58**(9), 5294–5299 (2009)
3. Ying-Chang, L., Yonghong, Z., Edward, C.Y.P., Anh, T.H.: Sensing-throughput tradeoff for cognitive radio networks. *IEEE Trans. Wirel. Commun.* **7**(4), 1326–1337 (2008)
4. Ian, F.A., Brandon, F.L., Ravikumar, B.: Cooperative spectrum sensing in cognitive radio networks: a survey. *Phys. Commun.* **4**(4), 40–62 (2011)
5. Danijela, C., Shridhar, M.M., Robert, W.B.: Implementation issues in spectrum sensing for cognitive radios. In: *Proceedings of the Asilomar Conference on Signals, Systems, and Computers*, vol. 1, pp. 772–776 (2004)
6. Liaoyuan, Z., Sean, M.: Spectrum sensing time optimization algorithm for spectrum efficiency maximization in the low-power cognitive radio ultra wideband system. In: *International Conference on Electronic Engineering and Computer Science* (2013). *IERI Procedia* **4**, 68–73 (2013)
7. Pei, E., Li, J.B., Cheng, F.: Sensing-throughput tradeoff for cognitive radio networks with additional primary transmission protection. *J. Comput. Inf. Sys.* **9**, 3767–3773 (2013)

# **Communication Engineering and Computing**

# Large Scale Fading Pre-coder for Massive MIMO Without Cell Cooperation

Tedros Salih<sup>1</sup>(✉), Elijah Mwangi<sup>2</sup>, and Kibet Langat<sup>3</sup>

<sup>1</sup> Department of Electrical Engineering, Pan African University,  
Po Box 62000, 00200 Nairobi, Kenya  
tedysal2001@gmail.com

<sup>2</sup> School of Engineering, University of Nairobi,  
Po Box 30197, Nairobi, Kenya  
elijah.mwangi@uonbi.ac.ke

<sup>3</sup> Department of Telecommunication and Information Engineering,  
Jomo Kenyatta University of Agriculture and Technology,  
Po Box 62000, Nairobi, Kenya  
kibetlp@jkuat.ac.ke

**Abstract.** In massive MIMO technology the channel is estimated using uplink training by sending an orthogonal pilot sequence from users to the base station. These sequences are re-used in the cell and also outside the cell. This gives rise to a channel estimation error referred to as pilot contamination. Large scale fading precoding which is based on the cooperation between cells has been proposed to mitigate pilot contamination. However this approach is known to limit in data transmission rate. In this paper, we propose a novel uplink training scheme to mitigate pilot contamination using a large scale fading precoding without the need of cooperation between cells. This achieves a higher transmission rate over existing method. The simulation results show that the proposed scheme improves 5% outage rate 10 times, over the existing method.

**Keywords:** Large scale fading precoding · Massive MIMO · Pilot contamination · Uplink training

## 1 Introduction

The effect of multipath fading, limited power at the transmitter and scarce spectrum makes the task of designing high data rate, high reliability wireless communication systems extremely challenging [1]. A solution to this challenge is the use of Multiple Input and Multiple Output (MIMO) technology that uses multiple antennas both for transmission and reception. When the number of antennas at the transmitter and receiver increases, the degree of freedom in propagation channel also increases and gives an improvement in the data rate and link reliability [2].

In point-to-point MIMO, the hardware complexity and power consumption of the signal processing units increase at both the transmitter and receiver sides [2]. In addition, it is also un-scalable and has unfavorable propagation properties. The unfavorable propagation can be minimized using the multi user MIMO (MU-MIMO)

technique in which the multiple antennas at the receiver side are divided into many independent terminal users. However, MU-MIMO is not a scalable technology since it was designed to have an equal number of service antennas and terminals and to operate in frequency division duplex mode [3]. It also uses dirty-paper coding/decoding, and requires the knowledge of the channel state information (CSI) at both ends of the link. This increases the computational complexity as the number of antennas increase [4].

The massive MIMO technology has been proposed as a solution to scalability. Massive MIMO is a multi-user MIMO technology where each base station (BS) is equipped with an array of many active antenna elements and utilizes these to communicate with several single-antenna terminals over the same frequency band [4–7].

In massive MIMO the BS needs channel state information to detect the signal that comes from all users. This CSI is obtained through uplink training. In uplink training multiple users send orthogonal signals that are known at the base station and then the base station estimates the channel based on the received pilot signal. Massive MIMO operates in a TDD system (time division duplex) [8, 9]. In the TDD system, the uplink channel and downlink channel are considered to be reciprocals for a given coherence time. This coherence time is divided into a number of time slots using the TDD protocol for uplink training, uplink data and downlink data. Due to the coherence time limitation the time slot that are given for uplink training are very short. If we increase the amount of time required for uplink training the amount of time required to send data for uplink and downlink will be decreased and causes the data rate to decrease. For this reason the orthogonal signals that are used for uplink training are limited and the users of different cells cannot have a unique assigned orthogonal signal rather the orthogonal signals are re-used within a cell or adjacent cell. Using of the same orthogonal signal for different users leads to imperfect of the channel estimation. This channel estimation error is referred as pilot contamination.

A pilot contamination mitigation scheme based on large scale fading precoding has been proposed in [10] and in [11]. It is based on the assumption that the signals from all terminals in all cells are accessible at each BS and that slow-fading coefficients are accessible to all the BS or alternatively to a network hub. Instead of mitigating interference caused by pilot contamination in estimating of the channel, each BS uses the pilot contamination for transmitting information to all terminals. This mitigation technique has a high computational complexity since it uses cooperation between cells. To avoid the cooperation between cells, we propose a new uplink training scheme that uses a large scale fading precoding within the cell. Our proposed uplink training scheme that avoids cell cooperation gives a high achievable sum rate compared to the scheme that uses cooperation between cells as reported in literature [11].

The rest of the paper is organized as follows. The system model is presented and discussed in Sect. 2. The pilot contamination analysis is presented in Sect. 3, and in Sect. 4, a discussion of the large scale fading precoding is discussed. In Sect. 5, a derivation of the achievable rate with a finite number of base station antennas is given. In Sect. 6, the computer simulation results that support our proposed method are presented. Finally, Sect. 7 gives a conclusion and suggestions for future extension of the work.

## 2 System Model

Consider  $L$  hexagonal cells, each consisting of one BS and  $K$  single antenna users. In each cell the BS consists of  $W$  distributed sub-array antennas and each sub-array has  $M$  antennas, as shown in Fig. 1.

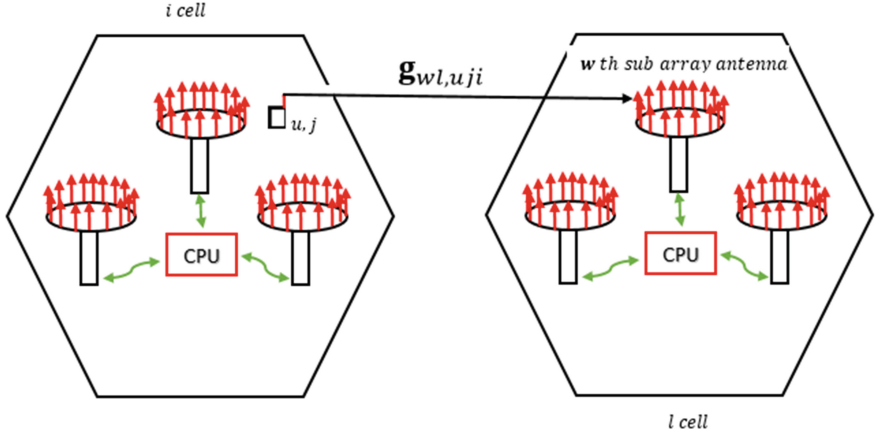


Fig. 1. Multi cell Massive MIMO

We design the uplink training so that each cell has  $U$  orthogonal signals, where  $U = \frac{K}{L}$  and since  $K > U$ , all  $U$  orthogonal signals are reused within the cell  $J$  times. We assume  $J, L$ , and  $W$  to be equal. Each orthogonal signal will have a length of  $\tau = K$  and the  $U$  orthogonal signals  $\{\phi_{1,l}, \phi_{2,l}, \dots, \phi_{U,l}\}$  have the following property:

- (i) For adjacent cells  $\phi_{n,l} \phi_{r,i}^\dagger = 0$
- (ii) Within the cell  $\phi_{n,l} \phi_{r,l}^\dagger = \delta_{nr}$ .

The propagation channel  $\mathbf{g}_{wl,uji}$  is modelled as the product of small scale fading (fast fading) and large scale fading (slow fading) [10–16]. The propagation channel coefficient from  $i$ -th cell of  $(u, j)$  user to the  $l$ -th cell of  $w$ -th sub-array sub array of antenna given as:

$$\mathbf{g}_{wl,uji} = \mathbf{h}_{mwl,uji} \sqrt{\beta_{wl,uji}}. \quad (1)$$

Where  $\mathbf{h}_{mwl,uji} \sim \mathcal{CN}(0, 1)$  small scale fading and  $\beta_{wl,uji}$  is a large scale fading. We consider the slow fading to be different at all  $w$  sub-array of the BS. However, since the distance between the antennas of the sub-array BS is very small compared to the distance of between the  $(u, j)$  user and the sub-array BS, the slow fading ( $\beta_{wl,uji}$ ) can be considered to be independent of  $M$  antennas of the sub-array of BS.

The received signal at the  $w$ -th sub-array of the BS that comes from users in the cell given by:

$$\mathbf{Y}_{l,w} = \sum_{i=1}^L \sum_{j=1}^L \sum_{u=1}^U \sqrt{\tau\rho_t} \mathbf{g}_{wl,uji} \phi_{u,l} + \mathbf{n}_{l,w}. \quad (2)$$

Where  $\mathbf{n}_{l,w} \in \mathbb{C}^{M \times U}$  is the additive noise with distribution  $\mathbf{n}_{l,w} \sim \mathcal{CN}(0, 1)$  and  $\rho_t$  is the average transmitted power at each user.

The  $w$ -th sub-array of BS estimates the channel  $\mathbf{g}_{wl,uji}$  from its own  $U$  users by multiplying Eq. (2) signal with  $\frac{\phi_{u,l}^\dagger}{\sqrt{\tau\rho_t}}$  i.e

$$\hat{\mathbf{g}}_{wl,uwl} = \mathbf{Y}_{l,w} \frac{\phi_{u,l}^\dagger}{\sqrt{\tau\rho_t}} = \mathbf{g}_{wl,uwl} + \sum_{j=1, j \neq w}^L \mathbf{g}_{wl,ujj} + \frac{\phi_{u,l}^\dagger}{\sqrt{\tau\rho_t}} \mathbf{n}_{l,w}. \quad (3)$$

The  $w$ -th sub-array of BS pre-code the desired signal  $s_{uwl}$  using a conjugated beamforming precoder  $T_{wu}$  [10–13]. The precoder  $T_{wu}$  given by:

$$T_{wu} = \frac{\hat{\mathbf{g}}_{wl,uwl}^\dagger}{\hat{\mathbf{g}}_{wl,uwl}}. \quad (4)$$

The  $w$ -th sub-array of BS sends the transmitted signal vector  $T_w$  to the intended user  $(u, j)$ . The vector  $T_w$  given by:

$$T_w = T_{wu}s_{uwl} = T_{w1}s_{1wl} + T_{w2}s_{2wl} + \dots + T_{wU}s_{Uwl}. \quad (5)$$

The sub-group user  $(u, j)$  receives the signal:

$$x_{ujl} = \sqrt{\rho_b} \sum_{i=1}^L \sum_{w=1}^L \mathbf{g}_{wl,uji} T_w + w_{ujl}, w_{ujl} \in \mathcal{CN}(0, \sigma_{ujl}). \quad (6)$$

$$x_{ujl} = \sqrt{\rho_b} \sum_{i=1}^L \sum_{w=1}^L \sum_{q=1}^U \mathbf{g}_{wl,uji} T_{wq} s_{qwl} + w_{ujl}. \quad (7)$$

Where  $\rho_b$  the average transmit power from  $w$ -th sub-array of BS.

### 3 Pilot Contamination Analysis

The analysis of the pilot contamination effect when the number of antennas approaches infinity has been deduced under the assumption of uplink training reuse in adjacent cells [10, 11, 13–15]. Here we analyze the pilot contamination effect when the uplink training is reused within cell but different in adjacent cells.

*Lemma 1.* Let  $x, y \in \mathbb{C}^{M \times 1}$  be two independent vectors with distribution  $\mathcal{CN}(0, c\mathbf{I})$ . Then;

$$\lim_{M \rightarrow \infty} \frac{x^\dagger y}{M} = 0 \text{ and } \lim_{M \rightarrow \infty} \frac{x^\dagger x}{M} = c. \quad (8)$$

Using the above Lemma when the number of antenna approaches infinity Eq. (7) becomes

$$\lim_{M \rightarrow \infty} \frac{x_{ujl}}{\sqrt{M}} = \lim_{M \rightarrow \infty} \sqrt{\rho_b} \sum_{i=1}^L \sum_{w=1}^L \sum_{q=1}^U \frac{\mathbf{g}_{wl,ujl} T_{wq} s_{qwl}}{\sqrt{M}} + \frac{w_{ujl}}{\sqrt{M}}. \quad (9)$$

$$\lim_{M \rightarrow \infty} \frac{\mathbf{g}_{wl,ujl} T_{wq}}{\sqrt{M}} = \lim_{M \rightarrow \infty} \left( \frac{\mathbf{g}_{wl,ujl} \hat{\mathbf{g}}_{wl,qwl}^\dagger}{M} \frac{1}{\frac{\|\hat{\mathbf{g}}_{wl,qwl}^\dagger\|}{\sqrt{M}}} \right). \quad (10)$$

$$\lim_{M \rightarrow \infty} \frac{\mathbf{g}_{wl,ujl} \hat{\mathbf{g}}_{wl,qwl}^\dagger}{M} = \begin{cases} \beta_{wl,ujl} & \text{for } q = u, i = l \\ 0 & \text{for } q \neq u \end{cases}. \quad (11)$$

$$\lim_{M \rightarrow \infty} \frac{\|\hat{\mathbf{g}}_{wl,qwl}^\dagger\|}{\sqrt{M}} = \mathcal{B}_{ujl} = \left( \sum_{j=1}^L \beta_{wl,ujj} + 1 \right)^{\frac{1}{2}}. \quad (12)$$

$$\lim_{M \rightarrow \infty} \frac{x_{ujl}}{\sqrt{M}} = \sqrt{\rho_b} X_{ujl} = \sqrt{\rho_b} \sum_{w=1}^L \beta'_{wl,ujl} s_{uwl}. \quad (13)$$

Where  $\beta'_{wl,ujl} = \frac{\beta_{wl,ujl}}{\mathcal{B}_{ujl}}$ .

From Eq. (13) we can observe two things. The first is that the signal that comes from adjacent cell vanish when  $M$  approaches infinity. The second is that the  $(u, j)$ -th user in the  $l$ -th receives not only the intended signal that comes from the BS of the  $l$ -th cell but also from other sub-arrays in the BS of the  $l$ -th cell. This causes signal interference. This interference happened due to pilot contamination error estimation of the channel at the sub array of BS of the  $l$ -th cell. The signal-to-interference-plus-noise ratio (SINR) is given in Eq. (14) and from this we can observe that the SINR is limited as  $M$  approaches infinity. But this limitation can be avoided by using large scale fading precoding.

$$SINR_{u1} = \frac{\beta_{1l,ujl}^2}{\beta_{2l,ujl}^2 + \beta_{3l,ujl}^2 + \dots + \beta_{L,ujl}^2} = \frac{\beta_{1l,ujl}^2}{\sum_{w=2}^L \beta_{wl,ujl}^2}. \quad (14)$$



### 4 Large Scale Fading Precoding

In this precoding scheme it is assumed that the slow fading coefficients are estimated. To mitigate the pilot contamination, the  $w$ -th sub-array can add to a pre-coder unknown variable form  $A_1, A_2 \dots A_L$  so as to get the received signal as:

$$\begin{aligned}
 X_{u1} &= \beta'_{1l,u1l} A_1 s_{u1l} + \beta'_{2l,u1l} A_2 s_{u2l} + \dots + \beta'_{Ll,u1l} A_L s_{uLl} = s_{u1l} \\
 X_{u2} &= \beta'_{1l,u2l} A_1 s_{u1l} + \beta'_{2l,u2l} A_2 s_{u2l} + \dots + \beta'_{Ll,u2l} A_L s_{uLl} = s_{u2l} \\
 &\vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \qquad \qquad \qquad \qquad \qquad \vdots \\
 X_{uL} &= \beta'_{1l,uLl} A_1 s_{u1l} + \beta'_{2l,uLl} A_2 s_{u2l} + \dots + \beta'_{Ll,uLl} A_L s_{uLl} = s_{uLl}
 \end{aligned} \tag{15}$$

Then Eq. (15) can be written in matrix form

$$\begin{pmatrix} \beta'_{1l,u1l} & \beta'_{2l,u1l} & \dots & \beta'_{Ll,u1l} \\ \beta'_{1l,u2l} & \beta'_{2l,u2l} & \dots & \beta'_{Ll,u2l} \\ \vdots & \vdots & \ddots & \vdots \\ \beta'_{1l,uLl} & \beta'_{2l,uLl} & \dots & \beta'_{Ll,uLl} \end{pmatrix} \begin{pmatrix} A_1 s_{u1l} \\ A_2 s_{u2l} \\ \vdots \\ A_L s_{uLl} \end{pmatrix} = \begin{pmatrix} s_{u1l} \\ s_{u2l} \\ \vdots \\ s_{uLl} \end{pmatrix}. \tag{16}$$

$$\begin{pmatrix} v_{u1l} \\ v_{u2l} \\ \vdots \\ v_{uLl} \end{pmatrix} = \begin{pmatrix} a_{1l,u1l} & a_{2l,u1l} & \dots & a_{Ll,u1l} \\ a_{1l,u2l} & a_{2l,u2l} & \dots & a_{Ll,u2l} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1l,uLl} & a_{2l,uLl} & \dots & a_{Ll,uLl} \end{pmatrix} \begin{pmatrix} s_{u1l} \\ s_{u2l} \\ \vdots \\ s_{uLl} \end{pmatrix}. \tag{17}$$

Where  $A_w s_{uwl} = v_{uwl}$  and

$$\begin{pmatrix} a_{1l,u1l} & a_{2l,u1l} & \dots & a_{Ll,u1l} \\ a_{1l,u2l} & a_{2l,u2l} & \dots & a_{Ll,u2l} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1l,uLl} & a_{2l,uLl} & \dots & a_{Ll,uLl} \end{pmatrix} = \begin{pmatrix} \beta'_{1l,u1l} & \beta'_{2l,u1l} & \dots & \beta'_{Ll,u1l} \\ \beta'_{1l,u2l} & \beta'_{2l,u2l} & \dots & \beta'_{Ll,u2l} \\ \vdots & \vdots & \ddots & \vdots \\ \beta'_{1l,uLl} & \beta'_{2l,uLl} & \dots & \beta'_{Ll,uLl} \end{pmatrix}^{-1}. \tag{18}$$

The  $w$ -th sub-array of BS pre-code the desired signal  $v_{uwl}$  using a conjugated beamforming  $T_{wu}$  and sends the vector  $T_w$  to the intended user  $(u, j)$ . This precoding scheme is called large scale fading precoding. The vector  $T_w$  given by:

$$T_w = T_{wu} v_{uwl} = T_{w1} v_{1wl} + T_{w2} v_{2wl} + \dots + T_{wU} v_{Uwl}. \tag{19}$$

The sub-group user  $(u, j)$  receives the signal

$$x_{ujl} = \sqrt{\rho_b} \sum_{i=1}^L \sum_{w=1}^L \sum_{q=1}^U \mathbf{g}_{wl,uji} T_{wq} v_{qwl} + \mathbf{w}_{ujl}. \quad (20)$$

$$\lim_{M \rightarrow \infty} \frac{x_{ujl}}{\sqrt{M}} = \sqrt{\rho_b} \sum_{w=1}^L \beta'_{wl,ujl} v_{uwl} = \sqrt{\rho_b} s_{ujl}. \quad (21)$$

## 5 Achievable Rates with Finite M

In this section we derived the achievable rate of the received signal when a finite number of antennas are used. We consider the MMSE estimator and the received signal as given by Eq. (2) then becomes;

$$\hat{\mathbf{g}}_{wl,uwl} = \mathbf{Y}_{l,w} \left( \Phi_{u,l}^\dagger \Phi'_{wl,uwl} \right) = \phi'_{wl,uwl} \sqrt{\tau} \rho_t \sum_{j=1}^L \mathbf{g}_{wl,ujl} + \mathbf{n}'_{l,w}$$

Where;

$$\phi'_{wl,uwl} = \frac{\sqrt{\tau} \rho_t \beta_{wl,uwl}}{\zeta_{wl,uwl}^2}, \quad \zeta_{wl,uwl}^2 = 1 + \rho_t \tau \sum_{z=1}^L \beta_{wl,uzl}$$

The vectors  $\mathbf{g}_{wl,uji}$  and  $\hat{\mathbf{g}}_{wl,uwl}$  have the following distributions;

$$\mathbf{g}_{wl,uji} \sim \mathcal{CN}(0, \beta_{wl,uji} \mathbf{I}_M). \quad (22)$$

$$\hat{\mathbf{g}}_{wl,uwl} \sim \mathcal{CN}\left(0, \left(\frac{\beta_{wl,ujl}}{\zeta_{wl,uwl}}\right)^2 \mathbf{I}_M\right). \quad (23)$$

The sub-group user  $(u, j)$  receives the signal:

$$x_{ujl} = \sqrt{\rho_b} \sum_{i=1}^L \sum_{w=1}^L \sum_{q=1}^U \mathbf{g}_{wl,uji} \frac{\hat{\mathbf{g}}_{wl,qwl}^\dagger}{\zeta_{wl,qwl}} v_{qwi} + \mathbf{w}_{ujl}. \quad (24)$$

We assume that the conjugated beamforming of  $T_{wq} = \frac{\hat{\mathbf{g}}_{wl,qwl}^\dagger}{\zeta_{wl,qwl}}$

After some manipulation, Eq. (24) can be written as:

$$x_{ujl} = F_1 + F_2 + F_3 + F_4 + F_5 + F_6. \quad (25)$$

Where;

$$\begin{aligned}
 F_1 &= s_{ujl} \sqrt{\rho_b} \sum_{w=1}^L \mathbf{a}_{wl,ujl} \mathbf{E} \left[ \mathbf{g}_{wl,ujl} \frac{\hat{\mathbf{g}}_{wl,uwl}^\dagger}{\zeta_{wl,uwl}} \right] \\
 F_2 &= s_{ujl} \sqrt{\rho_b} \sum_{w=1}^L \mathbf{a}_{wl,ujl} \left( \mathbf{g}_{wl,ujl} \frac{\hat{\mathbf{g}}_{wl,uwl}^\dagger}{\zeta_{wl,qwl}} - \mathbf{E} \left[ \mathbf{g}_{wl,ujl} \frac{\hat{\mathbf{g}}_{wl,uwl}^\dagger}{\zeta_{wl,uwl}} \right] \right) \\
 F_3 &= \sqrt{\rho_b} \sum_{w=1}^L \mathbf{g}_{wl,ujl} \frac{\hat{\mathbf{g}}_{wl,uwl}^\dagger}{\zeta_{wl,uwl}} \sum_{\substack{p=1 \\ p \neq j}}^L s_{upl} \mathbf{a}_{wl,upl} \\
 F_4 &= \sqrt{\rho_b} \sum_{w=1}^L \sum_{\substack{q=1 \\ q \neq u}}^U \mathbf{g}_{wl,ujl} T_{wq} v_{qwl} \\
 F_5 &= \sqrt{\rho_b} \sum_{\substack{i=1 \\ i \neq l}}^L \sum_{w=1}^L \sum_{q=1}^U \mathbf{g}_{wl,ujl} T_{wq} v_{qwi} \\
 F_6 &= \mathbf{w}_{ujl}
 \end{aligned}$$

The achievable rate given by:

$$\mathbf{C} = \log_2 \left( 1 + \frac{\mathbf{E} \left[ |\mathbf{F}_1|^2 \right]}{\mathbf{E} \left[ |\mathbf{F}_2|^2 \right] + \mathbf{E} \left[ |\mathbf{F}_3|^2 \right] + \mathbf{E} \left[ |\mathbf{F}_4|^2 \right] + \mathbf{E} \left[ |\mathbf{F}_5|^2 \right] + \mathbf{E} \left[ |\mathbf{F}_6|^2 \right]} \right). \tag{26}$$

Similar to [11] we can have;

$$|\mathbf{F}_1|^2 = \rho_b M^2 \left| \sum_{w=1}^L \frac{\rho_t \tau \beta_{wl,ujl} \beta_{wl,uwl} \mathbf{a}_{wl,ujl}}{1 + \rho_t \tau \sum_{z=1}^L \beta_{wl,uzl} \zeta_{wl,qwl}} \right|^2. \tag{27}$$

$$\mathbf{E} \left[ |\mathbf{F}_2|^2 \right] = \rho_b M \sum_{w=1}^L \left| \frac{\mathbf{a}_{wl,ujl}}{\zeta_{wl,qwl}} \right|^2 \frac{\rho_t \tau \beta_{wl,ujl} \beta_{wl,uwl}^2}{1 + \rho_t \tau \sum_{z=1}^L \beta_{wl,uzl}}. \tag{28}$$

$$\begin{aligned}
\mathbb{E}\left[|F_3|^2\right] &= \rho_b M \sum_{w=1}^L \sum_{\substack{p=1 \\ p \neq j}}^L \left| \frac{\mathbf{a}_{wl,upl}}{\zeta_{wl,uwl}} \right|^2 \frac{\rho_t \tau \beta_{wl,ujl} \beta_{wl,uwl}^2}{1 + \rho_t \tau \sum_{z=1}^L \beta_{wl,uzl}} \\
&+ \rho_b M^2 \sum_{\substack{p=1 \\ p \neq j}}^L \left| \sum_{w=1}^L \frac{\rho_t \tau \beta_{wl,ujl} \beta_{wl,uwl}}{1 + \rho_t \tau \sum_{z=1}^L \beta_{wl,uzl}} \frac{\mathbf{a}_{wl,upl}}{\zeta_{wl,uwl}} \right|^2.
\end{aligned} \tag{29}$$

$$\mathbb{E}\left[|F_4|^2\right] = \rho_b M \sum_{w=1}^L \sum_{\substack{q=1 \\ q \neq u}}^U \frac{\rho_t \tau \beta_{wl,ujl} \beta_{wl,qwl}^2}{1 + \rho_t \tau \sum_{z=1}^L \beta_{wl,qzl}} \left| \frac{\mathbf{v}_{qwl}}{\zeta_{wl,qwl}} \right|^2. \tag{30}$$

$$\mathbb{E}\left[|F_5|^2\right] = \rho_b M \sum_{i=1}^L \sum_{\substack{w=1 \\ i \neq l}}^L \sum_{q=1}^U \frac{\rho_t \tau \beta_{wl,ujl} \beta_{wl,qwl}^2}{1 + \rho_t \tau \sum_{z=1}^L \beta_{wl,qzl}} \left| \frac{\mathbf{v}_{qwi}}{\zeta_{wl,qwl}} \right|^2. \tag{31}$$

$$\mathbb{E}\left[|F_6|^2\right] = \text{var}(\mathbf{w}_{ujl}) = \sigma_{ujl}^2. \tag{32}$$

Finally, the achievable rate becomes;

$$\mathbf{C} = \log_2 \left( \mathbf{1} + \frac{\rho_b M^2 \left| \sum_{w=1}^L \frac{\rho_t \tau \beta_{wl,ujl} \beta_{wl,uwl}}{1 + \rho_t \tau \sum_{z=1}^L \beta_{wl,uzl}} \frac{\mathbf{a}_{wl,ujl}}{\zeta_{wl,uwl}} \right|^2}{M^2 \psi_1 + M \psi_2 + \sigma_{ujl}} \right). \tag{33}$$

Where;

$$\begin{aligned}
\psi_1 &= \rho_b \sum_{\substack{p=1 \\ p \neq j}}^L \left| \sum_{w=1}^L \frac{\rho_t \tau \beta_{wl,ujl} \beta_{wl,uwl}}{1 + \rho_t \tau \sum_{z=1}^L \beta_{wl,uzl}} \frac{\mathbf{a}_{wl,upl}}{\zeta_{wl,uwl}} \right|^2 \\
\psi_2 &= \rho_b \sum_{i=1}^L \sum_{w=1}^L \sum_{q=1}^U \frac{\rho_t \tau \beta_{wl,ujl} \beta_{wl,qwl}^2}{1 + \rho_t \tau \sum_{z=1}^L \beta_{wl,qzl}} \left| \frac{\mathbf{v}_{qwi}}{\zeta_{wl,qwl}} \right|^2
\end{aligned}$$

## 6 Simulation Result

In this section we consider a network with  $L = 7$  cells. Within each cell, BS has  $W = 7$  sub-array of antennas and the number of users in each cell is  $K = LW$ . For each cell

unique  $U = 7$  orthogonal signals are assigned and for users in each cell those orthogonal signals are reused. We use the 3GPP standard of Urban Macro model to generate large scale fading as given by Eq. (34) [17].

$$10 \log_{10} \beta_{wl,uji} = -139.5 - 35 \log_{10} d_{wl,uji} + \varphi_{wl,uji}. \tag{34}$$

Where  $d_{wl,uji}$  is the distance (in km) between the user and the base station, and  $\varphi_{wl,uji}$  is shadowing coefficient modeled as a Gaussian random variable with zero mean and variance 8 dB. We consider the cell radius to be  $r = 0.75\text{km}$  and the distance  $d_{wl,uji}$  of all users randomly distributed near the edge of the cell. We choose the bandwidth to be  $B = 20 \text{ MHz}$  and the noise variance at each receiver  $\sigma_{ujl}^2 = 92 \text{ dBm}$ . The average power of  $w$ -th sub-array of BS and at each user terminal are  $\rho_b = 48 \text{ dBm}$  and  $\rho_t = 23 \text{ dBm}$ , respectively.

We have simulated the cumulative distribution function (CDF) of the achievable sum rate Eq. (33) using MATLAB. The results in Figs. 2, 3 and 4 show the cumulative distribution function (CDF) of the achievable sum rate of  $LK$  users with  $M = 10^2, 10^3$ , and  $10^3$  respectively. From Figs. 2, 3 and 4 it can be noted that the large scale fading precoding with cell cooperation (LSFP-CC) achieves 5% outage rate around  $10^{-4}, 10^{-3}$ , and  $10^{-2}$  bits per channel respectively. When the large scale fading precoding with non-cooperation cell is used (LSFP-NCC) the achievable sum rate is improved to around  $10^{-3}, 10^{-2}$ , and  $10^{-1}$  bits per channel respectively. From those results, it can be observed that the purposed LSFP-NCC improve 5% outage rate 10 times compared to LSFP-CC. This is due to the large scale precoding which reduces the

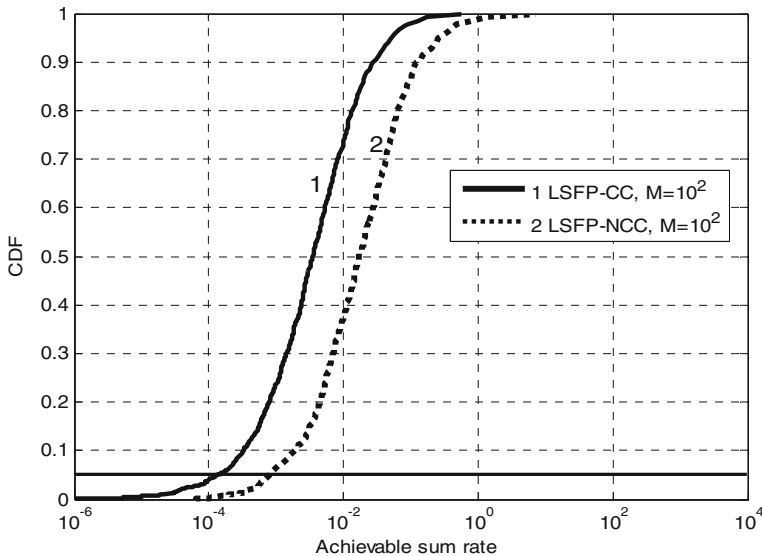
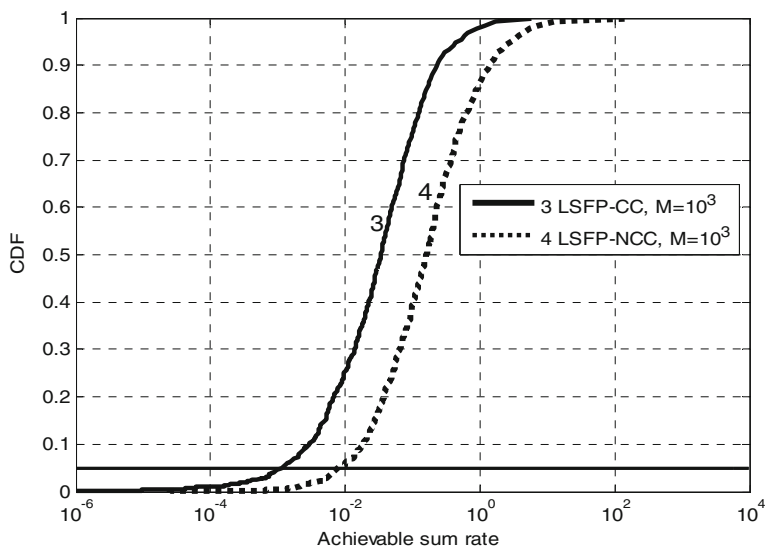
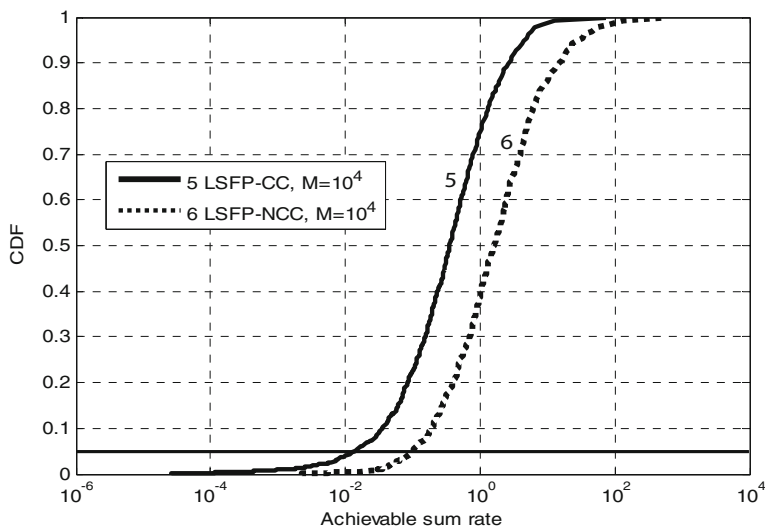


Fig. 2. The CDF of all received Achievable sum rate of  $LK$  users for  $M = 10^2$



**Fig. 3.** The CDF of all received Achievable sum rate of  $LK$  users for  $M = 10^3$



**Fig. 4.** The CDF of all received Achievable sum rate of  $LK$  users for  $M = 10^4$

interference that comes from the adjacent cells. The time complexity analysis shows that the proposed scheme required less time to execute than the existing method. This is due to each cell computes its own users but not the existing method which uses a network hub that computes to all  $L = 7$  cell users.

## 7 Conclusion

We have proposed a novel uplink training scheme that uses large scale fading precoding without cooperation between cells and the achievable rate of the scheme has been derived. The proposed scheme improves 5% outage rate 10 times over the existing method. This is a significant improvement, however it is achieved at the expense of sub array BS installation and power consumption.

We have assumed that users are located near the cell edges and but in practice user locations vary randomly within the cell. Further investigations shall be carried out to assess the effect of random user location on the improvement of the achievable sum rates.

## References

1. Biglier, E., Calderbank, R., Constantinides, A., Goldsmith, A.: Arogyaswami Paulraj, and H. Vincent Poor, MIMO Wireless Communications. Cambridge University Press, New York (2007)
2. Rusek, F., et al.: Scaling Up MIMO: opportunities and challenges with very large arrays. *IEEE Signal Process. Mag.* **30**(1), 40–60 (2013)
3. Larsson, E.G., Edfors, O., Tufvesson, F., Marzetta, T.L.: Massive MIMO for next generation wireless systems. *IEEE Commun. Mag.* **52**(2), 186–195 (2014)
4. Marzetta, T.L.: Massive MIMO: An Introduction. *Labs Technical Journal* **20**, 11–22 (2015)
5. Björnson, E., Larsson, E.G., Marzetta, T.L.: Massive MIMO: ten myths and one critical question. *IEEE Commun. Mag.* **54**(2), 114–123 (2016)
6. Ngo, H.Q., Larsson, E.G., Marzetta, T.L.: Massive MU-MIMO downlink TDD systems with linear precoding and downlink pilots. In: 2013 51st Annual Allerton Conference on Communication, Control, and Computing (Allerton), Monticello, IL, pp. 293–298 (2013)
7. Wang, L., Ngo, H.Q., Elkashlan, M., Duong, T.Q., Wong, K.K.: Massive MIMO in spectrum sharing networks: achievable rate and power efficiency. *IEEE Syst. J.* **99**, 1–12 (2015)
8. Lu, L., Li, G.Y., Swindlehurst, A.L., Ashikhmin, A., Zhang, R.: An overview of massive MIMO: benefits and challenges. *IEEE J. Sel. Top. Signal Process.* **8**(5), 742–758 (2014)
9. Xu, P., Wang, J., Wang, J.: Effect of pilot contamination on channel estimation in massive MIMO systems. In: 2013 International Conference on Wireless Communications & Signal Processing (WCSP), Hangzhou, pp. 1–6 (2013)
10. Ashikhmin, A., Marzetta, T.: Pilot contamination precoding in multi-cell large scale antenna systems. In: 2012 IEEE International Symposium on Information Theory Proceedings (ISIT), Cambridge, MA, pp. 1137–1141 (2012)
11. Ashikhmin, A., Marzetta, T.L., Li, L.: Interference reduction in multi-cell massive mimo systems I: Large-scale fading precoding and decoding. arXiv preprint [arXiv:1411.4182](https://arxiv.org/abs/1411.4182) (2014)
12. Ashikhmin, A., Marzetta, T.L., Li, L.: Interference reduction in multi-cell massive mimo systems II: Large-scale fading precoding and decoding, arXiv preprint [arXiv:1411.4182](https://arxiv.org/abs/1411.4182) (2014)
13. Marzetta, T.L.: Noncooperative cellular wireless with unlimited numbers of base station antennas. *IEEE Trans. Wirel. Commun.* **9**(11), 3590–3600 (2010)

14. Fernandes, F., Ashikhmin, A., Marzetta, T.L.: Inter-cell interference in noncooperative TDD large scale antenna systems. *IEEE J. Sel. Areas Commun.* **31**(2), 192–201 (2013)
15. Ngo, H.Q., Marzetta, T.L., Larsson, E.G.: Analysis of the pilot contamination effect in very large multicell multiuser MIMO systems for physical channel models. In: 2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Prague, Czech Republic, pp. 3464–3467 (2011)
16. Jose, J., Ashikhmin, A., Marzetta, T.L., Vishwanath, S.: Pilot contamination and precoding in multi-cell TDD systems. *IEEE Trans. Wirel. Commun.* **10**(8), 2640–2651 (2011)
17. 3GPP TR 25.996, Spatial channel model for multiple input multiple output (MIMO) simulations (2016)



# Performance Analysis of Symmetric and Asymmetric LTE Turbo Codes with Prioritisation and Regression Based Scaling

Y. Beeharry<sup>(✉)</sup>, Tulsı Pawan Fowdur, and K.M.S. Soyjaudah

Department of Electrical and Electronics Engineering,  
University of Mauritius, Réduit, Mauritius  
{y.beeharry, p.fowdur, ssoyjaudah}@uom.ac.mu

**Abstract.** Standards like Long Term Evolution (LTE) employ Turbo coded QAM systems in order to achieve high data rates. Despite the fact that several mechanisms have been proposed in order to enhance the error performance of Turbo coded QAM systems, there is still the need to come up with novel or hybrid systems which can contribute towards further improved error performances. In this paper, a comparative analysis has been performed between symmetric and asymmetric LTE Turbo codes with the incorporation of techniques such as prioritization and regression based extrinsic information scaling. Results demonstrate that significant enhancement in the error performance throughout the whole  $E_b/N_0$  range can be obtained with high order modulation when these techniques are used. With both symmetric and asymmetric LTE Turbo codes employing 64-QAM and a code-rate of 1/3, an average gain of 0.3 dB below BERs of  $10^{-1}$  is obtained over symmetric and asymmetric LTE Turbo codes.

**Keywords:** Prioritisation · Asymmetric LTE turbo codes · Scaling

## 1 Introduction

Ever since Berrou *et al.* came up with Turbo codes - the powerful and near Shannon-limit [1] error correcting code [2, 3], several standards in the realm of digital communications have approved their deployment. The Long Term Evolution (LTE) [4, 5] and CDMA 2000 [6] standards have exploited Turbo coded QAM systems with the aim of obtaining high data rates and reliable transmission. The significant impact of Turbo codes has led the research community to unearth several techniques such as asymmetric Turbo encoding [7–10], prioritisation in Quadrature Amplitude Modulation (QAM) constellation mapping [9, 11, 12], scaling of extrinsic information and iterative detection [13–15] which have contributed towards the improvement in error performance and computational complexity reduction. An overview of these different techniques is presented next.

Interesting modifications have been proposed for conventional Turbo codes over the years. In [10], the authors have studied the error performance of asymmetric Turbo codes and demonstrated an improved performance in the pre-error floor as well as the error floor regions. In view to proceed with the work of [10], an evaluation of the performance of asymmetric and symmetric Turbo codes in a channel with Additive White Gaussian Noise (AWGN) was done in [16]. The results of the simulations presented in [16], have demonstrated that asymmetric Turbo codes can provide better performance than symmetric Turbo codes in addition to lower complexity of the decoder, provided a proper selection of the generator polynomials is made. Additionally, in [7], the authors have proposed a novel kind of Turbo codes with un-identical encoders specifically for interleavers of medium size. This scheme has been devised using the selection algorithm for the generation of polynomials corresponding to the largest spread Quadrature Permutation Polynomial (QPP) interleaver.

The prioritised QAM constellation mapping impacts the error performance of Turbo codes positively. This is achieved by making use of the Unequal Error Protection (UEP) property of the QAM constellation such that an increased protection is provided to the systematic information bits. Lüders *et al.* have applied this technique in [9]. The authors of [11] have proposed a system combining prioritised constellation mapping, adaptive Sign Difference Ratio (SDR) based extrinsic information scaling, and Joint Source Channel Decoding (JSCD) in order to enhance the error performance of Turbo coded 64-QAM. The results demonstrate that combining these techniques can help achieve a gain of 2.5 dB on average compared to conventional system of Turbo coded 64-QAM at Bit Error Rates (BERs) above  $10^{-1}$ . As an extended work of [11], the work in [12] studies the error performance of LTE Turbo codes with the same combination of techniques. Additionally, the prioritised constellation mapping is investigated with both 16-QAM and 64-QAM. The results presented demonstrate that with 16-QAM, the combination of all the techniques can help achieve an average gain of 1.7 dB over the conventional scheme at BERs above  $10^{-1}$ . With 64-QAM, a performance gain of 3 dB on average is obtained compared to the conventional scheme.

The objective of extrinsic information scaling is to improve the performance of the Turbo decoder by scaling its extrinsic information using a scale factor. For instance, the authors of [14] have deployed a constant scale factor in order to enhance the Max-Log-MAP Turbo decoding algorithm, while in [17], a modified MAP algorithm which uses a constant scale factor has been proposed. Interestingly though, in [13], a scaling scheme extending the Sign Difference Ratio (SDR) mechanism in [15] to dynamically obtain a scaling factor at each iteration for each data block has been proposed. In [18], a novel early stopping mechanism and extrinsic information scaling based on regression analysis for LTE Turbo codes was proposed. The results demonstrate an improved error performance in addition to reduced number of average iterations as compared to both the conventional scheme and the one making use of SDR based stopping and scaling.

The work presented in this paper aims at analysing the performance of LTE Turbo coded QAM systems integrating: asymmetric encoders, prioritisation in QAM constellation mapping, and regression analysis based extrinsic information scaling. At the transmitter side, the symmetric/asymmetric encoder operates on the input systematic information to output the encoded bits. Prioritised constellation mapping is then

performed such that the highest priority is allocated to the systematic bits on the QAM constellation. Finally, a regression based extrinsic information scaling mechanism is used in the decoding mechanism at the receiver side. The performance of both symmetric and asymmetric Turbo codes was found to be almost similar but they outperformed the conventional LTE Turbo codes when prioritization and scaling were incorporated.

The organisation of this paper is as follows. The complete system model is described in Sect. 2. The simulation results and analysis are described in Sect. 3. Section 4 concludes the paper and provides some future works.

## 2 System Model

The system’s transmitter block diagram is shown in Fig. 1. The information bits are sent through the encoder in blocks of length equal to the interleaver size to output the systematic information  $S_0$  and the two parity information  $P_1$  and  $P_2$ . The asymmetric encoding is achieved with the different generator polynomials for RSC1 and RSC2 [7, 10, 16]. After the sub-block interleaving, a prioritization QAM constellation mapping is performed so as to provide more protection to the systematic information bits. The modulated symbols are then transmitted over the AWGN channel.

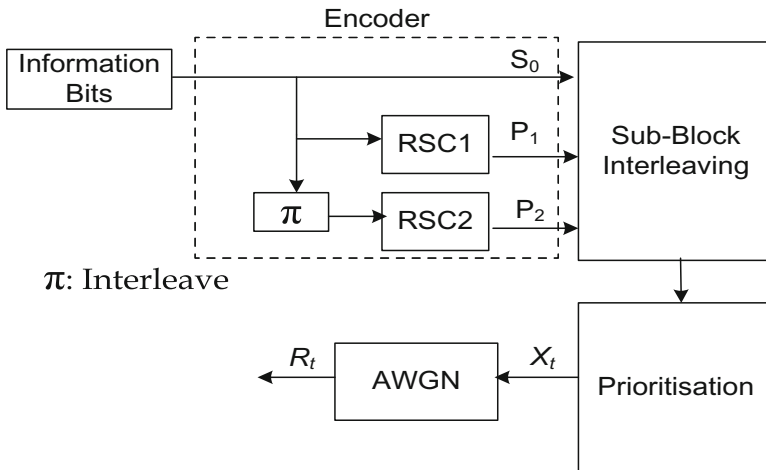
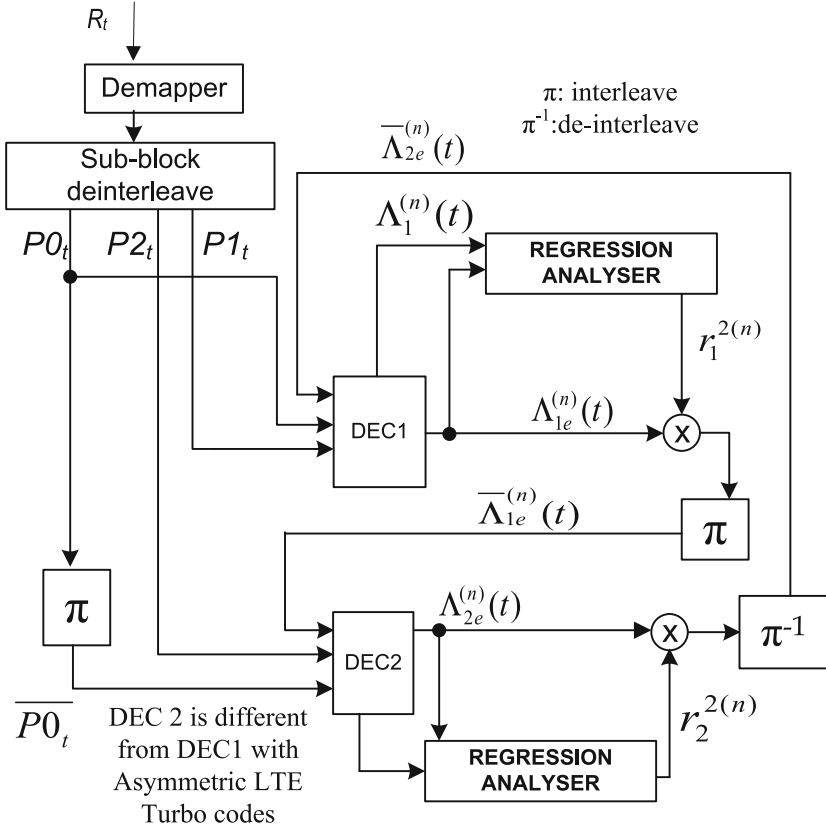


Fig. 1. Transmitter system.

The receiver side is shown in Fig. 2.  $P0_t, P1_t$  and  $P2_t$  are the soft information which are forwarded to the turbo decoders.  $P0_t$  corresponds to the systematic bits while  $P1_t$  and  $P2_t$  correspond to the parity bits.



**Fig. 2.** Receiver system [18].

The equations of the Max-Log-MAP algorithm [3, 19] for conventional turbo decoding are given next. The probability pertaining to the branch transition for DEC 1 at time instant  $t$ , starting in state  $l'$  and ending in state  $l$  with an input bit  $i$  ( $i = 0$  or  $1$ ) is:

$$\begin{aligned} \gamma_i^{1(i)}(l', l) &= \log \left[ p_i^1(i) \cdot \exp \left( - \frac{[P0_t - S_{0_i}]^2 + [P1_t - P_{1_i}]^2}{2\sigma^2} \right) \right] \\ &= \log [p_i^1(i)] - \left( \frac{[P0_t - S_{0_i}]^2 + [P1_t - P_{1_i}]^2}{2\sigma^2} \right) \end{aligned} \quad (1)$$

Where,

$P0_t$  and  $P1_t$  are the de-mapped soft-information bits corresponding to the bipolar equivalent of the transmitted systematic bits  $S_{0_i}$  and first parity bits,  $P_{1_i}$ , respectively;

$\log(p_i^1(i))$  is the natural logarithm of the a-priori probability corresponding to input bit  $i$  computed from the extrinsic information of the channel and fed to the first decoder;

$\sigma^2$  is the noise variance.

The forward recursive variable,  $\alpha_t^1(l)$ , at state  $l$  and time  $t$  for a decoder with  $M_S$  states is computed as:

$$\alpha_t^1(l) = \max\left(\alpha_{t-1}^1(l') + \gamma_t^{1(i)}(l', l)\right), \text{ for } 0 \leq l' \leq M_S - 1 \quad (2)$$

The backward recursive variable,  $\beta_t^1(l)$ , at state  $l$  and time  $t$  is computed as the following:

$$\beta_t^1(l) = \max\left(\beta_{t-1}^1(l') + \gamma_t^{1(i)}(l, l')\right), \text{ for } 0 \leq l' \leq M_S - 1 \quad (3)$$

The Log-Likelihood Ratio (LLR),  $\Lambda_1^{(n)}(t)$  at time  $t$  and  $n^{\text{th}}$  iteration for the first decoder is given as:

$$\begin{aligned} \Lambda_1^{(n)}(t) = & \max\left(\alpha_{t-1}^1(l') + \gamma_t^{1(1)}(l, l') + \beta_t^1(l)\right) \\ & - \max\left(\alpha_{t-1}^1(l') + \gamma_t^{1(0)}(l, l') + \beta_t^1(l)\right), \text{ for } 0 \leq l' \leq M_S - 1 \end{aligned} \quad (4)$$

The extrinsic information,  $\Lambda_{1e}^{(n)}(t)$  at time  $t$  and iteration  $n$  for the first decoder is given as:

$$\Lambda_{1e}^{(n)}(t) = \Lambda_1^{(n)}(t) - \frac{2}{\sigma^2} P_{0t} - \overline{\Lambda}_{2e}^{(n-1)}(t) \quad (5)$$

Where,

$\overline{\Lambda}_{2e}^{(n-1)}(t)$  is the de-interleaved extrinsic information at time  $t$  and the  $(n - 1)^{\text{th}}$  iteration for the second decoder.

The same decoding operations are performed by DEC 2 using the inputs  $\overline{P0}_t$ ,  $P_{2t}$  and  $p_t^2(i)$ .

Where,

$p_t^2(i)$  is the a-priori probability with input bit  $i$  and computed from the channel extrinsic information to be sent to the second decoder;

$\overline{P0}_t$  and  $P_{2t}$  are the de-mapped soft-information bits corresponding to the bipolar equivalent of the interleaved systematic bits  $\overline{S0}_t$  and second parity bits,  $P_{2t}$ , respectively.

Additionally, at the Turbo decoder side, an extrinsic information scaling mechanism based on regression analysis is deployed. The Regression Analyser [18], outputs  $r_d^{2(n)}$  which is the factor for scaling. This factor is computed by measuring the linear correlation between the *A-Posteriori* LLR  $\left(\Lambda_d^{(n)}(t)\right)$  and the Extrinsic LLR  $\left(\Lambda_{de}^{(n)}(t)\right)$ , where, the decoder number is  $d = \{1, 2\}$ ;  $r_d^{2(n)}$  is the factor used to scale the extrinsic information from decoder  $d$  at iteration  $n$ ;  $N$  is the packet length which is 6144 [4] in

this simulation;  $\Lambda_d^{(n)}(t)$  is the  $t^{\text{th}}$  *a-posteriori* LLR of the decoder  $d$  at time  $t$  and iteration  $n$ ; The mean *a-posteriori* LLR,  $\widehat{\Lambda}_d^{(n)}$ , of decoder  $d$  at iteration  $n$  and is computed as [18]:

$$\widehat{\Lambda}_d^{(n)} = \frac{1}{N} \sum_{t=1}^N \Lambda_d^{(n)}(t) \quad (6)$$

$\Lambda_{de}^{(n)}(t)$  is the  $t^{\text{th}}$  extrinsic LLR at iteration  $n$  of decoder  $d$ ;  $\widehat{\Lambda}_{de}^{(n)}$  is the mean extrinsic LLR at iteration  $n$  of decoder  $d$  and is computed as [18]:

$$\widehat{\Lambda}_{de}^{(n)} = \frac{1}{N} \sum_{t=1}^N \Lambda_{de}^{(n)}(t) \quad (7)$$

$n$  is the number of half-iterations and take values in  $\{1/2, 1, \dots, I_{\max}\}$  (maximum number of iterations)}.

The scale factor,  $r_d^{2(n)}$ , is computed as follows:

$$r_d^{2(n)} = \left( \frac{\sum (\Lambda_d^{(n)} - \widehat{\Lambda}_d^{(n)}) \times (\Lambda_{de}^{(n)} - \widehat{\Lambda}_{de}^{(n)})}{\sqrt{\sum (\Lambda_d^{(n)} - \widehat{\Lambda}_d^{(n)})^2 \times \sum (\Lambda_{de}^{(n)} - \widehat{\Lambda}_{de}^{(n)})^2}} \right)^2 \quad (8)$$

### 3 Simulation Results and Analysis

The performance of the following schemes for LTE Turbo codes have been evaluated through simulation:

1. Scheme 1: This scheme employs the symmetric LTE Turbo codes only.
2. Scheme 2: This scheme employs the asymmetric LTE Turbo codes only.
3. Scheme 3: This scheme employs the symmetric LTE Turbo codes with regression based scaling only.
4. Scheme 4: This scheme employs the asymmetric LTE Turbo codes only with scaling based on regression.
5. Scheme 5: This scheme employs the symmetric LTE Turbo codes with scaling based on regression and prioritization.
6. Scheme 6: This scheme employs the asymmetric LTE Turbo codes with scaling based on regression and prioritization.

The generator polynomial in octal for both upper and lower encoders of the symmetric LTE Turbo code is given as  $G = [1, g1/g2]$ , where  $g1 = 15$  and  $g2 = 13$ . The generator polynomial of the upper encoder in the asymmetric LTE Turbo code is identical to that of the symmetric LTE Turbo code while that for the lower encoder is given as  $G = [1, g1/g2]$ , where  $g1 = 13$  and  $g2 = 15$ . The parameters of the QPP interleaver are:  $f_1 = 263$ ;  $f_2 = 480$ ; packet size,  $N = 6144$  bits; maximum number of

iterations,  $I_{max} = 12$ ; Code-rates: 1/3 and 1/2; Modulations: 16-QAM and 64-QAM; Channel: complex AWGN.

Figure 3 shows the BER performance of the six schemes with code-rate = 1/3 and 16-QAM. It can be observed that Schemes 3 and 4 outperform all the other schemes for BERs below  $10^{-1}$ . Schemes 3 and 4 provide an average gain of 0.3 dB below BERs of  $10^{-1}$  as compared to Schemes 1 and 2 and an average gain of 0.25 dB below BERs of  $10^{-1}$  compared to Schemes 5 and 6. Schemes 5 and 6 outperform Schemes 1 and 2 only for BERs above  $10^{-4}$  and Schemes 3 and 4 only for BERs above  $10^{-1}$ .

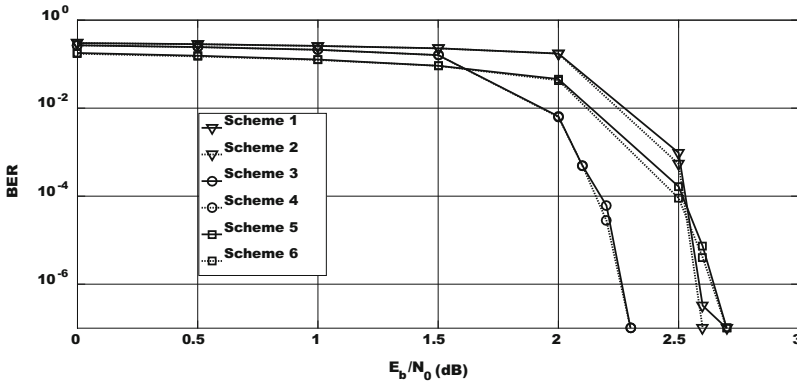


Fig. 3. BER performance of the schemes with code-rate = 1/3 and 16-QAM.

Figure 4 shows the BER performance for the schemes with code-rate = 1/2 and 16-QAM. It can be observed that Schemes 3 and 4 outperform all the other schemes for BERs below  $10^{-1}$ . Schemes 3 and 4 provide an average gain of 0.15 dB below BERs of  $10^{-1}$  as compared to Schemes 1 and 2 and an average gain of 0.2 dB below BERs of  $10^{-4}$  compared to Schemes 5 and 6. Schemes 5 and 6 outperform Schemes 1 and 2 only for BERs above  $10^{-5}$ .

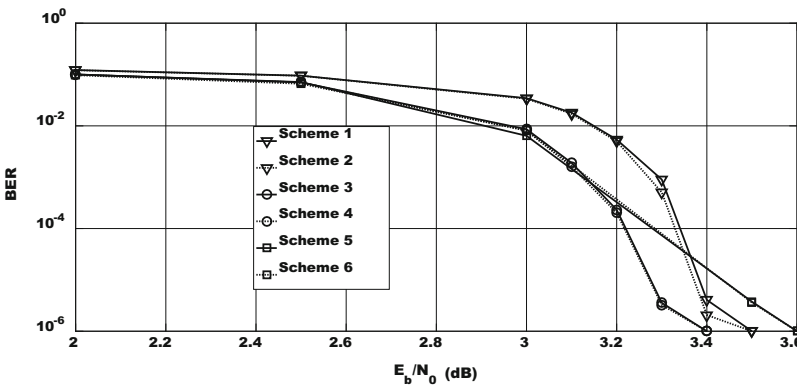


Fig. 4. BER performance for the schemes with 16-QAM and code-rate = 1/2.

Figure 5 shows the BER performance of the six schemes with code-rate = 1/3 and 64-QAM. It can be observed that Schemes 3, 4, 5 and 6 outperform Schemes 1 and 2 for BERs below  $10^{-1}$ . Schemes 3, 4, 5 and 6 provide an average gain of 0.3 dB below BERs of  $10^{-1}$  as compared to Schemes 1 and 2.

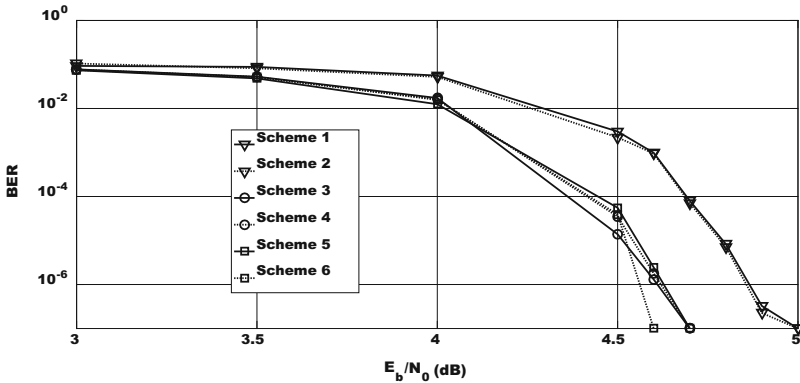


Fig. 5. BER performance of the schemes with code-rate = 1/3 and 64-QAM.

Figure 6 shows the BER performance for the schemes with code-rate = 1/2 and 64-QAM. It can be observed that Schemes 3, 4, 5 and 6 outperform Schemes 1 and 2 for BERs below  $10^{-1}$ . Schemes 3 and 4 provide an average gain of 0.4 dB below BERs of  $10^{-1}$  as compared to Schemes 1 and 2. Schemes 5 and 6 outperform Schemes 1 and 2 for BERs above  $10^{-6}$  and Schemes 3 and 4 only for BERs above  $10^{-3}$ .

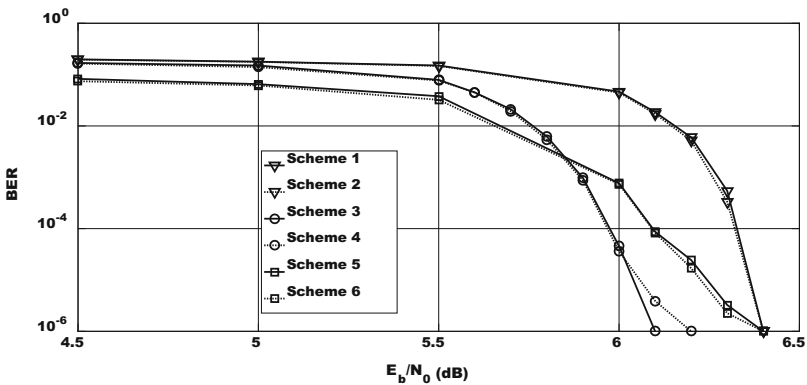


Fig. 6. BER performance for the schemes with 64-QAM and code-rate = 1/2.



## 4 Conclusion

In this paper, a comparative analysis has been performed between symmetric and asymmetric LTE Turbo codes with the incorporation of techniques such as prioritization and regression analysis based extrinsic information scaling. At the transmitter side, a symmetric/asymmetric encoder is used and a re-ordering technique is employed to prioritise the systematic information bits by placing them on the positions of the QAM constellation with highest protection. At the receiver side, an extrinsic information scaling mechanism based on regression analysis is used. Results demonstrate that significant enhancement in the error performance throughout the whole  $E_b/N_0$  range can be obtained with high order modulation schemes when prioritization and regression analysis based extrinsic information scaling techniques are used. With both symmetric and asymmetric LTE Turbo codes employing 64-QAM and a code-rate of  $1/3$ , an average gain of 0.3 dB below BERs of  $10^{-1}$  is obtained over symmetric and asymmetric LTE Turbo codes. The main contribution of this work is the performance analysis comparison of schemes with symmetric/asymmetric LTE Turbo encoders with prioritization and regression analysis based extrinsic information scaling. The comparison shows that the performance of both symmetric and asymmetric Turbo codes are similar but outperform the conventional LTE Turbo codes when prioritization and scaling are incorporated. Some interesting future works can be proposed based on the analysis in this work. A straightforward extension would be to assess the usability of the proposed combinations with non-binary Turbo codes. A more challenging future work would be to come up with a mathematical equation which would show the relationship between the error performance, modulation order, code-rates and channel quality. Further investigations can be performed for improving the performance of asymmetric Turbo codes with medium interleaver size.

**Acknowledgments.** The authors would like to thank the University of Mauritius for providing the necessary facilities for conducting this research and the Tertiary Education Commission for its financial support.

## References

1. Shannon, C.E.: A mathematical theory of communications. Part I Ann. Telecommun. Bell Syst. Tech. J. **27**, 379–423 (1948)
2. Berou, C., Glavieux, A., Thitimajshima, P.: Near Shannon limit error-correcting coding and decoding: Turbo-codes. In: IEEE International Conference on Communications, Geneva (1993)
3. Vucetic, B., Yuan, J.: Turbo Codes: Principles and Applications. Kluwer Academic Publications, Boston (2000)
4. 3GPP: Technical Specifications Rel. 8 (2009)
5. 3GPP2: CDMA 2000 High Rate Packet Data Air Interface Specification. (2006). [http://www.3gpp2.org/Public\\_html/specs/C.S0024-B\\_v1.0\\_060522.pdf](http://www.3gpp2.org/Public_html/specs/C.S0024-B_v1.0_060522.pdf). Accessed Nov 2012
6. Seisa, S., Toufik, I., Baker, M.: LTE - The UMTS Long Term Evolution: From Theory to Practice. John Wiley & Sons Ltd., New York (2009)

7. Cojocariu, D.T., Lazar, A.G.: Asymmetric turbo codes for LTE systems with medium frame length. In: Proceedings of SPAMEC, Cluj-Napoca, Romania (2011)
8. Khan, F.: LTE for 4G Mobile Broadband Air Interface Technologies and Performance. Cambridge University Press, New York (2009)
9. Lüders, H., Minwegen, A., Vary, P.: Improving UMTS LTE performance by UEP in high order modulation. In: 7th International Workshop on Multi-Carrier Systems & Solutions (MC-SS). Herrsching, Germany (2009)
10. Takeshita, Y., Collins, M., Massey, P., Costello, D.: A note on asymmetric turbo-codes. *IEEE Comm. Letters* **3**(3), 69–71 (1999)
11. Fowdur, T.P., Beeharry, Y., Soyjaudah, K.M.S.: Performance of turbo coded 64-QAM with joint source channel decoding, adaptive scaling and prioritised constellation mapping, CTRQ. In: 6th International Conference on Communication Theory, Reliability, and Quality of Service, Venice, Italy (2013)
12. Fowdur, T.P., Beeharry, Y., Soyjaudah, K.M.S.: Performance of LTE turbo codes with joint source channel decoding, adaptive scaling and prioritised QAM constellation mapping. *Int. J. Adv. Telecommun.* **6**(3&4), 143–152 (2013)
13. Lin, Y., Hung, W., Lin, W., Chen, T., Lu, E.: An efficient soft-input scaling for turbo decoding. In: IEEE International Conference on Sensor Networks, Ubiquitous and Trustworthy Computing Workshops (2006)
14. Vogt, J., Finger, A.: Improving the MAX-Log-MAP turbo decoder. *Electr. Lett.* **36**(23), 1937–1939 (2000)
15. Wu, Y., Woerner, B., Ebel, J.: A simple stopping criterion for turbo decoding. *IEEE Commun. Lett.* **4**(8), 258–260 (2000)
16. Kumar, M., Khedia, D.: Comparative performance evaluation of symmetric and asymmetric turbo codes for AWGN channel. *Int. J. Comput. Electron. Electr. Eng.* **2**(3), 1–6 (2012)
17. Gnanasekaran, T., Duraiswamy, K.: Performance of unequal error protection using MAP algorithm and modified MAP in AWGN and fading channel. *J. Comput. Sci.* **2**, 585–590 (2008)
18. Fowdur, T.P., Beeharry, Y., Soyjaudah, K.M.S.: A novel scaling and early stopping mechanism for LTE turbo code based on regression analysis. *Ann. Telecommun.* 1–20 (2016)
19. Fowdur, T.P., Soyjaudah, K.M.S.: Performance of joint source-channel decoding with iterative bit-combining and detection. *Ann. Telecommun.* **63**(7–8), 409–423 (2008)

# Wireless Body Area Network System Architecture for Real-Time Diabetes Monitoring

Geshwaree Huzooree<sup>1(✉)</sup>, Kavi Kumar Khedo<sup>2</sup>,  
and Noorjehan Joonas<sup>3</sup>

<sup>1</sup> Department of Information Technology, Charles Telfair Institute,  
Moka, Mauritius

geshwaree.huzooree@telfair.ac.mu

<sup>2</sup> Department of Computer Science, University of Mauritius, Reduit, Mauritius  
k.khedo@uom.ac.mu

<sup>3</sup> Ministry of Health and Quality of Life, Central Health Laboratory,  
Victoria Hospital, Candos, Mauritius  
njoonas@intnet.mu

**Abstract.** Technological advancement in Wireless Body Area Networks (WBANs) has gained much interest and became an emerging technology in the remote healthcare monitoring as WBANs are a collection of low-power, intelligent, miniaturized, lightweight, invasive or non-invasive bio sensors that operate in the proximity of a human body through wireless communication. They offer innovative and promising healthcare applications to improve the quality of life of patients by continuously monitoring human's physiological vital signs. Based on existing literature on WBANs, technologies and standards, this paper proposes a WBAN to support remote healthcare monitoring of diabetic patients by ensuring higher degree of confidence in Quality of Information (QoI) through real-time data. The proposed architecture includes several features which enables better patient monitoring. Moreover, a qualitative approach is used to evaluate the systems' architecture. Finally, the open research challenges and future works associated with WBANs design are discussed and evaluated.

**Keywords:** Wireless body area sensor network · Quality of Information · QoI · Wireless communication

## 1 Introduction

The high prevalence of chronic diseases is leading to a worldwide paradigm shift from reactive to preventive and affordable healthcare with wireless body area network systems (WBANs) playing a very important role in addressing the pre-vailing challenges in global healthcare [1]. The preventive healthcare system enables healthcare professionals to closely perform remote real-time monitoring, timely diagnosis and intervention of chronic diseases to ensure a cost-effective and patient-centric way to help maintain an optimal health status for patients.

WBAN consists of a collection of ultra-low-power, lightweight, invasive or non-invasive devices with wireless communication capabilities that operate in the proximity of a human body [2]. WBAN is composed of sensors which is responsible to collect, process and communicate patients' vital signals (e.g. temperature, heart rate, Electroencephalography (EEG), Electrocardiography (ECG), blood pressure, blood oxygen saturation levels, etc.) to the medical server for continuous, automated and unobtrusive health monitoring.

Due to constraints such as reliability, higher degree of confidence for Quality of Information (QoI) and non-deterministic sensor failures, designing a robust WBAN platform to continuously monitor the health conditions of diabetic patients is still challenging.

The paper is structured as follows: Sect. 2 presents an overview of related works on WBANs for monitoring of diabetic patients which are relevant to the context of this paper. Section 3 describes an overview of the proposed WBAN architecture. The design requirement, architecture design, sensor design communication architecture, data design and application design are detailed. Section 4 presents the evaluation of the WBAN design at the sensor level, mobile computing level and remote server level. Section 5 discusses the research challenges and future works. Finally, Sect. 6 concludes the contributions of this paper and high-lights potential future works directions.

## 2 Related Works on WBANs

The aim of WBANs is to simplify and improve speed, accuracy, and reliability of communication of sensors with the mobile phones and the remote server [3].

Mamykina et al. [4] developed a mobile application named MAHI that, through a Bluetooth enabled glucose meter and a Java-enabled cell phone, allows individuals with diabetes to capture rich media records (audio and video) indicating past actions and blood sugar levels, and to share and discuss these records with a diabetes educator through a website. However, the patients need to manually upload their glucose readings to the website for monitoring. The aspect of real-time monitoring is not supported and data quality issues have not been addressed.

Alhazbi et al. [5] developed a mobile application to help diabetes management using invasive glucose monitoring and diet management. The system has three components: patient module, server module and physician module. However, this research as well makes use of an invasive technique and focus is mostly geared towards diet management. The aspects of reliability, continuous monitoring and data quality issues have not been addressed.

Luo et al. [6] developed a personalized diabetes recommendation system, Gluco-Guide, for Type-2 Diabetes patients. The system aggregates a variety of lifestyle data via medical sensors and mobile devices, mines the data with a novel data-mining framework, and outputs personalized and timely recommendations to patients aimed to control their blood glucose level. However, the aspects of data quality issues have not been addressed.

Mougiakakou et al. [7] developed a mobile phone application as shown in Fig. 1, SMARTDIAB, for the self-management of people with Type 1 diabetes (T1DM) anytime and anywhere, which collects data from monitoring devices and regularly transfers them to a hospital web server, to be available to the physician. In case of an emergency, the individual can press a button, in order to transmit immediately his/her position to an emergency contacts. However, the reliability of the system has not been explored and the data quality issues have not been addressed.

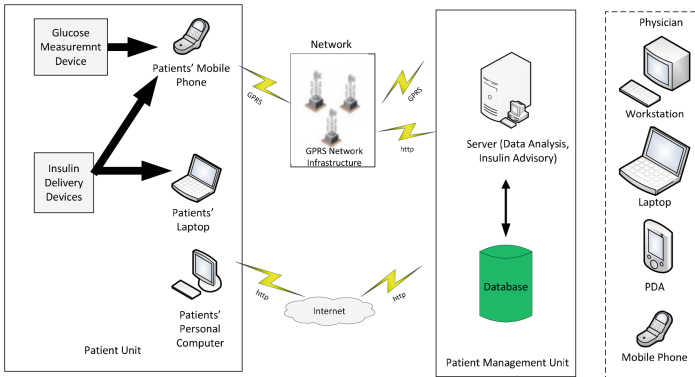


Fig. 1. SMARTDIAB architecture (Mougiakakou et al. [7]).

Despite the increased interest in the WBANs areas, there have been only a few studies related to the development of reliable WBANs system with high quality data assurance in terms of timeliness, accuracy, reliability, completeness, relevancy, usability and confidentiality to enable real-time and remote monitoring of patients' health conditions.

### 3 Proposed WBANs Platform

This section describes the WBANs platform which is designed and evaluated to perform remote monitoring of diabetic patients. The main aim of the proposed WBAN is to allow easy accessibility of diabetic patients' health condition to the physician from any remote location for continuous monitoring of the patients' condition and to take the necessary actions. Both invasive and non-invasive sensors are used and the sensed data is sent to the patient's android smartphone which then forwards it to the medical server whereby the physicians can directly have access.

#### 3.1 Design Requirement

In the proposed WBANs, the invasive and non-invasive sensor nodes collect the patients' physiological information such as the blood glucose level, temperature and blood pressure. Thus, data quality is highly crucial for correct processing, interpretation

of information as well as effective diagnosis, control, monitoring and treatment of chronic diseases [8]. Therefore, all vital signs for patients should transfer appropriately through the mobile infrastructure and the data should be sent to the patients and clinicians in a real-time environment. Any possible errors due network coverage, failure-prone sensors, signal strength, energy efficiency, latency, transmission reliability, delay performance, and failures of infrastructure components could lead to an unpredictable quality and reliability of patient monitoring [9–12]. Therefore, ensuring QoI, reliability and ease of use are key design requirements for the acceptance of the proposed WBANs system.

### 3.2 Architecture Design

As shown in Fig. 2, the proposed WBAN consists of 3 hierarchical tiers namely the sensor tier, the mobile computing tier and the remote server tier. The sensor tier interacts with the human body and mainly consists of both the invasive or non-invasive sensors that are responsible to sense the vital signs, represent the data (pre-processing, compression and filtering), and transfer the data over short-range wireless communication between the sensor and mobile computing tier. The mobile computing tier consists of the android smartphone which aggregates and stores the sensed data, processes the energy consumption, provides the healthcare monitoring interface to the patients for logging and also sends the physiological data to the remote server tier through GPRS networks at a specified time interval. The third tier is the remote server tier which consists of remote servers and doctors' smartphones. The remote server collects the patients' data from the GPRS network, categories and records the user's information on a database for further analysis, diagnosis and intervention.

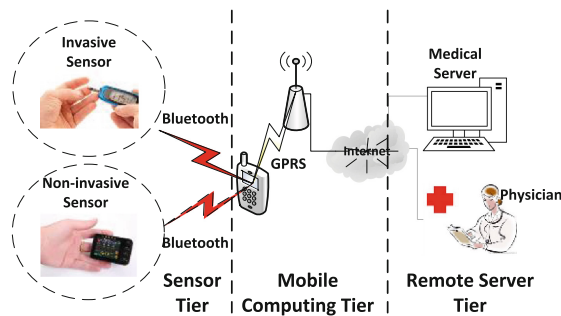


Fig. 2. High-level WBAN architecture.

### 3.3 Sensor Design

As shown in Fig. 3, the patients can measure their glucose blood level using either the invasive sensor (LifeScan's OneTouch Ultra 2 glucose meter) or the non-invasive sensor (Combo Gluco meter or GlucoTrack). The invasive glucometer measures the glucose blood level from the patient's blood sample whereas the non-invasive

glucometer can be placed directly on the finger for sensing the glucose level. The blood pressure is measured using Omron Automatic Blood Pressure and physical sensors such as pulse oximeter, weighing scale, cardiovascular and thermometer are used to measure other bio-information.



Fig. 3. Graphical user interface.

### 3.4 Communication Architecture and Protocols

In the proposed system, the invasive and non-invasive sensors record the data and send these readings to the mobile phone using existing communication protocols such as Bluetooth and ZigBee since they ensure security, robustness and privacy as shown in Table 1. The mobile phone acts as a gateway node which aggregates the traffic from the sensor nodes and forwards it to the access point. IEEE 802.11 /WiFi/GPRS is used as communication medium with the access point.

Table 1. Communication protocols.

Technology	Bluetooth	ZigBee
Operating frequency	2.4 GHz	868 MHz
RF data rate distance	1–3 Mbps	250 Kbps
Distance	10 to 100 m	10–200 m
Power consumption	< 30 mA	< 15 mA
Network topology	Scatter net	Star or Mesh
Robustness	Adaptive fast frequency hopping, FEC, fast ACK	DSSS, uses 16 channels ISM (2.4 GHz) band only
Security	64b/128b AES and application layer user defined	128b AES and application layer user defined

### 3.5 Data Design

The patients' physiological parameters sensed by the sensors have different range according to The American Diabetes Association [13] as shown in Table 2. The quality of the sensed data is captured via a collection of QoI attributes [8] that includes timeliness (delay from a given event), accuracy (data compared with a data referential), reliability, completeness (percentage of data missing at a given points), relevancy (impact of specific data on the decisions), usability (relevant and useful data for decision making) and confidentiality (confidential and secure data). Moreover, the frequency of the sensed data varies based on the risk level assigned by the physicians. The frequency can include different time intervals such as 15 min, 30 min, one hour, twice per day or once per day depending on the health condition of the patients.

**Table 2.** Physiological parameters.

Blood glucose parameter	Normal	Pre-diabetes	Diabetes
HbA1c (%)	< 5.7	5.7–6.4	≥ 6.4
Fasting plasma glucose (FPG)	≤ 99 mg/dl	100–125 mg/dl	≥ 126 mg/dl
Postprandial plasma glucose (PPG)	< 140 mg/dl	140–199 mg/dl	≥ 200 mg/dl
	Desirable	Borderline	High Risk
HDL Cholesterol	60 mg/dl	35–45 mg/dl	< 35 mg/dl
Blood Pressure	–	–	≥ 140/90
BMI	–	–	≥ 35

### 3.6 Application Design

The smartphone application is developed using C# using the Windows Mobile .net platform and Microsoft SQL Server Compact 3.5 is used because of its high security capability as confidentiality is of utmost importance. The application allow the patients to acquire, store, upload, download and display information (Fig. 2). The physicians can view patient's diabetes reports and assign an appropriate risk level based on the patient's condition whilst viewing their reports.

The primary goal of the proposed architecture design, sensor design, communication architecture and protocols, data design and application design is to ensure better QoI, reliability and ease of use on the system. Higher QoI is ensured by having the data processing logic at the sensor tier to integrate, correlate, and transform multiple sensor readings to avoid wrong decision making due to faulty readings, interferences and loss of data. Moreover, to increase reliability of the proposed system, the mobile computing tier processes the sensor readings according to the threshold configuration set by the physicians and trigger alerts in emergency cases. Finally, a color coding scheme is used to show the health status of the diabetic patients which will eventually increase ease of use of the system in helping them towards better tracking and self-monitoring.



### 4 Evaluation of the Proposed WBAN Design

The proposed WBAN is evaluated as shown in Table 3 on the three different levels such as (1) sensor tier (2) mobile computing tier and (3) remote server tier whereby several components from each tiers are addressed.

**Table 3.** WBANs evaluation of 3 Tier levels.

Sensor tier level	
Sensor technology	Both the invasive and non-invasive sensors that are used are physically comfortable from wear-ability point of view and they should not obstruct the patient’s normal activities
Data capture	Wearable sensor technology is used to ensure energy efficiency. The QoI attributes are also measured
Data representation	The data size to be transmitted is reduced through amplification/filtering in the aim of noise rejection, power consumption reduction, and removal of duplication before transmission. Compression techniques are also used to ensure network efficiency and power efficiency
Mobile computing tier level	
Data transfer	Data fusion techniques are used to reduce the amount of data to be transmitted and time required for transmission to ensure efficient energy aware communication and bandwidth utilization. Reliable communication protocols and specific mechanism are put into place to evaluate faulty readings, interferences and loss of data
Mobility	Patient should be able to stay connected anytime and anywhere within a reasonable range of 100 m
Mobile computing tier level	
Security and privacy	Transmitted data needs to be protected and integrity of received signal should be maintained
Fault tolerance	The remote server should continue to operate in the event of any failure for continuous monitoring of patients
Dependability	Patient-related data must be readily retrievable when node failure or data erasure happens
Data analysis/Interpretation/Decision making	The feedback of physicians relies on what the database shows. Therefore, the remote server has to validate the data integrity before making it available to the physicians or saving it to the medical record

The proposed WBAN exhibits great potential to improve the performance of the healthcare sector for diabetic patients by integrating smart sensors, wireless communication protocols and network technologies. This system improve the quality of life of diabetic patients as they will no longer need frequent hospital visits and they will be

location independent. Higher degree of confidence in QoI is maintained by correlating multiple sensor readings for reliable data capture, processing, transmission of patient’s physiological parameters so as to facilitate real-time diagnosis and intervention from physicians. However, future works are yet to be done concerning the privacy and security of the WBAN to ensure user acceptance. System acceptance should be further evaluated using both qualitative (usability, social acceptance and reliable data analysis) and quantitative (high degree of confidence in the QoI dimensions such as timeliness, accuracy, reliability, completeness, relevancy, usability and confidentiality) methods.

## 5 Research Challenges and Future Works

Since WBANs is an emerging technology, many issues are yet to be addressed and open future research directions. Common research challenges and future works are discussed in Table 4.

**Table 4.** Challenges and future work of WBAN.

Challenges	Details	Future work
QoI issues – (timeliness, accuracy, reliability, complete-ness, relevancy and usability)	Associated during the phases of: <ul style="list-style-type: none"> <li>• data collection (impacted by data collection rate, performance of body sensors, quantity of data to be pre-processed and transferred and quality of communication)</li> <li>• data processing (incorrect information during analysis and data enhancement)</li> <li>• data delivery (incorrect diagnosis, in-appropriate treatments and untimely decisions)</li> </ul>	Use of optimized data quality algorithms at the sensor tier, mobile computing tier and remote server tier to improve QoI in terms of timeliness, accuracy, reliability, completeness, relevancy and usability for a higher degree of confidence
Privacy/Security	Ensuring privacy and security during data collection and transmission from sensors to access point and from access point to remote server is highly critical for successful design and acceptance of WBANs	Advanced algorithms, authentication protocol, privacy-preserving mechanisms, encrypted and cryptography techniques should be developed to ensure data integrity and security

*(continued)*

**Table 4.** (continued)

Challenges	Details	Future work
User acceptance	Users are not willing to adopt new systems and technologies for multiple factors such as lack of trust, motivation, perceived usefulness and perceived ease of use	It is important to develop new Technology Acceptance Model to get precise feedback on how well such system are accepted
Mobility	Ensure long term monitoring and minimum interference while the patients can move anywhere and anytime is critical for the successful design	Energy aware protocol should be used to ensure nodes can negotiate their transmission power to a minimum

## 6 Conclusion

The WBANs has emerged as a major concern in the healthcare sector and still much issues are yet to be addressed for their successful development and deployment. In this paper, the design of a WBANs using both invasive and non-invasive sensors to continuously monitor diabetic conditions has been described. The proposed WBANs is used to bridge the gap of information between patients and physicians whereby the physicians have real-time and historical view of patient’s condition. This feature allows them to perform diagnosis and timely intervention remotely and take appropriate action in case of emergency. The patients can also better monitor and record their data into a more efficient manner as compared to the traditional way of frequent hospital visits. The next step of the research includes the implementation of the system by considering all the issues addressed and a thorough evaluation of the system to ensure user acceptance. The challenges and future work discussed in the paper should also be considered during the development and deployment of the system.

## References

1. Xing, J., Zhu, Y.: A survey on body area network. In: Proceedings of the 5th International Conference on Wireless Communications, Networking and Mobile Computing, pp. 1–4 (2009)
2. Dangi, K.G., Panda, S.P.: Challenges in wireless body area network-a survey. In: Proceedings of the International Conference on Optimization, Reliability, and Information Technology (ICROIT), pp. 204–207 (2014)
3. Movassaghi, S., Abolhasan, M., Lipman, J., Smith, D., Jamalipour, A.: Wireless body area networks: a survey. *IEEE Commun. Surv. Tutor.* **6**(3), 1658–1686 (2014)

4. Mamykina, L., Mynatt, E.D., Davidson, P.R., Greenblatt, D.: MAHI: investigation of social scaffolding for reflective thinking in diabetes management. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 477–486 (2008)
5. Alhazbi, S., Alkhateeb, M., Abdi, A., Janahi, A.: Mobile application for diabetes control in Qatar. In: Proceedings of the 8th International Conference on Computing Technology and Information Management (ICCM), pp. 763–766 (2010)
6. Petrella, R., Luo, Y., Ling, C., Schuurman, J., Petrella, R.: GlucoGuide: an intelligent type 2 diabetes solution using data mining and mobile computing. In: IEEE International Conference Data Mining Workshop, pp. 748–752 (2014)
7. Mougiakakou, S.G., Bartsocas, C.S., Bozas, E., Chaniotakis, N., Iliopoulou, D., Kouris, I., Pavlopoulos, S., Prountzou, A., Skevofilakas, M., Varotsis, K., Vazeou, A., Zarkogianni, K., Nikita, S.K., Member, S.: SMARTDIAB: a communication and information technology approach for the intelligent monitoring, management and follow-up of type 1 diabetes patients. *IEEE Trans. Inf Technol. Biomed.* **14**(3), 622–633 (2010)
8. Rodríguez, C.C.G., Riveill, M., Antipolis, S.: e-Health monitoring applications: What about data quality (2010)
9. Sachidananda, V., Khelil, A., Suri, N.: Quality of information in wireless sensor networks: a survey. In: Proceedings of the International Conference on Information Quality (ICIQ), vol. 1, pp. 193–207 (2010)
10. Varshney, U.: Mobile health: four emerging themes of research. *Decis. Support Syst.* **66**, 20–35 (2014)
11. Rafe, V., Hajvali, M.: A reliable architectural style for designing pervasive healthcare systems. *J. Med. Syst.* **38**(9), 1–10 (2014)
12. Ullah, S., Higgins, H., Braem, B., Latre, B., Blondia, C., Moerman, I., Saleem, S., Rahman, Z., Kwak, K.S.: A comprehensive survey of wireless body area networks. *J. Med. Syst.* **36**(3), 1065–1094 (2012)
13. Association, A.D.: Standards of medical care in diabetes. *Diab. Care* **31**, S12 (2008)

# A Secure Device and Service Discovery Protocol for Proximity Based Social Networks (PBSNs)

Asslinah Mocktoolah<sup>(✉)</sup>, Krishiv Askoolum, and Kavi Kumar Khedo

Department of Computer Science and Engineering,  
University of Mauritius, Moka, Mauritius

ashlinahmee@gmail.com,

krishiv.askooluml@umail.uom.ac.mu, k.khedo@uom.ac.mu

**Abstract.** Technical advances of Online Social Networking (OSN) have promoted new types of social interactions such as Proximity Based Social Networking (PBSN) where proximate users interact with each other using their mobile devices. However, personal details users share on these networks are not secured and messages exchanged are prone to different attacks and eavesdropping. Qualcomm introduces AllJoyn, which eases the development of such applications discovering with nearby devices and establishing a secure communication. In this paper, a secure device and service discovery protocol is introduced leveraging the AllJoyn framework. It allows mobile users to perform secure device advertising and discovery, authentication, encryption and decryption of exchanged messages. The security and performance of the proposed architecture are thoroughly analyzed and evaluated by using penetration tools and scenarios. Results showed that exchanged messages could not be captured and the system is not prone to attacks such as man-in-the-middle attacks.

**Keywords:** Proximity Based Social Networks · Secure discovery · Privacy · Device to device communication

## 1 Introduction

In the last decade, advances in information technologies have significantly altered the way humans interact. Our modern lifestyle has pushed its limits to a hyper-connected society where everyone is always online especially with the advent of online social networks. Technology has advanced so significantly that we can nowadays hold them in the palms of our hands and be connected to anyone regardless the place and time [1]. This emergence of smartphones has driven the mode of communication to a new level. However, on the other side, it has been outlined that our modern life is accompanied by a loneliness epidemic where face to face interactions are becoming more and more negligible. To merge the physical and virtual worlds, Proximity based Social Networking (PBSN) has been introduced helping users in close proximities to connect to each other based on interests using the geographical positions of the devices [2].

Unlike popular location based social networks such as Foursquare [3] which uses centralized servers to receive location updates from the mobile devices of users, PBSN applications are mostly based on a device-to-device (D2D) communication offering decentralized proximity-based services [4]. D2D communication enables direct communication between users within close proximities via short range radio such as Bluetooth or Wi-Fi direct proving as an innovative feature for the next generation cellular networks.

The uses of PBSN applications are not only limited to meeting new people located in proximity areas but these systems have aroused the interests of many advertisement enterprises. Companies make use of different location information of customers to discover common traffic patterns so as to understand and analyze customer buying behavior [5]. Businesses adopt these systems to promote their products or services to other related companies and these services can also be integrated into location based gaming applications [6]. Recently, the global craze Pokemon Go has been observed to be the most successful mobile game in the US attracting millions of users within few days after its release surpassing even the number of users of large Online Social Networking (OSN) applications [7].

However, it is noted that though there is a great usage of different types of PBSN applications based on the benefits offered, users nevertheless imperil their privacies [8]. Retrieving location updates from users' mobile devices constantly can enable undesirable persons to learn the location preferences of a user. Furthermore, sensitive information such as visiting pubs or clinics may not be wished to be shared to other users [9]. This information gathering can lead to stalking or other types of illicit activities such as being robbed or sexually assaulted. To ensure the continuity and success of PBSN applications, a secure device and service discovery protocol is needed. This will not only attract more users but also establish a level of trust between the registered users.

This paper introduces a new privacy preserving protocol which allows users in proximity to communicate to each other in a secure environment. The conversations or messages exchanged cannot be sniffed by intruders and also integrity and confidentiality is ensured in this protocol. The latter is built upon AllJoyn, an open source framework, which simplifies the development of proximity based applications and takes care of several complexities such as discovery, interface sharing, networking, etc. These features made AllJoyn the number one implementation choice over the other frameworks available for this protocol.

The remainder of this paper is structured as follows: Sect. 2 reviews the related works on PBSN systems highlighting the security and privacy provisions implemented in each. An overview of the different PBSN frameworks is given in Sect. 3 and in addition the features of the AllJoyn framework and also the security limitations are discussed. The detailed design of the secure device and service discovery protocol is presented in Sect. 4. Section 5 evaluates the performance of the proposed protocol followed by the interpretation of results. Some future works are proposed in Sect. 6 and the paper is closed on a concluding note in Sect. 7.

## 2 Related Works

Following the trend of PBSN, many studies have converged to simulating such systems concentrating on different protocols such as discoverability and communication. Given the numerous privacy challenges faced by PBSN users, many new researches were directed to creating privacy-preserving systems securing the discovery process and protecting the information shared between the proximity users. In this section, several such systems and protocols are discussed.

Even if these systems are based on proximity interactions, the architectures of such applications were based on centralized servers. MobiClique, however, is based on opportunistic connections between devices in proximity and enables decentralized ad hoc interactions between neighboring devices [10]. The neighborhood discovery procedure depends on the radio technology that the device is using such as Bluetooth device discovery or WiFi SSID. Once a new device has been discovered, MobliClique reaches the identification process in which the devices exchange their user identity information using a communication link. The last step refers to the data exchange phase, where messages are exchanged between each other. No security mechanisms are included in this design except the authentication of the users during the identification process. Furthermore, the authors ensure that this system gives a guarantee on privacy based on the forwarding rules when users exchange messages, that is, no other nodes will receive message contents unless they belong to a particular group.

EnCore, a mobile platform, based on secure encounters allows a privacy-preserving communication between pairs of mobile devices [11]. Users discover each other by D2D communication using the SDDR (Secure Device Discovery and Recognition) protocol. The discovery process is further secured by making use of the Diffie-Hellman (DH) public key and shared keys. All communication is protected by a message authentication code such as encounter key or event key. Furthermore, Encore ensures confidentiality and authenticity by encrypting the messages exchanged. Meetup is a similar encounter-based mobile application which uses visual authentication for any encounters by using the picture of the users for information exchange and verification. [12]. The Tor network is included to provide an anonymous communication between the users by obscuring their identity and location information. However, since the picture of the user is shared, the privacy of the user is somehow breached. For a secure PBSN application, any identifiable information should not be shared.

In addition of using unique identities of each device, PRISM introduces a new concept of privacy protection protocol matching the private information of two users without the risk of them disclosing their interests [13]. On top of their encrypted interests exchanged, some dummy interests known to both participants are also sent in plain text and encrypted values. The authors claimed that this new scheme makes the protocol resistant to different attacks such as man-in-the-middle attack and Sybil attacks since the index positions and their corresponding values are shuffled and same has to be deciphered before getting the real information. However, a complete evaluation of PRISM has not yet been done on a mobile application and one of a future work is to integrate an encryption based mechanism also. Prila comes up with a new privacy-preserving protocol where a CFO (carrier frequency offset) encryption technique is used to secure the

communication link between two users [14]. A multipath profile based location authentication is introduced in this paper using the DTW (Dynamic Time Warping) matching algorithm to cater for issues such as forged locations.

### 3 PBSN Frameworks

To ease the development of above mentioned PBSN applications, several frameworks have been designed and used. Different mechanisms such as discovery, device resource management and network maintenance are integrated in these frameworks, therefore, allowing developers to focus on the specific areas they are implementing the applications. For instance, Hum-Hub is a social networking framework allowing implementation of lightweight applications and concentrates mainly on communication and collaboration [15]. Elgg is a similar open-source platform ensuring flexibility and simplicity in the development of these types of applications [16]. Another such framework is Anahita which is known as the first open source solution following the real architecture of a social network [17]. It is based on knowledge sharing networks, collaboration environments, online learning and sharing of data.

AllJoyn, by Qualcomm, now managed by AllSeen Alliance, is another popular software framework offering several other advantages in addition of mobile application developments [18]. Automatic device discovery, dynamic network configuration and secure communication are some core features of AllJoyn. Developers can easily develop PBSN applications using AllJoyn without dealing with complexities of networking or discovery. However, the AllJoyn security integrated is not sufficient since it occurs only at the application level and no provisions have been made to maintain trust at the device level. Since users are more conscious of their privacy nowadays and would not expose their personal information to other people, a secure PBSN application will attract more user [9].

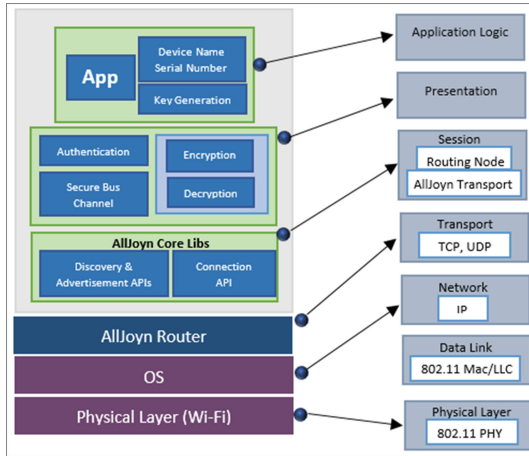
Therefore, in this paper, a new protocol is introduced leveraging the AllJoyn framework to ensure a secure discovery of users in proximity who wish to communicate with each other. The above mentioned features made AllJoyn the best option for implementing this secure device and service discovery protocol and is ideal since the participants are within close proximities. The next section outlines how this protocol was designed and implemented upon the AllJoyn framework.

### 4 Proposed Secure Device and Service Discovery Protocol

In order to test the protocol, a proximity messaging application has been implemented ensuring a secure device and service discovery of the nearby mobile devices and allowing a protected communication. The PBSN application is implemented upon the AllJoyn framework abstractions which already provide several services such as onboarding, session creation, device discovery and advertisement. By using the common AllJoyn service architecture, devices and applications can properly interoperate with each other to perform a specific functionality. The proposed architecture for the proximity message exchanger application is an overview of how the security modules



will operate in practice and are mapped through the OSI Layer and is designed as shown in Fig. 1 over the AllJoyn’s architecture.



**Fig. 1.** Proposed PBSN architectural design

The main components of the model are described as below:

## 4.1 Application Layer

### 4.1.1 Device Name and Serial Number

The Bluetooth name of the mobile device is used as the device name followed by the IMEI number of the phone. No two devices will have the same IMEI, and therefore it is used as a unique identifier for each device for service discovery. The Telephony Manager library is used to import the IMEI number of the device and this method is permission protected, therefore read access must be given in the androidManifest.xml file to retrieve this information.

### 4.1.2 Key Generation

Two random 256 bytes AES keys are generated on the node master side which will be used for authentication and encryption/decryption purposes respectively.

## 4.2 Presentation Layer

The presentation layer consists of three main phases of the secure protocol as explained below:

### 4.2.1 Secure Device and Service Discovery

The secure bus channel provides a secure device advertising and discovery mechanism. The distinctive feature of this solution is that it contributes in preserving the identity of the mobile users. The procedure is as follows:

A node master creates a message session and connects to the AllJoyn core library by sharing the name of the device (Bluetooth name) and the IMEI key which then connects to the AllJoyn router (Wi-Fi hotspot). To further preserve his identity, the user can choose to be anonymous by setting any name. The unique name of the device is automatically assigned when the application connects to the AllJoyn Bus which then validates if the object is secure. A channel is then hosted and a port is created via the BindSessionPort after which the node master can advertise the created channel to nearby mobile devices. The application discovers any nearby mobile devices whose status has been set to visible using the method FoundAdvertisedName. The node master can further choose which node will participate in the message session and only nodes having the appropriate rights can join. By simply turning the visibility to off, a device stops advertising itself.

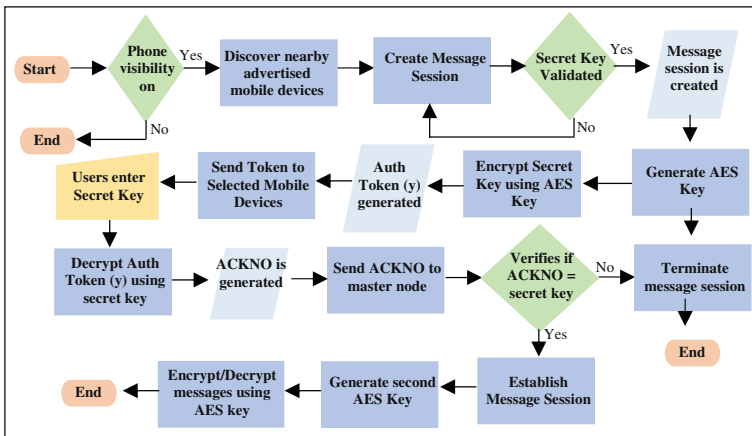


Fig. 2. Overall processes of the protocol

### 4.2.2 Authentication of User Devices

An authentication key and the message session name is prompted to be entered when a proximity message session is created and only nearby users having access rights to this session are aware of this secret key. To further strengthen the authentication scheme, strong passwords are required to protect against brute force attacks. After validating the password, a random 256 bytes AES key is generated. The AES key is used to encrypt the secret key generating an Auth\_Token(y) which is then sent to selected proximity devices. To proceed with the message session, the proximity devices are asked to enter the original authentication key used by the master node once they receive the Auth\_Token(y). The password is used to decrypt the Auth\_Token(y) which generates an ACKNO. The ACKNO is then sent back to the master node who ultimately verifies if the ACKNO matches to the username. If both are the same, the master node validates the connection by sending Auth\_Successful to the mobile users and finally they can send messages to each other.

### 4.2.3 Encryption and Decryption of Proximity Messages

Once authentication of the users has been validated, a second 256 bytes AES key is generated by the master node. This key is shared to the selected mobile users and is used for encryption and decryption of the messages together with the vector. The encryption process is as follows:

A Cipher object is created which calls the Cipher’s getInstance method. The name (“AES/CBC/PKCS7Padding”) is passed in the cipher and the encryption key is converted into an array of bytes. By passing this key and the algorithm (AES) as parameters, a SecretKeySpec is created and an IvParameterSpec object is created when the vector is passed as a parameter. The Cipher is encrypted with the Cipher.doFinal() method and is encoded using Base64. The result is converted into a string which is then sent as an encrypted message. The same steps are carried out for decryption purposes. For a better security, a re-authentication process is needed for mobile users who are idle for more than 20 min to re-active their sessions.

An overview of the processes of the protocol is displayed in Fig. 2.

## 5 Evaluation of Proposed Protocol

The proposed architecture was implemented as an application using Eclipse IDE JUNO and Android Development Tools and was tested in different Android devices. A Wi-Fi hotspot without internet access which was used as a router to connect to other mobile devices was setup. Each device could install the application and discover each other by using the AllJoyn discovery mechanism. The evaluation of the security algorithms were done in the below three phases:

### 5.1 Penetration Tools Evaluation

Existing tools such as Wireshark and zANTI have been used to perform penetration testing to check the robustness of the system and the reliability of the architecture by trying to break into the system. Wireshark is used to decode the encrypted messages that are exchanged and test if each packet contains any useful information. However, the only information captured by Wireshark relates to the device name which sent the message, the message session name and the size of the message as seen in Table 1 displaying the analysis of the packets captured by Wireshark. Neither the key used for authentication or cryptography technique nor the message contents could be captured.

**Table 1.** Packets intercepted by Wireshark

Message sent	Size (Bytes)	Wireshark detection	
		Session name	Message
M1 (from LG to Samsung device)	5	TestChatName	Not detected
M2 (from Samsung to LG device)	10	TestChatName	Not detected
M3 (from Samsung to LG device)	52	TestChatName	Not detected
M4 (from LG to Samsung device)	122	TestChatName	Not detected

zANTI was used to simulate man-in-the-middle (MITM) attacks on the secure PBSN application. Each message transmitted was checked by zANTI, but no information could be sniffed by this tool as well as illustrated in Fig. 3.

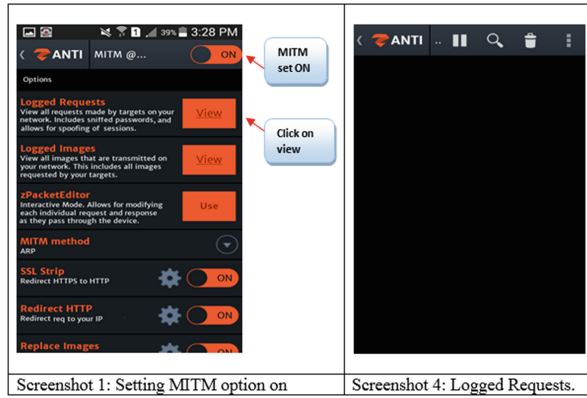


Fig. 3. MITM attacks by zANTI

### 5.2 Scenario Tests Evaluation

The secure PBSN system was assessed by using different scenarios. The first one relates to the extraction of the keys used for cryptography by malicious users which can end up into intercepting the entire discussions between devices. However, since a random 256-bytes generator is used, no hard-coded credentials are stored to perform authentication or cryptography. Encryption is usually regarded as not enough to ensure information confidentiality since attacks can be performed easily. AES with 256-bytes keys were used in this application which were derived using protocols such as PKCS7 (Public-Key Cryptography Standard) making the algorithms more robust to attacks. The data integrity of the system has also been tested. The results showed that the cryptography algorithms did not change the original message contents and the decrypted messages corresponded to the message sent as seen in Fig. 4.

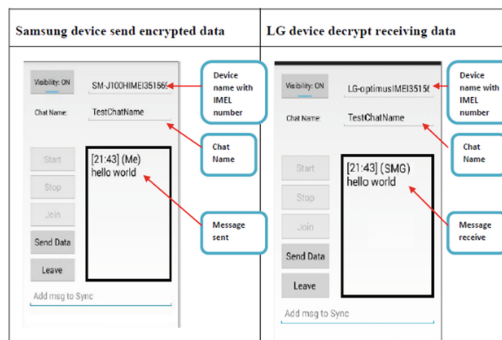


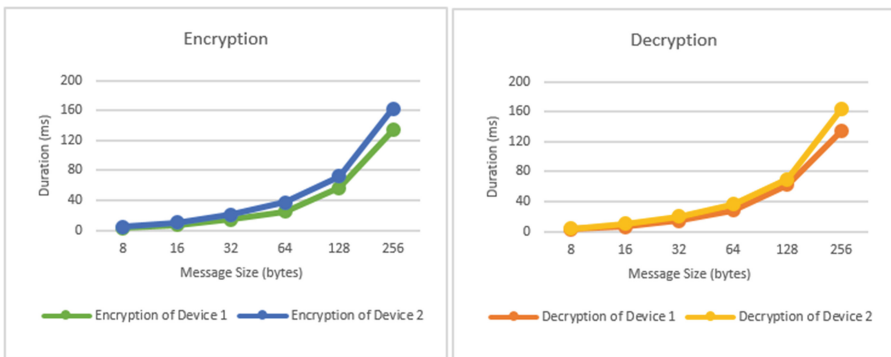
Fig. 4. System reliability

### 5.3 Performance Evaluation

Performance and efficiency of operations have direct impact upon the usability of the secure PBSN application in such a way that the user will not want to use the application or even bypass the security solution. The performance and power consumption are evaluated to the corresponding message sizes with respect to the different security techniques employed. The performance of the cryptography algorithms is analyzed by varying the messages sizes and measuring the time taken as illustrated in Fig. 5. It is observed that the time taken for encryption or decryption increases when the message sizes increase but energy consumption remains constant such that these operations do not affect the battery power even when varying message sizes as shown in Table 2.

**Table 2.** Encryption and Decryption Log File

Message sizes (bytes)	Samsung device		LG device	
	Battery level (mAH)	Duration (ms)	Battery level (mAH)	Duration (ms)
<b>Encryption</b>				
8	100	3	100	5
16	100	7	100	10
32	100	14	100	21
64	100	25	100	37
128	100	57	100	72
256	100	134	100	162
<b>Decryption</b>				
8	100	3	100	4
16	100	6	100	10
32	100	14	100	20
64	100	28	100	36
128	100	62	100	70
256	100	134	100	163



**Fig. 5.** Encryption and decryption log file

RKG was performed by both devices to ensure that the keys generated are truly random and not predictable. The outputs as displayed for the Samsung device in Table 3 showed that the AES keys produced were random, therefore achieving key uniqueness.

**Table 3.** Samsung galaxy J1 Log file sample.

No	Key generation (ms)			Generate AES KEY (256 bytes)
	Start time	End time	Duration	
1	671396	671397	1	E74AF90196B156CC.....
2	671397	671399	2	8863F36F49EBA89A.....
3	671399	671400	1	C7EFE0A26566AEFB.....
4	671400	671402	2	BA4635C6CB16EC48.....
5	671402	671403	1	F237F209A817850C.....
6	671403	671404	1	AB20780567742D0C.....
7	671535	671545	10	EBBF3C39222ECE08.....
8	671545	671546	1	39D39F4770A193C5.....
9	671546	671548	2	7CC755C005D4D361.....
10	671548	671549	1	EE41BC152C0F5C7D.....

## 6 Future Works

Though the current system attained the main objectives set such as authenticity, privacy, confidentiality and performance, some improvements can be integrated such as analyzing new approaches to protect the message session name. Most of the tests have been carried out using a limited number of devices, hence more tests should be performed using different types of devices for a better evaluation of the proposed protocol. Furthermore, load testing could be performed to evaluate the behavior of the system when more than two proximity users are sending messages on the system. Further tests must be done to test how the system might respond to other attacks such as evil twin attacks and Cafe Latte attacks.

## 7 Conclusion

In this paper, an efficient and effective PBSN system is proposed that allows mobile users to perform secure device advertising and discovery, authentication of trusted users and a protected communication by encrypting the messages exchanged. With the fast emergence of PBSN systems, security should be a hardline, and as per the results obtained with the proposed system, it can be concluded that the security of this system has been handled effectively. The PBSN system was tested using penetration testing tools and results have showed that no messages exchanged could be intercepted. The privacy and confidentiality of the messages have been preserved in addition to the data integrity by proving that messages exchanged were not altered during the cryptography

techniques. The performance tests of the application confirmed that it was quite fast to process each message in encryption and decryption operations and also results have shown that the energy consumption was constant.

## References

1. Raiwani, Y.P.: Internet of things: a new paradigm. *Int. J. Sci. Res. Publ.* **3**(4), 323–326 (2013)
2. Zhang, R., Zhang, Y., Sun, J., Yan, G.: Fine-grained private matching for proximity-based mobile social networking. In: *Proceedings IEEE INFOCOM, USA* (2012)
3. Foursquare (2015). <http://www.foursquare.com>
4. Lin, X., Andrews, J.G., Ghosh, A., Ratasuk, R.: An overview of 3GPP device-to-device proximity services. *IEEE Commun. Mag.* **52**(4), 40–48 (2014)
5. Ojala, T., Kruger, F., Kostakos, V., Valkama, V.: Two field trials on the efficiency of unsolicited Bluetooth proximity marketing. In: *Proceedings of Mobile and Ubiquitous Multimedia (MUM)*, pp. 1–4. ACM, Ulm, Germany (2012)
6. Bonchi, F., Castillo, C., Gionis, A., Jaimes, A.: Social network analysis and mining for business applications. *ACM Trans. Intell. Syst. Technol.* **2**(3), 22–37 (2011)
7. Allan, R.: Pokémon GO usage statistics say it's the biggest mobile game in U.S. history (2016). <https://www.surveymonkey.com/business/intelligence/pokemon-go-usage-statistics/>. Accessed 08 Aug 2016
8. Mocktoolah, A., Khedo, K.K.: Privacy challenges in proximity based social networking: techniques & Solutions. In: *Proceedings of Computing, Communication and Security (ICCCS)*, pp. 1–8. IEEE, Mauritius (2015)
9. Askoolum, K., Khedo, K.K.: Secure device and service discovery protocol for proximity based social networks (PBSNs), Master's thesis. University of Mauritius, Mauritius (2015)
10. Pietilainen, A.K., Oliver, E., LeBrun, J., Varghese, G., Diot, C.: Mobiclique: middleware for mobile social networking. In: *Proceedings of ACM Workshop Online Social Network*, pp. 49–54 (2009)
11. Aditya, P., Erdélyi, V., Lentz, M., Shi, E., Bhattacharjee, B., Druschel, P.: Encore: private, context-based communication for mobile social apps. In: *Proceedings of the 12th Annual International Conference on Mobile Systems, Applications, and Services*, pp. 135–148. ACM (2014)
12. Mohaisen, A., Kune, D.F., Vasserman, E., Kim, M., Kim, Y.: Secure encounter-based mobile social networks: requirements, designs, and tradeoffs. *IEEE Trans. Dependable Secure Comput.* **10**(6), 380–393 (2013)
13. Abbas, F., Rajput, U., Oh, H.: PRISM: privacy-aware interest sharing and matching in mobile social networks. *IEEE Access* **4**, 2594–2603 (2016)
14. Chen, Y., Wang, W., Zhang Q.: Privacy-preserving location authentication in WiFi with fine-grained physical layer information. In: *Global Communications Conference (GLOBE-COM)*, pp. 4827–4832. IEEE (2014)
15. HumHub (2016). <https://www.humhub.org>. Accessed 18 Jul 2016
16. Elgg (2016). <https://elgg.org>. Accessed 18 Jul 2016
17. Anahita (2016). <http://www.getanahita.com>. Accessed 18 Jul 2016

# Performance of Unequal Error Protection Schemes for Audio Transmission Over ADSL with Reed Solomon and Turbo Codes

Prateema Ragpot<sup>(✉)</sup>, Tulsi Pawan Fowdur, and K.M.S. Soyjaudah

Department of Electrical and Electronics Engineering,  
University of Mauritius, Réduit, Mauritius  
prati\_r910@hotmail.com, {p.fowdur,  
ssoyjaudah}@uom.ac.mu

**Abstract.** This paper performs a comparative analysis of Unequal Error Protection (UEP) schemes for audio transmission over Asymmetric Digital Subscriber Line (ADSL). The proposed UEP scheme exploits three different levels of protection. First at the Discrete Multi-tone Modulation (DMT) layer, second with prioritized retransmissions and finally by using different code rates. A comparative analysis between Reed-Solomon (RS) codes and Turbo Codes is performed using two different block sizes. For the Turbo codes, Sign Division Ratio (SDR)-based extrinsic information scaling is also applied to further enhance the performance. Results show that when UEP is applied at three different levels, a gain of 11 dB is obtained over a scheme which applies UEP at only two levels.

**Keywords:** Audio transmission · Unequal error protection · RS code · Turbo codes

## 1 Introduction

ADSL is the most common type of Digital Subscriber Line (DSL) technology for sending high-bandwidth information over standard telephone lines. Faster data transmission is enabled through a single connection. Users can download data and make voice calls simultaneously. The highest download speeds are usually 8 Mb, and the top upload speeds are usually below 1 Mb [1]. DSL now makes up 45.7% of fixed broadband subscriptions [2]. However, a significant load has been placed over current broadband infrastructures by the increase in multimedia traffic such as compressed audio over the internet [3]. Maintaining a certain level of Quality of Service (QoS) is imperative for the growth and usage of broadband applications. Error propagation terribly affects compressed audio data. So, ensuring a high QoS for the transmission of compressed audio over ADSL using conventional methods is very difficult. Retransmission is restricted due to tough delay constraints for services such as live audio streaming [4]. Thus various Unequal Error Protection (UEP) techniques have been developed. An analysis of such techniques is given next.



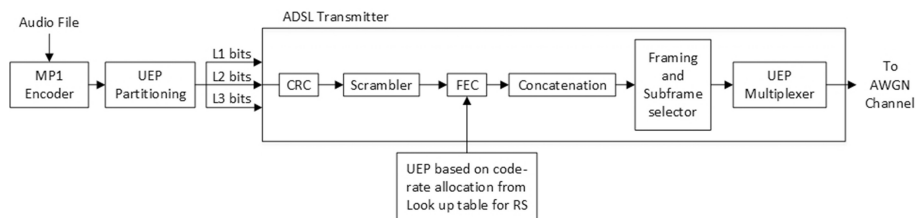
An optimized UEP scheme using Forward Error Correction (FEC) for transmission of embedded video bitstream over Additive White Gaussian Noise (AWGN) channel was applied in [5]. The error sensitivity and unequal importance of the bits were exploited. The bitstream was divided into high importance and low importance sub-streams based on their priority in the reconstruction of the video, and appropriate FEC was applied to them for protection against channel errors. In [6], an efficient UEP scheme to protect scalable video, coded with scalable extension of H.264/AVC using the genetic algorithm, over networks with packet loss was proposed. FEC schemes based on Reed Solomon codes were used for unequal error protection. The experimental results showed a significant gain of 1.15 dB in average when compared to conventional methods [5]. The principle of Turbo codes was analyzed in [7], and a new UEP Turbo codes was designed using a new puncturing scheme. The results showed that if the system complexity was not increased when the BER is lower than  $10^{-4}$ , the UEP Turbo codes achieved a coding gain of about 0.6 dB on the AWGN over the Equal Error Protection (EEP) codes. Finally, in [8], UEP using Turbo codes was applied to JPEG image transmission. The JPEG image was partitioned into DC components and AC components according to their respective sensitivity. A lower coding rate was applied to the highly sensitive DC components for better protection while the less sensitive AC components used a higher coding rate. The simulated results showed that the UEP schemes outperforms the EEP scheme in terms of bit error rate (BER) and peak signal to noise ratio (PSNR). In [9], a two level UEP scheme was proposed. First, the unequal importance of subbands generated by an MP1 codec was exploited. The bitstreams of the more important subbands were given more protection than the least important ones. Next, a sophisticated retransmission protocol was used where frames from the more important subbands were allowed a greater number of retransmissions than least important ones. The two level UEP scheme outperformed a conventional Equal Error Protection (EEP) scheme by an average gain of 28 dB.

This paper builds on the work in [9] and adds a third level of UEP by using variable code-rates for each  $E_b/N_0$  value. A look up table is derived for both RS and Turbo codes to provide UEP. The code-rates are varied accordingly and a performance comparison is done. For Turbo codes, extrinsic information scaling with SDR techniques is also performed to further improve the performance.

## 2 System Model

The proposed transmitter model is similar to the one in [9] except for the FEC block which accepts an additional input which is the look up table to vary the code rates. The modified system model is shown in Fig. 1.

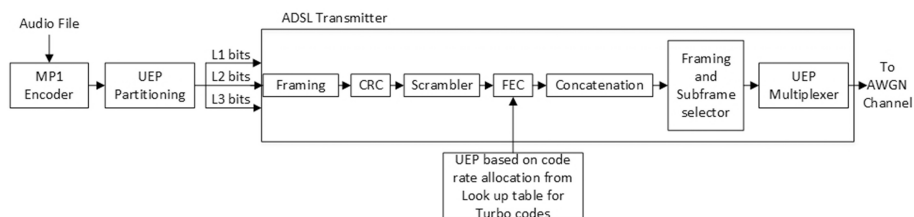
An audio file is encoded using the MP1 encoder. After the filter banks, a variance analyser is used to determine the variance of each subband. The UEP partitioning block uses this information to divide the encoded subbands into levels of different priority. The subbands having lower variance are considered less important than those with higher variance. The subbands have been divided in three levels, L1, L2 and L3. After the UEP partitioning, the levels are sent to the ADSL transmitter as in [9]. The bit-streams are broken into frames. The frames from each level are processed separately



**Fig. 1.** The proposed transmitter system for RS codes

by the various blocks. After the scrambler, the L1, L2 and L3 bitstreams go through the FEC block where different code rates are applied to the different bitstreams depending on the look up table.

For the Turbo codes, a framing block is added to determine the frame size of the L1, L2 and L3 frames to be transmitted as shown in Fig. 2. The frame size chosen depends on the interleaver block sizes for Turbo codes.



**Fig. 2.** The proposed transmitter system for Turbo codes

The RS encoder adds extra “redundant” bits to a block of data. For EEP, a message of length 112 is converted to a codeword of 120. Thus the code rate for L1, L2 and L3 bitstreams is set to 0.9333. For UEP in terms of code rate, a new message length of 96 is used in order to keep the total extra “redundant” bits being added to the whole system by the RS encoder same or less than that being added by the EEP system. Lower code rates are assigned to the L1 bitstream for maximum protection. The code rates for the L2 bitstream are either equal to that of the L1 bitstream or lower. The L3 bitstream is given the higher code rates. The codeword length is varied from 96 to 120 to obtain the different code rates being used by the look up table. The complete look up table is given in the next section.

For EEP the code rate of the turbo encoder for all three bitstreams (L1, L2 and L3) is set to 0.9333. It is ensured that total parity bits being added to the whole system by the turbo encoder is same or less than that being added by the RS encoder EEP system. Another lookup table with variable code rate is built for the turbo encoder.

To achieve the different code rates for the turbo encoder, puncturing matrices are used. Puncturing means that not all code bits being produced by the encoder are transmitted over the channel. This leads to a smaller amount of redundancy thus a higher code rate. The puncturing matrix consists of zeros and ones. The ones

correspond to keeping and the zeros to deleting the encoded bits. The puncturing in this work is performed is as follows:

Original unpunctured matrix with a code rate of 1/3:

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & & & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & \dots & & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & & & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} \rightarrow \begin{array}{l} \text{Systematic bits} \\ \text{Parity 1 bits} \\ \text{Parity 2 bits} \end{array}$$

The puncturing period is 96, thus the original unpunctured matrix consists of three rows of 96 bits set to 1.

The punctured matrix for EEP with a code rate of 0.9333 is as follows:

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & & & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 & 0 & \dots & & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & & & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix} \rightarrow \begin{array}{l} \text{Systematic bits} \\ \text{Parity 1 bits} \\ \text{Parity 2 bits} \end{array}$$

The systematics bits consists of ones, the first four bits of the second row are ones follows by 92 bits set to 0. The last row has 92 0 bits ending with four 1 bits.

One of the code rates for UEP used is 0.8571. Below is its punctured matrix used.

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & & & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 0 & \dots & & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & & & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

The systematics bits consists of ones, the first eight bits of the second row are ones follows by 88 bits set to 0. The last row has 88 0 bits ending with eight 1 bits.

The punctured data to be sent across the channel consists of the first row of bits in the punctured coded data matrix followed by the interleaved bits from rows two and three. The bits from rows two and three correspond to the non-zero elements in rows two and three of the puncture matrix.

Prioritized DMT is performed on the L1, L2 and L3 bitstreams as described in [9]. The lowest possible modulation orders of the DMT channels are filled with the L1 bits. The L2 and L3 bits occupy the successively higher modulation orders. The frame is then sent over the channel.

The proposed receiver model is the same one as in [9] with the addition of the look up table block as an input to the FEC block.

If a received frame is corrupted and if the limit for retransmission has not been exceeded, the receiver asks the transmitter to send the same frame again. The corrupted frame is stored. Equal gain combining is used to combine the corrupted frame with its retransmitted version to enhance the receiver performance. The ADSL decoder then decodes the frames. In the FEC decoding block, RS decoding and Turbo decoding depending on the UEP used is done using the look up table again.

For the Turbo code, erasure bits are inserted to the punctured received data using the puncture matrix. The bits are added in the deleted positions to match the positions of the punctured encoded bits at the transmitter.

To further enhance the performance of the system, SDR-based extrinsic information scaling is also applied. The reliability of the soft inputs and outputs of a turbo decoder increases as the number of iterations increases. Conventionally, the turbo decoder uses a fixed set of scaling factors. With the addition of the SDR-based extrinsic information scaling to the turbo decoder, the latter adaptively determines the scaling factor for each data block [10].

The same prioritized feedback mechanism as described in [9] has been used to generate the Negative Acknowledgement (NACK) packet types which allow adaptive multiplexing of frames at the transmitter.

### 3 Simulation Results and Analysis

The performances of the following schemes are compared:

Scheme 1: A two level UEP with RS is used.

Scheme 2: Prioritization with Turbo code is used.

Scheme 3: Prioritization at the DMT and retransmission stages in addition to variable code-rates with RS is used.

Scheme 4: A three level UEP with variable code-rates with Turbo codes is used.

Scheme 5: SDR-based extrinsic information scaling is added to the three level UEP scheme in scheme 5.

In all schemes, DMT Modulation is used with the ADSL system. Near End Cross-Talk (NEXT) and Far End Cross-Talk (FEXT) together with AWGN is used for the channel model. The retransmission limit is set to 1 with 4 users and 49 disturbers. The size of the audio file used is 36 KB with a rate of 44.1 kHz and maximum Segmented Signal to Noise Ratio (SSNR) of 42.27 dB.

Two different interleaver block sizes have been used for the Turbo codes – 672 and 1344. Same has been applied to the RS codes. The same audio file encoded at two different bitrates, namely 128 kbps and 256 kbps have been used. The other parameters are given in Table 1. The system has the same parameters for a bitrate of 256 kbps.

The look up tables used for RS and Turbo codes are as per Tables 2 and 3.

SSNR is being used as a measure as the performance of audio transmission is being investigated in all the schemes. The average gain in SSNR is calculated for the  $E_b/N_0$  range between 0 and 15 dB.

Figure 3 shows the graph of SSNR against  $E_b/N_0$  for the different schemes with the audio file at a bitrate of 128 kbps and block size of 672. The proposed scheme, Scheme 3 provides an average gain of 10.90 dB in SSNR over Scheme 1. Scheme 4 outperforms Scheme 2 by an average of 7.48 dB. Scheme 2 gives an average gain of 3.95 dB over Scheme 1. Scheme 5 exceeds Scheme 4 by an average gain of 0.96 dB in SSNR.

Figure 4 shows the graph of SSNR against  $E_b/N_0$  for the different schemes with the audio file at a bitrate of 128 kbps and block size of 1344. The new scheme, Scheme 3 exceeds Scheme 1 by an average of 11.59 dB in SSNR. Scheme 4 outperforms Scheme 2 by an average of 4.68 dB. Scheme 2 provides an average gain of 8.02 dB over Scheme 1. The average gain of Scheme 5 over Scheme 4 is 0.55 dB.

**Table 1.** Parameters for system using RS and turbo codes.

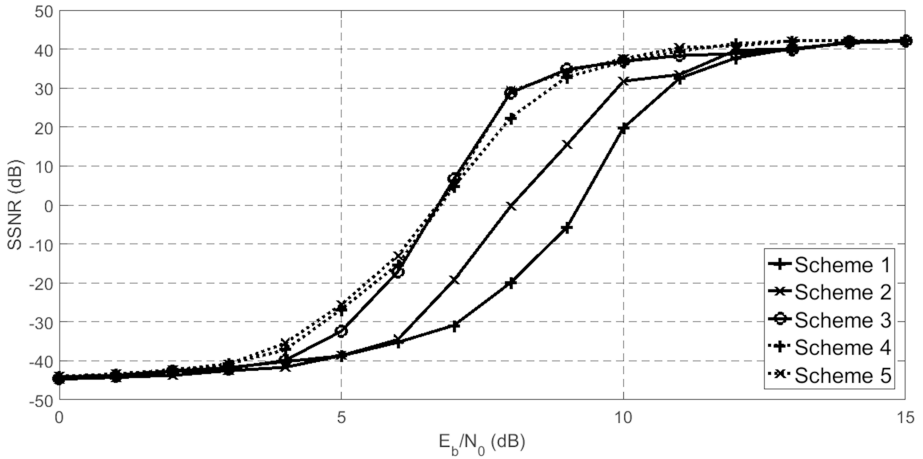
Parameters	RS codes		Turbo codes	
Bitrate/kbps	128		128	
Block size/bits per frame	672	1344	672	1344
Subframes in L1	45	48	45	48
Subframes in L2	66	66	66	66
Subframes in L3	129	132	129	132
Total number of subframes	240	246	240	246
Transmission budget(Total number of subframes × 2)	480	492	480	492
Transmitted and retransmitted subframes from each level	160	164	160	164
Maximum number of transmissions allocated to each subframes in L1	3	3	3	3
Maximum number of transmissions allocated to each subframe in L2	2	2	2	2
Maximum number of transmissions allocated to each subframe in L3	1	1	1	1

**Table 2.** RS code look up table.

$E_b/N_0$	Code rate for L1	Code rate for L2	Code rate for L3
0	0.9231	0.9231	0.9231
1	0.9057	0.9231	0.9412
2	0.9057	0.9231	0.9412
3	0.8421	0.9600	0.9796
4	0.9057	0.9057	0.9600
5	0.8276	0.9600	1.0000
6	0.8000	1.0000	1.0000
7	0.8000	1.0000	1.0000
8	0.8000	1.0000	1.0000
9	0.8000	1.0000	1.0000
10	0.8136	0.9796	1.0000
11	0.8000	1.0000	1.0000
12	0.8000	1.0000	1.0000
13	0.8727	0.9231	0.9796
14	0.8889	0.8889	1.0000
15	0.9057	0.9231	0.9412

Figure 5 shows the graph of SSNR against  $E_b/N_0$  for the different schemes with the audio file at a bitrate of 256 kbps and block size of 672. Scheme 3 yields an average gain of 11.51 dB in SSNR over Scheme 1. Scheme 4 outperforms Scheme 2 by an average of 6.83 dB. Scheme 2 exceeds Scheme 1 by an average of 5.23 dB. Scheme 5 improves upon Scheme 4 and gives an average gain of 0.25 dB.

Figure 6 shows the graph of SSNR against  $E_b/N_0$  for the different schemes with the audio file at a bitrate of 256 kbps and block size of 1344. Scheme 3 provides an



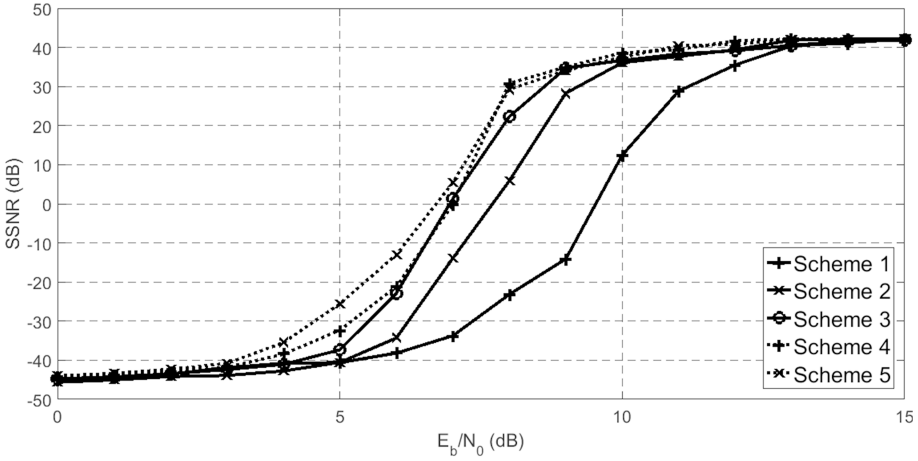
**Fig. 3.** Graph of SSNR against  $E_b/N_0$  for the different schemes with the audio file at a bitrate of 128 kbps and block size of 672

**Table 3.** Turbo codes look up table.

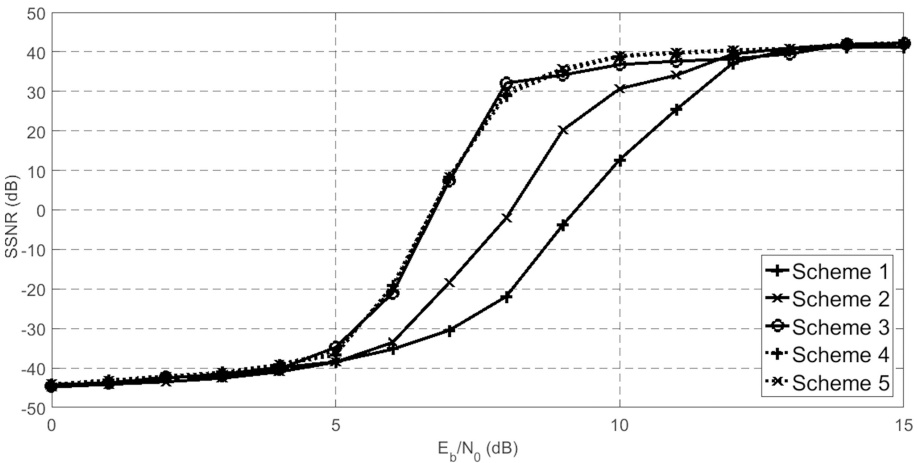
$E_b/N_0$	Code rate for L1	Code rate for L2	Code rate for L3
0	0.8571	0.9412	0.9796
1	0.9057	0.9057	0.9600
2	0.8571	0.9231	1.0000
3	0.8649	0.9505	0.9600
4	0.8649	0.9143	1.0000
5	0.8727	0.9057	1.0000
6	0.8889	0.9320	0.9505
7	0.8727	0.9412	0.9600
8	0.9057	0.9231	0.9412
9	0.8571	0.9600	0.9600
10	0.8889	0.9412	0.9412
11	0.8889	0.9143	0.9697
12	0.8727	0.9231	0.9796
13	0.8727	0.9231	0.9796
14	0.8727	0.9320	0.9697
15	0.8889	0.9057	0.9796

average gain of 12.20 dB in SSNR over Scheme 1. Scheme 4 outperforms Scheme 2 by an average of 5.92 dB. Scheme 2 exceeds Scheme 1 by an average of 7.18 dB. Scheme 5 yields an average gain of 0.59 dB over Scheme 4.

The gain obtained by Scheme 2 over Scheme 1 is mainly due to the change in FEC from RS code to Turbo codes. The addition of the lookup table further increases the SSNR for Schemes 3 and 4. The addition of the SDR-based extrinsic information scaling further enhances the overall system.



**Fig. 4.** Graph of SSNR against  $E_b/N_0$  for the different schemes with the audio file at a bitrate of 128 kbps and block size of 1344.

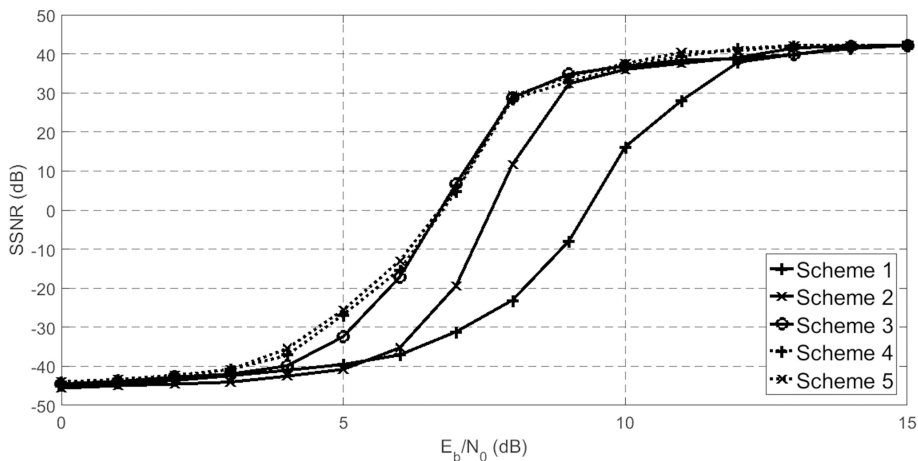


**Fig. 5.** Graph of SSNR against  $E_b/N_0$  for the different schemes with the audio file at a bitrate of 256 kbps and block size of 672.

The average throughput for the different schemes with the audio file at bitrate 128 kbps for both block sizes is given in Table 4. The throughput,  $T_p$  is calculated using following equation as in [9].

$$T_p = \frac{T_L}{T_{L1} + T_{L2} + T_{L3} + N_{ACK}} \tag{1}$$

The parameters of the equation are defined as follows:  $T_L$  is the total number of subframes in each buffer,  $T_{L1}$  is the total number of subframes transmitted for L1,  $T_{L2}$  is



**Fig. 6.** Graph of SSNR against  $E_b/N_0$  for the different schemes with the audio file at a bitrate of 256 kbps and block size of 1344.

**Table 4.** Throughput values.

	Block Size: 672		Block Size: 1344	
	$0 \leq E_b/N_0 < 10$	$10 \leq E_b/N_0 \leq 15$	$0 \leq E_b/N_0 < 10$	$10 \leq E_b/N_0 \leq 15$
Scheme 1	0.4987	0.4989	0.4987	0.4987
Scheme 2	0.4987	0.4990	0.4987	0.4989
Scheme 3	0.4925	0.4929	0.4868	0.4871
Scheme 4	0.4925	0.4929	0.4868	0.4871
Scheme 5	0.4925	0.4929	0.4868	0.4871

the total number of subframes transmitted for L2,  $T_{L3}$  is the total number of subframes transmitted for L3 and  $N_{ACK}$  is the overhead due to the NACK packets.

The same values are obtained for bitrate of 256 kbps. Schemes 2 to 5 have a negligible loss of throughput of about 1.2% with respect to Scheme 1 when the block size is 672. For a block size of 1344, the loss in throughput is about 2.4%.

## 4 Conclusion

In this paper, a three level UEP scheme for audio transmission over ADSL has been proposed. The first two UEP levels consisted of prioritization at the DMT stage with a sophisticated retransmission protocol. The third level consists of the addition of a look up table which significantly enhances the performance for both RS and Turbo codes. The proposed scheme gave a major gain in SSNR. An average gain of over 11 dB between Scheme 1 and 3 was seen with RS codes. For Turbo codes, Scheme 4 outperformed Scheme 2 by an average of over 7 dB. The gain in SSNR obtained was due to the variable rates used with the RS and Turbo codes. The complexity of the transmitter and receiver has increase as a look up table has been added to them. Future systems,



where more traffic needs to be transferred over limited bandwidth, could adopt the proposed scheme to enhance their QoS. As a future work, the proposed scheme could be optimized by optimizing the code rates and puncturing matrix used in Turbo codes.

## References

1. Guide to ADSL and Cable Broadband. <http://www.moneysupermarket.com/broadband/adsl/>. Accessed 07 Aug 2016
2. OECD broadband statistics update. <http://www.oecd.org/internet/broadband-statistics-update.htm>. Accessed 7 Aug 2016
3. Dischinger, M., Haeberlen, A., Gummadi, K.P., Saroiu, S.: Characterizing residential broadband networks. In: Paper presented at the IMC 2007 Internet Measurement Conference, San Diego (2007)
4. Zhang, Q., Kassam, S.A.: Hybrid ARQ with selective combining for fading channels. *IEEE J. Sel. Areas Commun.* **17**(5), 867–874 (1999)
5. Khan, M.A., Moinuddin A.A., Khan, E.: Unequal error protection of embedded video bitstream with optimized FEC. In: Paper presented at 2013 Annual IEEE India Conference (INDICON), India (2013)
6. Naghdinezhad, A., Hashemi, M.R., Fatemi, O.: Unequal error protection for the scalable extension of H.264/AVC using genetic algorithm. In: Sarbazi-Azad, H., Parhami, B., Miremadi, S.-G., Hessabi, S. (eds.) CSICC 2008. CCIS, vol. 6, pp. 194–202. Springer, Heidelberg (2008). doi:10.1007/978-3-540-89985-3\_24
7. Zhou, Z.D., Xu, C.: An improved unequal error protection turbo codes. In: Paper presented at International Conference on Wireless Communications, Networking and Mobile Computing, pp. 284–287 (2005)
8. Lakhdar, A.M., Meliani, R., Kandouci, M.: Research on unequal error protection with punctured turbo codes in jpeg image transmission system. *Serb. J. Electr. Eng.* **4**(1), 95–100 (2007)
9. Fowdur, T.P., Ragpot, P., Soyjaudah, K.M.S.: Enhanced audio transmission over ADSL using prioritised DMT modulation and retransmissions. In: Paper presented at EUROCON 2015 - International Conference on Computer as a Tool (EUROCON). IEEE, Salamanca, Spain (2015)
10. Lin, Y., Hung, W., Lin, W., Chen, T., Lu, E.: An efficient soft-input scaling scheme for turbo decoding. In: Paper Presented at IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing Workshops, Taichung, Taiwan (2006)

# Performance Analysis of Link Adaptation with MIMO and Varying Channel Estimation Schemes

Maryam Imran Sheik Mamode<sup>(✉)</sup> and Tulsi Pawan Fowdur

Department of Electrical and Electronic Engineering,  
University of Mauritius, Réduit, Mauritius  
maryamsheikmamode@hotmail.com, p.fowdur@uom.ac.mu

**Abstract.** The 3GPP (Third Generation Partnership Project) introduced Long Term Evolution (LTE), which is a 4G wireless technology used for broadband, derived from the 2G and 3G data technologies. LTE provides augmented capacity and speed. Among the different methods used for enhancing the performance of an LTE system is link adaptation, whereby the modulation, coding and other signal and protocol parameters are matched to the radio link conditions, leading to an improvement in transmission rates and/or bit error rates. This paper performs a comparative analysis of Link Adaptation with transmit diversity using  $2 \times 2$  Multiple Input Multiple Output (MIMO) as well as different Channel Estimation Techniques. Link Adaptation is performed at the modulation level where a particular modulation index is selected based on the Signal-to-Interference Noise Ratio (SINR) value to obtain the maximum possible data rate for a target BER. Results show that a better performance is achieved using slot average algorithm.

**Keywords:** Link adaptation · MIMO · Channel estimation

## 1 Introduction

The increase in need for mobile broadband services with greater data rates and Quality of Service (QoS) has contributed to the development of the LTE standard defined in 3GPP Releases 8 and 9. Early trials and commercial networks have shown that better mobile broadband user experience was delivered by LTE in real deployments [1]. However, the demand for very high capacity applications such as videoconferencing, multimedia streaming, P2P data sharing, multiplayer games and mobile TV, has taken exponential dimensions and has drastically increased the bandwidth requirements of cellular networks. By 2017, mobile data traffic is predicted to reach 11.2 EB/month (134 EBs annually) [2].

LTE is essentially a radio interface technology which uses Single Carrier-Frequency Division Multiple Access (SC-FDMA) in the uplink and Orthogonal Frequency Division Multiple Access (OFDMA) for the downlink. It also employs Multiple Input Multiple Output (MIMO) techniques to achieve antenna diversity [3, 4]. It supports peak data rates of up to 300Mbps for the downlink and 150 Mbps for the

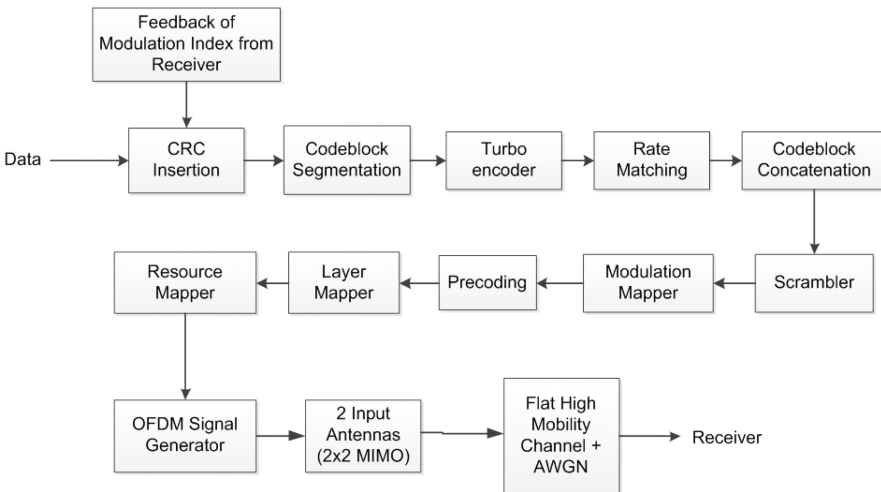
uplink. A key technique for achieving optimal bandwidth and power utilization is link adaptation. It is the process of selecting transmission parameters like modulation and coding (MCS) scheme along with the MIMO transmission rank and precoding, according to channel conditions on resources allotted by the scheduler [5]. The benefit of good link adaptation is also to increase the system capacity and provide a good coverage [6, 7].

Several papers have proposed advanced link adaptation techniques. The authors in [8] had developed a Carrier Aggregation technique employing an AMC scheme to further enhance the spectral efficiency. In [9], the author proposed a channel predictor for MCS selection in link adaptation. In [10], the researchers proposed a novel link adaptation method in order to raise the energy efficiency in addition to the enhanced spectral efficiency. Soft Sphere Detection (SSD) and Zero Forcing receivers are used to analyze CQI feedback in [11]. The authors in [12] have investigated the signal to Noise Ratio (SNR) to CQI mapping in LTE with various transmission modes (Table 1).

In this paper, a comparative analysis of link adaptation based on adaptive modulation index is performed with transmit diversity using  $2 \times 2$  MIMO and three different channel estimation techniques. The Channel estimation schemes used are interpolation, slot average and subframe average. The channel used is Flat High Mobility (FHM). Simulation results revealed that the scheme using slot average algorithm yields better performance than those using other channel estimation algorithms. Furthermore, the link adaptation scheme implemented achieved higher data rate for a target BER (Table 2).

## 2 System Model

The block diagram of a LTE downlink transmitter is shown in Fig. 1 [13].



**Fig. 1.** LTE downlink transmitter

The downlink transmitter consists of the channel coding process, the modulator, the MIMO transmission process, the resource mapping and OFDM modulation step [15]. The Cyclic Redundancy Check (CRC) has the function of detecting errors on transport blocks. The CRC parity bits generated using the whole transport block, are then attached to the latter (Table 3).

The main operation in channel coding is turbo coding, which is defined in the LTE standard. The turbo coding rate implemented in the algorithm is 1/3. Two conventional 8-state constituent encoders are concatenated in parallel, with an interleaver in between, in order to construct the turbo code. The LTE interleaver used is founded upon a Quadratic Polynomial Permutation (QPP) scheme. The following polynomials indicate the trellis structure of the constituent encoder:

$$G_0(z) = 1 + z^{-2} + z^{-3}. \quad (1)$$

$$G_1(z) = 1 + z^{-1} + z^{-3}. \quad (2)$$

The modulation techniques employed in the LTE standard consist of QPSK (Quadrature Phase Shift Keying), 16 QAM (Quadrature Amplitude Modulation), and 64 QAM. Precoding comprises of the adaptation of the signal transmitted to the actual Channel State Information. The precoding matrix is selected from a group of pre-defined matrices which are derived from the number of antenna ports and rank. In case two antenna ports are used for transmission, a codebook of four matrices is stipulated while for four antenna ports, a codebook size of 16 is defined. In layer mapping, data is selected and forwarded into every substream to realize a specific MIMO mode among those described for downlink transmission [13]. Transmit diversity, spatial multiplexing and beamforming form part of the specified MIMO transmission methods in downlink.

The various scenarios for the downlink for LTE Release 8 are reflected in the different transmission modes (TMs) as shown in table below.

**Table 1.** Downlink transmission modes for LTE release 8 [14]

Transmission mode	Downlink transmission system
Mode 1	Single Antenna Port (SISO or MIMO)
Mode 2	Transmit Diversity
Mode 3	Open-Loop Spatial Multiplexing
Mode 4	Closed Loop Spatial Multiplexing
Mode 5	Multi-User MIMO
Mode 6	Closed-Loop Rank-1 Spatial Multiplexing
Mode 7	Single Antenna Port Beamforming
Mode 8	Dual Layer Beamforming

In the Release 9 standard, for the base station as well as for the User Equipment (UE), four antennas can be used. As from Release 10, up to eight antennas can be used in the downlink as shown in Table below.

**Table 2.** Additional downlink transmission modes for LTE release 10

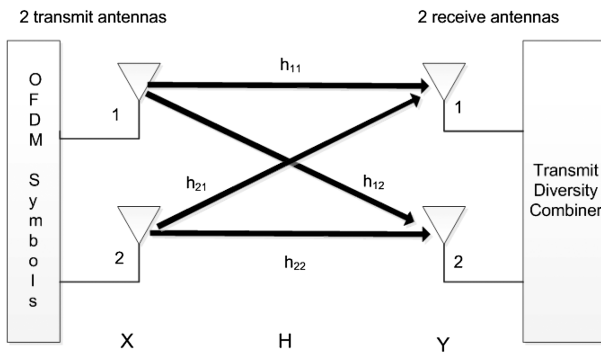
Transmission mode	Downlink transmission system
Mode 9	8 Layer Transmission
Mode 10	8 Layer Transmission

Transmit diversity capitalizes on diversity gains and enhances link quality by using multiple antennas to transmit redundant data [13]. Transmit diversity does not increase the data rate but makes the system more robust against channel fading. Figure 2 illustrates the transmit diversity model. The two key transmit diversity techniques described by LTE are  $2 \times 2$  and  $4 \times 4$  Space Frequency Block Code (SFBC) technique. The two transmit diversity methods utilize one data stream. The receiver obtains the signal  $Y$ , which is the input vector  $X$  multiplied by the transmission matrix  $H$ .

$$Y = H * X. \tag{3}$$

where

$$H = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix}. \tag{4}$$



**Fig. 2.** Transmit diversity model

The components of the resource grid are positioned in spaces defined in the standard by the resource element mapping process. This process is undertaken mainly by constructing indices to the resource grid matrix and allocating different types of information inside the grid. Primary Synchronization Signal (PSS) and Secondary Synchronization Signal (SSS) are found in either subframe 0 or subframe 5 of the six main resource blocks surrounding the DC subcarrier. The Cell-Specific Reference signals (CSRs) are located in symbols 0 and 5 of each slot, and they are separated by six subcarriers in the frequency domain. The generation of the OFDM signal generation takes place on the resource grid. OFDM signal generation consists of performing inverse Fourier transform followed by Cyclic Prefix (CP) addition to yield a stream of OFDM symbols in time and the latter are transmitted through the antennas. Data are gathered in the FFT buffer and reordering is performed in order to remove the DC subcarrier before the IFFT operation.

N last samples from the IFFT output are prepended to the start of the buffer by the CP addition process. In the first OFDM symbol, the value of N is not similar to that in other OFDM symbols. The CP values in the last six symbols of the slot are somewhat smaller than the length of the CP in the first OFDM symbol.

Figure 3 represents the block diagram of a LTE downlink receiver [15].

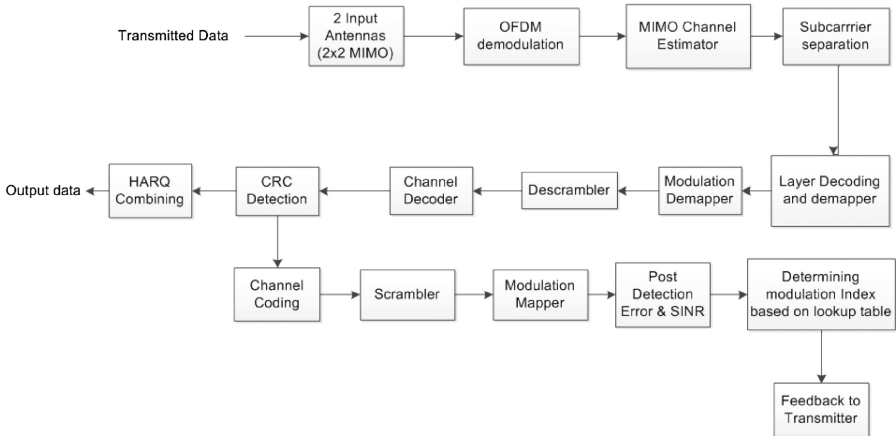


Fig. 3. LTE downlink receiver

The downlink receiver blocks correspond to the transmitter part. They are a group of processing chains with the specified purpose of de-allocating subcarriers from their corresponding frame/subframe, suppressing the reference signals and calculating sub-carrier estimation based on a maximum likelihood method. The downlink receiver also comprises of an OFDM demodulator and a MIMO channel estimator with the task of differentiating the antenna transmissions whose linear combinations have been obtained at the receiver antennas. The function of the channel estimator is to determine the error performance at the receiver. The channel estimation process is performed by analyzing the CSR symbols (pilot symbols) which are located at regular intervals in the resource grid [16]. It can be noted that the CSR signals hold a unique place on a particular antenna port and thus data transmission at these places on a different antenna port is not allowed.

The channel estimation process can be separated into two steps: First, a partial channel frequency response is calculated over elements of the resource grid associated with pilot symbols. Then the gridResponse function is called to expand the channel matrix over the entire resource grid.

The channel frequency responses (CE) estimate is determined as follows:

$$CE = \frac{\text{receive pilots}}{\text{transmitted pilots}} \tag{5}$$

At each subcarrier and at every OFDM symbol in a particular subframe, the process of grid response expansion is carried out. It is implemented by averaging or interpolation.

There are three channel estimate expansion algorithms namely interpolation, average slot and subframe average. In order to execute the expansion process, a group of parameters is needed which consists of the limited set of channel responses calculated over CSR, the resource grid's dimensions, the expansion mode selected as well as a vector specifying the position of the CSR signals relative to the edge of the resource block. The three channel estimate expansion algorithms boast of a dynamic behavior.

The interpolation process is performed with subcarriers in the frequency domain in OFDM symbols that contain CSR signals. A time domain interpolation step is executed on OFDM symbols to determine the channel response over the whole resource grid for subcarriers [13]. An averaging process is performed among OFDM symbols in every slot in the Average Slot expansion scheme, in place of a time domain interpolation. A frequency domain interpolation is carried out after getting the number of subcarriers separating the CSR signals. Lastly, the same channel response is executed on the OFDM symbols of all slots. Thus, the whole resource grid channel response is determined. In the first slot of a particular subframe, the channel response is constant while in the second slot the channel response number is different for all the OFDM symbols [13].

The Subframe Average channel estimate expansion algorithm is performed over a whole subframe, based on frequency domain interpolation and averaging in time among OFDM symbols. It is important to determine if there are two or four OFDM symbols which contain CSR signals in particular subframe, in order to correctly utilize the subframe average algorithm. The CSR signals are averaged in case there are four OFDM symbols. Two symbols are combined if there is only one OFDM symbol per slot. This, for both cases, there is a CSR signal for every three subcarrier. Afterwards, frequency domain interpolation is carried out and the equivalent channel response is executed on all the 14 OFDM symbols. The subframe average algorithm is the least dynamic among the three algorithms mentioned [13].

The original subcarriers are recovered by a layer de-mapper and a layer decoding matrix after the channel estimation. MIMO decoding is carried out and afterwards modulation demapping, descrambling, channel decoding, CRC analysis and HARQ combining is performed to yield the received bitstream.

In order to obtain the link adaptation algorithm, the LTE system was simulated with QPSK, 16 QAM and 64 QAM without any adaptation. The average SINR was computed for a specific BER value corresponding to a given SNR. For the link adaptation scheme, the output obtained from the CRC detection block is encoded, scrambled and modulated. Post-detection error is computed on the data stream obtained and the corresponding SINR is determined. The modulation index is selected based on the SINR obtained.

The algorithm for selection of Modulation Index can be illustrated as follows:

1. If  $\text{SINR} > x$ ,
2. Modulation Index = 2
  3. If  $\text{SINR} > y$ ,
  4. Modulation Index = 3
    5. Else if  $\text{SINR} \leq x$ ,
    6. Modulation Index = 1

Where  $x$  represents the average SINR obtained for a target BER for QPSK and  $y$  represents the average SINR obtained for the same target BER for 16 QAM.

The following lookup table was computed.  $2 \times 2$  MIMO was used with a target BER of  $9 \times 10^{-4}$ .

**Table 3.** Lookup table for link adaptation

Average SINR			
Modulation type	Interpolation	Slot average	Subframe average
QPSK	7.52	8.78	8.46
16 QAM	12.37	12.51	12.15
64 QAM	20.33	17.82	17.42

### 3 Simulation Results and Analysis

The performances of the following schemes are compared:

Scheme 1: In this scheme,  $2 \times 2$  MIMO with transmit diversity is used with Interpolation. The performance is compared using QPSK, 16 QAM, 64 QAM and link adaptation.

Scheme 2: In this scheme,  $2 \times 2$  MIMO with transmit diversity is used with Slot Average. The performance is compared using QPSK, 16 QAM, 64 QAM and link adaptation.

Scheme 3: In this scheme,  $2 \times 2$  MIMO with transmit diversity is used with Subframe Average. The performance is compared using QPSK, 16 QAM, 64 QAM and link adaptation.

The parameters used in all schemes are given in Table 4.

Figures 4 and 5 illustrate the BER and Data Rate achieved with interpolation algorithm. From the BER plot, it can be observed that the link adaptation scheme reaches the target BER of  $9 \times 10^{-4}$  at an SNR of 12 dB. It can also be noticed that the 16 QAM and 64 QAM curves yield higher data rate with greater BER than the QPSK curve at any given SNR. From the Data Rate plot, it can be noticed that the link adaptation scheme achieves higher data rate for the target BER.

**Table 4.** Parameters for schemes

Parameters	
Modulation scheme	QPSK, 16 QAM, 64 QAM
Channel model	Flat High Mobility (FHM)
Signal to Noise Ratio	0–23 dB
Antenna configuration	$2 \times 2$ transmit diversity
Correlation level	Low
Channel bandwidth	10 MHz
Channel estimation	Interpolation, Slot Average, Subframe Average
Maximum number of bits	$5e7$
Code rate	$1/3$



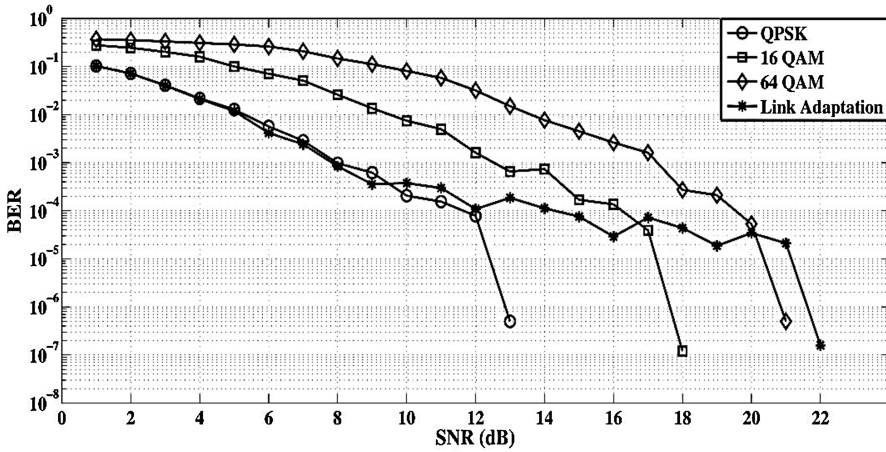


Fig. 4. BER plot for interpolation

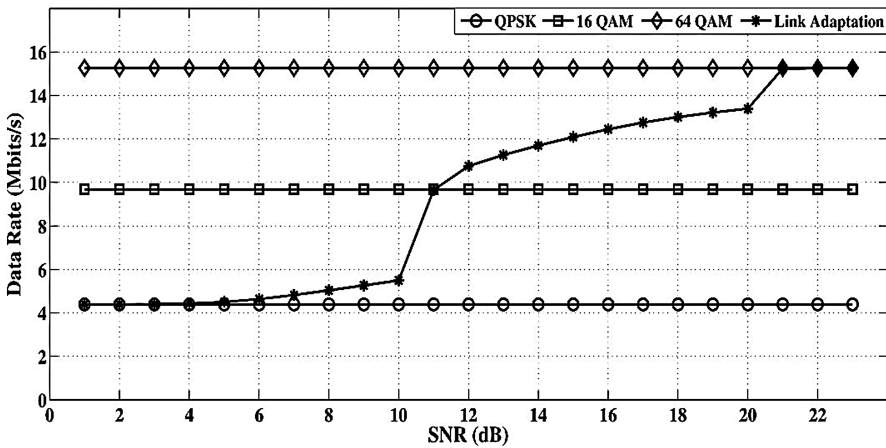


Fig. 5. Data rate plot for interpolation

Figures 6 and 7 show the BER and Data Rate attained using Slot Average Algorithm. From the BER plot, it can be observed that the link adaptation scheme reaches the target BER of  $9 \times 10^{-4}$  at an SNR of 10 dB, thus achieving the target BER at a lower SNR than the Interpolation scheme. Moreover, from the Data Rate plot, it can be noticed that the link adaptation scheme achieves better data rate at lower SNR as compared to that using the Interpolation algorithm.

Figures 8 and 9 shows the BER and Data Rate obtained using Subframe Average Algorithm. From the BER plot, it can be observed that the link adaptation scheme reaches the target BER of  $9 \times 10^{-4}$  at an SNR of 10 dB, but is less stable than the

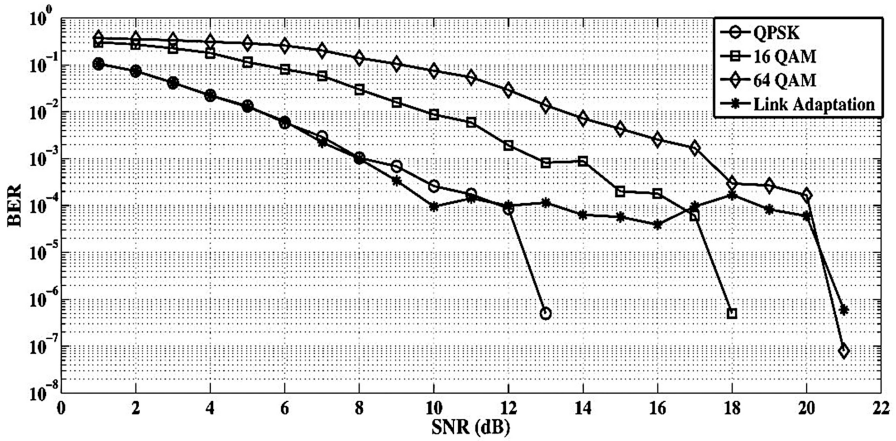


Fig. 6. BER plot for Slot average

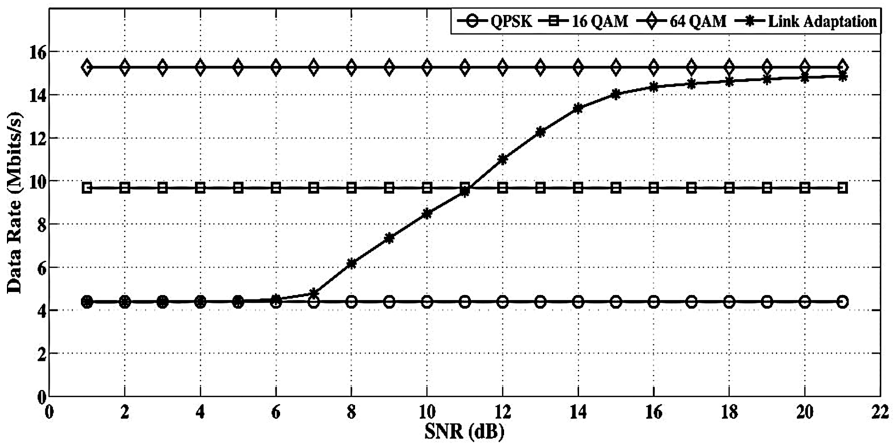


Fig. 7. Data rate plot for slot average

Interpolation and Slot Average algorithms. The Data Rate plot is almost similar to that of the Slot Average algorithm

Figure 10 and 11 show the BER and Data rate plots for all three algorithms. From the BER plot, it can be observed that the slot average algorithm yields a better and more stable performance as compared to the interpolation and subframe average algorithms.

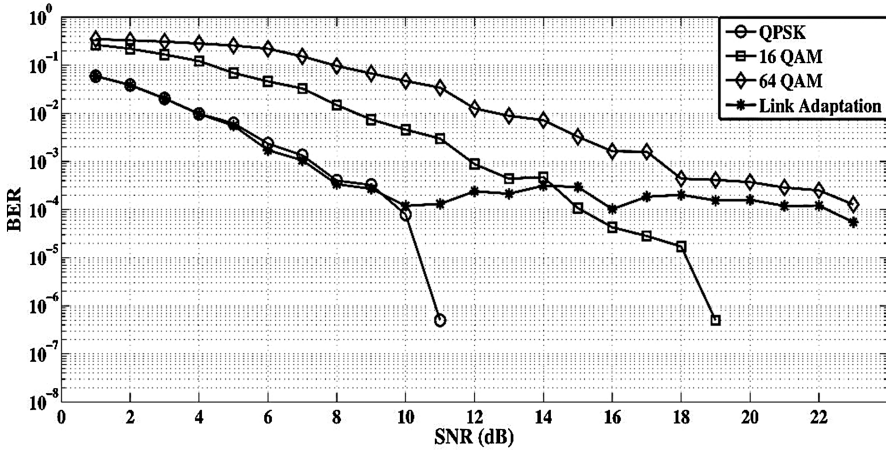


Fig. 8. BER plot for subframe average

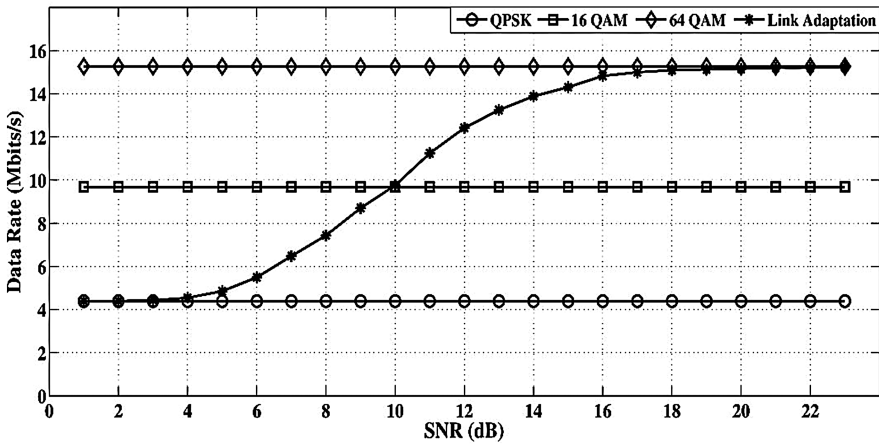


Fig. 9. Data rate plot for subframe average

The data rate plot shows that the Subframe Average algorithm reaches a higher data rate more quickly than the other 2 algorithms, at the expense of lower BER at a higher SNR value as well as a less stable system performance. The slot average algorithm provides a smoother increase in data rate as compared to the interpolation algorithm.

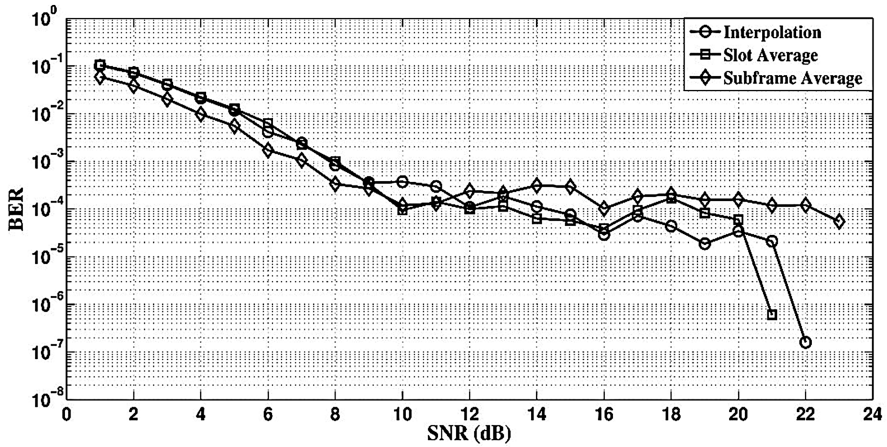


Fig. 10. BER plot for all three algorithms

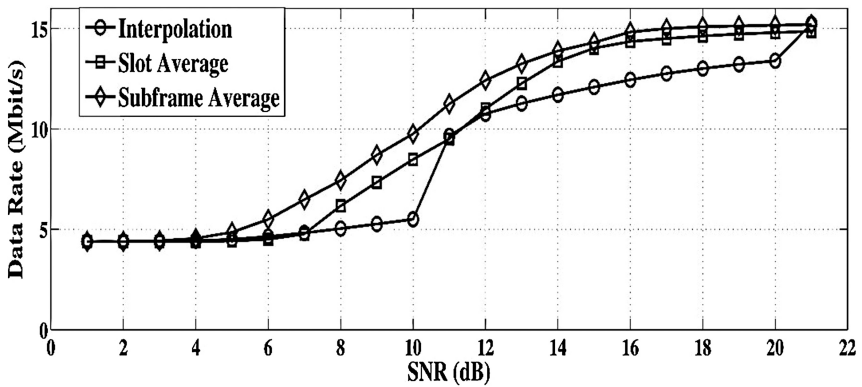


Fig. 11. Data rate plot for all three algorithms

## 4 Conclusion

In this paper, three Channel estimation techniques namely interpolation, slot average and subframe average, have been investigated for LTE systems and a link adaptation scheme has been proposed and simulated with all three algorithms. Results have demonstrated that the link adaptation method used has increased the data rate for a target BER at any given SNR. Furthermore, simulation results have shown that the subframe algorithm is less stable than the other two algorithms and that the slot average channel estimation technique achieves the target BER at a lower SNR than the interpolation algorithm.

## References

1. Nokia-Siemens Networks: LTE-Advanced - The advanced LTE toolbox for more efficient delivery of better user experience. Technical White Paper (2011)
2. Gartner Forecasts 59 Percent Mobile Data Growth Worldwide in 2015 <http://www.gartner.com/newsroom/id/3098617>. Accessed 15 August 2016
3. The race to 5G: Inside the fight for the future of mobile as we know it <http://www.techrepublic.com/article/does-the-world-really-need-5g/>. Accessed 15 August 2016
4. 3GPP: Technical Specifications Rel.8 (2009)
5. Khan, F.: LTE for 4G Mobile Broadband. Cambridge University Press, New York (2009)
6. 3GPP: LTE Release 12 and Beyond (2012). [http://www.3gpp.org/IMG/pdf/lte\\_africa\\_2013\\_3gpp\\_lte\\_release\\_12.pdf](http://www.3gpp.org/IMG/pdf/lte_africa_2013_3gpp_lte_release_12.pdf). Accessed 15 August 2015
7. 3GPP: LTE – Advanced (2013). <http://www.3gpp.org/technologies/keywords-acronyms/97-lte-advanced>. Accessed 15 August 2015
8. Shayea, I., Ismail, M., Nordin, R.: Downlink spectral efficiency evaluation with carrier aggregation in LTE-advanced system employing adaptive modulation and coding schemes. In: IEEE 11th Malaysia International Conference on Communications (2013)
9. Sahlin H.: Channel prediction for link adaptation in LTE uplink. In: IEEE Vehicular Technology Conference (2012)
10. Mustafa, I.S., Chee, K.N., Nor, K.N., Borhanuddin, M.A., Aduwati, S.: CQI-MCS mapping for green LTE downlink transmission. *Asia-Pac. Adv. Netw. J.* **36**(3), 74–82 (2013)
11. Maina, I.: Channel quality indicator feedback in long term evolution (LTE) system. *IOSR J. Electron. Commun. Eng.* **9**(2), 14–19 (2014)
12. Mohammad, T.K., Nafiz, I.B.H., Nayeemul, M.H., Shah, A., Musfiqur, R.: Downlink SNR to CQI mapping for different multiple antenna techniques in LTE. *Int. J. Inf. Electron. Eng.* **2** (5), 757–760 (2012)
13. Zarrinkoub, H.: Understanding LTE with Matlab. Wiley, New York (2014)
14. Maximizing LTE performance through MIMO optimization. <http://rfsolutions.pctel.com/artifacts/MIMOWhitePaperRevB-FINAL.pdf>. Accessed 15 August 2015
15. Ravindra, K., Manohar, S.S., Govindaswamy, U.S.: Long term evolution downlink physical layer simulation in matlab and simulink. In: *International Journal of Future Computer and Communication*, vol. 1(2) (2012)
16. Moghaddam, S.S., Saremi, H.: Joint LS estimation and ML detection. Telecommunication Infrastructure Company, Iran (2011)

# A Context-Aware Mobile Learning System Using Dynamic Content Adaptation for Personalized Learning

Brita Curum<sup>(✉)</sup>, Nigel Chellapermal, and Kavi Kumar Khedo

Faculty of Engineering, University of Mauritius, Reduit, Mauritius  
brita\_curum@hotmail.com, nig\_chels@hotmail.com,  
k.khedo@uom.ac.mu

**Abstract.** Mobile technologies nowadays contribute to greater options to how learners can enhance their learning styles, leaving behind the traditional learning setup. Indeed mobile learning has become an integral part of our everyday life. Mobile learning refers to the use of mobile devices whereby learning is supported independent of the location and activity of the learner. This learning concept considers continual changes of contexts such as locations and the time available to learn within the mobile environment. Given that the current state of the use of contextual information in mobile learning is not extended, this paper put forward a context aware mobile learning system upon which parameters such as intrinsic and extrinsic context are used in order to improve mobile learning experience of the users. This work aims to synthesize a context-aware mobile learning system using dynamic content adaptation to establish personalized learning contents delivery on portable devices.

**Keywords:** Mobile learning · Context-awareness · Content adaptation · Personalized-learning

## 1 Introduction

The rapid proliferation of mobile technology cannot be overlooked. This pervasive technology has brought about considerable changes in the educational sector. The concept of mobile learning is regarded as a tool for students to engage themselves in learning anytime and anywhere. The transparent characteristics of mobile devices have led researchers to focus on its potentials [1]. Hence, the development of quality learning contents to meet the needs of learners is obvious [2].

Context-aware sensing is a key feature of the upcoming technology development, which aims to provide pertinent services and information to end users, based on their situational conditions [3]. Consequently, an attempt to build and deploy efficiently-built context-aware applications using dynamically-generated data obtained from real circumstances raises further investigations. The goal is to develop a resourceful and wiser mobile application without neglecting core features such as contextual data in dynamic environments to bring forward a proactive learning environment.

This paper introduces a context-aware mobile learning system using dynamic content adaptation for personalized learning. In Sect. 2 some existing mobile learning system is presented. Contextual information and adaptability are discussed in Sect. 3. Section 4 describes the architecture and design of the proposed context-aware mobile system. Section 5 of the paper discusses about the results obtained after evaluating context integration combined with content adaptation. Finally, in Sect. 6 some ideas and improvements are discussed to further enhance the system and Sect. 7 draws some concluding remarks in this area.

## 2 Mobile Learning Systems

Mobile learning refers to learning which takes place using portable devices while on the move [4]. The ubiquitous mobile technologies are now widely identifiable not only for their portability but also for the use of contextual information to deliver appropriate learning contents [5]. Below are some examples of mobile learning system already implemented and in use by educational outlets.

WizIQ [6] is an educational platform that allows participation in live classes without constraints of time and location. It provides exclusive features like screen sharing and chat to enhance collaboration between students and teachers and notifications of important deadlines, tests and class schedules are made available to users. Open edX [7] is the open-source version of the Massive Open Online Course (MOOC). Based on the learner's previous knowledge, options are made available to choose the preferred topic. However, it is not fully mobile compatible. This is a serious drawback of the application.

Mobl21 [8] is a mobile learning platform that allows for an unstructured way of learning. Layout presentation of the application across various devices is neat and clean. Therefore, focus from the learner is apparent and meets some of the requirements for cognitive maturity. However, limited exploitation is seen regarding repetitiveness of contents or confused statements to extract cognitive maturity at its best. Docebo [9] is a cloud-based, e-learning solution provider which can be accessed through the mobile browser. In addition to providing online contents to learners, tutors can track records of course contents, view progression of learners and manage reports.

## 3 Context-Awareness and Dynamic Content Adaptation

Context-awareness [10] is any information that can be used to characterize the situation of an entity (person, place or object). Three important characteristics which determine context are where you actually are, who you are with, and what resources are in proximity. A detailed evaluation on context-aware mechanisms to improve cognitive user load showed that technologies requires more work to fully understand human reaction for satisfying interaction [11]. Thus, an investigation on having devices which can sense and acknowledge user context information is essential [11, 12].

A common approach to classify context instances is to distinguish between the context dimensions. These dimensions are referred to as external or extrinsic and

internal or intrinsic [11, 13]. External dimensions refer to context that can be measured by hardware sensors such as light, sound, touch, temperature and many others. Internal dimension is specified by the user or captured by monitoring user interactions, such as the learner's goals, tasks to complete and emotional state. Importance is given to external context instances most of the time [11, 14]. Nonetheless, some efforts to use internal context information have been carried out recently [15]. Information about the user himself, his nearby context entities and adaptation to this context is lagging behind. Table 1 classifies context into four main categories.

**Table 1.** The four main categories of context.

Context category	Description
Learners' context	The learner's identity or a unique identifier from which we can obtain pieces of related information such as user preferences, roles, status, needs, objectives, relationships with peers, and so on
Status or physical context	Properties, which can be perceived by the user. It includes noise level, temperature, lighting, device information, connectivity etc
Time context	Time of the day, day of week, holidays, month, year, date
Location context	Spatial or geographical data. Defines the actual location of the user (in bus, at home, in a public place and so on)

Additionally, context data is important for successful content adaptations in mobile learning systems. Adaptability can be defined as a technique to regenerate a requested content's presentation to adjust to the device capabilities for optimized user experiences [16, 17]. Content adaptations are split into two types which are static and dynamic adaptation. By definition static adaptation pre-processes and stores different versions of the content in contrast to dynamic adaptation where contents are adapted in real-time [18]. Authors [19] stated that there is very little attention on context information and adaptive learning with personalization is missing.

## 4 Proposed Context-Aware Mobile Learning System (WLP)

In this section, we present a context-aware system, Web-Based Learning Platform (WLP) [20] that adapts the learning contents based on the contexts of the user. To perform content adaptation on the system for personalized learning, two important parameters are considered firstly context-awareness such as the user maturity, cognitive load of the learner and device configuration is used to retrieve context data and secondly content adaptation such as resource, device and network adaptation is accountable for suitable adaptation of the learning materials in the application.



## **4.1 Context-Awareness**

### **4.1.1 User Maturity**

The learners' ages are grouped into three categories (junior, 11–17 years old; adult, 18–45 years old; and senior, 46–65 years old). Junior and senior learners are presented with courses which carries basic explanations while adults are presented with explicitly detailed contents. Seniors are presented with only the essential topics of study as their pace in learning reduces as they grow older.

### **4.1.2 Device Configuration**

Responsive design is one important aspect to consider while designing an application since devices on the market are available in different sizes. The learning materials should adapt to the varied device characteristics. Hence, HTML5, CSS3 and JavaScript languages are used in the system to cater for this change in screen resolution and browser compatibility.

### **4.1.3 Cognitive Load**

In an attempt to encourage learner activities through optimal learning curve, cognitive load is one important factor to consider. The higher the cognitive load, the lower is the user satisfaction in learning [21]. In WLP, short quizzes are provided depending on the learner's age group. Course beginners are provided with simple multiple choice questions with "Yes" or "No" options while intermediate learners are provided with more complex questions with the possibility to choose from four options. The difficulty of the questions, depending on the quiz level and the user age group, is increased after each correct answer given by the learner.

## **4.2 Content Adaptation**

### **4.2.1 Resource Adaptation**

The application checks for CPU and RAM usages to adapt itself in such a way that these resources are used at their optimum level in order to reduce the loading time during the learning sessions. Some 3D content is implemented on WLP which is later tested on different CPU and RAM ranges. The difference in the behavior of the system in accordance of the adaptation levels is noted. Low processor and memory makes contents lag during the loading time whereas using a high processor and memory the content loads flawlessly during user interaction.

### **4.2.2 Device Adaptation**

Device adaptation is one major feature for using diverse platforms such as Android or IOS platform. Besides, browsers are considered so that contents from WLP are presented in the desirable format across varied screen resolutions and browsers.

### **4.2.3 Network Adaptation**

WLP used different bandwidth from cable or WIFI connection to ensure that a constant connection speed is kept when browsing the contents. With an internet speed ranging from 0 kbps to 100 kbps, contents are displayed only as text so that users can learn

though at a lower bandwidth while with a speed range of 101 kbps to 250 kbps, content will be textual with some basic images loaded in JPEG format. This will enable users to have some visual content for their learning process. A connection above 1000 kbps (1 Megabyte), results in contents displayed in high definition (HD) quality video (480p or 720p) with textual and pictorial elements. Figure 1 below summarizes the implementation structure of the WLP model.

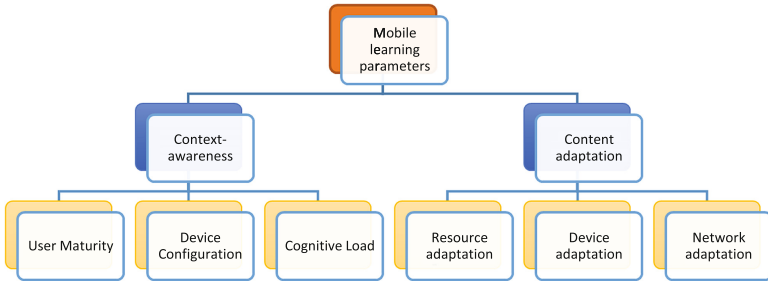


Fig. 1. Implementation structure for the WLP model

4.3 Architecture of WLP

The web system was implemented on PHP’s Lavarel Framework and Angular JS by Google. Its data-binding feature allows for automatic updates of views based on changes that occur through the model when a learner interacts with the system. The WLP platform is attached to a database (MySQL) for information storage. Apache is the web server used to control multiple requests made from the devices using context information. Based on the information obtained, a response is delivered to the learner after identifying the devices’ specifications in terms of CPU, RAM and browser installed. This will provide for the adaptation of the different contents on the mobile devices. The application is portable on Microsoft Windows, IOS and Android operating systems. A content adaptation mechanism is developed so that users are provided with proper contents during their learning process. Figure 2 represents the architecture of the proposed system.

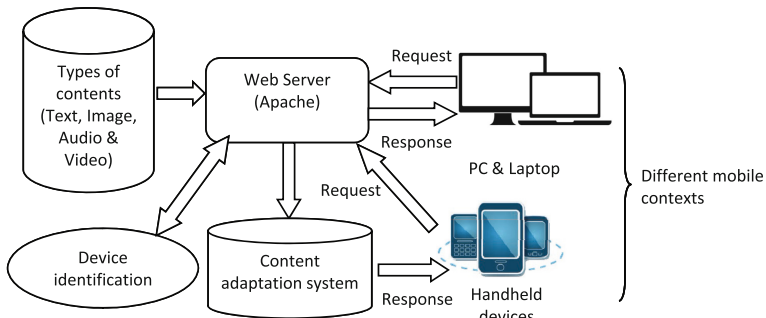


Fig. 2. The system architecture of the WLP

## 5 Results and Discussions

The learning contents are adapted based on the three age group assumed and defined. A back-end section in the system has been implemented to allow administrator of the application to save contents per age groups as shown in Fig. 3.

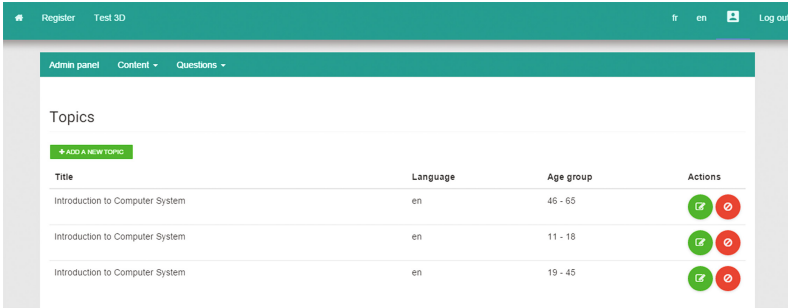


Fig. 3. Different contents saved under the three age groups assumed and defined

Once the user is registered with details including his/her age, he/she is presented with contents appropriate for his/her corresponding maturity. The below screenshot reveals the scenario of a user aged 50. The video loaded is of small length including simple diagrams to transfer the knowledge more conveniently (Fig. 4).

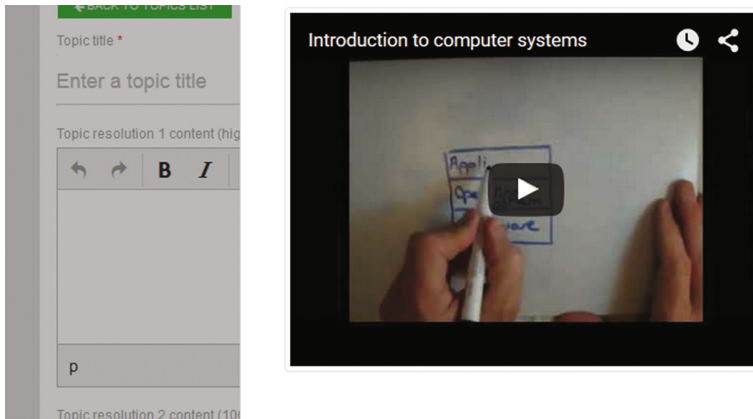


Fig. 4. Short-length video loaded respecting the user maturity factor (50 years of age).

Moreover, the different browsers chosen to test the application revealed good results with WLP. The application loaded with no bugs in its display and thus a fine preview of the contents was made. The layout is well-defined in terms of UX/UI designs and in conformance with W3C standards (Table 2).

**Table 2.** Device configuration result in WLP

Step	Prerequisite	Expected result	Actual result
1	Application tested on Mozilla Firefox	Application layout to fit perfectly the window	Yes
2	Application tested on Google Chrome	Application layout to fit perfectly the window	Yes
3	Application tested on Safari	Application layout to fit perfectly the window	Yes

The simplicity in the quizzes has been ideally composed per age group and level of study to reduce intrinsic load. In this way, the learner's conscious focus stays on the questions and helps in recalling prior knowledge to provide good answers. Depending on the types of resolutions available through the proposed dynamically adapted resources, the system is accustomed to the delivery of adapted contents based on technical issues such as brightness of screens, speed of video loading, lagging of contents and so on. Good use of media queries was made to acknowledge responsive design of the learning materials. In this way, text contents once appeared larger with only crucial information made visible on smaller-screen devices. The use of wireless communication or cable networks displayed adjusted contents based on the availability of the bandwidth. Table 3 represents the scenarios tested to confirm network adaptability.

**Table 3.** Network adaptability result in WLP

Step	Prerequisite (connection speed)	Expected result	Actual result
1	78 kbps	Content is displayed only in textual format	Yes
2	110 kbps	Content is displayed both in textual and pictorial format	Yes
3	350 kbps	Content is displayed in textual format and more images are loaded than in previous connection speed	Yes
4	568 kbps	Content is displayed in low quality video (240p or 360p) with text and images	Yes
5	825 kbps	Content is displayed in low quality video (240p or 360p) with text and images	Yes
6	1207 kbps	Video content loaded in standard quality (480p or 720p) with to text and high quality images	Yes

In addition to the network speed, the quality of content streamed should also be dependent on the processor of the mobile device and the current location of the user. Concepts which are very difficult to explain without the use of graphical representation should be made mandatory even in its lower resolutions on low network bandwidth (less than 100 kbps). These are ideas which have not yet been implemented and will definitely contribute to improve the system.

## 6 Future Works

WLP uses different context information in order to adapt learning contents. Some further evaluation is required to assess the extent of the positive impact on the learner experience. Automatic content adaptation based on the different use cases of user maturity instead of inserting contents manually will be highly in support to a dynamically improved system. Bio sensors can contribute to effective delivery of more personalized data based on the learner's context information. For instance, capturing the eye movement of a user could bring significant amendments while reading. Moreover, the system further needs an in-depth investigation on pedagogical factors such as human behavioral aspects and psychology to reach a higher potential [22]. Related adopted techniques such as the Human Computer Interaction (HCI) or instructional design principles should also be explored. No adaptation is seen based on physical context data such as noise level, light intensity, time and so on. A valid check would be to connect earphones in noisy surroundings while travelling before loading content as audio or video. No detection of earphones would load the contents in text format. The time context can be explicitly defined as per the time of the year. For instance, the amount of information to be tailored during the vacations should focus in reducing the cognitive load for effective study to happen.

## 7 Conclusion

This paper focused on an adaptive approach to output personalized learning contents by adapting to the learners' context data such as their maturity level and to the resource, device, and network contexts. The dynamic experiments conducted revealed so far a good impact on the user's learning experience. It is obvious that the clear presentation achieved in loading the contents across heterogeneous devices would retain the users' attention and allows good readability. Fluctuations in connectivity often interrupt students from learning. This issue is solved by adapting its contents to the flow in network resources that even with low connection the learner is displayed with adequate amount of information to complete his/her study. A revised version of the WLP may be implemented using additional context information which may result into a more flexible learning lifestyle for students.

## References

1. Hwang, G.J., Wu, P.H., Zhuang, Y.Y., Huang, Y.M.: Effects of the inquiry-based mobile learning model on the cognitive load and learning achievement of students. *Interact. Learn. Environ.* **21**(4), 338–354 (2013)
2. Jalil, A., Beer, M., Crowther, P.: Pedagogical requirements for mobile learning: a review on MOBIlearn task model. *J. Interact. Media Educ.* **12**, 1–17 (2015)
3. Chen, H., Finin, T., Joshi, A.: A context broker for building smart meeting rooms. In: *Proceedings of the Knowledge Representation and Ontology for Autonomous Systems Symposium*. AAAI Spring Symposium, (2004)

4. Kukulska-Hulme, A., Traxler, J.: Evaluating mobile learning: reflections on current practice (2005)
5. Walker, K.: Introduction: mapping the landscape of mobile learning. In: Big issues in mobile learning: report of a workshop by the kaleidoscope network of excellence mobile learning initiative. University of Nottingham (2006)
6. Singh, H.: E-learning queen, 21 March 2013. <http://elearnqueen.blogspot.com/2013/03/interview-with-harman-singh-wiziq.html>
7. Harvard, MIT (2012). <https://open.edx.org/about-open-edx>
8. Emantras, Wikipedia (2015). <https://en.wikipedia.org/wiki/Emantras>
9. Erba (2005). <https://en.wikipedia.org/wiki/DoceboLMS>
10. Dey, A., Abowd, G., Salber, D.: A conceptual framework and a toolkit for supporting the rapid prototyping of context-aware applications. *Hum. Comput. Interact.* **16**(2–4), 97–166 (2001)
11. Curum, B., Khedo, K.K.: Improving user cognitive processes in mobile learning platforms through context-awareness. In: International Conference on Computing Communication and Security (ICCCS) (2015)
12. Loa, J.-J., Chana, Y.-C., Yehb, S.-W.: Designing an adaptive web-based learning system based on students' cognitive styles identified online. *Comput. Educ.* **58**(1), 209–222 (2012)
13. Prekop, P., Burnett, M.: Activities, context and ubiquitous computing. *Comput. Commun.* **26**(11), 1168–1176 (2003)
14. Soualah-Alila, F., Mendes, F., Nicolle, C.: A context-based adaptation in mobile learning. *IEEE Comput. Soc. Tech. Committee Learn. Technol. (TCLT)* **15**(4), 5 (2013)
15. Hofer, T., Schwinger, W., Pichler, M., Leonhartsberger, G., Altmann, J.: Context-awareness on mobile devices – the hydrogen approach. In: Proceedings of the 36th Annual Hawaii International Conference on System Sciences, pp. 292–302 (2002)
16. Chen, R.C.S., Yang, S.J.H., Zhang, J.: Enhancing the precision of content analysis in content adaptation using entropy-based fuzzy reasoning. *Expert Syst. Appl.* **37**(8), 5706–5719 (2010)
17. Yang, S.J.H., Zhang, J., Huang, A.F.M.: Applying web page adaptation technique to the augmentation of mobile learning. *J. Res. Pract. Technol. Enhanced Learn.* **3**(3), 253–273 (2008)
18. Fudzee, M.F., Abawajy, J.: A classification for content adaptation system. In: Paper Presented at the Proceedings of the 10th International Conference on Information Integration and Web-based Applications and Services (2008)
19. Zafar, A., Hasan, H.S.: Towards contextual mobile learning. *Int. J. Mod. Educ. Comput. Sci. (IJMECS)* **6**(12), 20–25 (2014)
20. Chellapermal, N.: Content Adaptation on Mobile Learning Platforms for Improved Mobile Learning Experience. Department of Engineering, University of Mauritius (2015)
21. Segall, N., Doolen, T.L., Porter, J.D.: A usability comparison of PDA-based quizzes and paper-and-pencil quizzes. *Comput. Educ.* **45**(4), 417–432 (2005)
22. Wang, T.H.: Developing web-based assessment strategies for facilitating junior high school students to perform self-regulated learning in an e-Learning environment. *Comput. Educ.* **57**(2), 1801–1812 (2011)

# Towards Improving the Security of Low-Interaction Honeypots: Insights from a Comparative Analysis

Abubakar Zakari<sup>1(✉)</sup>, Abdulmalik Ahmad Lawan<sup>1(✉)</sup>,  
and Girish Bekaroo<sup>2</sup>

<sup>1</sup> Department of Computer Science,  
Kano University of Science and Technology, Wudil, Kano, Nigeria  
abubakar.zakari@yahoo.com, aaltofa2000@gmail.com

<sup>2</sup> School of Science and Technology, Middlesex University  
(Mauritius Branch Campus), Vacoas, Mauritius  
g.bekaroo@mdx.ac.mu

**Abstract.** The recent increase in the number of security attacks by cyber-criminals on small businesses meant that security remained a concern for such organizations. In many such cases, detecting the attackers remained a challenge. A common tool to augment existing attack detection mechanisms within networks involves the use of honeypot systems. A fundamental feature of low-interaction honeypots is to be able to lure intruders, but the effectiveness of such systems has nevertheless been affected by various constraints. To be able to secure honeypots systems, it is important to firstly determine its requirements, before taking appropriate actions to ensure that the identified requirements have been achieved. This paper critically examines how existing low-interaction honeypot systems abide to major requirements before recommending how their security could be improved.

**Keywords:** Low-interaction honeypots · Deception in Depth (DiD) · Deceptiveness · Intelligence · Robustness

## 1 Introduction

Even though the massive growth of the internet over the past years provided various benefits to end users and businesses, security remained a concern [1, 2]. It has recently been reported that there has been a huge rise in attacks as cyber-criminals have been targeting small businesses [2]. In many such attacks, detection of the attackers remained a challenge. Although security systems like intrusion detection systems (IDS), firewalls, intrusion prevention systems (IPS) have been existent since many years to enhance the security of networks; various issues were raised with regards to detection of new attacks [3]. A common tool to augment existing attack detection mechanisms within networks is honeypot and by using such systems, new attacks could be uncovered, assault patterns might be revealed, and the precise thought processes of the intruder could be studied [4].

A key purpose of a honeypot is to serve as a decoy used to lure intruders in order to accumulate important information about the intruder and technique of attack that was used to compromise the system. The gathered information could then be used by the organization to trace back the attacker and to also improve its internal defense mechanisms. Honeybot systems can be developed for two reasons purposes, namely, production and research, and can either be of low-interaction or high-interaction [5]. Low-interaction honeypots simulate some portion of the operating system for instance the network stack, while focusing on services that cannot be utilized by the intruder to adventure the real system. This type of honeypot normally implements only the Internet protocols to permit interaction with intruder while making the latter believe the real system is being compromised [6]. On the other hand, high interaction honeypots are complete production similar systems that have a full set of services and permit an intruder a great deal of scope throughout the intrusion. Generally, high interaction honeypots are challenging to recognize and are costly to maintain [7]. The costs of high interaction honeypots is a barrier to their adoption by small businesses and also most firms do not need high-interaction honeypots that captures massive amounts of data [8].

A fundamental feature of low-interaction honeypots is to be able to lure intruders, but the effectiveness of such systems has nevertheless been affected by various constraints. In the past, attackers have been using OS fingerprint techniques such as NMAP, Xprobe to remotely attack and distinguish honeypot from a real system [9]. Moreover, data security experts have been increasing focus on defensive strategies while neglecting offensive strategies [5] thereby increasing its vulnerability to intruders. If a honeypot is discovered by an intruder, its purpose is defeated and no advantage is provided to the organization adopting it. As such, improvement of existing honeypot systems is needed. To be able to secure network systems in general, it is important to firstly determine its requirements, before taking appropriate actions to ensure that the identified requirements have been achieved [10].

In terms of related work, although various studies focused on improving honeypot systems against newly identified vulnerabilities and attacks, limited published literature is available on the comparison related to how existing low-interaction honeypots adhere to their security requirements. As such, this paper critically examines how existing low-interaction honeypot systems abide to key identified requirements before recommending how their security could be improved. The analysis and recommendations provided in this paper could be used by researchers and experts in their endeavor to improve the security of low-interaction honeypot systems to eventually benefit businesses. The paper is structured in the following manner: In Sect. 2, the security requirements of honeypots are investigated before reviewing the existing honeypot systems in Sect. 3. Results from a comparative analysis is provided in Sect. 4 followed by recommendations on how the security of low-interaction honeypots could be improved in Sect. 5. Finally, the work is concluded in Sect. 6.



## 2 Key Requirements of Honeypots

A major security requirement of honeypot systems is its deceptiveness [11, 12]. Deceptiveness involves obscuring valuable data in bland-looking files, and set up honeypots that divert attackers from the real assets whereby leading them to false intellectual property, or causing them to trip alarms [13, 14]. In short, deception involves misleading attacker into believing something that is false. Among the deceptive techniques, camouflage involves disguising the network infrastructure by making it a moving target, changing addresses, infrastructure topologies, and available resources daily. In other words, camouflaging take steps to prevent attackers from seeing the same infrastructure twice [15]. Disinformation is another process which involves diverting or confusing attackers with false information [14]. In this process as well the hacker is supplied with fake successes, responses, files, and assets to exploit. Also, disinformation poised that any false information given must not be easily disprovable. Moreover, work has also been done to categorize the sophistication of deceptive discipline into different levels, namely, static, dynamic and adaptive deception [16]. Static deception has been referred as constant execution of an often uncontrollable trait whereas dynamic deception is implemented upon activated response to some stimuli. Adaptive deception in turn adjusts and reacts to a situation while also and employing cognitive assessment before, during and after the fact.

Another essential requirement of honeypots systems are their robustness and fault tolerance [6, 17]. A system is said to be fault-tolerant if it is able to automatically recover from errors or faults while also being able to eradicate faults without suffering from an externally perceivable failure [18]. This essential ability ensures that honeypots are able to recuperate while at the same time guarantees robustness as the system is able to also cope with errors during execution [6].

Furthermore, intelligence of honeypot systems is also important. Intelligence enables honeypots to gain actionable insights by gathering threat intelligence feeds and adversary indicators that define and describe trends, tendencies, methods, and actions taken by attackers [19]. Intelligent honeypot pretend to surrender to one form of attack in order to suppress a second, less-obvious defense. With series of attacks on different levels of relevance and context, intelligent honeypot continue to ramp up their threat intelligence capabilities while increasing their effectiveness with regards to intelligence-led deceptions.

## 3 Low-Interaction Honeypot Systems

Different low-interaction honeypot systems have been proposed and this section reviews the common ones.

### 3.1 Honeyahole

Principally designed and developed to escape from honeypot hunting, honeyahole implements three phases, namely, collection, redirection and deception in order to

gather four types of attacking information to build up the blacklist [5]. The honeypot has two redirection techniques embedded to dynamically send incoming traffic to a production or a deception server in the same redirection phase.

### 3.2 Honeywall

Honeywall is a honeypot that helps in deploying honeynet with ease by automating the process of deployment [1]. This honeypot can also capture and analyze traffic (both inbound and outbound) of honeynet activity. An identified vulnerability of this honeypot involves construction of a traffic stream that consists of strings matching snort\_inline's rewriting database before verifying whether all packets are received unmodified [20].

### 3.3 Honeyd

Honeyd is a low-interaction, open source honeypot, and can be deployed in various platforms (Windows, UNIX) [21]. It can emulate operating systems at TCP/IP stack level and also monitor all UDP and TCP based ports (as shown in Fig. 1). A few vulnerabilities of this type of honeypot were also found. Studies showed that Honeyd can be detected remotely using fingerprint attacks [22, 23] and using timing analysis of ICMP ECHO request [4]. In an attempt to improve the identified limitations, work has been done to create a new camouflaged Honeyd by modifying the original honeypot in addition to the underlying operating system support in order to permit high-fidelity emulation of events [24].

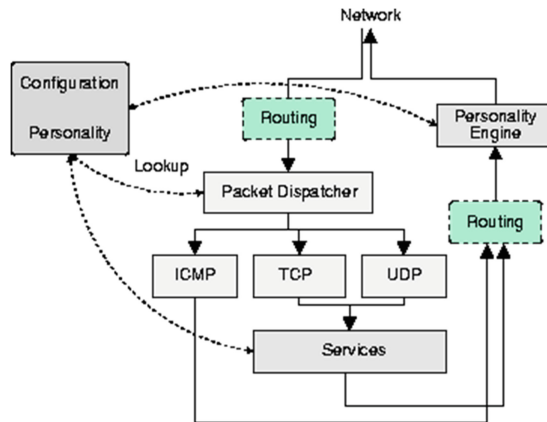


Fig. 1. Honeyd infrastructure [24]

### 3.4 Honeytrap

Being a low-interaction honeypot that operates by observing attacks against network services, honeytrap aims at collecting malware in an automated manner [13]. It permits the collection of traffic information for pre-closed ports by opening them in which an access is observed. As a limitation, this honeypot is not able to capture details pertaining to the activities of an attacker unless a second attempt is made to the same honeypot [25].

### 3.5 Nepenthes

As a honeypot used for malware collection, Nepenthes inherits the scalability of low-interaction honeypots and its flexibility enables it to emulate vulnerabilities of different operating systems on a single machine during a single attack [26]. In terms of limitations, Nepenthes is only capable to collect autonomously propagating malware and that malicious software which spread using hitlist to find vulnerable systems are hard to detect [27]. Furthermore, since Nepenthes emulates huge number of vulnerabilities, it makes it easy to detect by attacker, because many TCP ports are open in the process [27].

## 4 Comparative Analysis

Literature analysis reveals that the focus of the different reviewed low-interaction honeypots vary in terms of characteristics and abilities. For Honeyd, detection using OS fingerprinting attacks reduces its deceptiveness, camouflaging and robustness capabilities [22, 23], although it has the ability to emulate sensitive features of operating system and gather vital information of attacker as well, thus illustrating intelligence [23]. On the other hand, Honeywall depicts robustness and intelligence due to its ability to capture and analyze traffic (both inbound and out bound) of honeynet activity [19]. Moreover, this honeypot has also been portrayed as intelligent as it is able to deceive and link the attacker to the honeypot system [20]. Honeytrap, in turn, does not collect details of attacker on the first attempt the attack is made but rather relies on further attack attempts to be able to correctly detect the attacker which is a big disadvantage in terms of intelligence [25]. Honeyhole is basically designed to escape honeypot hunting with focus on camouflaging, deceptiveness and robustness [5]. Finally, although Nepenthes is portrayed as a highly deceptive honeypot, it has different vulnerabilities thus making it easy to detect by an attacker [27]. Moreover, due to its inability to detect malwares that propagates using a hitlist makes its intelligence and robustness undermined [27]. Comparisons are summarized in Table 1 to show how existing low-interaction honeypot systems have been portrayed by literature.

From Table 1, it could be seen that deceptiveness has been the major focus of the reviewed low-interaction honeypots towards misleading attackers. The five honeypots showed to have this capability although a few vulnerabilities have been highlighted especially relating to the use of Honeyd and Nepenthes. Robustness, which relates to coping with errors during execution, was identified as an important requirement of

**Table 1.** Comparative analysis

Item	Deceptiveness	Robustness	Intelligence
Honeyd	✓	✗	✓
Honeywall	✓	✓	✓
Honeytrap	✓	✓	✓
Honeyanole	✓	✓	✗
Nepenthes	✓	✗	✗

honeypots. However, literature showed that this requirement has not been a major focus of a few honeypots. Among the three requirements, intelligence seems to be the least focused aspect of low-interaction honeypots. Among the 5 honeypots that were compared, 3 of them did not portray intelligence abilities, which is important to better deceive attackers while also accurately obtaining their details.

## 5 Recommendations

Results showed that the key requirements needing attention are robustness and intelligence of low-interaction honeypot systems. Robustness can be improved using redundancy or collaborative honeypot systems such that in case one of them fails, others remain operational. Faults within existing honeypots could also be isolated by further testing such systems. On the other hand, different works have been conducted to improve the intelligence of low-interaction honeypots. First of all, honeypot systems can embed intelligence by learning the moves of attacker in addition to tools used to compromise systems [28]. Also, honeypots systems can be made dynamic whereby having the capability to learn about network environments and infrastructures before autonomously deploying individual honeypots based on current layout [19]. Furthermore, after deployment, such systems should be able to repeatedly monitor network changes and update configurations accordingly [19]. Additionally, the Deception-in-Depth (DiD) concept of operation could be utilized [29]. DiD utilizes the layering approach with three different layers in the proposed model aimed at strengthening the defense of honeypot systems [30]. Within the model, the honeypot asset is represented in the innermost layer whereas the honeypot is positioned in the middle layer of the model. The purpose of the outermost layer is to improve deception using techniques including fake access points.

## 6 Conclusions

This paper examined how existing low-interaction honeypot systems abide to their requirements before recommending how their security could be improved. Three important requirements of honeypot systems were identified namely deceptiveness, robustness and intelligence. Among these requirements, existing low-interaction honeypots seem to focus on deceptiveness with reduced attention given to their robustness and intelligence. As such, more work is needed towards improving

robustness and intelligence of low-interaction honeypot systems so as to improve their overall effectiveness. As future work, the proposed recommendations could be practically evaluated to assess their effectiveness. Moreover, a framework could be proposed focusing on the three requirements investigated in this study to help businesses in their endeavor to prevent attackers from detecting, exploiting, and deceiving honeypot systems and assets.

## References

1. Chakrabarti, A., Manimaran, G.: Internet infrastructure security: a taxonomy. *IEEE Netw.* **16** (6), 13–21 (2002)
2. Tiwari, R., Jain, A.: Design and analysis of distributed honeypot system. *Int. J. Comput. Appl.* **55**(13), 20–23 (2012)
3. Smith, M.: Huge rise in hack attacks as cyber-criminals target small businesses (2016). <https://www.theguardian.com/small-business-network/2016/feb/08/huge-rise-hack-attacks-cyber-criminals-target-small-businesses>. Accessed 20 Aug 2016
4. Yang, Y., Yang, H., Mi, J.: Design of distributed honeypot system based on intrusion tracking. In: *IEEE 3rd International Conference on Communication Software and Networks (ICCSN)*, pp. 196–198. IEEE (2011)
5. Mulkamala, S., Yendrapalli, K., Basnet, R., Shankarapani, M.K., Sung, A.H.: Detection of virtual environments and low interaction honeypots. In: *Information Assurance and Security Workshop, IAW 2007. IEEE SMC*, pp. 92–98. IEEE (2007)
6. Shiue, L., Kao, S.: Countermeasure for detection of honeypot deployment. In: *IEEE International Conference on Computer and Communication Engineering, ICCCE 2008*, pp. 595–599. IEEE (2008)
7. Mohammadi, S., Nikkahan, B.: A fault tolerance honeypots network for securing E-government. In: *IEEE International e-Conference on Advanced Science and Technology, AST 2009*, pp. 13–17. IEEE (2009)
8. Defibaugh-Chavez, P., Veeraghattam, R., Kannappa, M., Mulkamala, S., Sung, A.H.: Network based detection of virtual environments and low interaction honeypots. In: *2006 IEEE Information Assurance Workshop*, pp. 283–289. IEEE (2006)
9. Brown, B.: How to make a honeypot network security system pay off (2007). <http://www.networkworld.com/article/2296754/lan-wan/how-to-make-a-honeypot-network-security-system-pay-off.html>. Accessed 10 Aug 2016
10. Valli, C.: Honeyd-A OS fingerprinting artifice. In: *Proceedings of 1st Australian Computer Network and Information Forensics Conference* (2003)
11. Bishop, M.: What is computer security? *IEEE Secur. Priv.* **1**(1), 67–69 (2003)
12. Cohen, F.: The use of deception techniques: honeypots and decoys. *Handb. Inf. Secur.* **3**(1), 646–655 (2006)
13. Zhang, F., Zhou, S., Qin, Z., Liu, J.: Honeypot: a supplemented active defense system for network security. In: *Proceedings of the Fourth International Conference on Parallel and Distributed Computing, Applications and Technologies, PDCAT 2003*, pp. 231–235 (2003)
14. Provos, N.: A virtual honeypot framework. In: *USENIX Security Symposium*, vol. 173, pp. 1–14 (2004)
15. Rowe, N.: Deception in defense of computer systems from cyber attack. In: *Cyber Warfare and Cyber Terrorism* (2008)

16. Fu, X., Yu, W., Cheng, D., Tan, X., Streff, K., Graham, S.: On recognizing virtual honeypots and countermeasures. In: 2006 2nd IEEE International Symposium on Dependable, Autonomic and Secure Computing, pp. 211–218. IEEE (2006)
17. Yek, S., Australia, W.: Measuring the effectiveness of deception in a wireless honeypot. In: Australian Computer, Network and Information Forensics Conference (2003)
18. Nikkhahan, B., Aghdam, A., Sohrabi, S.: E-government security: a honeynet approach. *Int. J. Adv. Sci. Technol.* **5**, 75–84 (2009)
19. Avizienis, A., Kelly, J.: Fault tolerance by design diversity: concepts and experiments. *Computer* **17**(8), 67–80 (1984)
20. Zakaria, W., Kiah, M.: A review on artificial intelligence techniques for developing intelligent honeypot. In: 2012 8th International Conference on Computing Technology and Information Management (ICCM), pp. 696–701. IEEE (2012)
21. Provos, N., Holz, T.: *Virtual Honeypots: From Botnet Tracking to Intrusion Detection*. Pearson Education, Upper Saddle River (2007)
22. Krutz, R., Vines, R.: *The CEH Prep Guide: The Comprehensive Guide to Certified Ethical Hacking* (With CD). Wiley, Hoboken (2007)
23. NOSTROMO: Techniques in OS-Fingerprinting. Hagenberg (2005)
24. Boyle, A.: *A Remote OS Identification Primer*. SANS (2001)
25. Fu, X., Graham, B., Cheng, D., Bettati, R., Zhao, W.: Camouflaging virtual honeypots. Texas A&M University (2005)
26. Song, J., Takakura, H., Okabe, Y.: Cooperation of intelligent honeypots to detect unknown malicious codes. In: IEEE WOMBAT Workshop on Information Security Threats Data Collection and Sharing, WISTDCS 2008. IEEE (2008)
27. Kumar, S., Sehgal, R., Singh, P., Chaudhary, A.: Nepenthes honeypots based botnet detection. *J. Adv Inf. Technol.* **3**(4), 215–221 (2012)
28. Baecher, P., Koetter, M., Holz, T., Dornseif, M., Freiling, F.: The nepenthes platform: an efficient approach to collect malware. In: International Workshop on Recent Advances in Intrusion Detection (2006)
29. Gupta, N.: Improving the effectiveness of deceptive honeynets through an empirical learning approach. In: 3rd Australian Information Warfare and Security Conference (2002)
30. Yek, S.: Implementing network defence using deception in a wireless honeypot. In: Australian Computer, Network and Information Forensics Conference (2004)
31. Gerwehr, S., Anderson, R.: Employing deception in INFOSEC (2000). <http://www.cert.org/research/isw/isw2000/papers/26.pdf>. Accessed 10 Aug 2016

# Implementation of Driver Drowsiness Detection Application in Mauritius

Karishma Ramodhine<sup>(✉)</sup> and Shireen Panchoo

Department of Innovative Technologies and Engineering,  
University of Technology, La Tour Koenig, Port Louis, Mauritius  
karishmaramodhine@gmail.com,  
s.panchoo@umail.utm.ac.mu

**Abstract.** Driver drowsiness has become a major cause of traffic accidents. As a matter of fact, studies show that around one quarter of all serious motorway accidents are attributable to sleepy driver. One of the common ways to detect sleepy driver is by tracking the eye region since the major symptoms of fatigue such as frequent eye blinks and difficulty keeping eyes open, are visible through the driver's eyes. Hence, it has become a necessity to proffer an investigation about the development of a mobile application system to detect the driver's tiredness through Real-Time Eye. Naïve Bayes classifier algorithm and an alarm system have been implemented to proactively predict if there is any sign of drowsiness symptoms. The assessment of a population of 25 people were assessed which shows young people are mostly vulnerable to falling asleep on the wheel and that the notification alarm has also proved to be effective.

**Keywords:** Real-time eye detection · OpenCV · Drowsiness · Prediction · Image processing · Naive Bayes classifier

## 1 Introduction

Driver drowsiness refers to a progressive diminishing alertness, vigilance and concentration or attention on the road and, traffic demands and as a result, there is a large number of vehicle accidents. Drowsiness remains an inevitable state of mind resulting from sleep. By taking into consideration the analysis, these numbers are underestimated and up to 6,000 fatal crashes each year may be caused by drowsy drivers [1, 2]. The development of technologies for detecting or preventing drowsiness at the wheel are major challenges in the field of accident avoidance system [3].

Drowsiness can harm the driver either mentally or physically. However, if the tired and sleep deprived driver with frequent eye blinking or repeatedly yawning reactions, drowsiness can prove to be deadly. However, driving under drowsiness influence can reduce vigilance leading to high speed crashes as well as decreases the correctness of decision-making [4]. Usually, the vehicle operator needs to either have a pre-drive nap or a mid-drive nap or have an interacting partner, in order not to be a victim of the tragedy of driving drowsy. Today many applications such as facial expression recognition, human-computer interaction and driver fatigue detection systems, are using eye states as inputs [5] for detecting drowsiness.

With the lack of mobile software to predict drowsiness in Mauritius, techniques need to be devised to counteract its affects [6] due to the constant growth in the number of drowsiness's victims on wheels. Hence, the aim of this paper is to develop a mobile application that will predict the level of driver's drowsiness through historical data. The real-time eye detection system is based on real-time acquisition of the driver's face images and template matching method is applied to extract the eye's stress level and activates an alarm in case drowsiness symptoms are detected. The frequencies of eye blinks on that particular day and time are captured and stored for computation of drowsiness' prediction.

## 2 Literature Review

Drowsiness is the intermediate state between alert wakefulness and sleep [7] which results from persistent sleepiness accompanied by lethargy, weakness, and lack of mental agility. It might indicate the sleeping disorders or the medical disorders of a person of which sleep apnea, insomnia, and narcolepsy are well-known. The latter can alternate rapidly between different levels within a matter of seconds [8]. By employing these kinds of applications in academic and commercial fields, eye state detection has been a heated debate since past decade and has drawn great attention among researchers [5]. Face monitoring systems can be classified into two main approaches. In the first one, the fatigue and distraction of the driver are detected by solely processing the eye region where parameters such as iris, eye perimeter, eye edge and eye lids are being employed. The eyes are the most significant region of the face as symptoms of fatigue and distraction are visible [9]. In the second approach, detection of symptoms of fatigue and distraction are not made only from the eyes but from other regions of the entire face [9] such as symptoms related to the eye region, mouth region and head which includes other symptoms such as yawning, head nodding and head orientation [10].

### 2.1 Algorithm for Detection of Eye Fatigue and Distraction

Different algorithm can be used with different techniques such as face detection algorithm which is used to identify the location and scale of all the faces in the image [11], template matching algorithm implementing Eigen spaces with the help of image-processing and viola jones algorithm working on OpenCV. One of the algorithm is PERCLOS, which is a measure of total time where the person's eyes are 80% or more closed and shows slow eyelid closures instead of blinks [12–15]. The eyelid distance and eye blink rate of a drowsy person differ from those of a normal person. The number of blinks detected in 30 s is called the blink frequency. A blink action will be detected as a transition among 3 states: closing, closed and opening. According to prior researches the most appropriate technique is to use the drivers' physiological characteristics as a mean of detecting drivers' drowsiness [16] and we have also seen that drowsiness can be detected by monitoring drivers' head movement, eye blink rate, yawning frequency by using mobile phone camera or a mobile stand [17]. Moreover, there exist many techniques, used by researchers to detect drivers' drowsiness which can be sensing of driver operation and vehicle response, sensing and monitoring the



physiological characteristics and driver’s response [18, 19]. The PERCLOS formula is illustrated below.

$$PERCLOS = \frac{N_{blink}}{30 \times S} \tag{1}$$

Where N blink, the number of eye blinks in the recent 30 s time window and S, the sampling rate.

Naive Bayes algorithm is a classification technique that can be used to proactively predict drowsiness since it is easy and fast to predict class of test data set and useful for very large data sets. It uses Bayes’ Theorem, a formula that calculates the probability of an event occurring given the probability of another event that has already occurred. If B represents the dependent event and A represents the prior event, Bayes’ Theorem can be stated as follows:

$$P(\mathbf{B}|\mathbf{A}) = \frac{P(\mathbf{A}|\mathbf{B})P(\mathbf{B})}{P(\mathbf{A})} \tag{2}$$

### 2.2 Assessment of Prior Software

Drowsiness is an increasing issue, thus, an analysis was carried out to identify different software with their respective functionalities (Table 1).

**Table 1.** Functionalities of existing eye software

Functionalities							
Prior software	Anti-sleep driver	Stay awake	Drive awake	Anti drowse	Driver alarm	Stay awake Pro	Coffewake
Blue light	√	×	×	×	×	×	×
Reset button	×	√	×	×	×	×	×
Eye detection & redirection	×	×	√	×	×	×	×
Noises at intervals of time	×	×	×	√	×	×	×
Eye detection & monitoring with alarm	×	×	×	×	√	×	×
Repeat words with alarm	×	×	×	×	×	√	×
Burst like vibrations	×	×	×	×	×	×	√

Based on the above analysis, we learnt that there is an inadequacy of software with limited functions. For instance, “Blue Light” software has been developed for anti-sleep drivers and “Eye Detection & Monitoring with alarm” has been developed for driver’s alarm. Thus, the need to predict drowsiness and waking up the driver is significant.

### 2.3 Prototype of Fatigue Detection Tools

OpenCV is basically a library of functions written in C/C++ [20]. The program only requires around 70 mB of RAM to run in real-time and it is free of cost [20]. Matlab and OpenCV support Windows, Linux, MacOS [21] and any device that can run on C/C++. OpenCV involves more coding lines for similar function that is written in MATLAB. Matlab is built on Java which is built upon C [20, 22]. However, Matlab is time consuming to interpret all the code to turn it into Java and then finally executes the code. E.g. 3–4 frames analysed per second. Also, it requires over a gig of RAM to run through video [23] as it uses a lot of system resources. EmguCV is another cross-platform “NET” wrapper to the OpenCV image processing library [24]. It can be launched on Mono platforms such as Android, Mac OS X, IOS, LINUX PC. EmguCV improves the ease of use and is readable which shows a clear presentation of the C# code and is considered to be a better image processing library [25].

## 3 Methodology

To develop the mobile application to allow drowsiness prediction, Android Studio and OpenCV (Open Source Computer Vision) were used due to the performance speed, that is, it allows a fast execution of codes and more image processing where at least 30 fps result in real-time detection. While MATLAB development is restricted to its own environment, and supports many Integrated Development Environment (IDE), OpenCV fits well the requirements for development of the android application. The proposed system basically works on three modules namely eye detection, fatigue detection and drowsiness prediction. Each module is responsible for a specific operation on data. The detail information about modules is evaluated in Fig. 1.

In the proposed mobile application, the eye detection module is responsible for capturing real-time images of the eye region and process the inputs using `haarcascade_eye_tree_eyeglasses.xml` classifier. In the first stage, the original color information of the detected eye region is converted into gray scale because the value of pixels in that region is relatively lower than the other face’s region. The classifier computes the mean gray value along X axis and roughly locates the eye region. Secondly, the fatigue detection module detects whether the eye(s) is closed or is blinking. If driver closes the eyes, then the system generates an alarm and stores the rate or frequencies of eye(s) blinking or closure at that specific day and time. Finally, for a given period of time, the chance that the driver could fall asleep on wheels can be predicted in the drowsiness prediction module. The module computes this probability by using Naïve Bayes Theorem. The android application scans the human eyes in real-time and records the eyes blink frequencies in an external file. In case one or both eyes are closed, an alarm sound is triggered. Thus, the android application is able to predict the probability that the driver will be drowsy on the present day and time. With the aim of implementing an application to detect drowsiness, a population of 25 licensed drivers in Mauritius aged between 18–29, 30–50 and 51–70 were assessed through a questionnaire.

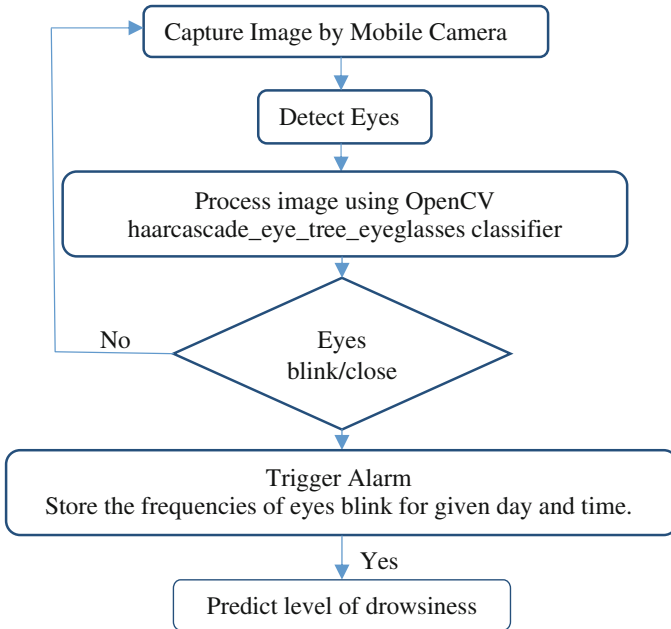
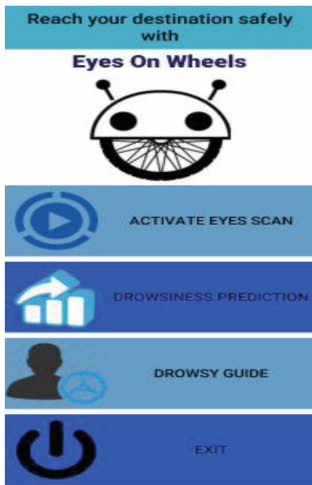


Fig. 1. Flowchart of the drowsiness prediction mobile application.

## 4 Results/Discussion

### 4.1 The Software – Inputs and Features



The android application, ‘Eyes on Wheels’, has as main feature ‘Activate Eyes Scan’ to scan the driver’s eyes in real-time. The front mobile camera is switched on and the OpenCV classifier `haarcascade_eye_tree_eyeglasses.xml` is loaded. The driver’s face image is monitored in real-time from which frames are captured and processed. Moreover, from each face’s frame captured, the eyes state (closed or opened) is monitored and an alarm sound is automatically triggered if the eyes are closed. The eye blink frequencies are captured together with the day of week and time and stored in the storage file. This dataset is used to calculate the probability that a driver is drowsy on a given day and time by using Naïve Bayes algorithm. Furthermore, drowsiness prediction is listed as a secondary feature in the application. This feature calculates the level of drowsiness from past captured

eyes blink frequencies over a weekly period and predicts the driver on current day and time where level of drowsiness 0%–49% is NORMAL, 50%–70% is MODERATE and above 70% is RISKY. The application software interface shows 92% of the audience were for the dynamic interface. 43.5% concluded that the notification alarm helped them in staying awake.

#### 4.2 Assessment of Software

The android application is deployed on a mobile device where the time taken process, detect and trigger alarm sound as well as the system performance were taken into consideration which are illustrated in Table 2.

**Table 2.** Performance of the mobile application.

Performance evaluation of the system	Average time taken (ms)
Time taken to load haarcascade_eye_tree_eyeglasses classifier in the system	109
Time taken to process the captured image	844
Time taken to detect closed eyes(s) after processing captured image	4
Time taken to trigger alarm after detecting closed eye(s)	3
Time taken to update drowsiness frequency file	6

Table 3 represents the 2-dimension array where the frequencies of eyes blinks and closures are recorded. The application records the frequencies of eyes blinks and closures based on following conditions:

**Table 3.** Drowsiness frequency table

No.	Day	Time	No. of eyes blinks/closures	Drowsy	Time	Categories
1	Mon	Night	2	No	0:00 to 5:00	Early morning
2	Tues	Evening	3	Yes	5:00 to 11:30	Morning
3	Thurs	Evening	3	Yes	11:30 to 13:00	Noon
4	Fri	Night	12	Yes	13:00 to 18:00	Afternoon
5	Tues	Evening	2	No	18:00 to 20:00	Evening
6	Fri	Night	8	Yes	20:00 to 22:00	Night
7	Sat	Night	2	No	22:00 to 24:00	Late Night
8	Sat	Late Night	14	Yes		
9	Fri	Night	10	Yes		
10	Tues	Evening	2	No		
11	Fri	Night	9	Yes		
12	Sat	Late Night	5	Yes		
13	Tues	Afternoon	2	No		
14	Thurs	Evening	5	Yes		
15	Fri	Night	2	No		

- (a) Condition 1: If the number of eyes’ blinks are greater than 1 within 30 s, then this is measured as a symptom of drowsiness.
- (b) Condition 2: If the eyes remain closed for more than 2 s, then this is measured as symptom of drowsiness.

The frequencies values and the current day and time are recorded in the frequency table as shown in Table 3. Hence, the higher the eyes’ blinking frequencies means the higher the driver’s tiredness. From the stored information in frequency table, the drowsiness prediction module computes the driver’s drowsiness level and displays an on-screen message. From Table 3, based on the information collected we can deduce that the driver has drowsiness symptoms mostly on Fridays’ nights. However, on Mondays and Tuesdays’ evenings, we can deduce that the driver is in a normal driving state. For a driver to be drowsy, the number of eyes blinks/closures on that given day and time must be greater or equal to 3.

For example, suppose driver X decides to take the wheels on a Friday’s night. The application calculates driver X’s drowsiness level by using Naïve Bayes Theorem. Table 4 shows training of dataset.

**Table 4.** Training of dataset of drowsiness on a Friday night

Day	Drowsy=Yes	Drowsy=No	Total	Time	Drowsy=Yes	Drowsy=No	Total
Fri	4/9	1/6	5/15	Night	4/9	3/6	7/15

$$P(\text{Drowsy}=\text{Yes} \mid \text{Day}=\text{Friday}, \text{Time}=\text{Night}) = \frac{\frac{4}{9} \times \frac{4}{9} \times \frac{9}{15}}{\frac{5}{15} \times \frac{7}{15}} = 0.761904\dots$$

As per the calculation, the level of drowsiness on a Friday night is 76% which is RISKY. Thus, by taking into consideration a population of 25 people, 67% of fatigue accidents are related to the age range below 29 years old, 84% of accidents are related to the age range between 30 to 50 years old and 64% are related to 51 to 70 years old people where 52% of the people were having health disorders. Moreover, 43.5% of the tested population concluded that the application was effective due to the notification alarm which helped them in staying awake. Interestingly, 31% of the drivers found that the application needed more improvement as there is a 1–2 s(s) delay when returning the processed image to the front camera view. Furthermore, as a pre-requirement, the front camera should be positioned in a fixed orientation, that is, landscape in order to process the image. Insomnia, overstressed and overloaded works and drugs/medicine consumption are the reasons causing fatigue on the road. The survey clearly states that 65.5% man are more likely to fall asleep while driving unlike women who remains alert on roads.

## 5 Conclusion

In this paper, the android application developed for the driver fatigue detection, makes the driver more vigilant on road, since he/she is already aware of his/her drowsiness state whether it is normal, moderate or risky. The accuracy of the eye detection decreases when the lighting condition varies in intensity to focus to the camera lenses. However, the improvement of the accuracy and efficiency of eyes detection when the amount of light vary from high to low or vice versa, the image processing module can be enhanced by implementing IR illuminator. This allows the camera to see in complete darkness. Finally, to reduce the delays from displaying images to front camera, an alternative solution can be used, that is, instead of returning whole face image, the system can only output the eyes' parts.

## References

1. Masten, S.V., Stutts, J.C., Martell, C.A.: Association for the Advancement of Automotive Medicine, Chicago, IL, October 2006
2. Tefft, B.C.: Prevalence of motor vehicle crashes involving drowsy drivers. AAA Foundation for Traffic Safety, Washington, DC, USA (2009–2013). AAA Foundation for Traffic Safety 2014, 19 October 2015
3. Juvale, H.B., Mahajan, A.S., Bhagwat, A.A., Badiger, V.T., Bhutkar, G.D., Dhabe, P.S., Dhore, M.L.: Drowsy detection and alarming system (DroDeASys). In: Proceedings of the World Congress on Engineering and Computer Science, 2007 WCECS 2007 San Francisco, USA, 24–26 October 2007
4. Chang, L.Y., Mannering, F.: Analysis of injury severity and vehicle occupancy in truck non truck-involved accidents. *Accident Anal. Prevent* **31**, 579–592 (1999)
5. IEEE Conference on Intelligent Transportation System, 1997, ITSC 1997, 9–12 November 1997
6. Road Transport and Road Traffic Accident Statistics (Mauritius)
7. Crowley, K.E., Johns, M.W., Chapman, R.J., Tucker, A.J., Patterson, J.: An Ocular Measure of Drowsiness and the EEG Changes with Sleep Deprivation. Sleep Diagnostics Pty Ltd, Swinburne University of Technology, Melbourne, Australia
8. Johns, M.W.: Drowsy Driving and the Law. Tasmania Law Reform Institute re: Issus Paper No 12, p. 3, November 2007
9. *Real-Time Imag.* **8**, 357–377 (2002). doi:[10.1006/rtim.2002.0279](https://doi.org/10.1006/rtim.2002.0279)
10. Sigari, M.-H., Fathy, M., Soryani, M.: *Intl. J. Vehicular Technol.* **2013**, 11 p. (2013). Article ID 263943
11. Bharathi, C.J.: Detection of drowsiness in human eye using SVM. *Intl. J. Innovative Res. Comput. Commun. Eng.* **2**(2), February 2014
12. von Jan, T., Karnahl, T., Seifert, K., Hilgenstock, J., Zobel, R.: Don't Sleep and Drive—VW's Fatigue Detection Technology. Centre for Automotive Safety Research, Adelaide University, Adelaide, Australia (2005)
13. Ji, Q., Yang, X.: Real-time eye, gaze, and face pose tracking for monitoring driver vigilance. *Real-Time Imag.* **8**(5), 357–377 (2002)

14. Batista, J.: A drowsiness and point of attention monitoring system for driver vigilance. In: Proceedings of the 10th International IEEE Conference on Intelligent Transportation Systems (ITSC 2007), Seattle, Washington DC, USA, pp. 702–708, October 2007
15. de la Escalera, A., Flores, M.J., Armingol, J.M.: Driverdrowsiness warning system using visual information for both diurnal and nocturnal illumination conditions. *EURASIP J. Adv. Sig. Process.* **2010** (2010). Article ID 438205
16. Weirwille, W.W.: Overview of research on driver drowsiness definition and driver drowsiness detection. In: 14th International Technical Conference on Enhanced Safety of Vehicles 1994, pp. 23–26 (1994)
17. He, J., Roberson, S., Fields, B., Peng, J., Cielocha, S., Coltea J.: Fatigue detection using smartphones USA 2013. *J. Ergon.* **3**. ISSN:2165-7556 JER
18. Parmar, S.H., Jajal, M., Brijbhan, Y.P.: Drowsy driver warning system using image processing. *Intl. J. Eng. Develop. Res.* ISSN: 2321-9939
19. Podder, S., Roy, S.: Driver's drowsiness detection using eye status to improve the road safety. *Intl. J. Innovative Res. Comput. Commun. Eng.* **1**(7), 1–8 (2013)
20. Murukesh, C., Padmanabhan, P.: Drowsiness Detection for Drivers Using Computer Vision Chennai, vol. 12 (2015). India, E-ISSN 2224-3402
21. Jain, P., Chopra, N., Gupta, V.: Automatic license plate recognition using OpenCV. *Intl. J. Comput. Appl. Technol. Res.* **3** (2014). New Delhi, India
22. Garg, D., Sharma, A.K.: Face recognition. *IOSR J. Eng. (IOSRJEN)* **2**(7) 128–133 (2012). ISSN: 2250-3021, M. Tech Scholar, Dissertation Guide Department Of Information and Technology Lingaya's University, Nachauli, Faridabad, INDIA
23. Daxini, N., Sharma, S., Patel, R.: Real time animal detection system using HAAR like feature. *Intl. J. Innovative Res. Comput. Commun. Eng.* **3**(6), 1–6 (2015). Research Scholar Gujarat Technological University, Ahmedabad, India
24. Pingale, R.P., Rokade: S.M. *Intl. J. Sci. Res. Publ.* **5**, 5 May 2015. ISSN 2250-3153
25. Beisenbek, B., Naizabayeva, L.K., Hernsoo, H.: Design and implementation of human-computer interaction system based on real-time tracking of the objects, Almaty (2013)

# A Hybrid Three-Phased Approach in Requirement Elicitation

Abubakar Zakari<sup>1(✉)</sup>, Abdulmalik Ahmad Lawan<sup>1</sup>,  
and Girish Bekaroo<sup>2</sup>

<sup>1</sup> Department of Computer Science, Kano University of Science  
and Technology, Wudil, Kano, Nigeria

abubakar.zakari@yahoo.com, aaltofa2000@gmail.com

<sup>2</sup> School of Science and Technology, Middlesex University  
(Mauritius Branch Campus), Vacoas, Mauritius  
g.bekaroo@mdx.ac.mu

**Abstract.** Requirement elicitation is one of the most important activities in requirement engineering and allocating limited amount of time in this activity is considered to significantly contribute towards failure of software projects. Having quality requirements is also greatly influenced by the techniques utilized during requirement elicitation process. The adoption of a single requirement elicitation technique within software development projects has various drawbacks. As solution, hybrid techniques are being considered as the way towards comprehensive requirements engineering. This paper investigates the hybrid requirement elicitation technique to tackle the challenges developers are facing in the process of software development. In this paper, the combination of 3 requirement elicitation techniques, namely use of questionnaire, interview and prototyping in a unified framework is investigated during the implementation of an online educational system.

**Keywords:** Requirement elicitation · Requirement engineering · Hybrid requirement elicitation · Software development life-cycle (SDLC)

## 1 Introduction

Requirement engineering is a segment of software engineering that is responsible for identifying the real functions and limitations of software system. By identifying both user and system requirements through the respective stakeholders of a system, it leads to a quality deliverance of a software system [1]. During recent years, researchers actively attempted to improve quality in the initial stage of the software development life cycle (SDLC). Automation of requirement engineering process became of high importance, but despite all the effort, requirements engineering (RE) still remain a tough problem to automate because of its human-centered nature [2].

Requirements are highly important in understanding, managing and controlling costs in software projects and their identification are considered as vital towards success of software projects [3, 4]. Although requirements need to be sufficiently complete, consistent and testable, the most neglected practice in SDLC is to document



them [3]. Moreover, using requirement documents that have errors as a reference in other projects can adversely cause further errors in the final product. A study showed that companies spent nearly 10% out of total time allocated on a project in requirement gathering and on completion of the project many companies realized that about 50–80% of their budget is gone on rework because time spent on requirement gathering was not enough [3]. Furthermore, projects where adequate time was spent on requirement gathering was found to have high success rate in comparison to projects that were allocated less time for requirement gathering [4].

In software requirement engineering, two types of requirements are gathered, namely, functional and non-functional requirements [5–8]. Functional requirement specifies something the system should while non-functional requirements relate to the operation of a system [2, 9, 10]. The core activities in requirement engineering are:

- i. **Requirement Elicitation:** Requirement elicitation is the process of gathering data from the user or stakeholders to the system developer [11, 12].
- ii. **Requirement Analysis:** All information gathered in the requirement elicitation process are analyzed and broken-down for the understanding of stakeholders needs [13].
- iii. **Requirement Implementation:** Requirement implementation is the stage where the software is coded and executed.
- iv. **Requirement Documentation:** The elicited data are documented for use during the implementation of the software [12].
- v. **Requirement Validation:** Also called as requirement verification [14], this process ensures that requirement documents are complete with unambiguity and users/stakeholders are satisfied with the requirement specification.

## 1.1 Requirement Elicitation

Getting quality requirements is of high importance to any software development project, which is directly proportional to the success of that project irrespective of the methodology utilized [14]. As such, requirement elicitation is vital in software development process [2]. It is the process of understanding the problems a proposed system will address through seeking, understanding and full disclosure of the needs of users and stakeholders, so as to communicate those needs to the developers [15]. There are two types of requirements elicitation techniques, namely, the direct approach and indirect approach. The direct approach techniques are based on case-study, interview, and prototyping [16]. The indirect approach is used in cases where information and data are cannot be easily retrieved. The techniques under in-direct approach involve use of questionnaire and document analysis, among others. Figures and statistics are utilized in this approach to clarify things.

## 1.2 Requirement Elicitation Techniques

As discussed earlier, requirement elicitation is the stage where the system developer gets to understand the problems of a proposed system [15, 17]. Different techniques are used in the process where the first one is interview. The main aim of interview is to

investigate and understand the requirement engineering process [18, 19]. In an interview, the users/stakeholders need to be interviewed first [20] and the interviewer will discuss the requirement of the product (system) with the user/stakeholder to get the overall view of the whole system. This technique has also been identified as the most utilized one because it mandates face to face interaction between the interviewer and the users/stakeholders and information can be driven quickly [21]. Survey is another technique and is used to gather requirements from users/stakeholders that may reside at different locations [22]. This technique is also utilized to analyze data from larger population of people than interviews [21]. With questionnaire, information can be obtained from a large group of people to get different views from the users/stakeholders [6, 12]. Another technique is observation which involves observing how people do their work practically. This technique can help in getting complex requirements that interviews cannot reveal [24]. Brainstorming is another technique where an individual member is free to express his/her idea about a product (system) to help bring about new ideas and solutions to a problem [25]. Finally, prototyping involves developing a version of the product (system) in order to get feedback from users/stakeholders.

## 2 Related Work

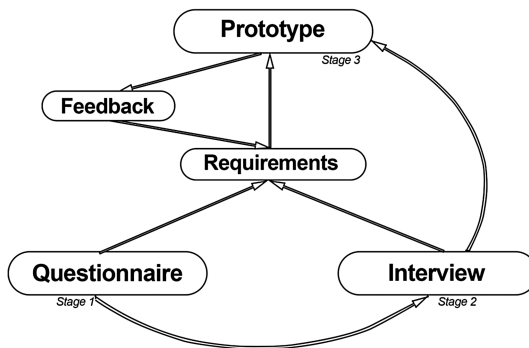
Hickey and Davis [26] presented a mathematical model of requirements elicitation that provided understanding of what analysts need to perform during elicitation, how elicitation techniques should be selected, and clue on improving likelihood that the system conform to customers' needs. The same work suggested that future models should capture the critical roles played by knowledge in both elicitation and elicitation technique selection. Another study [11] provided an overview on requirement elicitation techniques while comparing the strength of various requirement elicitation tools based on various parameters. The cons of adopting single requirement elicitation technique were highlighted in the same study. Basir et al. [27] constructed a framework for eliciting requirements that are considered as hidden or embedded and whose omission might cause software failure (i.e. tacit requirements). A hybrid framework was designed by integrating a reputable process and model of tacit requirements elicitation. Furthermore, 15 expert interviews were conducted to explore current practices in requirements engineering in three industries developing hybrid products [2]. Results of the same study showed that most components of hybrid products are developed independently from each other while involving high-level of technological integration of the elements and because of that, hybrid techniques was suggested as the way toward comprehensive requirements engineering. To improve effectiveness and efficiency of requirement elicitation different studies have been conducted using the hybrid approach [28–30]. Rooksby et al. [31] developed a hybrid process to fast-track consensual problem definition in large-scale systems with multiple stakeholders when eliciting requirement. Additionally, hybrid approaches was also highlighted to be effective in agile software development [32].

As hybrid techniques has been suggested as the way towards comprehensive requirements engineering, this paper proposes a novel hybrid requirement elicitation

technique to tackle the challenges developers have in the process of software development [4, 14, 33].

### 3 The Proposed Hybrid Requirement Elicitation Approach

This study attempted the combination of 3 requirement elicitation techniques, namely use of questionnaire, interview and prototyping in a unified framework that is expected to strengthen the process of requirement elicitation. The approach used operates as follows and is depicted in Fig. 1:



**Fig. 1.** A three-phased hybrid approach to requirement elicitation

**Stage 1.** In the first phase, the aim is to get information from large group of people so as to get different views from the users/stakeholders using questionnaires. Information collected is then analyzed to get insightful information on the key questions that need to be asked to the main stakeholders of the system.

**Stage 2.** In this stage, an interview is conducted to further refine the requirements driven from the questionnaire in phase 1, which will help in building the first prototype of the system at the later stage.

**Stage 3.** After acquiring information from users/stakeholders using questionnaires and interview, the requirement driven from interview will help in developing the first prototype of the product (system). This first prototype is to give the users/stakeholders the practical experience of the product (system) and their feedback will help in developing the final prototype.

## 4 Application of the Three-Phased Hybrid Approach

The conceptualized three-phased hybrid approach was utilized when building a system called NailClassroom<sup>1</sup> to elicit requirements from users/stakeholders. NailClassroom is an online educational system that makes communication and data sharing between lecturers and students easy and fast. In order to achieve these requirements, different factors pertaining to web design were also implemented [34, 35]. Application of the approach is described as follows:

### 4.1 Participants

**Stage 1 Participants.** A questionnaire was formulated and administered to 550 students and 50 lecturers from public universities in Kano State Nigeria. A valid response of 150 students and 20 lecturers was recorded. The demographic details of the participants are given in Tables 1–3.

**Table 1.** Age split of the questionnaire respondents

Age	Count	Percentage
18–22	67	39.4
23–27	64	37.6
28–32	20	11.8
33–37	8	4.7
38–42	4	2.4
Above 42	7	4.1
Total	170	100.0

**Table 2.** Gender split of the questionnaire respondents

Gender	Count	Percentage
Male	114	67.1
Female	56	32.9
Total	170	100.0

**Table 3.** Academic level of the questionnaire respondents

Level	Frequency	Percentage
100	17	10.0
200	44	25.9
300	39	22.9
400	70	41.2
Total	170	100.0

<sup>1</sup> NailClassroom. Available at: <http://nailclassroom.com/>.

**Table 4.** Departments of the questionnaire respondents

Department	Frequency	Percentage
Mathematics	98	57.6
Computer Science	72	42.4
Total	170	100.0

**Stage 2 Participants.** Information obtained from the questionnaire helped in identifying the real users/stakeholders and in narrowing vital questions to be used in our interview to get our requirement right. In Stage 2, 5 lecturers from the Computer Science department, 2 lecturers from Mathematics department and 5 students all from Kano University of Science and Technology, Wudil were interviewed.

**Stage 3 Participants.** Requirements finalized in phase 2 helped in building the first prototype of the system. The first prototype was developed and tested on 20 students and 5 lecturers at Kano University of Science and Technology. The feedback accumulated from the first prototype was used to make required changes on the requirement document; which helped in developing the second prototype. This prototype is presently adopted by more than 500 students across Nigerian universities.

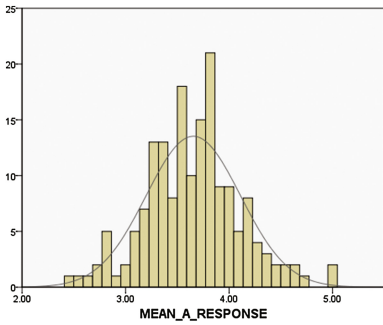
## 4.2 Analysis of Proposed Approach

Among the elicitation techniques utilized, interview and questionnaire were found to be effective in getting ambiguous and complex requirements from users/stakeholders. From phase 1 towards phase 3, requirements were further fine-tuned and the stakeholders claimed to be more involved in the process. Among the three techniques, prototyping was found to be more effective as the stakeholders would be able to obtain the look and feel of the system. However, to confirm whether elicited requirements were correctly collected, the final prototype was validated by the same users involved during each of the 3 stages of elicitation. In the process, data was collected pertaining to collaboration/syllabus (part A) and ease of use of the implemented final prototype (B) was collected.

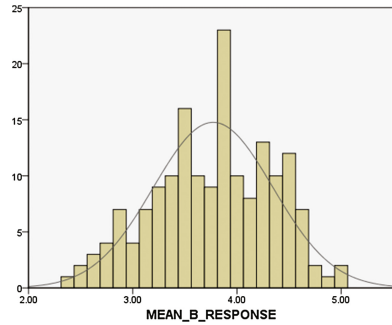
## 5 Results and Discussions

Results from the three stages, namely, use of questionnaire, interview and prototype are given as follows (Figs. 2, 3, 4, 5 and Tables 4, 5):

**Stage 1 Result: Questionnaire**



**Fig. 2.** Collaboration/syllabus validity



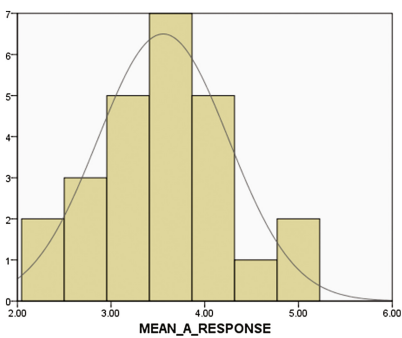
**Fig. 3.** Ease of use validity

**Stage 2 Result: Interview**

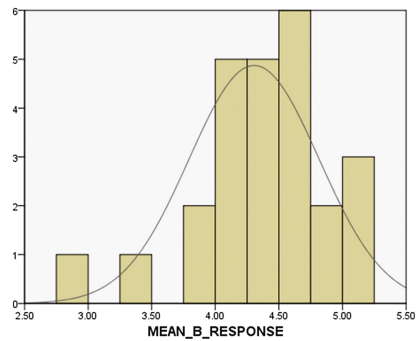
**Table 5.** Interview responses

Part A:		
	Positive response	Negative response
Students	41.7%	0.0%
Lecturers	58.3%	0.0%
Total	100.0%	0.0%
Part B:		
	Positive response	Negative response
Students	41.7%	0.0%
Lecturers	50.0%	8.3%
Total	91.7%	8.3%

**Stage 3 Result: Prototype Feedback**



**Fig. 4.** Collaboration/syllabus validity



**Fig. 5.** Ease of use validity

Results from the three phases revealed a high positivity in terms of validity of requirements. A few negative responses were also gathered especially from users who had more expectations from the system in terms of look and feel, although a prototype was used as part of validation. Overall, findings of the study made it clear that hybrid approach to requirement elicitation is the way forward to requirements elicitation, which could be smoothened by the proposed three-phased hybrid requirement elicitation technique in this study. The accuracy of requirements collected varies in each stage of our hybrid technique but showed to improve from Stage 1 until the last stage. There is no such thing as 100% accurate requirement but getting a successful system running according to user requirements, is a tangible and reliable indicator that a developer should consider. The proposed hybrid approach is thus expected to provide software developers a feasible framework toward generating accurate requirements.

## 6 Conclusions

This paper investigated the hybrid requirement elicitation technique involving the combination of 3 such techniques, namely use of questionnaire, interview and prototyping in a unified framework. The proposed approach was investigated during the implementation of an online educational system called NailClassroom. Results revealed a high positivity in terms of validity of requirements by participants from the three phases. Results also confirmed that hybrid approach to requirement elicitation is the way forward to requirements elicitation as accuracy of gathered requirements improved from first stage to the final one. As future work, the same approach could be further investigated in different types and size of software development projects.

## References

1. Lee, Y., Zhao, W.: Domain requirements elicitation and analysis-an ontology-based approach. In: International Conference on Computational Science, pp. 805–813 (2006)
2. Berkovich, M., Leimeister, J.M., Krcmar, H.: An empirical exploration of requirements engineering for hybrid products. In: 17th European Conference on Information Systems, pp. 1–13 (2009)
3. Mrayat, O.I.A., Norwawi, N.M., Basir, N.: Requirements elicitation techniques: comparative study. *Int. J. Recent Dev. Eng. Technol.* **1**(3), 1–10 (2013)
4. Shehzad, K., Awan, M., Rizvi, S., Khiyal, M.: A hybrid technique based on standard SRS modules for software requirement prioritization. In: Informing Science and IT Education Conference, pp. 279–294 (2014)
5. Pa, N., Zin, A.: Requirement elicitation: identifying the communication challenges between developer and customer. *Int. J. New Comput. Archit. Appl. (IJNCAA)* **1**(2), 371–383 (2011)
6. Marupaka, T., Raju, C., Tudigani, A.: Requirement engineering – monitoring elicitation technique for end product software. *Int. J. Emerg. Trends Technol. Comput. Sci. (IJETTCS)* **3**(1), 1–6 (2014)
7. Saranya, R.: Survey on security measures of software requirement engineering. *Int. J. Comput. Appl.* **90**(17), 12–19 (2014)

8. Nisar, S., Nawaz, M., Sirshar, M.: Review analysis on requirement elicitation and its issues. *Int. J. Comput. Commun. Syst. Eng. (IJCCSE)* **2**(3), 484–489 (2015)
9. Besrou, S., Rahim, L., Dominic, P.: The study of available techniques for existing requirements engineering challenges based on literature review evidences. *Res. J. Appl. Sci., Eng. Technol.* **8**(19), 2082–2091 (2014)
10. Yousof, M., Asger, M., Bokhari, M.: A systematic approach for requirements elicitation techniques selection: a review. *Int. J. Adv. Res. Comput. Sci. Softw. Eng.* **5**(4), 1399–1403 (2015)
11. Abbasi, M., Jabeen, J., Hafeez, Y., Batool, D., Fareen, N.: Assessment of requirement elicitation tools and techniques by various parameters. *Softw. Eng.* **3**(2), 7–11 (2015)
12. Zapata, C., Losada, B., Calderón, G.: An approach for using procedure manuals as a source for requirements elicitation. In: *IEEE XXXVIII Conferencia Latinoamericana En Informatica (CLEI)*. IEEE (2012)
13. Viller, S., Sommerville, I.: Social analysis in the requirements engineering process: from ethnography to method. In: *IEEE International Symposium on Requirements Engineering* (1999)
14. Swarnalatha, K., Srinivasan, G., Bhandary, P., Kishore, P., Rakesh, R.: Requirement elicitation in web applications: challenges. *Int. J. Res. Comput. Commun. Technol. Adv. Technol.* **3**(3), 382–386 (2014)
15. Anwar, F., Razali, R.: A practical guideline of selecting stakeholders for requirements elicitation—an empirical study. *Int. J. Softw. Eng. Appl.* **9**(2), 95–106 (2015)
16. Khan, S., Dulloo, A., Verma, M.: Systematic review of requirement elicitation techniques. *Int. J. Inf. Comput. Technol.* **4**(2), 133–138 (2014)
17. Zave, P.: Classification of research efforts in requirements engineering. *ACM Comput. Surv. (CSUR)* **29**(4), 315–321 (1997)
18. Zowghi, D., Coulin, C.: Requirements elicitation: a survey of techniques, approaches, and tools. In: *Engineering and Managing Software Requirements*, pp. 19–46 (2005)
19. Hoffer, J., George, J., Valacich, J.: *Modern Systems Analysis and Design*, 5th edn. Pearson International Edition, New Jersey (2008)
20. Ismail, N., Razak, M., Zakariah, Z., Alias, N., Aziz, M.: E-learning continuance intention among higher learning institution students’. In: *Procedia-Social and Behavioral Sciences*, Malaysia, vol. 67, pp. 409–415 (2012)
21. Ur Rehman, T., Khan, M., Riaz, N.: Analysis of requirement engineering processes, tools/techniques and methodologies. *Int. J. Inf. Technol. Comput. Sci. (IJITCS)* **5**(3), 40 (2013)
22. Driscoll, D.: Introduction to primary research: observations, surveys, and interviews. In: *Writing Spaces: Readings on Writing*, vol. 2, pp. 153–174 (2011)
23. Ogwueleka, F.: Requirement elicitation problems in software development—a case study of a GSM service provider. *Indian J. Innovations Dev.* **1**(8), 599–605 (2012)
24. Mulla, N., Girase, S.: Comparison of various elicitation techniques and requirement prioritisation techniques. *Int. J. Eng. Res. Technol.* **1**(3), 1–8 (2012)
25. Mulla, N., Girase, S.: A new approach to requirement elicitation based on stakeholder recommendation and collaborative filtering. *Int. J. Soft. Eng. Appl.* **3**(3), 51 (2012)
26. Hickey, A., Davis, A.: Requirements elicitation and elicitation technique selection: model for two knowledge-intensive software development processes. In: *IEEE 36th Annual Hawaii International Conference on System Sciences*, p. 10 (2003)
27. Basir, B., Salam, R.: Tacit requirements elicitation framework. *ARPN J. Eng. Appl. Sci.* **10**(2), 572–578 (2015)



28. Murali, V., Sinha, N., Torlak, E., Chandra, S.: What gives? a hybrid algorithm for error trace explanation. In: Working Conference on Verified Software: Theories, Tools, and Experiments, pp. 270–286 (2014)
29. Albert, M., Ravi, T.: Structural software testing: hybrid algorithm for optimal test sequence selection during regression testing. *Int. J. Eng. Technol. (IJET)* **7**(1), 270–279 (2015)
30. Harrold, M., Steimann, F., Tip, F., Zeller, A.: Fault prediction, localization, and repair. In: Dagstuhl Seminar, vol. 13061 (2013)
31. Rooksby, J., Sommerville, I., Pidd, M.: A hybrid approach to upstream requirements: IBIS and cognitive mapping. In: *Rationale Management in Software Engineering*, pp. 137–154 (2006)
32. Paetsch, F., Eberlein, A., Maurer, F.: Requirements engineering and agile software development. In: *12th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprise*, pp. 308–313 (2003)
33. Achimugu, P., Selamat, A., Ibrahim, R., Mahrin, M.: A systematic literature review of software requirements prioritization research. *Inf. Softw. Technol.* **56**(6), 568–585 (2014)
34. Martinez, M.: Key design considerations for personalized learning on the web. *Educ. Technol. Soc.* **4**(1), 26–40 (2001)
35. Bekaroo, G., Bokhoree, C., Pattinson, C.: Impacts of ICT on the natural ecosystem: a grassroot analysis for promoting socio-environmental sustainability. *Renew. Sustain. Energy Rev.* **57**, 1580–1595 (2016)

# A Framework to Reduce the Testing Time of a Mobile Application Using an Automation Tool

V. Hurbungs<sup>(✉)</sup>, B. Dookheea, and Y.K. Suttroogun

Department of Computer Science and Engineering, Faculty of Engineering,  
University of Mauritius, Reduit, Mauritius  
v.hurbungs@uom.ac.mu,  
{bhoomita.dookheea,yogin.suttroogun}@uemail.uom.ac.mu

**Abstract.** Nowadays, with the ever-increasing demand for a wide range of mobile applications, mobile test automation is required to provide cost-effective, high quality and reliable mobile applications. Compared to manual testing which is time consuming and requires more resources in terms of people and time, automation aims at improving the test process. There are several mobile automation applications available on the market but there no proper mechanisms or guidelines that define the steps or best practices to adopt when performing mobile application development with the view of improving the testing process in terms of time taken. This research work investigates latest mobile application testing tools and guidelines, and proposes a framework to reduce the testing time of an android mobile application. In the proposed framework, five mobile application development guidelines that are common to android applications have been considered and tested using an automated testing tool.

**Keywords:** Testing · Guidelines · Android · Mobile · Framework

## 1 Introduction

According to International Data Corporation (IDC), Android is the most popular mobile operating system from 2012 to 2015 with a market share of 82.8% [1]. Android is also powerful and supports a large number of applications. In July 2013, more than 50 billion of applications have been downloaded from Google Play [2]. Android developers find it very competing to create amazing applications with great features. Mobile application testing is therefore very important since the process improves mobile application usability. It helps reduce time to market and increase productivity as a high quality application is deployed at the end of the process. If testing is not properly conducted, there will be bugs which may cause the application to crash. Also, developers face difficulties while testing their mobile applications as they do not adhere to a standard development process.

Mobile application testing is therefore a key component in building reliable mobile applications. The objective is to ensure that mobile users have a positive experience when they use a mobile application. However, mobile application testing presents

unique challenges since there are additional parameters that need to be considered compared to desktop applications. For example, there exists a variety of devices with different specifications such as operating system, screen size, memory, standards, network and communication interfaces. According to [3], the testing strategy for mobile applications can be based on the recommendations mentioned in Table 1.

**Table 1.** Testing strategy for mobile application testing [3]

Recommendation	Description
Take advantage of a device emulator	Emulated devices are less costly and can perform testing quickly and efficiently. A device emulator can offer options such as network bypass, network options, support for scripting language, and switching between different device profiles
Invest in a real device cloud	A cloud environment provides possibility to access real device anytime
Automate wherever possible	Emulators or real-devices with scripting support can provide test cases playback which saves testing time

The work done in this paper is therefore based on the last recommendation, which is mobile application test automation.

### 1.1 Mobile Application Test Automation

Automation testing is the process during which a tester runs test scripts using testing tools. Tools are used to help in the testing process by reducing test cycles and by increasing productivity and quality. Also, manual testing processes are replicated at a greater speed. Technically, automation [4] is described as the process that helps reduce manual effort with the use of tools and strategies. The automation process includes the following stages [5]: Test Tool Selection, Define scope of Automation, Planning, Design and Development, Test Execution and Maintenance. Automated testing is therefore more efficient compared to manual testing in various ways as automation increases efficiency and effectiveness. The best way to test mobile applications is via software testing tools as they will ease the process of testing and will result in a smaller delay for the application to be sent to the market and deployed by end users. With the use of automated tools, a testing framework needs to be established in order to facilitate the testing process.

According to [6], there are 2 ways to automate mobile testing of applications: (1) **Record and Playback Approach:** In this approach, some set of actions are performed on the user interface (UI) and the interactions are recorded automatically, tested and can be played back as many times as necessary. It is less prone to human errors and is not time-consuming compared to hand-written test-scripts. It is mostly used to test Graphical User Interfaces. (2) **Automatic test Exercises:** This approach is best suitable

for smoke and sanity testing for mobile application. With the use of automatic test exercises [6], there is no need for tests, configurations or setups. It is used to test the user interface logic (example: clicking buttons, swiping). After uploading the application, instant results are obtained.

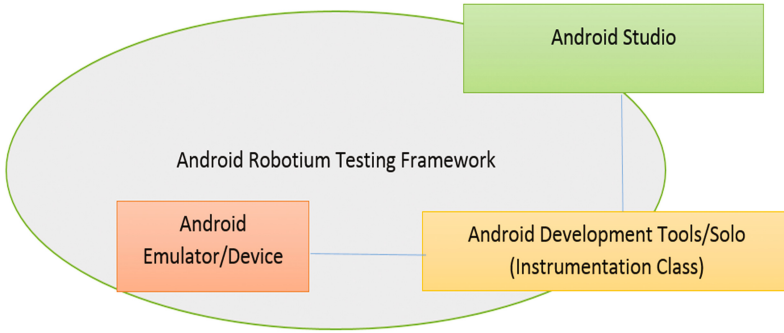
## 1.2 Existing Mobile Application Testing Frameworks

Although Android has gained popularity, the field for testing android mobile applications is still being explored. One approach is to use a record and replay approach using a tool called RERAN [7], which does not provide the ability to understand and manipulate test scripts. In [8], they proposed a capture-and-replay method whereby user interactions are captured and altered into Robotium test-scripts. The events can be replayed using the user's assertions documented and compared to the intended output. In [9], a GUI crawling method which uses AndroidRipper as a testing tool was used to imitate real user events. In [10], testers developed an online Android application testing platform named Testdroid. Recorded tests are run in parallel on a variety of handsets.

In [11], another approach for android application testing which laid more emphasis on GUI bugs was proposed. First, a study was conducted to understand the different bugs. Based on the latter, techniques were identified to detect bugs through automatic test case generation. In [12], they proposed an automated input generator using Robotium testing framework in order to test for valid inputs and compared this framework with a randomized testing approach. In [13], they proposed a test automation framework MobTAF whereby testing is done directly on mobile devices without connecting to a computer.

## 2 The Mobile Application Testing Framework

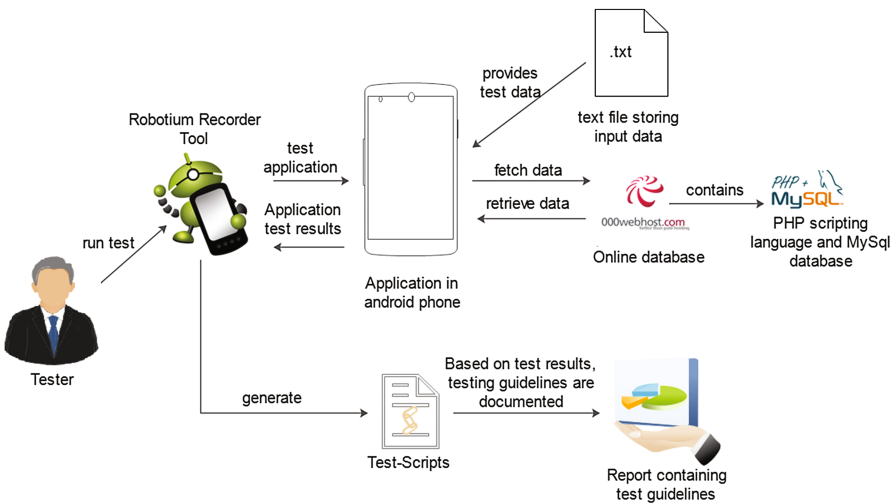
Although there exist various automation tools which help with testing of mobile applications, they are not convenient for mobile developers due to a lack of functionalities. As a solution, a framework that can easily set developers' needs has been proposed. Five guidelines, which help in enhancing the testing process, have been identified. (1) It is best to make use of error-prompt messages rather than toast messages which usually take long time to appear and disappear thereby constantly irritating users. (2) Instead of using multiple screens (activities in android studio), the use of dialog boxes makes it easier to test the interface. (3) Constant scrolling through the interface is quite tiresome and might cause the application to crash. A search bar can help avoiding such issues. (4) Providing users with an overview of the format to use for inputs is user-friendly but restricting their freedom to write is very inconvenient; it is therefore best to use a grey-out text whereby users are able to enter data according to a format without erasing the previous one. (5) Rotation is an issue as components may get distorted as the orientation changes. Grid layout format is recommended compared to Relative layout.



**Fig. 1.** Robotium android framework using android studio

Robotium using Android Studio has been used to implement the proposed framework (Fig. 1).

A car rental mobile application with features such as new member registration, login, browsing for cars and renting cars has been developed to evaluate the proposed framework. Based on the car rental prototype and the above guidelines, two android applications have been developed and tested using the Robotium Recorder tool. A normal application (without any guidelines) and a specialized prototype (with appropriate guidelines) have been implemented. The time taken for testing the normal application as well as the testing difficulties have been determined. A set of guidelines was then devised to amend the interface for rapid testing with respect to the issues faced by the user. The suggestions for improving testing are implemented as a framework in the specialized application and the testing results are recorded. Both testing results for the normal prototype and the specialized application are eventually compared and a summary report is generated (Fig. 2).



**Fig. 2.** System architecture

During the testing phase, the tester will run the car rental application with the use of Robotium Recorder tool. Input data will be automatically generated as these will be stored in a text file in the android phone. Also, while the application is in use, information from the online database will be accessed. As soon as the testing is over, Robotium tool automatically generates test-scripts. These test-cases can be replayed again as many times as required as Robotium Recorder provides the capture and replay feature. The time-taken for the test completion can be noted. Based on the test-report, a set of guidelines is documented for decreasing the testing time. Another car rental application has been developed using these testing guidelines and has been tested in the same manner.

### 3 Results

#### 3.1 Using Error Prompt Messages Rather than Toast Messages

All applications use error messages to notify users about their actions. The login screen is tested without inputting data. The time taken for toast messages or error prompt messages to appear and disappear has been derived from the Robotium recorder test script. Multiple tests have been carried out and recorded and the average has been determined. The graph below shows that it is best to use error prompt messages as the time taken is less compared to toast messages (Fig. 3).

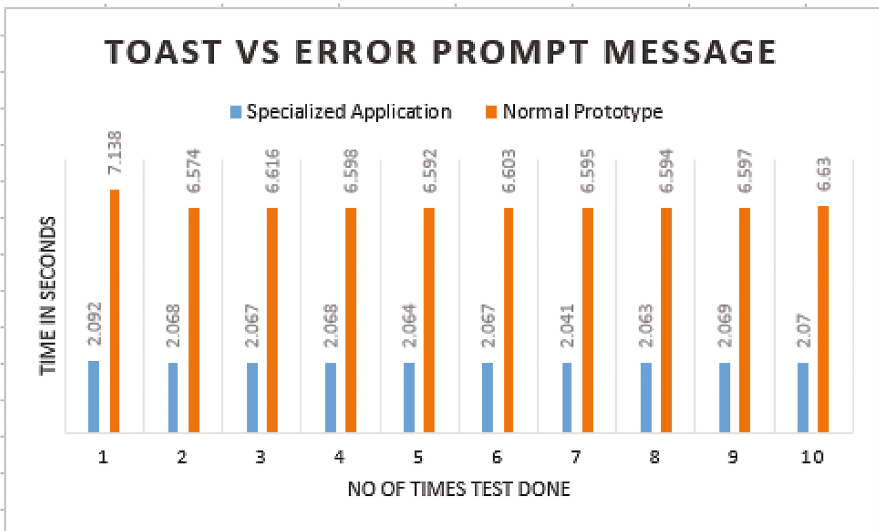


Fig. 3. Toast v/s error prompt message

### 3.2 Guideline 2: Using Dialog Box Rather than Activity (Screen)

The normal prototype has two registration screens compared to the specialized application. When a user clicks on “Sign Up” button, a dialog-box will appear where all required registration details are filled and appropriate error messages are displayed.

Based on Fig. 4, it can be deduced that testing is much simpler with the use of dialog box rather than an application having multiple screens.

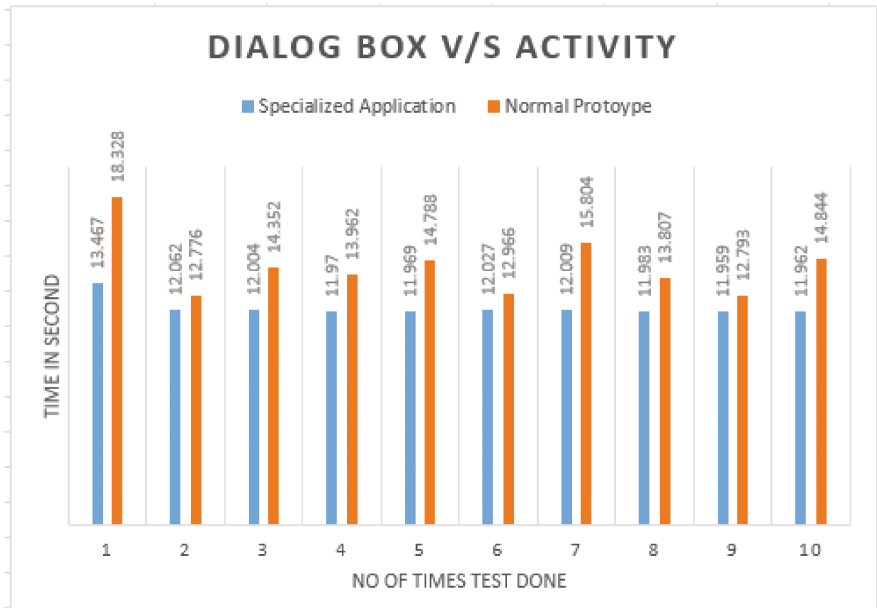


Fig. 4. Dialog box v/s activity

### 3.3 Using Search Rather than Scrolling Through a List

The specialized application implements a search bar option in the first gallery that helps avoid continuous scrolling when search for a specific keyword. Scrolling through a list might cause the application to crash since the mobile application must display a lot of information in a small space area. Instead of scrolling down the list, the user can simply type the keyword in the search bar and get the results rapidly. Search is therefore more user-friendly and also easier to test (Fig. 5).

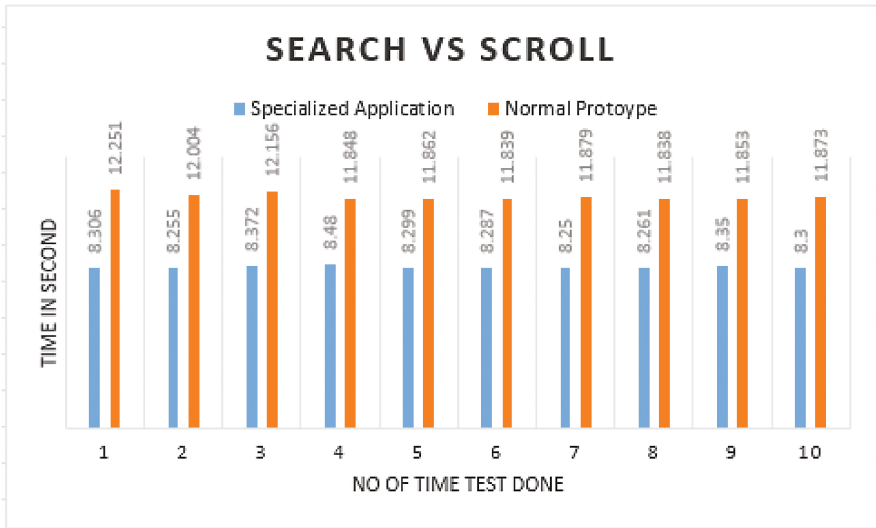


Fig. 5. Search v/s scroll

#### 3.4 Guideline 4: Using Date Hint Format v/s Text Format

In the normal prototype, the field for entering date has to be cleared before the user enters the date he wants. Also, the date format is not easy to remember and the user may forget the correct date format. In the specialized application, the date format is explicitly defined in the specialized application and it greys out without the need of erasing text before inputting data. The use of date format grey-out feature makes the use of the application much easier and it is less time consuming to test (Fig. 6).

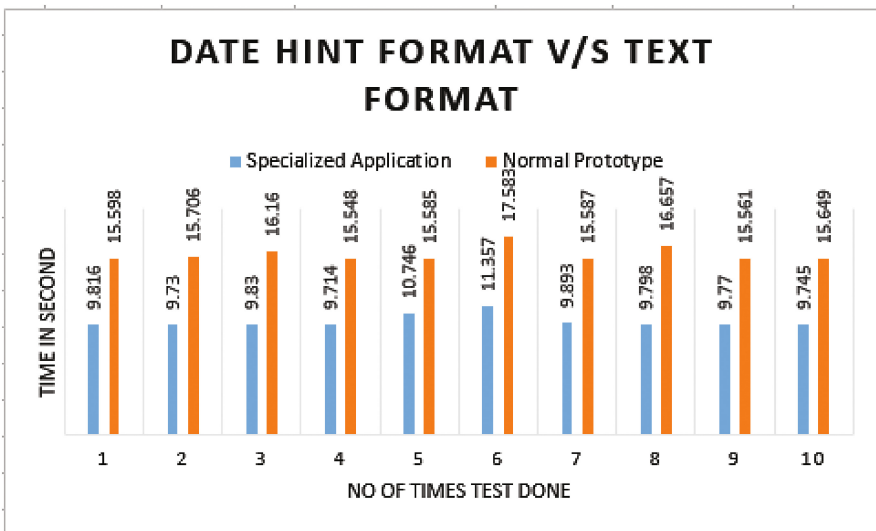


Fig. 6. Date hint format v/s text format



### 3.5 Guideline 5: Using Grid Layout and Scroll View Instead of Relative Layout

The login screen is similar to the one in the normal prototype but is different in its layout style as it uses a Grid layout as well as Scroll to ensure no distortion while rotation. Distortions occur when orientations are changed with the use of relative layout as shown in Fig. 7.

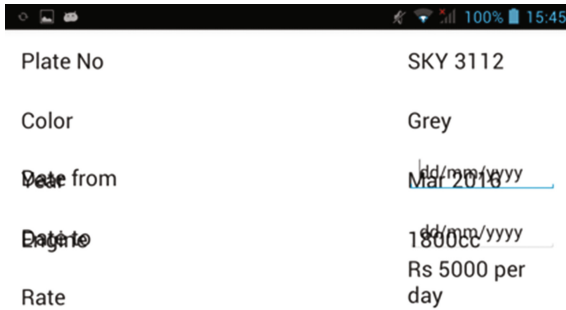


Fig. 7. Relative layout with distortion

### 3.6 Guideline 6: Comparison for Testing All Interfaces for Normal Prototype and Specialized Prototype

The total time taken to test the Normal prototype is 49.258 s while the total time take to test the Specialized application is 27.649 s (Fig. 8).

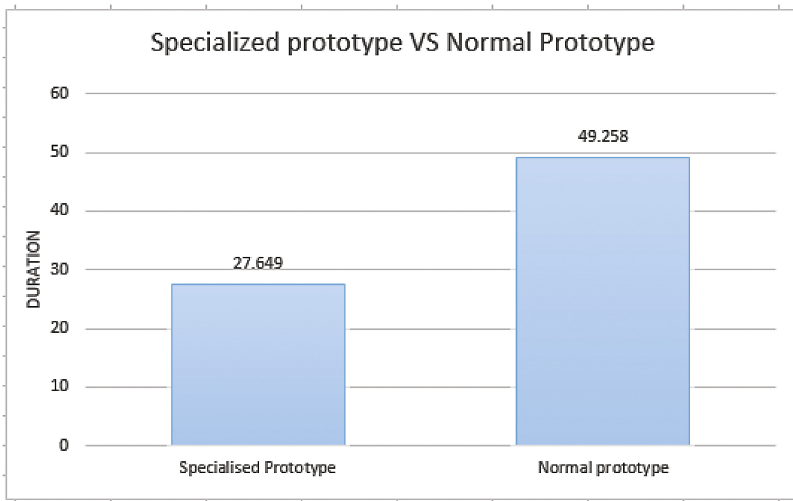


Fig. 8. Normal prototype v/s specialized prototype

Table 2 below summarises the test results in terms of the average time taken for both prototypes.

**Table 2.** Summary for test results of normal prototype v/s specialized prototype

Screens of application	Normal prototype	Specialized application	Average time for testing normal prototype (in seconds)	Average time for testing specialized prototype (in seconds)
Login screen	Use toast messages to display error messages	Use error message prompts to display error messages	6.6537	2.0669
Registration screen	2 registration screens used	Display registration details in a single screen	14.442	12.1352
Gallery screen	To search, user has to scroll down the list	Use a search bar for browsing cars	11.887	8.316
Rent screen	To enter date, the user has to type the date format	The date format is provided.	15.9634	10.0399
Screen rotation	Use of relative layout which is subject to distortion	Use grid layout and scroll view to avoid distortion	Distortion	No distortion

## 4 Recommendations

Based on the testing process and test results, Android mobile application testing seems to be quite tedious and time-consuming even with the use of an automated testing tool. In order to alleviate common issues that make testing difficult, the following guidelines for reducing testing life cycles with respect to the design of the mobile interfaces are therefore recommended:

1. Toast messages take a long time to appear and are irritating for a user to constantly see toast messages. In order to avoid the issue of receiving toasts, error prompts can be used.
2. The minimal use of screens (activities in android studio) makes things easier to test and is also convenient for users. Instead, dialog boxes can be used.

3. Constantly scrolling through the interface may cause the application to crash. Adding a search bar will increase user-friendliness and will prevent the application from crashing.
4. Giving users an overview of what format to use for date inputs is very user-friendly. Use a greyed-out text, users will be able to input date without erasing the previous one and the format will be known.
5. Rotation is an important issue in the normal prototype as components get distorted as the orientation changes. Proper layout formats such as Grid Layout can be applied to avoid distortions.

## 5 Conclusion

The aim of the project was to develop guidelines that will help in reducing testing life cycles of android applications. The system has been implemented successfully in the Android development environment. Results demonstrate reduction in the testing process by comparing the test results of a normal prototype to those of a specialized application which takes into consideration user interface guidelines. The testing framework designed is efficient and reliable as it decreases the testing time for android application to a large extent with the use of properly designed user interfaces. Based on the test results obtained, the total time taken to test the normal application is 49.258 s while the total time taken to test the specialized application is 27.649 s. It can therefore be concluded that the proposed framework reduced the testing time by about 50%. This reduction in testing time is therefore very significant for the quicker deployment of mobile applications on the android market.

However, the number of test cases Robotium Recorder offers is limited that is it is not free-licensed. Also, it cannot be exported to other platforms such as IOS. It is only limited to android applications. Good coding practices can also reduce the testing time and a future work may combine both user interface guidelines with the way developers write mobile application codes. Future works include assessing the framework performance under realistic network conditions and evaluation in terms of rendering times, CPU, battery and memory.

**Acknowledgments.** The authors would like to thank the University of Mauritius for providing the necessary facilities and services for conducting this research.

## References

1. IDC: Smartphone OS Market Share. [www.idc.com](http://www.idc.com), <http://www.idc.com/prodserv/smartphone-os-market-share.jsp>. Accessed Sept 2015
2. Esbjörnsson, L.: Android GUI Testing: a comparative study of open source Android GUI testing frameworks (2015)

3. Keynote: Testing Strategies and Tactics for Mobile Applications. [www.keynote.com](http://www.keynote.com), <http://www.keynote.com/resources/white-papers/testing-strategies-tactics-for-mobile-applications>. Accessed Aug 2016
4. Vennapoosa, C.: Automated Testing Advantages, Disadvantages and Guidelines (2005)
5. Automation Testing vs. Manual Testing[Software Testing Class. [Softwaretestingclass.com](http://www.softwaretestingclass.com), <http://www.softwaretestingclass.com/automation-testing-vs-manual-testing/>. Accessed Oct 2015
6. Helppi, V.-V.: The basics of test automation for apps, games and the mobile web. Smashing Magazine (2015)
7. Gomez, L., Neamtui, I., Azim, T., Milstein, T.: RERAN: timing- and touch-sensitive record and replay for Android. In: 2013 35th International Conference on ICSE, San Francisco, CA (2013)
8. Liu, C.H., Lu, C.Y., Cheng, S.J., Chang, K.Y., Hsiao, Y.C., Chu, W.M.: Capture-replay testing for android applications. In: 2014 International Symposium on Computer, Consumer and Control (IS3C), Taichung (2014)
9. Amalfitano, D., Fasolino, A.R., Tramontana, P.: A GUI crawling-based technique for android mobile application testing. In: Proceedings of the Fourth IEEE International Conference on Software Testing, Verification and Validation Workshops (ICSTW), Berlin (2011)
10. Kaasila, J., Ferreira, D.K.V., Ojala, T.: Testdroid: automated remote UI testing on Android. In: Proceedings of the 11th International Conference on Mobile and Ubiquitous Multimedia (MUM 2012), Germany (2012)
11. Hu, C., Neamtui, I.: Automating GUI testing for Android applications. In: Sixth International Workshop on Automation of Software Test (AST 2011), USA (2011)
12. Dutia, S.N., Tae, H.O., Young, H.O.: Developing automated input generator for android mobile device to evaluate malware behavior. In: Proceedings of the 4th Annual ACM Conference on Research in Information Technology (2015)
13. Nagowah, L., Sowamber, G.: A novel approach of automation testing on mobile devices. In: 2012 International Conference on Computer and Information Science (ICCIS), Kuala Lumpur, vol. 2 (2012)

# Author Index

## A

Ah King, Robert T.F., 71, 107, 132, 144  
Askoolum, Krishiv, 272

## B

Beeharry, Y., 252  
Bekaroo, Girish, 314, 331  
Binns, R., 107, 184  
Bissessur, Yasdeo, 63, 220  
Bodireddy, Kartheek, 201  
Bojkovic, Zoran, 25  
Boodoo, Sajaad, 220  
Busawon, K., 107, 184

## C

Calchand, S.G., 175  
Chellapermal, Nigel, 305  
Curum, Brita, 305

## D

Dahal, Keshav, 3  
Derick, M., 184  
Dookheea, B., 341

## E

Essackjee, Ismaël Adam, 71

## F

Fowdur, Tulsi Pawan, 252, 283, 293

## G

Gandham, Srinivasa Rao, 201  
Gao, Hui, 44  
Goolaup, Sarjoosing, 94

## H

Hoolaus, Raaju, 63  
Hurbungs, V., 341  
Huzooree, Geshwaree, 262

## I

Islam, Mohammad Fardeen, 165

## J

Jahmeerbacus, M.I., 119  
Joonas, Noorjehan, 262  
Jugurnauth, Rameshwar, 154

## K

Kamucha, George, 192  
Khedo, Kavi Kumar, 262, 272, 305  
Kiragu, Henry, 192  
Kumar, Boya Pradeep, 201

## L

Langat, Kibet, 239  
Latchoomun, L., 107  
Lawan, Abdulmalik Ahmad, 314, 331

## M

Makhanbet, Meruyert, 44  
Marappa Naiken, Sarah, 132, 144  
Mavungu, Masiala, 230  
Mawooa, D., 107  
Milovanovic, Dragorad, 25  
Mishra, Bhupesh Kumar, 3  
Mocktoolah, Asslinah, 272  
Murdan, Anshu Prakash, 154, 165  
Mwangi, Elijah, 192, 210, 239

## N

Nel, A.L., 230  
Ngocho, Boniface M., 210  
Nirsimloo, Ravishwara Rakesh, 154

## O

Oree, Vishwamitra, 165

**P**

Paidimarry, Chandra Sekhar, [201](#)  
Panchoo, Shireen, [322](#)  
Paurobally, Roshun, [220](#)  
Pervez, Zeeshan, [3](#)

**R**

Ragpot, Prateema, [283](#)  
Rajesh, M., [184](#)  
Rajkumarsingh, B., [81](#)  
Rajkumarsingh, Bhimsen, [94](#)  
Ramjug-Ballgobin, R., [175](#)  
Ramodhine, Karishma, [322](#)  
Rani, C., [184](#)  
Ruhomaun, M.S., [119](#)

**S**

Salih, Tedros, [239](#)  
Sayed Hassen, S.Z., [119](#)  
Sewraj, K., [119](#)  
Sheik Mamode, Maryam Imran, [293](#)  
Sinthamrongruk, Thepparit, [3](#)  
Sokappadu, B.N., [81](#)  
Soyjaudah, K.M.S., [252](#), [283](#)  
Suraweera, Himal A., [44](#)  
Suttroogun, Y.K., [341](#)

**Z**

Zakari, Abubakar, [314](#), [331](#)  
Zhang, Xuewei, [44](#)