



SENSORCOMM 2015

The Ninth International Conference on Sensor Technologies and Applications

ISBN: 978-1-61208-425-1

August 23 - 28, 2015

Venice, Italy

SENSORCOMM 2015 Editors

Sergey Yurish, IFSA, Spain

Jaime Lloret Mauri, Polytechnic University of Valencia, Spain

SENSORCOMM 2015

Foreword

The Ninth International Conference on Sensor Technologies and Applications (SENSORCOMM 2015), held between August 23-28, 2015 in Venice, Italy, continued a series of events covering related topics on theory and practice on wired and wireless sensors and sensor networks.

Sensors and sensor networks have become a highly active research area because of their potential of providing diverse services to broad range of applications, not only on science and engineering, but equally importantly on issues related to critical infrastructure protection and security, health care, the environment, energy, food safety, and the potential impact on the quality of all areas of life.

Sensor networks and sensor-based systems support many applications today on the ground. Underwater operations and applications are quite limited by comparison. Most applications refer to remotely controlled submersibles and wide-area data collection systems at a coarse granularity.

Underwater sensor networks have many potential applications such a seismic imaging of undersea oilfields as a representative application. Oceanographic research is also based on the advances in underwater data collection systems.

There are specific technical aspects to realize underwater applications which cannot be borrowed from the ground-based sensors net research. Radio is not suitable for underwater systems because of extremely limited propagation. Acoustic telemetry could be used in underwater communication; however off-the-shelf acoustic modems are not recommended for underwater sensor networks with hundreds of nodes because they were designed for long-range and expensive. As the speed of light (radio) is five orders of magnitude higher than the speed of sound, there are fundamental implications of time synchronization and propagation delays for localization. Additionally, existing communication protocols are not designed to deal with long sleep times and they can't shut down and quickly restart.

In wireless sensor and micro-sensor networks, energy consumption is a key factor for the sensor lifetime and accuracy of information. Protocols and mechanisms have been proposed for energy optimization considering various communication factors and types of applications. Conserving energy and optimizing energy consumption are challenges in wireless sensor networks, requiring energy-adaptive protocols, self-organization, and balanced forwarding mechanisms.

We take here the opportunity to warmly thank all the members of the SENSORCOMM 2015 Technical Program Committee, as well as the numerous reviewers. The creation of such a high quality conference program would not have been possible without their involvement. We also kindly thank all the authors who dedicated much of their time and efforts to contribute to SENSORCOMM 2015. We truly believe that, thanks to all these efforts, the final conference program consisted of top quality contributions.

Also, this event could not have been a reality without the support of many individuals, organizations, and sponsors. We are grateful to the members of the SENSORCOMM 2015

organizing committee for their help in handling the logistics and for their work to make this professional meeting a success.

We hope that SENSORCOMM 2015 was a successful international forum for the exchange of ideas and results between academia and industry and for the promotion of progress in the area of sensor technologies and applications.

We are convinced that the participants found the event useful and communications very open. We hope Venice provided a pleasant environment during the conference and everyone saved some time for exploring this beautiful city.

SENSORCOMM 2015 Chairs:

SENSORCOMM Advisory Chairs

Jean Philippe Vasseur, Cisco Systems, Inc., France

Jaime Lloret Mauri, Polytechnic University of Valencia, Spain

Jens Martin Hovem, Norwegian University of Science and Technology, Norway

Pascal Lorenz, University of Haute Alsace, France

Sergey Yurish, IFSA, Spain

SENSORCOMM 2015 Industry Liaison Chairs

Sarfraz Khokhar, Cisco Systems, Inc., USA

Michael Niedermayer, Siemens AG, Germany

Harkirat Singh, Samsung Electronics Co., Korea

SENSORCOMM 2015 Research/Industry Chairs

Hristo Djidjev, Los Alamos National Laboratory, USA Teng Rui, National Institute of Information and Communication Technology, Japan

S. Biju Kumar, Philips Research - Eindhoven, The Netherlands

Ivan Ndip, Fraunhofer-Institut fuer Zuverlaessigkeit und Mikrointegration / Technische Universitaet (TU) Berlin, Germany

SENSORCOMM 2015 Special Area Chairs

Embedded systems

Joshua Ellul, University of Malta, Malta

Security

Yenumula Reddy, Grambling State University, USA

Body networks

Alessandro Pozzebo, Università degli Studi di Siena, Italy

Underwater systems

Mylène Toulgoat, Communications Research Centre - Ottawa, Canada

Applications

Elena Gaura, Coventry University, UK

Social sensing

Abdulrahman Alarifi, King Abdulaziz City for Science and Technology (KACST), Kingdom of Saudi Arabia

Atmospheric Icing and Sensing

Muhammad Shakeel Virk, Narvik University College, Norway

SENSORCOMM 2015

Committee

SENSORCOMM Advisory Chairs

Jean Philippe Vasseur, Cisco Systems, Inc., France
Jaime Lloret Mauri, Polytechnic University of Valencia, Spain
Jens Martin Hovem, Norwegian University of Science and Technology, Norway
Pascal Lorenz, University of Haute Alsace, France
Sergey Yurish, IFSA, Spain

SENSORCOMM 2015 Industry Liaison Chairs

Sarfraz Khokhar, Cisco Systems, Inc., USA
Michael Niedermayer, Siemens AG, Germany
Harkirat Singh, Samsung Electronics Co., Korea

SENSORCOMM 2015 Research/Industry Chairs

Hristo Djidjev, Los Alamos National Laboratory, USA
Teng Rui, National Institute of Information and Communication Technology, Japan
S. Biju Kumar, Philips Research - Eindhoven, The Netherlands
Ivan Ndip, Fraunhofer-Institut fuer Zuverlaessigkeit und Mikrointegration / Technische Universitaet (TU) Berlin, Germany

SENSORCOMM 2015 Special Area Chairs

Embedded systems

Joshua Ellul, University of Malta, Malta

Security

Yenumula Reddy, Grambling State University, USA

Body networks

Alessandro Pozzebo, Università degli Studi di Siena, Italy

Underwater systems

Mylène Toulgoat, Communications Research Centre - Ottawa, Canada

Applications

Elena Gaura, Coventry University, UK

Social sensing

Abdulrahman Alarifi, King Abdulaziz City for Science and Technology (KACST), Kingdom of Saudi Arabia

Atmospheric Icing and Sensing

Muhammad Shakeel Virk, Narvik University College, Norway

SENSORCOMM 2015 Technical Program Committee

Abdulrahman Alarifi, Computer Research Institute - King Abdulaziz City for Science and Technology, Saudi Arabia

Carlo Alberto Boano, Graz University of Technology, Austria

Mothanna Alkubely, Université de Technologie de Compiègne, France

Boushra Alkubily, Université de Technologie de Compiègne, France

Maykel Alonso Arce, CEIT and Tecnun (University of Navarra), Spain

Mansour Alsaleh, King Abdulaziz City for Science and Technology (KACST), Saudi Arabia

Tariq Alsboui, Manchester Metropolitan University, UK

Adil Al-Yasiri, University of Salford, UK

Isabelle Augé-Blum, INSA Lyon - Laboratoire CITI -Villeurbanne, France

Reza Azarderakhsh, The University of Western Ontario, Canada

Sebastian Bader, Mid Sweden University, Sweden

Dominique Barthel, Orange Labs Division R&D - Meylan, France

Novella Bartolini, "Sapienza" University of Rome, Italy

Majid Bayani Abbasy, Universidad Nacional de Costa Rica, Costa Rica

Rezaul K Begg, Victoria University, Australia

Paolo Bellavista, University of Bologna, Italy

Stephan Benecke, Technische Universität Berlin | Fraunhofer Institut Zuverlässigkeit und

Mikrointegration, Germany

Ali Berrached, University of Houston-Downtown, USA

Karabi Biswas, Indian Institute of Technology - Kharagpur, India

Alessandro Bogliolo, University of Urbino, Italy

David Boyle, Imperial College London, UK

Lina Brito, Madeira Interactive Technologies Institute (M-ITI) - University of Madeira, Portugal

Ioannis Broustis, AT&T Labs Research, U.S.A

Tiziana Calamoneri, "La Sapienza" Università di Roma, Italy

Maria-Dolores Cano Baños, Technical University of Cartagena, Spain

Juan Vicente Capella Hernández, Universidad Politécnica de Valencia, Spain

Alexandru Caracas, IBM Research - Zurich, Switzerland

Berta Carballido Villaverde, United Technologies Research Centre, Ireland

Matteo Ceriotti, University of Duisburg-Essen, Germany

Amitava Chatterjee, Jadavpur University, India

Harsha Chenji, University of Texas at Dallas, USA

Sang H. Choi, NASA Langley Research Center, USA

Hugo Coll Ferri, Polytechnic University of Valencia, Spain

Daniel Curiac, "Politehnica" University of Timisoara, Romania

David Cuartielles, Malmö University, Sweden

Juarez Bento da Silva, Universidade Federal de Santa Catarina, Brazil

Debabrata Das, International Institute of Information Technology - Bangalore, India

Danco Davcev, University for Information Science & Technology "St. Paul the Apostle" - Ohrid, Republic of Macedonia

Eli De Poorter, Ghent University - iMinds, Belgium

Jerker Delsing, Lulea University of Technology, Sweden

Behnam Dezfouli, University of Iowa, USA
Akshaye Dhawan, Ursinus College, USA
Vincenzo Di Lecce, Politecnico di Bari, Italy
Mari Carmen Domingo, Barcelona Tech University, Spain
Wan Du, Nanyang Technological University (NTU), Singapore
Sylvain Durand, LIRMM/Université Montpellier II, France
Imad H. Elhadj, American University of Beirut, Lebanon
Joshua Ellul, University of Malta, Malta
Vincent Emeakaroha, University College Cork, Ireland
Xiang Fei, Coventry University, UK
Paulo Felisberto, Institut for Systems and Robotics-Lisbon / Universidade do Algarve, Portugal
Gianluigi Ferrari, University of Parma, Italy
Armando Ferro Vázquez, Universidad del País Vasco - Bilbao, Spain
Paul Fortier, University of Massachusetts Dartmouth, USA
Leonardo Franco, Universidad de Malaga, Spain
Pedro Furtado, University of Coimbra, Portugal
Miguel Garcia Pineda, Polytechnic University of Valencia, Spain
Elena Gaura, Coventry University, UK
Hamid Gharavi, National Institute of Standards and. Technology (NIST) - Gaithersburg, USA
Chris Gniady, University of Arizona, USA
Stephane Grumbach, INRIA, France
Jayavardhana Gubbi, University of Melbourne, Australia
Jianlin Guo, Mitsubishi Electric Research Laboratories - Cambridge, USA
Malka N. Halgamuge, The University of Melbourne, Australia
Mohammad Hammoudeh, Manchester Metropolitan University, UK
Vincent Huang, Ericsson Research - Stockholm, Sweden
Muhammad Ali Imran, University of Surrey, U.K.
Abhaya Induruwa, Canterbury Christ Church University, UK
Vasanth Iyer, International Institute of Information Technology, India
Shaghayegh Jaber, Islamic Azad University - Tehran, Iran
Imad Jawhar, United Arab Emirates University - Al Ain, UAE
Zhen Jiang, West Chester University, USA
Miao Jin, University of Louisiana at Lafayette, U.S.A.
Adrian Kacso, University of Siegen, Germany
Kyoung-Don Kang, Binghamton University, USA
Reinhardt Karnapke, Brandenburg University of Technology Cottbus – Senftenberg, Germany
Dimitrios A. Karras, Chalkis Institute of Technology, Hellas
Fotis Kerasiotis, University of Patras / Rio-Patras, Greece
Yaser Khamayseh, Jordan University of Science and Technology, Jordan
Abdelmajid Khelil, Huawei European Research Center, Germany
Sarfraz Khokhar, Cisco Systems Inc., USA
Kwangsoo Kim, Electronics and Telecommunications Research Institute (ETRI), Korea
Boris Kovalerchuk, Central Washington University, USA
Thorsten Kramp, IBM Research Zurich, Switzerland
Dilip Krishnaswamy, Qualcomm Research - San Diego, U.S.A.
Danny Krizanc, Wesleyan University - Middletown, USA
Sisil Kumarawadu, University of Moratuwa, Sri Lanka
Erlend Larsen, The Norwegian Defence Research Establishment (FFI) - Kjeller, Norway

Seongsoo Lee, Soongsil University - Seoul, Korea
Pierre Leone, University of Geneva, Switzerland
Jacek Lewandowski, Coventry University, UK
Xiuqi Li, University of North Carolina at Pembroke, U.S.A
Chiu-Kuo Liang, Chung Hua University - Hsinchu, Taiwan
Qilian Liang, University of Texas at Arlington, USA
Weifa Liang, Australian National University - Canberra, Australia
Chun-Lung Lin, Industrial Technology Research Institute (ITRI) Hsinchu, Taiwan
Thomas Lindh, STH/KTH - Stockholm, Sweden
André Luiz Lins de Aquino, Federal University of Ouro Preto, Brazil
Aihua Liu, Qingdao Institute of Bioenergy & Bioprocess Technology - Chinese Academy of Sciences, China
Hai Liu, Hong Kong Baptist University, Hong Kong
Jaime Lloret Mauri, Polytechnic University of Valencia, Spain
Elsa María Macías López, University of Las Palmas de Gran Canaria, Spain
Abdallah Makhoul, Université de Besancon - Belfort, France
Gianfranco Manes, University of Florence, Italy
Vladimir Marbukh, NIST, USA
Andrew Markham, University of Oxford, UK
José Ramiro Martínez de Dios, University of Seville, Spain
Francisco Martins, University of Lisbon, Portugal
Kovatsch Matthias, ETH Zurich, Switzerland
Natarajan Meghanathan, Jackson State University, USA
Fabien Mieyeville, Ecole Centrale Lyon - INL (Institute of Nanotechnology of Lyon), France
Nader Faisal Jaafar Mohamed, UAEU, UAE
Jose M. Moya, Universidad Politécnica de Madrid, Spain
Mohammad Mozumdar, California State University, Long Beach, USA
Abderrahmen Mtibaa, Texas A&M University, U.S.A.
Umair N. Mughal, Narvik University College, Norway
Mohamad Y. Mustafa, Narvik University College, Norway
Leonid Naimark, BAE Systems, USA
Deok Hee Nam, Wilberforce University, USA
Mahmuda Naznin, Bangladesh University of Engineering and Technology - Dhaka, Bangladesh
Ivan Ndip, Fraunhofer-Institut fuer Zuverlaessigkeit und Mikrointegration / Technische Universitaet (TU) Berlin, Germany
Sarmistha Neogy, Jadavpur University, India
Edith C.-H. Ngai, Uppsala University, Sweden
Michael Niedermayer, Siemens AG, Germany
Gregory O'Hare, University College Dublin (UCD), Ireland
Brendan O'Flynn, Tyndall National Institute/University College Cork, Ireland
Cyril Onwubiko, Research Series Ltd., UK
Knut Øvsthus, Bergen University College, Norway
Carlos Enrique Palau Salvador, Universidad Politecnica de Valencia, Spain
Sung-Joon Park, Gangneung-Wonju National University, South Korea
Lorena Parra Boronat, Universitat Politècnica de Valencia (UPV), Spain
Eros Pasero, Politecnico di Torino, Italy
Animesh Pathak, Inria Paris-Rocquencourt, France
Leonidas Perlepes, University of Thessaly, Greece
Dirk Pesch, Cork institute of Technology, Ireland

Patrick Pons, CNRS-LAAS, France
Miodrag Potkonjak, University of California - Los Angeles, USA
Shrisha Rao, International Institute of Information Technology - Bangalore, India
Shahid Raza, Swedish Institute of Computer Science (SICS) – Stockholm, Sweden
Yenumula Reddy, Grambling State University, USA
Mark Reed, Yale School of Engineering & Applied Science, USA
Juan Reig Pascual, Polytechnic University of Valencia, Spain
Càndid Reig, University of Valencia, Spain
Tor Arne Reinen, SINTEF ICT, Norway
Biljana Risteska Stojkoska, University "Ss. Cyril and Methodius", Macedonia
Joel Rodrigues, University of Beira Interior, Portugal
Laura Margarita Rodríguez Peralta, Universidad Popular Autónoma del Estado de Puebla - UPAEP, Puebla, México
Lorenzo Rubio-Arjona, Universidad Politécnica de Valencia, Spain
Ulrich Rückert, Bielefeld University, Germany
Teng Rui, National Institute of Information and Communication Technology, Japan
Jorge Sá Silva, University of Coimbra, Portugal
Tapio Saarelainen, Army Academy, Finland
Sicari Sabrina, University of Insubria, Italy
Husnain Saeed, National University of Sciences & Technology (NUST), Pakistan
Addisson Salazar, Polytechnic University of Valencia, Spain
Ioakeim Samaras, Industrial Systems Institute / Intracom-Telecom - Software Development Center, Thessaloniki, Greece
Francisco Javier Sánchez Bolumar, ADIF, Spain
Olga Saukh, ETH Zurich, Switzerland
Kamran Sayrafian, National Institute of Standards & Technology, USA
Elad Michael Schiller, Chalmers University of Technology, Sweden
Leo Selavo, University of Latvia, Latvia
Sandra Sendra Compte, Polytechnic University of Valencia, Spain
Kuei-Ping Shih, Tamkang University - Taipei, Taiwan
Simone Silvestri, Sapienza University of Rome, Italy
Francesco Simeone, University "Sapienza" of Roma / INFN-Roma, Italy
Jasvinder Singh, Cork Institute of Technology, Ireland
Andrzej Skowron, University of Warsaw, Poland
K M Sunjiv Soyjaudah, University of Mauritius, Mauritius
Arvind K. Srivastava, NanoSonix Inc. - Skokie, USA
Grigore Stamatescu, University Politehnica of Bucharest, Romania
Yannis Stamatiou, University of Patras, Greece
Razvan Stanica, National Polytechnic Institute of Toulouse, France
Jack Stankovic, University of Virginia, USA
Kris Steenhaut, Vrije Universiteit Brussel, Belgium
Julinda Stefa, Sapienza University of Rome, Italy
David Stork, Rambus Labs, USA
David Sundaram, University of Auckland, New Zealand
Ewa Szynekiewicz, Warsaw University of Technology, Poland
Muhammad Tariq, Waseda University - Tokyo, Japan
Kerry Taylor, CSIRO Canberra, Australia
Rolf Thomasius, Technische Universität Berlin, Germany

Vassilis Triantafillou, Technological Educational Institution of Messolonghi, Greece
Neeta Trivedi, Aeronautical Development Establishment- Bangalore, India
Wilfried Uhring, University of Strasbourg, France
Fabrice Valois, INRIA SWING / CITI, INSA-Lyon, France
Jean Philippe Vasseur, Cisco Systems, Inc., France
Roberto Verdone, Università di Bologna, Italy
Manuela Vieira, UNINOVA/ISEL, Portugal
Muhammad Shakeel Virk, Narvik University College, Norway
Michael Walsh, Tyndall National Institute, Ireland
Wenwu Wang, University of Surrey, UK
You-Chiun Wang, National Sun Yat-sen University, Taiwan
Zhibo Wang, Wuhan University, China
Chih-Yu Wen, National Chung Hsing University - Taichung, Taiwan
Andreas Wombacher, University of Twente, Netherlands
Fang-Jing Wu, Institute for Infocomm Research (I2R), Singapore
Hui Wu, University of New South Wales, Australia
Yizhong Xu, Narvik University College, Norway
Dongfang Yang, National Research Council Canada - London, Canada
Chih-Wei Yi, National Chiao Tung University, Taiwan
Nicolas H. Younan, Mississippi State University, USA
Sergey Y. Yurish, IFSA - Barcelona, Spain
Yifeng Zhou, Communications Research Centre, Canada
Tanveer Zia, Charles Sturt University, Australia

Copyright Information

For your reference, this is the text governing the copyright release for material published by IARIA.

The copyright release is a transfer of publication rights, which allows IARIA and its partners to drive the dissemination of the published material. This allows IARIA to give articles increased visibility via distribution, inclusion in libraries, and arrangements for submission to indexes.

I, the undersigned, declare that the article is original, and that I represent the authors of this article in the copyright release matters. If this work has been done as work-for-hire, I have obtained all necessary clearances to execute a copyright release. I hereby irrevocably transfer exclusive copyright for this material to IARIA. I give IARIA permission to reproduce the work in any media format such as, but not limited to, print, digital, or electronic. I give IARIA permission to distribute the materials without restriction to any institutions or individuals. I give IARIA permission to submit the work for inclusion in article repositories as IARIA sees fit.

I, the undersigned, declare that to the best of my knowledge, the article does not contain libelous or otherwise unlawful contents or invading the right of privacy or infringing on a proprietary right.

Following the copyright release, any circulated version of the article must bear the copyright notice and any header and footer information that IARIA applies to the published article.

IARIA grants royalty-free permission to the authors to disseminate the work, under the above provisions, for any academic, commercial, or industrial use. IARIA grants royalty-free permission to any individuals or institutions to make the article available electronically, online, or in print.

IARIA acknowledges that rights to any algorithm, process, procedure, apparatus, or articles of manufacture remain with the authors and their employers.

I, the undersigned, understand that IARIA will not be liable, in contract, tort (including, without limitation, negligence), pre-contract or other representations (other than fraudulent misrepresentations) or otherwise in connection with the publication of my work.

Exception to the above is made for work-for-hire performed while employed by the government. In that case, copyright to the material remains with the said government. The rightful owners (authors and government entity) grant unlimited and unrestricted permission to IARIA, IARIA's contractors, and IARIA's partners to further distribute the work.

Table of Contents

A Real-Time Bridge Scouring Monitoring System Based on Accelerometer Sensors <i>Chih-Chyau Yang, Yi-Jie Hsieh, Ssu-Ying Chen, Wen-Ching Chen, Chih-Ting Kuo, Chen-Chia Chen, Chien-Ming Wu, and Chun-Ming Huang</i>	1
Using Energy Budgets to Reach Lifetime Goals while Compensating Dynamic Effects <i>Andre Sieber, Jorg Nolte, and Reinhardt Karnapke</i>	7
An Optimized Temperature Sensing Period for Battery Lifetime in Wireless Sensor Network <i>Seongman Jang, Keonhee Cho, Tacklim Lee, Byeongkwan Kang, and Sehyun Park</i>	14
Design and Implementation of Indoor Position Estimation System using Drone for Industrial Security <i>Sanghoon Lee, Seonki Jeon, Myeong-in Choi, Byeongkwan Kang, and Sehyun Park</i>	19
Edge-Based Technique for Ultra-Fast Gating of Large Array Imagers <i>Octavian Maciu, Wilfried Uhring, Jean-Pierre Le Normand, Jean-Baptiste Kammerer, and Foudil Dadouche</i>	24
Design Methodology of TDC on Low Cost FPGA Targets <i>Foudil Dadouche, Thimothe Turko, Wilfried Uhring, Imane Malass, Jeremy Bartringer, and Jean-Pierre Le Normand</i>	30
Temperature Sensor for Hydro Generator Bearings using Thermally Regenerated Fiber Bragg Gratings <i>Erlon Vagner da Silva, Uilian Jose Dreyer, Kleiton de Moraes Souza, Cicero Martelli, Valmir de Oliveira, Hypolito Kalinowski, and Jean Carlos Cardozo da Silva</i>	36
A Novel Elliptically-Slotted Patch Antenna-based Biosensor Design <i>Sunday Ekpo, Vijayalakshmi Velusamy, and Rupak Kharel</i>	42
Integrated Smart Glove for Hand Motion Monitoring <i>Brendan O'Flynn, Javier Torres Sanchez, James Connolly, Joan Condell, Kevin Curran, Philip Gardiner, and Barry Downes</i>	46
Ultra-miniature, Computationally Efficient Diffractive Visual-bar-position Sensor <i>Mehjabin Monjur, Leonidas Spinoulas, Patrick R. Gill, and David G. Stork</i>	52
A Distributed Scheduling Algorithm to Improve Lifetime in Wireless Sensor Network based on Geometric Placement of Sensors with Coverage and Connectivity Constraints <i>Diery Ngom, Pascal Lorenz, and Bamba Gueye</i>	58
An Algorithm to Evaluate and Build Schedules for a Distributed Sensor System with Respect to Clock Synchronization <i>Andreas Puhm, Michael Kramer, and Martin Horauer</i>	65

Self-Stabilizing Structures for Data Gathering in Wireless Sensor Networks <i>Sandra Beyer, Stefan Lohs, Jorg Nolte, Reinhardt Karnapke, and Gerry Siegemund</i>	72
Wireless Sensor Networks in Structural Health Monitoring: a Modular Approach <i>Fabio Angeletti, Mario Paoli, Ugo Maria Colesanti, and Andrea Vitaletti</i>	78
Building the O-Life Franco-Lebanese Environmental Observatory Using Sensor Web Enablement Framework: Challenges and First Approach <i>Hicham Hajj Hassan, Anne Laurent, Nicolas Arnaud, Olivier Lobry, Laurent Drapeau, and Carla Khater</i>	82
A Virtual Force Movement Scheme for Sensor Deployment in Directional Sensor Networks <i>Chiu-Kuo Liang and Yu-Shu Lo</i>	86
Classification of Human Interactions with Tools Using a Tool-mounted Wireless Sensor Node to Support Sustainable Manufacturing <i>Andreas Tilhein, The Duy Nguyen, Stephan Benecke, Eduard Wagner, Jorg Kruger, and Klaus Dieter Lang</i>	92
Evidential Network for Multi-Sensor Fusion in an Uncertain Environment <i>Eric Villeneuve, Francois Peres, Cedrik Beler, and Vicente Gonzalez-Prida</i>	98
Near Real-Time Oceanographic Data Management through Sensor Web Enablement (SWE) Standards <i>Elena Partescano, Alessandra Giorgetti, and Alberto Brosich</i>	104
Feature Selection and Interpretation of GSR and ECG Sensor Data in Biofeedback Stress Monitoring <i>Tom Gedeon, Xuanying Zhu, Leana Copeland, and Nandita Sharma</i>	106
Using Mutual Charge Scheme to Measure Salinity of Ice <i>Umair Najeeb Mughal and Bhushan Nikumbh</i>	112
Infrared Thermal Signature Evaluation of a Pure Ice Block <i>Taimur Rashid, Hassan A. Khawaja, Kare Edvarsen, and Umair N. Mughal</i>	117

A Real-Time Bridge Scouring Monitoring System Based on Accelerometer Sensors

Chih-Chyau Yang, Yi-Jie Hsieh, Ssu-Ying Chen, Wen-Ching Chen, Chih-Ting Kuo, Chen-Chia Chen,
Chien-Ming Wu, and Chun-Ming Huang
Applied Research Laboratories, National Chip Implementation Center,
Hsinchu City, Taiwan.
Email: ccyang@cic.narl.org.tw

Abstract—With the fast global climate change, many bridge structures are facing the nature disasters such as earthquakes and floods. The damage of bridges can cause the severe cost of human life and property. The heavy rain that comes with typhoon occurs in July and August in Taiwan causes the bridge scour and makes the damage or collapse for bridges. Since scour is one of the major causes for bridge failure, how to monitor the bridge scour becomes an important task in Taiwan. This paper presents a real-time bridge scour monitoring system based on accelerometer sensors. The presented sensor network consists of a gateway node and under-water sensor nodes with the wired RS-485 communication protocol. The proposed master-slave architecture of the bridge scour monitoring system owns the scalability and flexibility for mass deployment. This technique has the potential for further widespread implementation in the field. The experimental results show our sensor system can detect the bridge scour effectively with our proposed scour detection algorithm in real time.

Keywords—bridge; bridge scour; sensor network; accelerometer

I. INTRODUCTION

Bridges are the important pivots of traffic and the damage of bridges can cause the severe cost of human life and property. With the fast global climate change, many bridge structures are facing the nature disasters such as earthquakes and floods. These nature disasters cause lots of bridge collapse or destruction and thus endanger our daily life.

In Taiwan, many bridges have exceeded their 50-year life span, while many highway bridges are more than 20 years-old [1]. The strength of these old bridges is no longer affordable to the severe nature disasters. In other words, the bridges in Taiwan are likely to suffer from the damage. Scour is one of the major causes for bridge failure [2]. The heavy rains brought by the typhoon in July and August in Taiwan can cause the bridge scour and makes the damage or collapse for bridges. Thus, how to monitor the bridge health and real-time diagnose the bridge structure becomes an important task in Taiwan.

Bridge scour has been extensively studied in the world for more than a hundred years. Many methodologies and instruments have been employed to measure and monitor the local pier scour depth, such as bricks, sonar, radar sensor, Time-Domain Reflectometry (TDR), Fiber Bragg Grating (FBG) sensor and accelerometer sensor. The bricks sensors [3][4] are buried in the certain location of the sand before the rain season. After the floods, the bricks are

digged out and the number of the remained bricks is calculated. Thus, the bridge scour depth can be obtained. This method can only be used only one-time and the scour detection cannot be real-time detected. The sonar and radar sensors [4] provide contactless measurement of streambed scouring near bridge pier and abutments, and usually used to show the final status of streambed after a flood. One of disadvantages of the sonar and radar is that they have limit for measuring status of streambed in real time as rush water contained sands, even rocks during a flood. The TDR [5][6][7] measures the reflections that results from a fast-rising step pulse travelling through a measurement cable. The depth of soil-water interface is determined by counting the round trip travel time of the pulse. However, the major drawback of TDR is that accuracy of TDR is strongly dependent on the environment temperature and humidity. Monitoring the scour depth by the FBG [5] is dependent on number of FBG elements. However, the cost of monitoring of the scour depth by FBG technique is higher than that of existing methods [5]. The costs of Radar and TDR are expensive due to high-speed hardware requirement. For example, a commercial TDR (Campbell Scientific Inc., TDR100) was used to real-time monitor scour evolution, and its price is high. For FBG, optical devices such as laser, photo detectors and the optical fibers are very expensive. In addition, most of the existing methods used for scour detection are expensive and complicated, which is a major challenge for mass deployment to a lot of bridge piers. The frequency response with Fast Fourier Transform (FFT) and the time domain response with the root mean square (RMS) values of the accelerometer [8][9] are used to detect the scour. Since the accelerometer [8] does not sense the vibration data by the flow directly, the result of scour depth may be inaccurate due to the unpredictable interferences in the complicated under-water environment. Besides, in order to obtain the frequency response result [9], it may consume the large computations to get the bridge scour.

This paper presents a sensor network with accelerometer sensors to real-time detect the bridge scour with our proposed simple scour detection algorithm. The presented sensor network consists of a gateway node and under-water sensor nodes with the wired RS-485 communication protocol. The proposed master-slave architecture of the bridge scour

monitoring system owns the scalability and flexibility for mass deployment. This technique has the potential for further widespread implementation in the field. The experimental results show our sensor system can detect the bridge scour effectively with our proposed scour detection algorithm in real time.

In Section II, the operation and the proposed algorithm of scour bridge detection are introduced. In Section III, the overview of our proposed architecture is presented. In Section IV, the experimental setups and the experimental results are illustrated. Finally, we conclude this paper in Section V.

II. THE OPERATION AND PROPOSED ALGORITHM OF BRIDGE SCOUR DETECTION

In this paper, the accelerometer sensor system is presented to real-time detect the bridge scour. The accelerometers are buried into the sand of riverbed in advance. During the season of typhoon, the heavy rain that comes with typhoon causes the river full of water. The sand of the riverbed is scoured and it causes the accelerometers exposed. The accelerometers are scoured and thus vibrated due to the river water flow. The accelerometer owns the characteristics of low-cost, high sensitivity, small form factor, and low power compared with those in other instruments. With the accelerometers, the vibration can be detected easily no matter the river water is clean or mixed with sand. With the accelerometers, it is easy to setup in the field without the direction alignment.

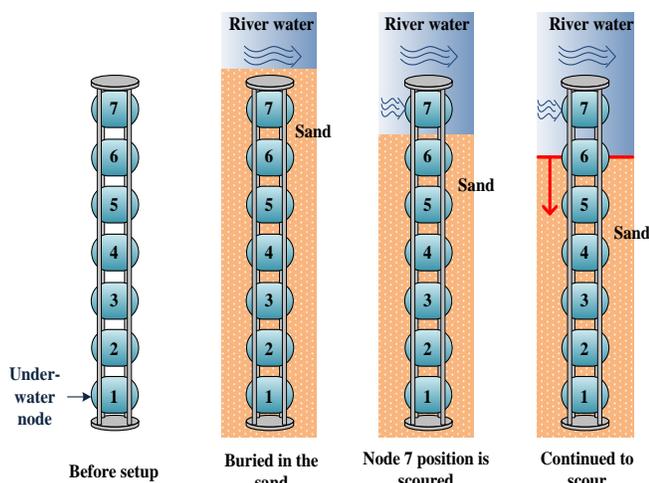


Figure 1. The operation of under-water sensor node to detect the bridge scour

The main purpose of the under-water sensors is to monitor the scouring condition of the bridge pier and riverbed. Figure 1 shows the concept of the scour detection

with accelerometer sensors. The under-water sensors are arranged equidistance and vertically fixed on the steel shelf. The under-water sensors are then buried deeply in the riverbed close to the bridge pier. In the normal condition, the sand of the riverbed can fully cover the under-sensor, and the sensor nodes are in a steady state condition. When the water of the river becomes rapid due to the storm or heavy rain, it washes away part of the riverbed and the sensors originally buried in the sand becomes exposed and vibrated due to the scouring. The deeper the riverbed gets scoured, the more sensors are exposed. The vibration data of each sensor will be real-time sent to the data logger through the Ethernet and the host program help to identify the scouring degree. To keep track of the scouring condition of the riverbed in long terms, it can provide reference information of the stability of the bridge pier, and achieve the purpose of disaster prevention.

TABLE I. THE ALGORITHM FOR BRIDGE SCOUR DETECTION LOOP

Scour Detection Loop	
1	LOOP
2	FOR t=0 TO (N-1)
3	Node(Ax, Ay, Az) = GetAcclInfo() ;
4	(μ_x, μ_y, μ_z) = GetMean (Ax, Ay, Az);
5	$\Delta\mu_x = \mu_x - \mu_{x0} $; $\Delta\mu_y = \mu_y - \mu_{y0} $; $\Delta\mu_z = \mu_z - \mu_{z0} $;
6	IF ($\Delta\mu_x > \mu_{ThD}$) OR ($\Delta\mu_y > \mu_{ThD}$) OR ($\Delta\mu_z > \mu_{ThD}$)
7	Node_Scoured =ON;
8	ELSE
9	Node_Scoured =OFF;
10	IF (i >= Alarm_ThD) THEN
11	AlarmTrigger() ;
12	GOTO LOOP;

Table 1 illustrates our proposed algorithm to detect the bridge scour with accelerometers. The algorithm consists of scour detection loop for each sensor node. Each accelerometer executes the scour detection loop. In this detection loop, the host program acquires the N-point accelerometer value. These N-point values are averaged and then subtracted by the initial accelerometer value to obtain the absolute difference value. If the absolute difference value between current accelerometer and initial accelerometer exceeds the threshold value, then the sensor node is labeled as the scoured status. Otherwise, the sensor node is labeled as the status of non-scoured. If the position of the scoured sensor node exceeds the position of alarm threshold, the alarm is triggered. Note that the threshold value and the position of alarm threshold are obtained from the experiment in the lab. Figure 2 illustrates the flow chart of the proposed algorithm for bridge scour detection loop. Note that the threshold value for the absolute difference of accelerometer values is obtained from the experiments in the laboratory.

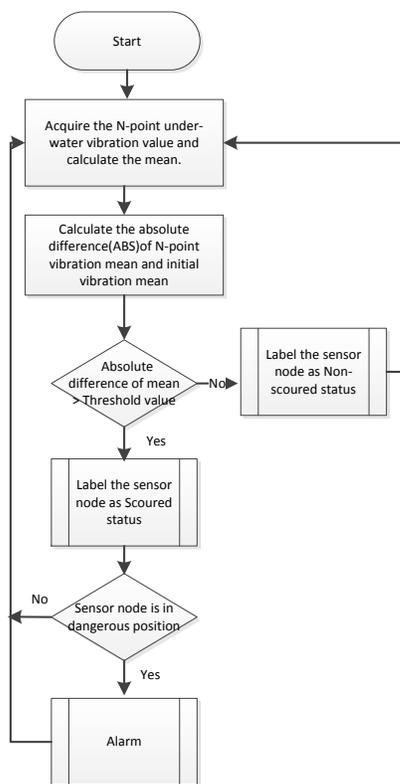


Figure 2. The flow chart to detect the bridge scour with accelerometers

III. DEVELOPMENT OF REAL-TIME BRIDGE PIER SCOUR MONITORING SYSTEM

Architecture of bridge scouring monitoring system

The architecture of real-time bridge scour monitoring system shown in Figure 3 is based on master-slave configuration. A master sends commands to slave for controlling sensor nodes and accessing sensor data. The data logger communicates with gateway through Power over Ethernet (POE) switch. The data logger sends a command to the gateway. When the gateway receives command, the gateway converted Ethernet command to RS485 command. After converting nodes, the gateway broadcasts it to all the sensor nodes. Since the command packet contains unique sensor ID, only the specific sensor node returns the sensor data to the data logger. We adopt the accelerometer sensor module in our sensor node. The POE switch is connected with 48V battery (3 packs in series for 48V with individual 16V lithium iron phosphate battery).

Gateway and Sensor nodes

The gateway node is comprised of two stacked PCBs – a power module and a core module (see Figure 4). The top board is the power module, which operates as a DC-DC converter for generating 1.2~5V outputs from the 48V input. An Ethernet PHY (TI, DP83640) is used to send/receive Ethernet data from POE switch, and send/receive the signals and power to sensor nodes through RS485 interface (ADI,

ADM2682E). The core module is composed of a Cortex-R4 Mico Controller Unit (MCU, TI, RM48L952) and a FPGA (Xilinx, Spartan-6). Ethernet data and RS485 data are processed by the Cortex-R4 MCU and the FPGA, respectively. The FPGA mainly is used to translate the sensor data from serial format to parallel format. Three signals (Int, Rdy, En) are used to control the operation between the FPGA and the Cortex-R4 MCU. The FPGA receives the sensor data in 8-bit as a unit. After the FPGA collects 8-bit data, the FPGA deposits to register, then send Int signal to the Cortex-R4 MCU, and then notifies the Cortex-R4 MCU to receive sensor data. After the Cortex-R4 MCU receives 8-bit data, the Cortex-R4 MCU sets the Rdy signal to send it to the FPGA. The FPGA En Signal is set to “0” to indicate that the sensor data has been transferred completely.

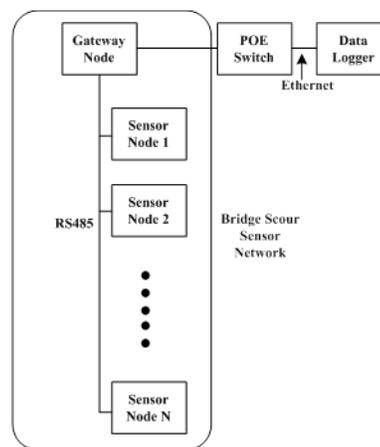


Figure 3. The architecture of real-time bridge scouring monitoring system



Figure 4. The pictures of printed circuit boards of gateway (left) and sensor node (right).

The configuration of sensor node is similar to that in the gateway node. The Cortex-R4 MCU is used to access sensor data through Serial Peripheral Interface (SPI) interface and the Field Programmable Gate Array (FPGA) is used to process RS485 data. The block diagram of the FPGA in sensor node is shown in Figure 5.

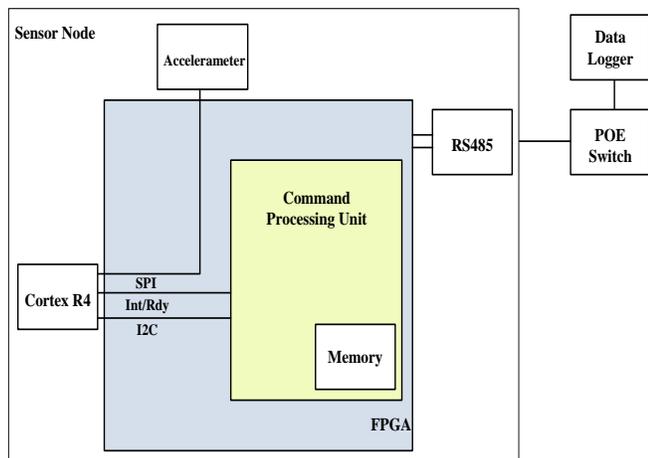


Figure 5. Block diagram of FPGA in sensor node

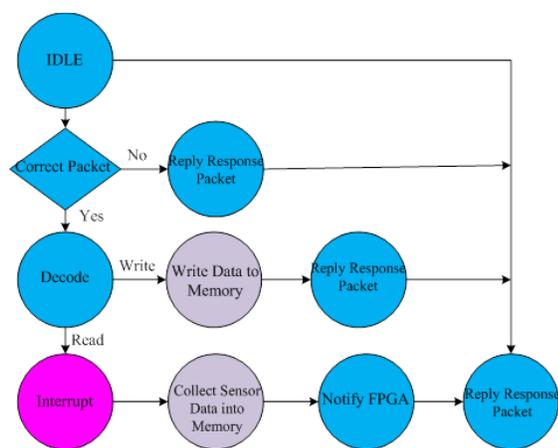


Figure 6. Processing sequence of FPGA and MCU

The FPGA parses receives commands, executes part of commands, and responses to the data logger. The Cortex-R4 MCU takes charge of collecting sensor data. Figure 6 describes processing sequence of the FPGA and the Cortex-R4 MCU. In Figure 6, the steps with blue color are tasks of the FPGA, those steps with purple color are memory related tasks, and those with red color are the tasks of the Cortex-R4 MCU. In the case that the data logger requests sensor data, the FPGA will receive a Read command. The FPGA then parses and decodes the command and is aware that cooperation with the Cortex-R4 MCU is necessary. The FPGA puts this command in memory and notifies the Cortex-R4 MCU with an interrupt. The Cortex-R4 MCU reads command from memory via I2C interface, and then collects sensor data and stores them in memory. After the data collection is done, the Cortex-R4 MCU notifies the FPGA by a General-Purpose Input/output (GPIO) signal. Then, the FPGA reads data from memory and generates response to the data logger.

Accelerometer sensor module

The core module of the sensor node which is connected to an accelerometer (ADI, ADXL345) module which is used in this study is widely available online. Figure 7 shows the top-view and bottom-view pictures of the accelerometer module. The accelerometer is read by Cortex-R4 MCU via the SPI interface. The sensor data is then sent back to data logger.

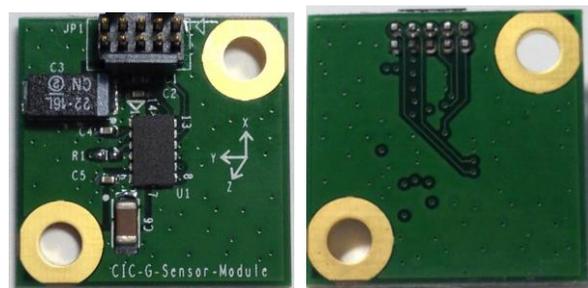


Figure 7. Top and bottom-view pictures of the accelerometer module

IV. THE EXPERIMENTAL SETUP AND RESULTS

The accelerometer module is fixed on thin metal strip with thickness of 0.3 mm, as shown in Figure 8. The accelerometer module is filled with silicon to be water-proof. Figure 9 shows the picture of setup of real-time bridge scouring monitoring system. The accelerometer sensor module is installed along the pier model. The 48V battery, control circuits of gateway and sensors nodes and cables are setup near the laboratory flume.

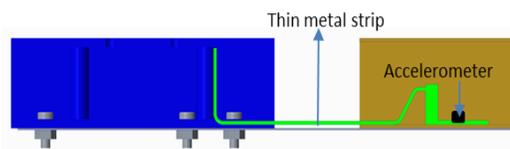


Figure 8. The drawing of house for accelerometer module

The monitoring bridge scour erosion detection is carried out in a recirculating laboratory flume (length = 36m, width = 1 m, depth = 1.1 m) at Hydrotech Research Institute of National Taiwan University, Taiwan [10]. The layout of the flume and experimental setup are shown in Figure 10. A false test bed has a sediment recess (length = 2.8 m, width = 1 m, depth = 0.3 m) which is filled by nearly uniform sediment. A 15-cm-diameter hollow cylindrical pier made of plexiglas is located at the middle of the recess. An inlet valve and a tailgate are used to regulate depths of flow and flow speed.

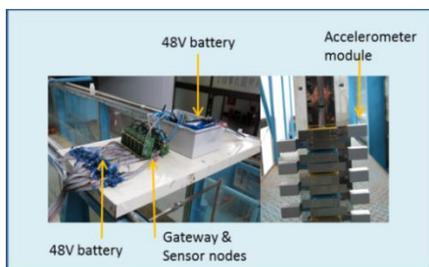


Figure 9. The photos of setup of real-time bridge pier scouring monitoring system.

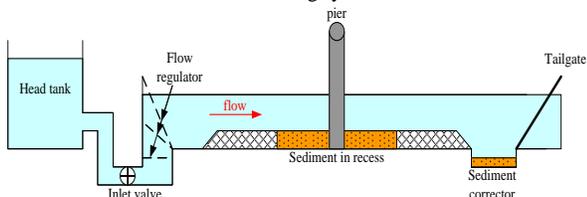


Figure 10. Partial layout of recirculating laboratory flume

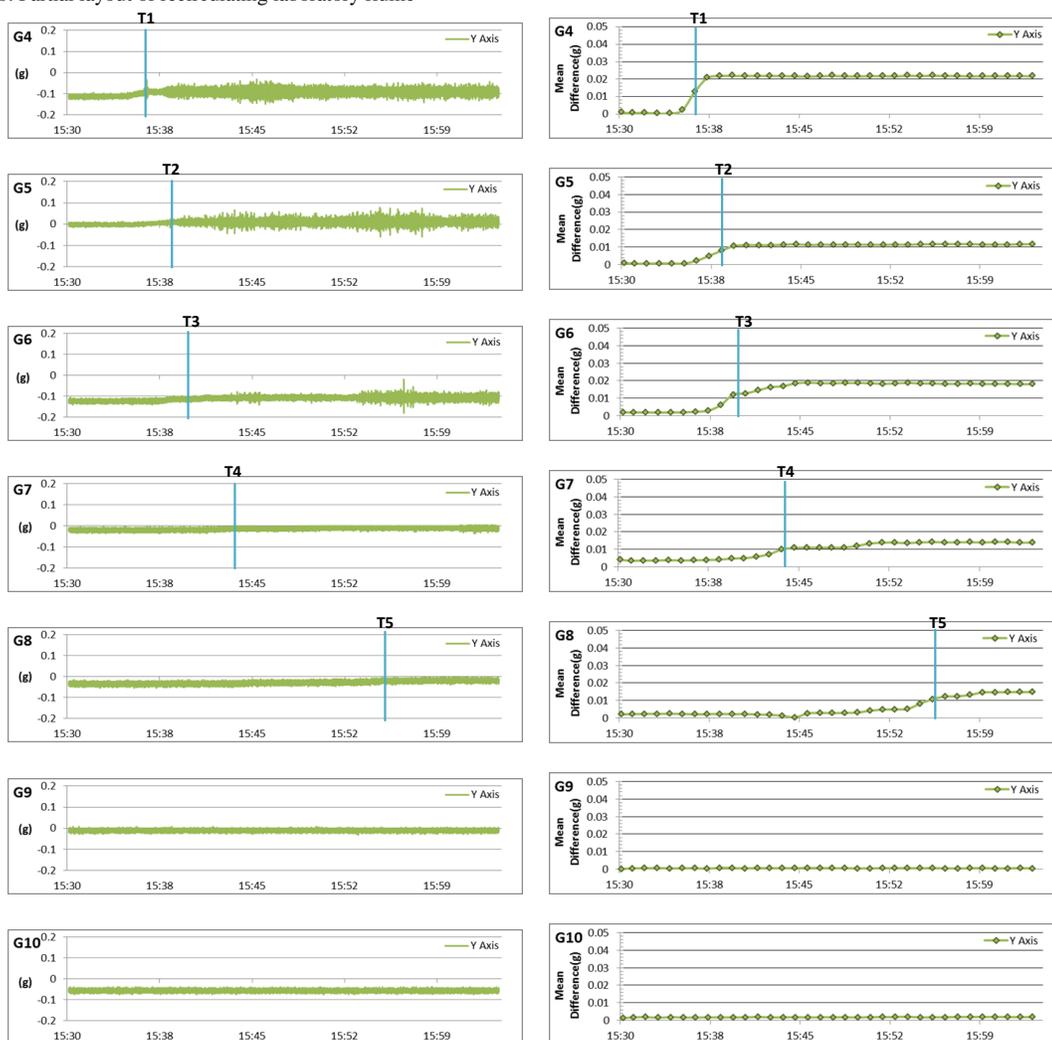


Figure 11. The experimental results of time-domain vibration raw data and absolute difference vibration data

Figure 11 shows the experimental results. The left figure shows the time-domain vibration raw data while the right figure shows the absolute difference vibration data.

First, we discuss the time-domain vibration data obtained in the scouring experiment. At first, the sensors of G4, G5, G6, G7, G8, G9 and G10 are buried in the sand. At the time of T_i , the G_k sensor starts to be scoured by water and exposed from the sand; T_i refers to the times $\{T_1, T_2, T_3, T_4, T_5\}$ and G_k refers to the sensor nodes $\{G_4, G_5, G_6, G_7, G_8\}$, respectively. The accelerometer data of G7 and G8 starts to change little from T_4 and T_5 respectively. We find the the G7 and G8 are scoured and exposed from the sand; however it's not easy to obtain the correct scouring information from the vibration raw data in the time domain. For the sensor of G9 and G10, they are buried in the sand in this experiment, so that we cannot observe the change of the vibration data for these two accelerometers.

The right figure of Figure 11 shows the absolute difference vibration data. We utilize our proposed algorithm in Table I to detect the bridge scour. The threshold absolute difference of vibration data is set to 0.01 according to the experiment results. At the time of T_i , the G_k sensor starts to be scoured by the water and starts to be exposed from the sand. The value of absolute difference vibration data for G_k sensor node starts to be larger than the threshold value from T_i . ; T_i refers to the times $\{T_1, T_2, T_3, T_4, T_5\}$ and G_k refers to the sensor nodes $\{G_4, G_5, G_6, G_7, G_8\}$, respectively. By using the proposed algorithm shown at Table 1 and the proposed flow chart shown in Figure 2, the bridge scour detection can be easily realized. This sensor system and proposed algorithm will be utilized for the mass production deployment in the field in the near future.

V. CONCLUSIONS

Bridges are the important pivots of traffic, and the damage of bridges can cause the severe cost of human life and property. The heavy rain that comes with typhoon occurs in July and August in Taiwan often causes the bridge scour and makes the damage or collapse for bridges. Since scour is one of the major causes for bridge failure, how to monitor the bridge scour becomes an important task in Taiwan. This paper presents a real-time bridge scour monitoring system based on accelerometer sensors. The proposed sensor network consists of a gateway node and under-water sensor nodes with the wired RS-485 communication protocol. With the proposed scour detection algorithm, the system can detect the bridge scour effectively in real time. The proposed master-slave architecture of bridge pier scour monitoring system has scalability and flexibility for mass deployment. This technique has the potential for future widespread implementation in the field.

REFERENCE

- [1] T. Y. Liu, W.L. Chiang, C.W. Chen, W.K. Hsu, L.C. Lu and T. J. Chu, "Identification and Monitoring of bridge health from ambient vibration data," *Journal of Vibration and Control*, , Nov. 2010, pp. 1-15.
- [2] J.A. Puleo and J. T. Hayden, "A Near Real-time Scour Monitoring System at Indian River Inlet, Delaware, USA," *IEEE Bilox-Marine Technology for Our Future: Global and Local Challenges* , Oct. 2009, pp. 1-10.
- [3] J. Lu, J. Hong, C. Su, C. Wang, and J. Lai, "Field Measurements and Simulation of Bridge Scour Depth Variations during Floods," *Journal of Hydraulic Engineering*, 2008, pp. 810–821.
- [4] H. C. Yang and C. C. Su, "Real-time River Bed Scour Monitoring and Synchronous Maximum Depth Data Collected During Typhoon Soulik," *Hydrological Processes*, vol. 29, 2015, pp. 1056-1063.
- [5] L. J. Prendergast, "A review of bridge scour monitoring techniques," *Journal of Rock Mechanics and Geotechnical Engineering*, 2014, pp. 138-149.
- [6] H. Wang, S. C. Hsieh, C. Lin, and C. Y. Wang, "Forensic Diagnosis on Flood-Induced Bridge Failure. I: Determination of the Possible Causes of Failure," *Journal of Performance of Constructed Facilities*, vol. 28, 2014, pp. 76-78.
- [7] J. Tao, X. Yu, and X. B. Yu, "Real-time TDR Field Bridge Scour Monitoring System," *Structures Congress*, 2013, pp. 2996-3009.
- [8] J. L. Briaud, S. Hurlebaus, K. Chang, C. Yao, H. Sharma, O. Yu, et al., "Realtime Monitoring of bridge scour using remote monitoring technology," *FHWA/TX-11/0-6060-1*, Texas Transportation Institute, Austin, USA, 2011.
- [9] M. Fisher, S. Atamturktur, A. Khan, "A Novel Vibration-based Monitoring Technique for Bridge Pier and Abutment Scour," *Structure Health Monitoring*, vol. 12, 2013, pp. 114-125.
- [10] Y. B. Lin, J.-S. Lai, K. C. Chang, W.-Y. Chang, F.-Z. Lee, and Y.-C. Tan, "Using MEMS Sensor in the Bridge Scour Monitoring System," *Journal of the Chinese Institute of Engineers*, vol. 33, Jan, 2010, pp. 25-35.

Using Energy Budgets to Reach Lifetime Goals while Compensating Dynamic Effects

André Sieber, Jörg Nolte

Distributed Systems/Operating Systems Group
Brandenburg University of Technology Cottbus-Senftenberg
Cottbus, Germany
Email: {as, jon}@informatik.tu-cottbus.de

Reinhardt Karnapke

Communication and Operating Systems Group
Technische Universität Berlin
Berlin, Germany
Email: karnapke@tu-berlin.de

Abstract—Nodes within sensor networks often have tight bound goals for the lifetime while running from a non-renewable energy source. Variations within the hardware or induced by the software complicate the prediction of the energy consumption. Additionally, batteries are vulnerable to temperature and non-linear effects. To reach certain lifetime goals under these influences without sacrificing energy due to pessimistic estimations, online energy management is necessary. In this paper, we present policies to control the behavior of applications and devices using energy budgets. First experiments yield promising results, with nodes reaching their lifetime goals while maintaining a high application quality.

Keywords—Wireless Sensor Networks; Energy Management; Energy Awareness; Lifetime Goals

I. INTRODUCTION

Wireless sensor networks are being used in a wide range of applications, with environmental monitoring and/or hazard detection among the most prominent ones. In many of these monitoring applications, the sensor networks should run for a certain time without direct human interaction. This requirement may result from the network being placed in a hardly reachable area, in a dangerous area (e.g., in a steelworks) or because human interference should be kept to an absolute minimum (e.g., when monitoring nesting animals). For these applications, the sensor networks are usually deployed with a certain time-to-live in mind, until either the next maintenance is due and batteries can be changed, or until the whole monitoring period (e.g., the breeding season) is over.

When application developers design networks with lifetime goals in mind, they carefully choose the appropriate sensing equipment, devise a communication pattern and select batteries that supply enough power. However, in some cases the usage of sufficient batteries is impossible, e.g., due to the form factor. If the sensor nodes need to fit in a certain area, the size of the batteries is limited. Even if there is no limiting form factor, there are a number of influences which can not be calculated easily. Sensor node hardware is subject to production variances, resulting in different power consumption characteristics even for components manufactured together. The same is true for batteries, only in reverse: The amount of energy they supply varies. Moreover, supply and consumption can both vary, depending on environmental conditions like temperature, the battery voltage level, or the dynamic voltage converter efficiency. Even more, batteries suffer from non-linear effects like rate-capacity and the recovery effect [1].

Sensor network applications that fail to take variances and run-time-effects into account may lead to early node failures

due to depleted energy. To prevent this, the required energy is often overestimated on purpose. The amount of energy for reaching the lifetime goal is calculated and a safety margin (e.g., 20%) is added afterwards. Please note that these 20% are calculated using the nodes with the highest assumed consumption. When, for example, the communication represents the main energy drain, the nodes closest to the sink would be used as basis, as they need to forward messages more often than outer nodes. Even though this results in wasted energy, it is often thought of as being necessary. However, the quality of service supplied by the application can be improved at least on some nodes, if it is possible to use all available energy. In the example, some outer nodes could sample/sense more often to increase the quality of the data, e.g., by calculating averages or max values, without increasing the network load.

In order to avoid energy being wasted, dynamic energy management is necessary, with a manager that can use various handles to influence the energy consumption on each node individually, based on the available energy. The application duty cycle can be changed or the MAC/routing timings adapted. Sensors can be used with varying levels of detail or granularity. However, this requires the possibility of isolating tasks and devices. To enable the management to limit their consumption, we introduce fine grained energy budgets. These budgets are assigned to individual (sub-)tasks, provide a certain amount of energy, and enable the energy management to plan how to distribute the scarce resources.

Even though the main goal is to reach a certain lifetime with the whole network, the methods presented in this paper are primarily concerned with the options available on each individual node. The reasons for this are twofold:

- 1) Reaching the lifetime goal with the whole network requires each node to reach the lifetime goal.
- 2) Network wide strategies require communication, which in itself induces more energy consumption. However, the presented approach does not obstruct a global strategy, e.g., adaption of routes to distribute the communication load. The presented approach can even compensate an increase or decrease in communication load, as it only represents another dynamic influence.

Figure 1 shows an example network with a grid topology containing 11 times 11 nodes. Each node has a variable sensing interval and may need to forward messages from neighboring nodes. The routing topology is a tree, resulting in a higher communication load on nodes closer to the sink. Therefore, these nodes need to adjust their sensing interval in order to

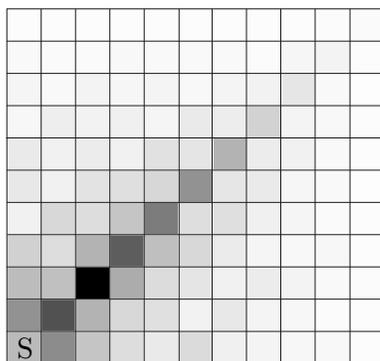


Figure 1. Difference of Sensing Interval due to balancing the load with a non-uniform network load. Sensing interval ranges from 1.71s (white) to 3.36s (black)

conserve energy and reach the lifetime goal. In the example, outer nodes sample every 1.71 seconds while the inner nodes sample less often, the longest interval is nearly twice as long (3.36 seconds).

The rest of this paper is structured as follows: In Section II related work is presented. Section III focuses on the energy budgets principles and our approaches to react on changing energy demands and extreme situations. In Section IV early evaluation results are shown. Finally, a conclusion is given in Section V.

II. RELATED WORK

Energy management on sensor nodes traditionally faces three questions:

- 1) How much energy is available within the energy source (usually translating to the state-of-charge of the battery)?
- 2) How much energy is consumed by the (different parts of the) application?
- 3) How much of the available energy can be spent and how is it partitioned between consuming parts to aid the application goal?

There are three kinds of approaches to answer the first question. Hardware based approaches like fuel gauges, e.g., [2] or smart battery monitors like [3] supply accurate information about the remaining energy in batteries but in turn increase the system complexity and energy consumption. Model based approaches, e.g., [4], [5], use chemical or analytic models to predict the behavior of the energy source. They are very accurate but suffer from high complexity and computation costs, which renders them unusable for deeply embedded systems. They also often rely on complex parameters which have to be retrieved via experiments and thus depend on a certain type of battery. Measurement based approaches use easily observable parameters, often the voltage, to estimate the state of charge. They can be based on tables generated in advance [6] or observe the relative voltage decline [7]. Due to the limited resources and energy constraints in wireless sensor networks, we favor the approach presented in [7].

Tracking the consumed energy and thus answering the second question is possible both in hard- and software. Hardware approaches cover Coulomb counters and smart battery systems

[3], *Sensor Node Management Devices* [8] and specifically designed measurement devices [9], [10]. All hardware approaches introduce an overhead in device costs and energy consumption. Software approaches rely on the observation of certain events to account for the consumption. To be able to observe these events, the code has to be modified with hooks to call the accounting functions. The events can either be based on functional application blocks [11], [12] or on device driver actions [13], [14]. The latter are based on taking time stamps when devices change their state to obtain the duration of each state and calculate the consumed energy. While the hardware based approaches deliver precise results, we favor a software based approach using time stamps [15], as it is capable of delivering not only the global (node wide) consumption but can also distinguish between individual system elements, including software components. In contrast to other approaches, it is capable of taking varying consumption due to voltage changes of the battery and varying efficiency of potentially used voltage converters into account.

There are various approaches for managing the energy consumption and answering the third question. Table I shows an overview over existing approaches and a simple categorization based on the platform they are ran on, the scope of the approach and the controlled entity/granularity. The management can be done either with a direct influence on system and application parameters (upper half) or indirect, using a limited resource distributed among application parts (lower half).

Providing the application with energy awareness by using an indirect approach enables the use of different adaption strategies. Instead of providing a handle to the management, the application itself can manage its consumption. Apart from using individual service levels and controlling timers, more complex schemes, like limiting the number of forwarded messages, are possible. This makes indirect approaches more flexible. Based on the concept of resource containers (RC) [27], energy capsules [28] and energy containers [29] provide information about the energy usage of (sub-)tasks, but do not take advantage for energy management from them. The indirect approaches of Cinder [21] and ECOSystem [22] utilize this concept to store information about energy availability.

Apart from the approaches that are based on setting the service level, EPOS [17], Nemesis [25] and SORA [26] also suffer from insufficient isolation between application parts, resulting in a potentially uncontrolled impact of a changed

TABLE I. OVERVIEW OF ENERGY MANAGEMENT APPROACHES. DIRECT APPROACHES IN UPPER HALF, INDIRECT IN LOWER HALF.

	platform	scope	controlled entity
Odyssey [16]	PC	local	Service Level
EPOS [17]	WSN	local	Task
Energy Levels [6]	WSN	local	Service Level
Eon [18]	WSN	local	Timer/Service Level
EMA [19]	WSN	local	Timer
IDEA [20]	WSN	neighborhood	Service Level
Cinder [21]	mobile phone	local	energy distribution
ECOSystem [22]	PC	local	energy distribution
Pixi OS [23]	WSN	local	energy distribution
Virtual Battery [24]	WSN	local	energy distribution
Nemesis OS [25]	PC	local	price
SORA [26]	WSN	global	price

demand, up to starvation. Only Virtual Battery [24] and Pixi OS [23] resolve short-term energy shortages by allowing debts. Most approaches lack the flexibility to resolve long-term changes of the energy demand of individual application parts. Cinder, EMA [19] and Pixi OS have a high computational overhead as they work continuously instead of periodically. Approaches that do not only consider local information but include neighborhood (IDEA [20]) and global information (SORA [26]) suffer from increased energy consumption due to the communication overhead. Additionally, the preparation effort depends on the used energy accounting and the parameters needed by the management. IDEA, SORA and Energy Levels have a relatively high demand.

All existing approaches have different benefits and drawbacks. Especially the reaction to changes in the energy demand of single application(-parts) is often inadequately addressed. However, the management must be capable of dynamically adjusting single application(-parts). Fine-grained management makes isolation between single applications or application parts necessary to prevent starvation. Indirect management enables energy-aware applications, but needs to provide mechanisms of direct management to aid the applications through functions for calculating duty cycles and service levels. Using a periodic approach reduces overhead and the necessary planning horizon of the applications.

III. MANAGEMENT USING ENERGY BUDGETS

Using information about the energy consumption of the system, it is possible to enforce limits to meet desired discharge rates. Applications should react to reach lifetime goals, therefore they must be provided with information about the available energy. With different application parts and goals competing for the energy, a mechanism which separates them and provides a local view of the remaining energy for a single part is necessary. However, the allocated energy also needs to be associated with a time window. This information can be provided using energy budgets, which are also based on the concept of resource containers [27], but are only used to limit a single resource, namely energy, here.

An energy budget is an abstract reservoir for energy. It represents the right to a certain amount of energy. The system energy (or parts of it) is/are divided between the budgets. An abstract budget B is defined by its currently stored amount b and the validity interval $[t_{start}, t_{end}]$ of b . The demand of a budget is defined by the minimal energy min needed and the maximal energy max consumable by its associated consumers during $[t_{start}, t_{end}]$. They give the system a hint on reasonable values for filling Budget B at t_{start} (1).

$$B = (b, [t_{start}, t_{end}], min, max) \quad (1)$$

As most activities within a sensor node live rather long, the validity interval is more or less a constant refresh interval at whose begin the budget is refilled. This also divides the system consumption and reduces the prediction horizon for the application. The essential requirement for a reliable operation of the application is that the minimal demand for all budgets is satisfied in every interval. In scenarios using harvesting, this requirement may not be fulfilled due to the unsteady energy income. In times with no income, nodes must be able to perform only essential tasks and wait for more income. The

remaining energy can be distributed among the budgets in various ways (discussed later). If no energy harvesting is available, the network's maintainer must be informed immediately if this demand can not be satisfied, as the functionality of the node or even the whole network can no longer be guaranteed.

Obtaining reasonable values for min and max can be hard, as it involves the determination of boundaries of the energy consumption of individual application parts. Apart from careful calculation, these values could be obtained by simulation or simply by experiments. This process could also be automated to a certain degree by including a learning phase. But if the demand fluctuates or is based on spontaneous events, min and max may represent the demand inadequately. Thus, the management must be able to adjust the distributed amount of energy and other forms of cooperation, e.g., lending energy, can be utilized by the budgets (see below).

Using multiple budgets in parallel is not a problem, as long as different devices are used. Shared resources like the CPU need to be accounted by using requests. Fairness, as called for in [29], can be achieved by logging the input, e.g., in the form of transmitted bytes, and then dividing the total consumed energy by the total number of bytes transmitted. Then, each task can be charged with this value times the number of bytes it transmitted. If all requests always result in the same consumption, this approach can be simplified to counting the number of requests. It can also be applied to buffered operations, like writing onto an SD-card. Then, the energy would be subtracted from the budget before the actual operation takes place.

Energy budgets can be mapped to different entities. In some cases, they might be associated with a device or a certain device mode. This also enables task- or activity granularity. On a larger scale, they can also be applied to whole execution paths of a program, e.g., a message that is created, a routing decision made, the MAC involved and the radio hardware activated.

The basic concept of energy budgets and their behavior can be extended and policies implemented in numerous ways. Apart from the way the energy is distributed among the budgets, there are different ways to approach special situations when the energy is insufficient or not consumed. Which policy or combination should be used depends mainly on the requirements of the application. Our initial evaluation shows that with different policies, the behavior of the system varies widely in different situations.

A. Energy Allowance

The main goal of energy distribution is to reach the lifetime goal while distributing the energy in a fair way. In this case, fair means that for all budgets at least their minimum requirement is allocated and no budget is favored unintentionally. There may be user defined utilities (U_i) as well as boundary conditions (min_i and max_i) which have to be taken into account and can lead to an uneven distribution of energy. The energy distribution into budgets boils down to a resource allocation problem, where resources are allocated based on utilities (2).

$$max \sum U_i(E_i) \quad (2)$$

subject to:

$$min_i \leq E_i \leq max_i$$

$$\sum E_i \leq E$$

The utility function U_i represents the amount of useful work that can be gained by adding more energy to the budget.

This allocation problem can be solved with brute force or with linear optimization, but these approaches are hard to realize on resource constrained sensor nodes. Additionally, the formal problem description can lead to a solution where budgets with the highest utility get all the energy. Even though this is a valid solution, it is an unwanted one. An additional constraint is necessary which takes the utility as weight for a fair mix of all budgets. This is possible with a linear utility function. By distributing the minimum demand first, the problem can be eased. The additional energy available can then be scattered, e.g., in percent of the max requirement of each task (3):

$$b_i = b_i + \frac{(max_i - b_i) * U_i}{\sum (max - b) * U} * E \quad (3)$$

This heuristic can lead to situations where the share of a budget is higher than its max_i . Then, the amount exceeding the maximum remains in the global energy pool and equation 3 is applied again.

Since the available energy as well as the demand and utility of application parts can change, the allowance must be done frequently. The allowance frequency again depends on the application demand and scenario. With increased consumption, the frequency should be also increased, to be able to react to changes in the available and requested energy in a timely manner.

B. Extreme Situations

In some cases, the calculation of the remaining energy might reveal that there is not enough energy left to fulfill the requirements of all tasks until the end of the lifetime goal. Then, the choice remains to either divide the energy strictly and not fulfill the requirements in some intervals, or to continue as before, accepting a possible early node failure but informing the network maintainer. In other cases, a budget might have received more energy than it can consume. For the associated task, this is not a problem. However, the feedback to the accounting might lead to wrong decisions, as the observed state of charge (SoC) suddenly is higher than anticipated. Once the task starts consuming all allocated energy within its budget, the battery is drained faster than the management expects. In a third case, a task might suddenly need more energy than in the previous intervals, e.g., for additional message retransmissions. If this amount is higher than the budget allocated by the management, the task might steal energy from a different task if its priority is higher. It could also borrow energy from the system or another task that does not need all of its budget in this interval, but would have to return the favor later.

C. The Duty of the Application

All of the methods described above are used to distribute energy, and give the application incentives to change its behavior if there is not enough or more than the required amount of energy available. The application itself needs to react to these

incentives, for example by adapting its duty cycle. Devices that are not directly influenced must also be taken into account by the application (e.g., the radio). As described above, the easiest way to guarantee fairness is to calculate costs for shared resources based on the number of bytes or requests.

The number of budgets used within the system depends mainly on the application. While it is possible to use only a single budget for a full sensor network application, including sensing and communication, the isolation of application parts eases the adaption decision, as only the relevant parts must be considered.

As literature shows, there are two common ways to adapt to a changed energy availability. First, the application's duty cycle can be changed within a tolerable range. Second, the application can adapt by using distinct service levels. While the first is appropriate for sensing tasks, the latter allows more complex changes to the behavior of the application. As the necessary information is provided by the budgets, both can be implemented easily.

The duty cycle τ is computed based on the remaining time d before the budget is refilled, the consumption of the last execution c and the energy available b_i (4).

$$\tau = \lfloor \frac{d * c}{b_i} \rfloor \quad (4)$$

The service level σ is calculated as ratio of how much the budget is filled compared to min and max (5).

$$\sigma = \begin{cases} 1 & \text{if } b_i \leq min_i \\ N & \text{if } b_i \geq max_i \\ \frac{N * (b_i - min)}{max - min} & \text{else} \end{cases} \quad (5)$$

IV. EVALUATION AND RESULTS

In order to evaluate and compare the different approaches, we used simulations and experiments with real sensor nodes. The experiments were realized using the so-called FeuerWhere nodes [30]. These nodes feature a MSP430 micro controller and three transceivers, one operating at 868 MHz and two operating at 2.4 GHz. We used REFLEX [31], an operating system for deeply embedded systems, as basis for the integration of our energy budgets. As real experiments are better suited to prove that our approach works, we focus only on those in the following evaluations.

A. Cost of Energy Allowance

To evaluate the computational overhead of the allowance heuristic, the runtime of the algorithm with different options was measured. Figure 2 shows the results.

While using only the heuristic (equation 3) has the lowest runtime costs, it alone can not guarantee the minimal demand (A). Distributing the minimal demand of each budget first (B) increases the runtime only slightly but fulfills the requirements. To increase the dynamic between the budgets, it is feasible to take the consumed and requested energy of the budgets into account (C). Distributed between the minimal demand and the heuristic, the runtime increases considerably, as the additional step is a modified version of the heuristic itself.

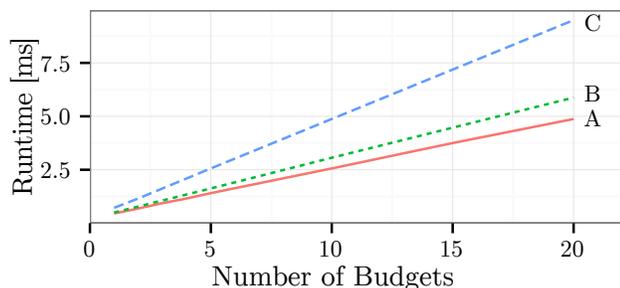


Figure 2. Runtime of the distribution algorithm for MSP430 running at 16MHz. (A) Based on percentage of maximal requested energy. (B) Additionally distribution of minimal requested energy. (C) Distribution of consumed/requested energy.

With increased number of budgets the runtime rises linearly. As the allocation of energy only happens every few hours in most scenarios, the overhead is negligible even when a lot of budgets are in use.

B. Influence of different policies

To evaluate the influence of the policies in situations where the energy demand changes, we used two scenarios. Two tasks were involved, one sampling and one transmitting messages. The experiments ran for three phases with 50 seconds per phase. The sending task wanted to transmit a message every ten seconds while the sampling interval was adjusted dynamically.

1) *Scenario 1:* In the first scenario the amount of energy allocated to the sending task in each phase was only sufficient to send four of the five desired messages, forcing it to adapt according to the chosen policy: Isolation, adaption, lending, or stealing.

Figure 3 shows the results. The vertical lines represent the transmitted messages, the line that starts horizontally is the sampling interval. When using isolation as policy, there is no way the sending task can obtain enough energy. Therefore, every fifth message is not transmitted. There is no influence on the sampling task.

When the adaption policy is used, the sending task receives the

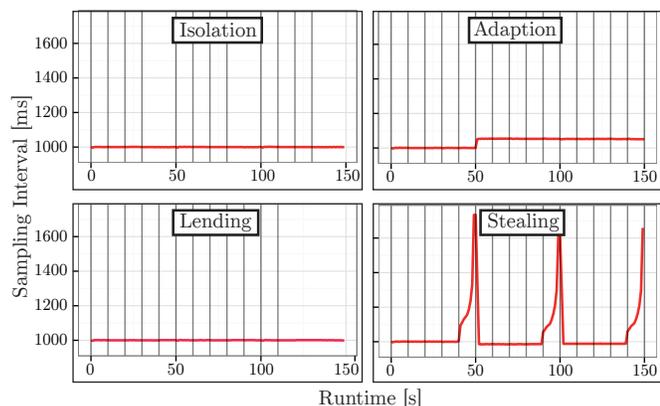


Figure 3. Influence of different policies when more energy is needed than allocated

minimum energy it requires to fulfill its assignment, enabling it to transmit all five messages in each interval. However, the energy consumption is higher, the remaining energy lower, which results in an adjustment of the sampling interval for phases two and three. Sampling less frequent conserves the energy required for the sending task.

The third part of the figure shows the policy lending, where the sending task borrows energy from the system. However, what is borrowed needs to be paid back, resulting in no transmissions at the end of the third phase. The sampling is not influenced in any way.

Stealing, the fourth policy, takes the energy required for sending directly and at the time it is required from the sampling task, making huge changes in the sampling interval necessary near the end of each phase.

The choice between policies is up to the application programmer, each variant has its advantages and disadvantages. Adaption makes sure that all parts of the sensor node continue to function, albeit some of them are under the influence of others. Lending ensures that only the part that requires too much energy may fail early, while all others continue to function as planned. This may enable a timely response from the network maintainer. Lending and stealing might also be combined.

2) *Scenario 2:* In the second scenario the energy consumption of the sending task was lower, resulting in a too large budget. Figure 4 shows the results for four different policies: unlimited savings, isolation, adaption and no savings.

When a task may store an unlimited amount of energy, all the excess of energy remains with that task and there is no influence on the other tasks (upper left). When the isolation policy is used, a budget that has not been completely spent needs less energy to be filled again. Therefore, the remaining system energy is higher and more energy can be distributed among all other tasks, in this case the sampling task (upper right). When using adaption, the budget for the sending task is reduced to the amount it used, making an even smaller amount of energy necessary to fill it in phase two, leading to more system energy and more energy available for the sampling task in phase two. However, once the adaption is complete, the energy available for the sampling task is reduced again in phase 3 and would remain the same in future phases (lower

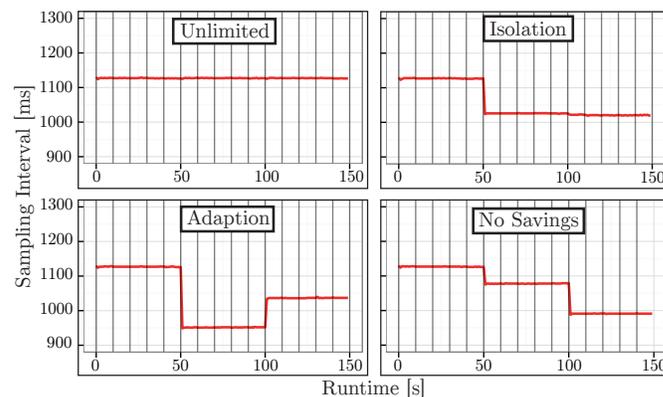


Figure 4. Influences of the choice of policy if a task needs less energy than assumed

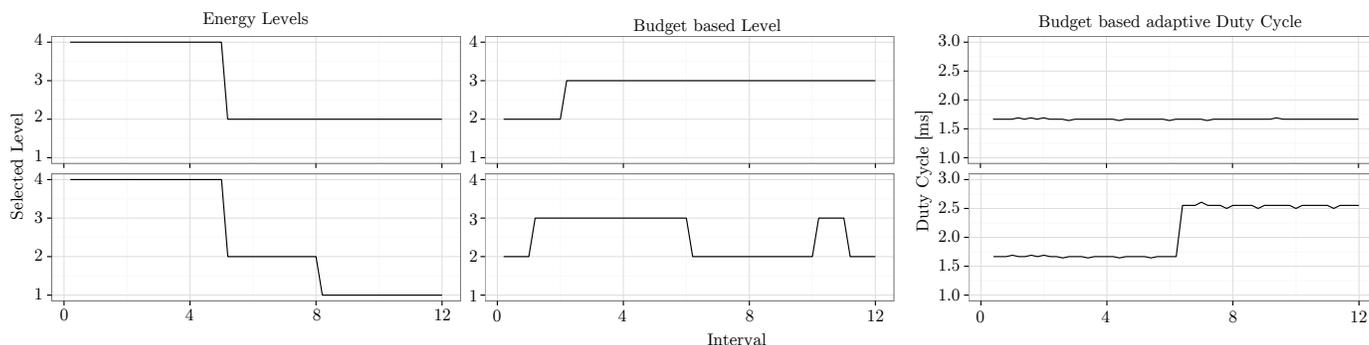


Figure 5. Experimental comparison of management approaches. *Energy Levels* computes feasible level assignment based on available energy and level utility. *Budget based Level* assigns level based on available energy in budget. Budget based duty cycle is computed based on available energy and runtime cost. Lower half shows experiment where in interval 6 the available energy was reduced by 20%.

left). The policy of no savings removes all remaining energy from a budget at the end of a phase, also resulting in a higher system energy and more energy available for the sampling task. As the too-high budget is filled at every phase, there is also some energy to "reclaim" at the end of each phase (lower right).

The comparison shows that allowing a task to store an unlimited amount of energy in its budget is a bad idea, as the stored energy is effectively lost, unless there comes a time when the task needs a tremendous amount of energy. The other three policies differ mainly in the point in time at which they reclaim the unused energy, but they all lead to an improvement for the sampling task.

C. System behavior

To evaluate the management behavior, a comparison of our energy budgets and the *Energy Levels* approach was conducted. The application periodically sent messages to a sink and featured four service levels distinguished by the sensor duty cycle. The energy consumption of the different service levels varies between $100\mu\text{A}$ and $270\mu\text{A}$, resulting in potential runtime between 32 and 11 months with conventional alkaline batteries. Periodically, one of the approaches is used to compute the service level for the following interval. To increase the comparability, both managers were based on the same energy accounting. *Energy Levels* uses a simplex algorithm to compute a feasible service level allocation based on the available energy, the consumption of each of the 4 used levels, and a utility ranging from one to four for each level (see [6] for the description of the optimization problem). As described in [6], the highest possible level is selected to provide the best QoS. The energy budget version (*Budget based Level*) used one budget and selected the service level based on the available energy within the budget (see 5). Additionally, a variant based on adaptive computation of the duty cycle based on equation 4 was tested using two budgets to isolate the sensing expenses from the communication.

The top half of Figure 5 shows the results achieved by each variant with a steady declining energy reserve. As *Energy Levels* selects the highest possible level, it starts with the highest level and later falls down to second last one. In contrast, the energy budget variant starts with the second last level, but shortly increases to and maintains the next higher

level. While this approach is more conservative, it does not depend on future energy availability. This is also true for the adaptive duty cycle, which is based on the energy within the associated budget and its consumption. The computed sampling interval is steady around 1.66s and thus slightly higher than level 3 (fixed 1.5s interval).

The lower half of Figure 5 shows the impact of a varying energy availability on the three variants. In the middle of the experiment, the available energy is reduced by 20% (e.g., due to reduced battery capacity). As the *Energy Levels* starts with the highest level, it must use the lowest level in the last third, because the optimistic approach used more energy at the beginning. The *Budget based Level* consumes the energy balanced with the runtime and thus reduces the service level only by one, resulting in less drastic changes. The duty cycle of the third variant adapts after the changed energy availability to 2.55s and thus is lower than level 2 (fixed 3s interval).

Both variants based on the energy budgets not only deliver a more homogeneous application quality, but the battery may also benefit from a reduced rate-capacity effect due to the the overall lower load of the conservative approach.

D. Interaction with the battery

The results of an experiment which included feedback from the battery are shown in Figure 6. It shows how the energy is divided among three budgets associated with three application parts. Sensing covers all energy expenses involved with the

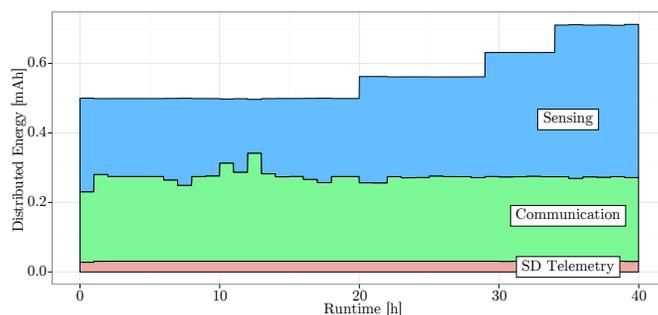


Figure 6. Distributed energy to different application parts. Battery Manager increases available energy over time due to higher capacity than expected.

data acquisition based on a dynamically calculated duty cycle. Communication includes a fixed communication interval every 2 minutes and SD Telemetry writes system information onto an SD-Card periodically.

In the course of the experiment, the communication effort varies and influences the energy distributed to sensing and its possible duty cycle. Additionally, the battery contains more energy than expected, resulting in an increased availability. As Communication and SD telemetry need a nearly fixed amount, sensing and, thus, the application duty cycle profits.

V. CONCLUSION

In this paper, we have presented our approach of using fine grained energy budgets to reach lifetime goals in wireless sensor networks. *Energy budgets* present a framework for dealing with varying demands in sensor networks. We evaluated the approach using FeuerWhere sensor nodes and presented promising first results for different kinds of policies. Additionally, we compared our approach with an existing approach and showed how the battery feedback influences the energy management.

In the future we plan to continue the evaluation of our approach by using different policies in a real deployment.

REFERENCES

- [1] T. Reddy, *Linden's Handbook of Batteries*, 4th Edition. Mcgraw-hill, 2010.
- [2] Texas Instruments, "Datasheet bq26500 single cell li-ion and li-pol battery gas gauge." [Online]. Available: <http://www.ti.com> (Accessed July 10, 2015)
- [3] Maxim Integrated, "Datasheet ds2438 smart battery monitor." [Online]. Available: <http://www.maximintegrated.com> (Accessed July 10, 2015)
- [4] M. Doyle, T. F. Fuller, and J. Newman, "Modeling of galvanostatic charge and discharge of the lithium/polymer/insertion cell," *Journal of the Electrochemical Society*, vol. 140, no. 6, 1993, pp. 1526–1533.
- [5] C.-F. Chiasserini and R. R. Rao, "Energy efficient battery management," *Selected Areas in Communications, IEEE Journal on*, vol. 19, no. 7, 2001, pp. 1235–1245.
- [6] A. Lachenmann, P. J. Marrón, D. Minder, and K. Rothermel, "Meeting lifetime goals with energy levels," in *SenSys*, vol. 7, 2007, pp. 131–144.
- [7] A. Sieber and J. Nolte, "Utilizing voltage decline for reaching lifetime goals," Paderborn, Germany, Tech. Rep., 2011.
- [8] A. Hergenröder, J. Horneber, D. Meier, P. Armbruster, and M. Zitterbart, "Distributed energy measurements in wireless sensor networks," in *Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems*. ACM, 2009, pp. 299–300.
- [9] X. Jiang, P. Dutta, D. Culler, and I. Stoica, "Micro power meter for energy monitoring of wireless sensor networks at scale," in *Proceedings of the 6th international conference on Information processing in sensor networks*. ACM, 2007, pp. 186–195.
- [10] P. Dutta, M. Feldmeier, J. Paradiso, and D. Culler, "Energy metering for free: Augmenting switching regulators for real-time monitoring," in *Information Processing in Sensor Networks, 2008. IPSN'08. International Conference on*. IEEE, 2008, pp. 283–294.
- [11] A. Lachenmann, P. J. Marrón, D. Minder, and K. Rothermel, "Meeting lifetime goals with energy levels," in *Proceedings of the 5th international conference on Embedded networked sensor systems*. ACM, 2007, pp. 131–144.
- [12] A. Castagnetti, A. Pegatoquet, C. Belleudy, and M. Auguin, "An efficient state of charge prediction model for solar harvesting wsn platforms," in *Systems, Signals and Image Processing (IWSSIP), 2012 19th International Conference on*. IEEE, 2012, pp. 122–125.
- [13] S. Kellner and F. Bellosa, "Energy accounting support in tinyos," *PIK-Praxis der Informationsverarbeitung und Kommunikation*, vol. 32, no. 2, 2009, pp. 105–109.
- [14] A. Dunkels, F. Osterlind, N. Tsiftes, and Z. He, "Software-based on-line energy estimation for sensor nodes," in *Proceedings of the 4th workshop on Embedded networked sensors*. ACM, 2007, pp. 28–32.
- [15] A. Sieber and J. Nolte, "Online device-level energy accounting for wireless sensor nodes," in *Proceedings of the 10th European conference on Wireless Sensor Networks, 2013*, pp. 149–164.
- [16] J. Flinn and M. Satyanarayanan, "Energy-aware adaptation for mobile applications," in *Proceedings of the Seventeenth ACM Symposium on Operating Systems Principles*, ser. SOSP '99. New York, NY, USA: ACM, 1999, pp. 48–63.
- [17] G. R. Wiedenhof, L. F. Wanner, G. Gracioli, and A. A. Fröhlich, "Power management in the epos system," *ACM SIGOPS Operating Systems Review*, vol. 42, no. 6, 2008, pp. 71–80.
- [18] J. Sorber, A. Kostadinov, M. Garber, M. Brennan, M. D. Corner, and E. D. Berger, "Eon: a language and runtime system for perpetual systems," in *Proceedings of the 5th international conference on Embedded networked sensor systems*. ACM, 2007, pp. 161–174.
- [19] X. Jiang and J. Taneja, "Energy management for wireless sensor networks," CS270 Project Report, Spring 2007.
- [20] G. W. Challen, J. Waterman, and M. Welsh, "Idea: Integrated distributed energy awareness for wireless sensor networks," in *Proceedings of the 8th international conference on Mobile systems, applications, and services*. ACM, 2010, pp. 35–48.
- [21] A. Roy, S. M. Rumble, R. Stutsman, P. Levis, D. Mazières, and N. Zeldovich, "Energy management in mobile devices with the cinder operating system," in *Proceedings of the sixth conference on Computer systems*. ACM, 2011, pp. 139–152.
- [22] H. Zeng, C. S. Ellis, A. R. Lebeck, and A. Vahdat, "Ecosystem: Managing energy as a first class operating system resource," in *ACM SIGPLAN Notices*, vol. 37, no. 10. ACM, 2002, pp. 123–132.
- [23] K. Lorincz, B.-r. Chen, J. Waterman, G. Werner-Allen, and M. Welsh, "Resource aware programming in the pixie os," in *Proceedings of the 6th ACM conference on Embedded network sensor systems*. ACM, 2008, pp. 211–224.
- [24] Q. Cao, D. Fesehaye, N. Pham, Y. Sarwar, and T. Abdelzaher, "Virtual battery: An energy reserve abstraction for embedded sensor networks," in *Real-Time Systems Symposium, 2008*. IEEE, 2008, pp. 123–133.
- [25] R. Neugebauer and D. McAuley, "Energy is just another resource: Energy accounting and energy pricing in the nemesis os," in *Hot Topics in Operating Systems, 2001. Proceedings of the Eighth Workshop on*. IEEE, 2001, pp. 67–72.
- [26] G. Mainland, D. C. Parkes, and M. Welsh, "Decentralized, adaptive resource allocation for sensor networks," in *Proceedings of the 2nd conference on Symposium on Networked Systems Design & Implementation-Volume 2*. USENIX Association, 2005, pp. 315–328.
- [27] G. Banga, P. Druschel, and J. C. Mogul, "Resource containers: A new facility for resource management in server systems," in *OSDI*, vol. 99, 1999, pp. 45–58.
- [28] A. Dunkels, J. Eriksson, N. Finne, and N. Tsiftes, "Powertrace: Network-level power profiling for low-power wireless networks," *Swedish Institute of Computer Science*, 2011.
- [29] S. Kellner, "Flexible online energy accounting in tinyos," in *Real-World Wireless Sensor Networks*. Springer, 2010, pp. 62–73.
- [30] K. Piotrowski, S. Ortmann, and P. Langendörfer, "Multi-radio wireless sensor node for mobile biomedical monitoring," *Biomed Tech*, vol. 57, 2012, p. 1.
- [31] K. Walther, R. Karnapke, and J. Nolte, "An existing complete house control system based on the reflex operating system: Implementation and experiences over a period of 4 years," in *Proceedings of 13th IEEE Conference on Emerging Technologies and Factory Automation, 2008*, pp. 40–45.

An Optimized Temperature Sensing Period for Battery Lifetime in Wireless Sensor Network

Seongman Jang, Keonhee Cho, Tacklim Lee, Byeongkwan Kang, and Sehyun Park

School of Electrical and Electronics Engineering

Chung-Ang University

Seoul, Republic of Korea

e-mail: ac2221, thckwall, tacklim34, byeongkwan, shpark@cau.ac.kr

Abstract— The battery lifetime is the most important issue in a wireless sensor network (WSN). As the battery lifetime is restrictive in WSN, energy efficiency is very important. In case of outdoor monitoring, the data of temperature has a feature which changes slowly. If the value of temperature does not change constantly, the network communication has to extend sensing period because it is the most energy consuming aspect of the battery. In this paper, we propose optimized temperature sensing period (OTSP) for efficient energy consumption. The main ideas of OTSP are that the network communication find successive stable section through proposed algorithm and the network communication reduce the wireless temperature network period in stable section. We experimented on OTSP in the rooftop of the building in order to remain unaffected by environments. The result of experiment indicates that OTSP reduces the energy consumption and extends the battery lifetime as compared with temperature sensing system.

Keywords- battery lifetime; energy consumption; wireless sensor network; temperature sensor

I. INTRODUCTION

WSN is used to monitor the various areas of the environment for implementing an IoT-based technology as used for the purpose of collecting and transmitting the data. Applications of WSN are used in the domain ecology, climate, environment, industry, transport, fire monitoring.

Most of these measurement of the wireless network environment is supplied with power through the battery. However, current battery technology cannot sufficiently supply the power in WSN. In addition, wireless communication constantly needs to consume the battery. As a result, research was done to minimize the WSN period, because the network communication approximately 20 percent of the battery's total energy consumption.

The data in WSN is divided into two types of feature, that is, the discontinuous and the continuous type of data. The discontinuous type of data is unstable, and the continuous type of data is stable. If the outdoor temperature is stable for a long time, the data in WSN is useless as we need to gather information about various data to optimize services. However, existing WSN system gathers all information although it is useless. To handle this problems, WSN has to minimize in stable section.

In this paper, we proposed the system to maximize the battery lifetime using OTSP. The wireless temperature sensor network has both types of unstable feature and stable feature.

This paper is organized as follows: in Section II, we discuss the works related to existing WSN; in Section III, We present the algorithm and important schemes; in Section IV, we show the implement and the results; finally, some concluding remarks and directions for future work are given in Section V.

II. RELATED WORK

WSN is required to extend battery lifetime. Improved battery lifetime in WSN has been proposed in several papers.

A. Study on management of energy consumption to improve battery lifetime in WSN

One of the energy management solution in WSN is network communication control, which is able to improve battery lifetime by analyzing data pattern. Wenguo et al [1] introduced the relationship between the energy utilization rate and the maximal network hop number. Akshay et al [2] used various models of the alkaline battery for maximizing its utilization and its lifetime. Suh et al [3] Software-based self-test has emerged as an effective strategy for online testing of processors integrated in bad-safety critical applications. However, these studies mostly focus on energy saving without utilizing the information on the accuracy.

B. Study on sleep mode to manage energy consumption in WSN

As a similar method used in this paper, studies using sleep mode in WSN are performed [4]. Several studies use sleep mode for energy management. Two terminal reliability of WSN is calculated, using Monte Carlo Simulations. Then via simulation results the effects of sleep probability on consumed energy and network reliability are considered [5]. However, sleep mode introduced in these studies has not been applied in environment monitoring system in WSN. Therefore, to improve battery lifetime in WSN, network communication control using sleep mode is needed

III. AN OPTIMIZED TEMPERATURE SENSING PERIOD

The proposed paper designs the OTSP system with temperature sensor and wireless communication interface. Before presenting the proposed system with system algorithm

and important scheme, the proposed paper discusses the problem of temperature sensing systems.

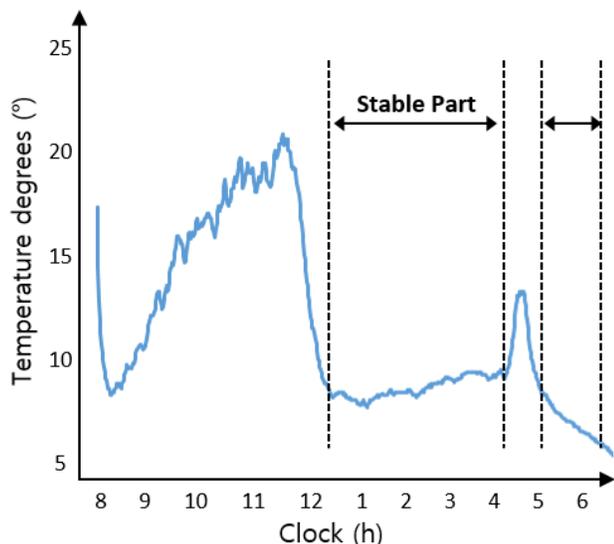


Figure 1. Data of temperature graph for daytime

The proposed system controls the network communication to minimize sensing period. The value of temperature is changed into environments (e.g., sunrise, sunset, wind, clouds, etc.) during daytime. The temperature variation during a certain period can be divided into two types. One type is much temperature variation, which is important to apply service models. The other type is little temperature variation, which is useless if it maintains for a long time. As shown in Fig. 1, the temperature graph for daytime has little temperature variation which is called stable part. If sensing period extends to sleep mode for stable part, it can get improved battery lifetime by saving network communication in WSN. Therefore, it is necessary to properly set the value according to temperature characteristics.

A. An optimized temperature sensing period Algorithm

Fig. 2 illustrates a flowchart of an optimized temperature sensing period algorithm. Low data change is received from temperature degrees through embedded system per proposed method. The successive stable section denotes that the current data is stable state.

The proposed optimized sensing period algorithm adjusts sensing period based on the successive stable section of temperature, which are inputted via embedded systems. The value of temperature degrees at the latest is $x_{Temp.n}$, the value of sensing time intervals that is a gap between previous sensing time is Δx , the value of temperature variation that has been changed differences degrees at the latest is $a_{Temp.n}$, the time that has been aggregate successive stable section is $S_{a.n}$. The procedures are composed of three steps.

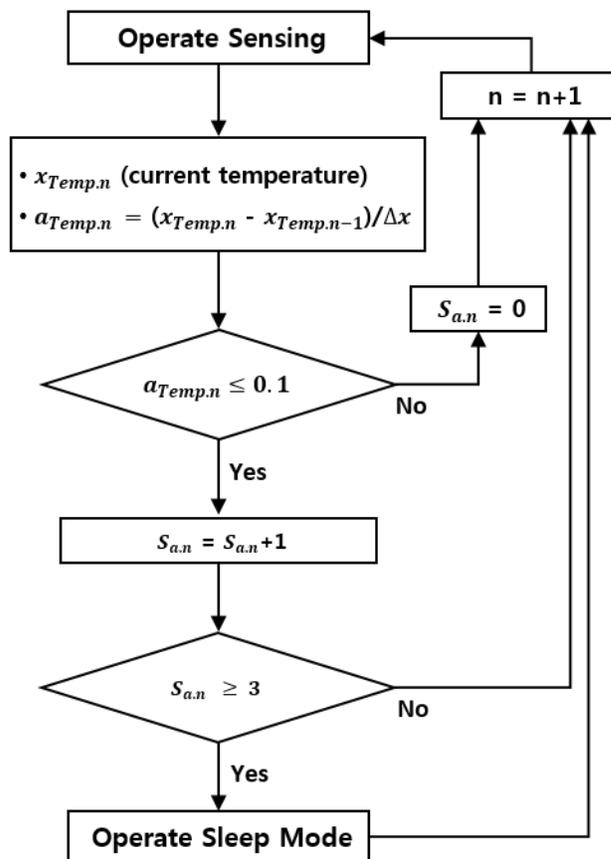


Figure 2. A flowchart an optimized temperature sensing period algorithm.

Step 1. A value of current temperature degrees through sensor is $x_{Temp.n}$. Then, calculates Δx which is sensing time intervals at the latest and $a_{Temp.n}$ which is temperature variation.

Step 2. Check whether $|a_{Temp.n}|$ is less than or equal to 0.1 that a measure effects of little temperature variation and wind and clouds, than class stable section and add 1 point in $S_{a.n}$. If $|a_{Temp.n}|$ is greater than 0.1, than initialize $S_{a.n}$ and operate sensing.

Step 3. Check whether $S_{a.n}$ is greater than or equal to 3, then operate sleep mode. If $S_{a.n}$ is less than 3, then operate sensing.

It is possible to derive the optimized sensing period, which can save energy at the maximum without useless sensing section through the proposed algorithm.

The proposed system can reduce energy consumption via interaction with the information about temperature variation.

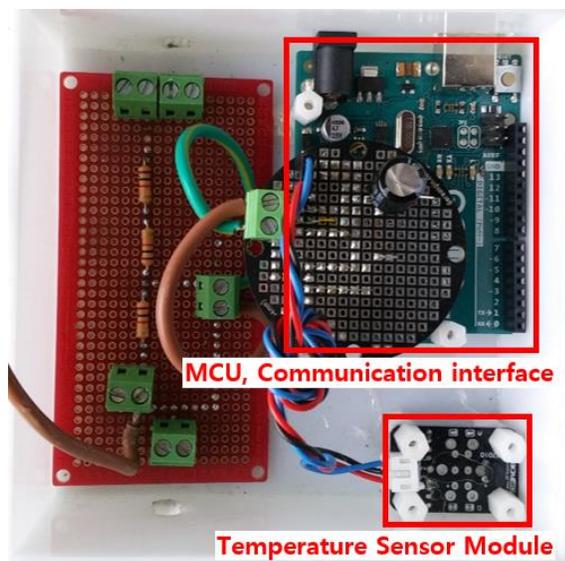


Figure 3. A prototype of an optimized temperature sensing period.

IV. IMPLEMENTATION AND EXPERIMENTAL RESULTS

A. Hardware of an optimized temperature sensing period

Fig. 3 shows a prototype of an optimized temperature sensing period. The main processor is a 16-bit micro controller unit (MCU). It is used for performing the main tasks such as complex event processing, processing of priority based scheduling algorithm, the provision of sensing period, pattern generation, and situation analysis. A ZigBee transceiver is used for communication with other networked devices. We used a 250 kbps/2.4 GHz ZigBee transceiver of Arduino Module. The metering circuit plays a role in measurement of the temperature degrees. The system is composed of temperature sensor module, MCU and communication interface. We operate parallel two systems, which are temperature sensing system and OTSP system in order to compare energy efficient system.

B. Experiment

We deployed OTSP system on the rooftop of the building in order to remain unaffected by environments. We analyzed

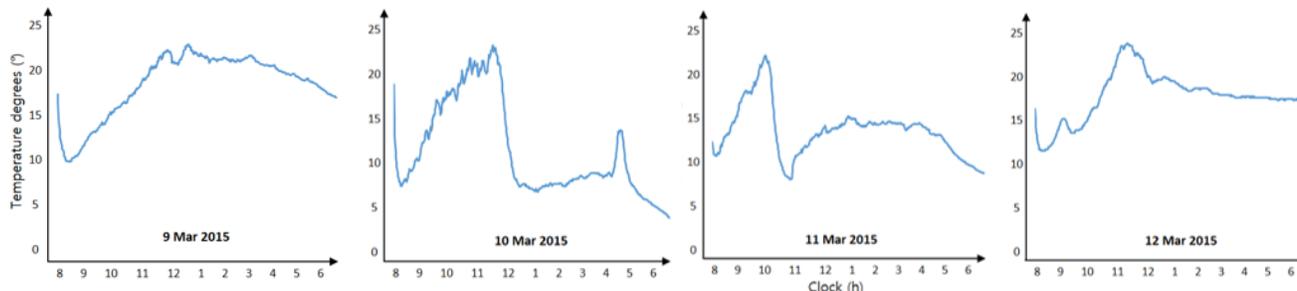


Figure 5. Daytime temperature graph from 8 A.M to 6 P.M for four days

proposed system according to two different methods. The results presented in the following were collected from 8:00 A.M to 6:00 P.M for four days.

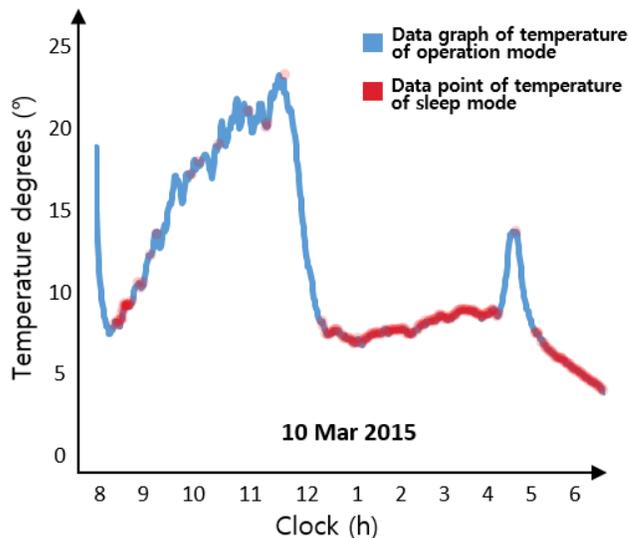


Figure 4. Daytime temperature graph from 8 A.M to 6 P.M and point of sleep mode in stable section.

We conducted experiment with two different systems that are the temperature sensing system and OTSP system. The temperature sensing system performs sensing of remaining 1 minute period. The OTSP system performs scheduling of operation mode/sleep mode according to previous temperature variation. The sensing period on sleep mode is remaining 2 minutes. Fig. 4 shows the sleep mode using OTSP in stable sections. The result show that almost point of sleep mode lie at stable section. Therefore, the OTSP system has little effect of services.

C. Result of an optimized temperature sensing period

Figs. 5 and 6 show the sensing counts for four days. As shown in Fig. 6, the value of network communication is reduced by 30.1 percent through OTSP system.

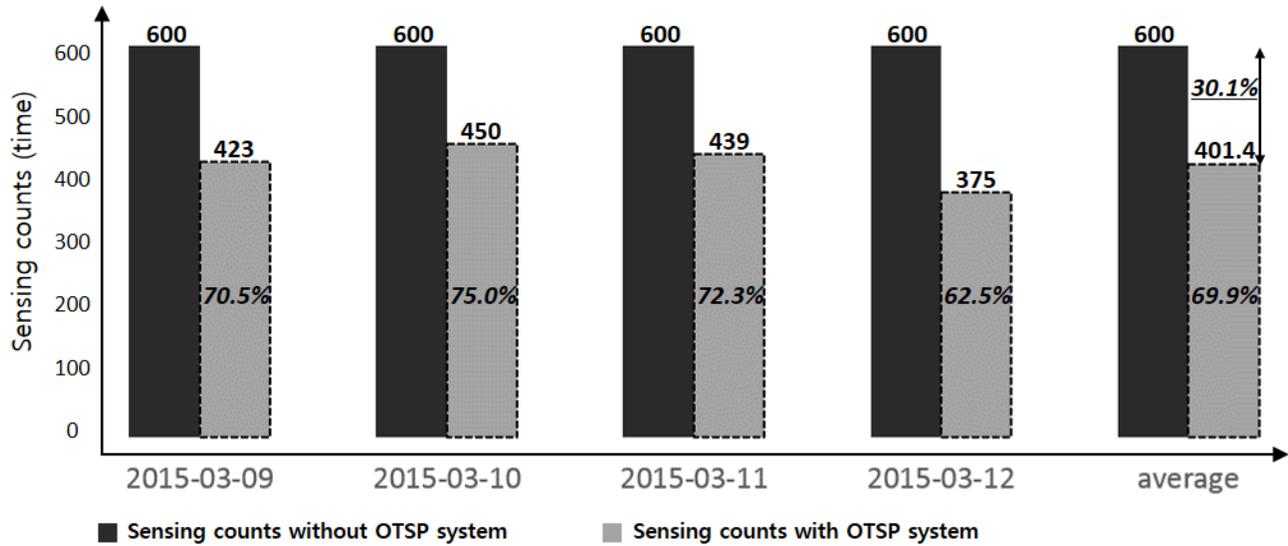


Figure 6. Comparison of system sensing counts for four days

Fig. 6 shows the sensing counts for four days. The result shows that the average sensing counts of temperature reduce.

D. Result of Battery Lifetime

1) Average sensing counts for 1 day

- Sensing counts without OTSP : 600 times/day
- Sensing counts with OTSP : 419.28 times/day

Simple experiment was performed to evaluate the above optimized temperature sensing period for improved battery lifetime. The performance of the proposed OTSP depends on average sensing counts during four days.

Therefore, the effectiveness of the above an optimized temperature sensing period can be evaluated by analyzing the improvement of battery lifetime of WSN using (1) to (4).

$$S_d = C_o \times C_S \times T_o \quad (1)$$

Where S_d is the sensing consumption current for 1 day (Ah/day); C_o is the operation current; C_S is sensing counts (/day); and T_o is the operation time (h).

$$S_y = C_o \times C_S \times T_o \times 365 \quad (2)$$

$$= S_d \times 365 \quad (3)$$

Where S_y is the sensing consumption current for 1 year (Ah/year).

$$S = C_n + C_m \quad (4)$$

Where C_n is the network node current; and C_m is the MCU spending current.

$$L_B = C_B / S_y \quad (5)$$

Where L_B is the battery lifetime (year); and C_B is the battery capacity;

a) Average Current Consumption on Operation Modes

- Transceiver Module : 50mA (0.6sec per 1 time)
- MCU : 81.33mAh/year

2) Sensing consumption current without OTSP for 1 year

a) consumption current of network node current

- Operating time : 0.6sec
- $S_d = 50\text{mA} \times 600 \text{ time/day} \times 0.6\text{sec} = 5\text{mAh/day}$ (6)
- $S_y = 5\text{mAh} \times 365\text{day} = 1825\text{mAh/year}$ (7)

b) Sensing consumption current

- $S = 1825\text{mAh/year} + 81.33\text{mAh/year}$ (8)
- $= 1906.33\text{mAh/year}$

c) Battery lifetime

- $L_B = 2200\text{mAh} / 1906.33 = 1.154 \text{ year}$ (9)

3) Sensing consumption current with OTSP for 1 year

a) consumption current of network node current

- Operating time : 0.6sec
- $S_d = 50\text{mA} \times 401 \text{ time/day} \times 0.6\text{sec}$ (10)
- $= 3.34\text{mAh/day}$
- $S_y = 3.34\text{mAh} \times 365\text{day} = 1219.7\text{mAh/year}$ (11)

b) Sensing consumption current

- $S = 1219.7\text{mAh/year} + 81.33\text{mAh/year}$ (12)
- $= 1301.04\text{mAh/year}$

c) Battery lifetime

- $L_B = 2200\text{mAh} / 1301.04 = 1.691 \text{ year}$ (13)

4) Benefit of battery lifetime using OTSP.

a) Compared with battery lifetime

- $n(\text{year}) = 1.691 - 1.154 = 0.537\text{year} = 196 \text{ day}$ (14)

As a result, the proposed OTSP improve the total battery lifetime by 196 days.

V. CONCLUSION

Saving energy has become one of the most important issues in WSN. The network communication approximately 20 percent of the battery's total energy consumption. The main ideas of OTSP are that the network communication find successive stable section through proposed algorithm and the network communication reduce the wireless temperature network period in stable section. An optimized temperature sensing period algorithm was proposed. The proposed OTSP can be easily modified and cost effectively implemented. We deployed OTSP system on the rooftop of the building in order to remain unaffected by environments, and measured the sensing counts to evaluate the performance of battery lifetime. The proposed OTSP improves the total battery lifetime by 196 days. It is expected that this work will contribute to providing guidance on the development of WSN.

ACKNOWLEDGMENT

This work was supported by the MSIP(Ministry of Science, ICT and Future Planning), Korea, under the ITRC(Information Technology Research Center) support program (IITP-2015-H8501-15-1018) supervised by the IITP(Institute for Information & communications Technology Promotion), Korean Home Operation and Mashup Environment Development and Evaluation (NIA-2015-E1505-1) supported by NIA(National Information Society Agency), and the Human Resources Development (No.20154030200860) of the Korea Institute of Energy Technology Evaluation and Planning (KETEP) grant funded by the Korea government Ministry of Trade, Industry and Energy.

REFERENCES

- [1] Y. Wenguo and G. Tiande, "The Non-uniform Property of Energy Consumption and its Solution to the Wireless Sensor Network," Education Technology and Computer Science (ETCS), 2010 Second International Workshop on, vol.2, Mar 2010, pp. 186-192.
- [2] N. Akshay, M. P. Kumar, B. Harish, and S. Dhanorkar, "An efficient approach for sensor deployments in wireless sensor network," Emerging Trends in Robotics and Communication Technologies (INTERACT), 2010 International Conference on, Dec 2010, pp. 350-355.
- [3] B. Suh, C. Won, and S. W. Kim, "Minimizing sleep duration time for energy harvesting wireless sensor networks," IEEE Sensors, Oct 2009, pp. 555-559.
- [4] K. Lutz and A. Konig, "Minimizing power consumption in wireless sensor networks by duty-cycled reconfigurable sensor electronics," Intelligent Solutions in Embedded Systems (WISES), 2010 8th Workshop on, pp.97-102, July 2010
- [5] Mahani and Ali, "A tradeoff in wireless sensor networks," Communication Systems, Networks & Digital Signal Processing (CSNDSP), 2012 8th International Symposium on, pp.18-20, July 2012

Design and Implementation of Indoor Position Estimation System using Drone for Industrial Security

Sanghoon Lee, Seonki Jeon, Myeong-in Choi, Byeongkwan Kang, and Sehyun Park,

School of Electrical and Electronics Engineering
 Chung-Ang University
 Seoul, Republic of Korea

e-mail: leessan0, junsk0108, auddlscjswo, byeongkwan, shpark@cau.ac.kr

Abstract—Indoor Position Estimation System (IPES) via RSSI and Trilateration using three fixed nodes is method to identify the location of the moving node in the building. The location of the moving node can be estimated within a range of relatively small error by using the three fixed nodes. This paper suggests that replacing one of the three fixed nodes with the drone located outside can be used to estimate the location of the moving node in the building. And finally, it is proved that the person could be scanned by a drone that is outside the building.

Keywords-LBS; Indoor Positioning; RSSI; Trilateration; Drone

I. INTRODUCTION

Current Location-Based Services (LBSs), about Outside Position Estimation, are available because of GPS technology. It came with a lot of development in the present, and now that is applied to the majority of the smart phones or mobile devices. On the contrary, Indoor Position Estimation cannot use GPS technology because of interference from buildings. Therefore, Indoor Position Estimation uses the Wireless LAN (WLAN), WiBro, RFID, Zigbee, and etc. Among them, the utilization of indoor position estimation using the WLAN is better than the others. One of the best thing among the WLAN infrastructure components is WiFi, because WiFi is continuously developing new versions, so WiFi's data transfer rate was gradually improved. It can be used valuably for inside position estimation.

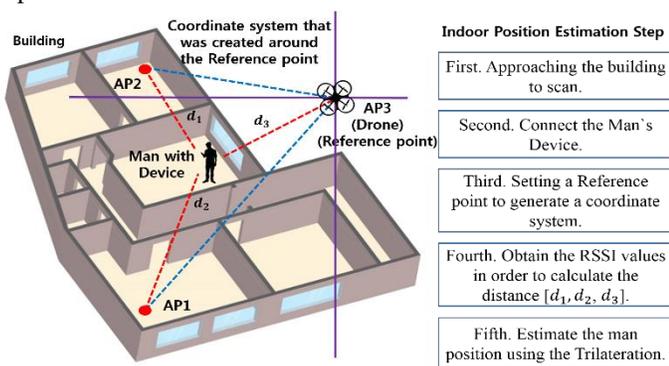


Figure 1. Method of scanning the inside of the building using the drone.

In this paper, two indoor fixed nodes and the third indoor fixed nodes are replaced by a drone from outside, and these fixed nodes are possible to estimate the location of the moving nodes in a building. To do this, three fixed nodes and one

moving node are connected to the WiFi, and Signal loss calculated by the Friis formula using the signal intensity got via RSSI. After then, the mobile device estimates the distance from each fixed nodes to moving node. Finally, drone estimates the location of the moving node based on the position of the three fixed nodes using the trilateration.

This paper is organized as follow. In Section 2, it presents a method for estimating the location using the moving node in three fixed nodes and one moving node through the existing RSSI method. In Section 3, we introduce the test bed to experiment with one fixed node replaced with a drone and test procedures. In Section 4, we present the experimental values obtained from our experiments on the test bed in tables and graphs. Finally, in Section 5, using the drone of the outside would conclude it possible to scan for devices inside the building by analyzing the values obtained in the experiment.

II. INDOOR POSITION ESTIMATION SYSTEM

A. RSSI measurement

RSSI denotes the received signal strength, and the units of RSSI using 'dB'. This is the same in WiFi. And, the high value of RSSI expressed as a negative number is clear signal intensity. As the moving distance of the signal is long, the strength of signal is weakened. With this feature, the signal strength received from the AP is used to estimate the distance from AP. The accuracy of position estimation using the RSSI is dependent on the high accuracy RSSI value and the environment of the measurement. Therefore, Position estimation method using the RSSI is showing relatively low reliability in an accuracy and a stability. Nevertheless the position estimating method using the RSSI uses widely, because the number of places WiFi Access zone has increased than before.

An important feature of wireless signal transmission is that the signal strength decreases as the distance increases. The value of the distance measurement of the RSSI is determined by changing of the signal intensity corresponding to the moving distance of the signal. It will obtain a moving distance between the respective nodes from the change of the RSSI value according to the moving distance of the signal. Researchers have done some effective researches about signals in different transmission environment [1], and conclude some good empirical formula.

$$L_d = L_1 + 10 \times \rho \times \log d + v \tag{1}$$

$$L_1 = 10 \log G_1 G_r \left(\frac{c}{4\pi f}\right)^2 \tag{2}$$

G_t is Transmitting antenna gain, G_r is Receiver antenna gain, c is the velocity of light, f is carrier frequency, ρ is Channel attenuation coefficient (value 2~6), v is the Gaussian random variable which considered the shadow effect, then $v \sim N(0, \delta^2)$, d is the distance, L_d is the channel loss after the distance d . In practice, we get the relation between RSSI and the distance through the measurement showing the relation for transmission power and receiving power by the following formula [2].

$$P_R = \frac{P_T}{r^n} \quad (3)$$

After conversion was

$$P_R(\text{db}) = A - 10 \times n \log r \quad (4)$$

P_R is the receiving power of the wireless signal, P_T is transmission power, n is the Propagation factor, r is the distance between Transceiver Unit. A is the receiving signal power when the signal transmit 1 meter. The numerical value of constant A and n determined the relation between receiving signal strength and signal transmission distance.

B. Trilateration.

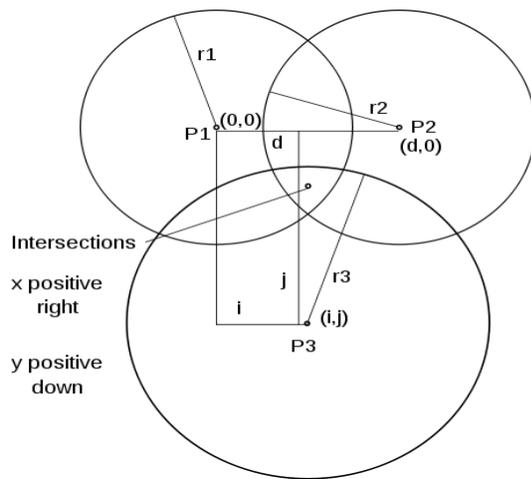


Figure 2. Trilateration.

The Trilateration is a method to obtain the relative position of the object using the triangular geometry as with triangulation. Unlike Triangulation, using one of the sides and two included angles of a triangle, Trilateration need two or more reference points, the distance between the target and the respective reference point to find the location of the target [3]. Only using Trilateration requires at least three reference points in order to accurately determine the relative position of the two-dimensional surface. In Fig. 2. The plane $z = 0$, shows the three sphere centers, P_1 , P_2 , and P_3 ; their x , y -coordinates; and the three sphere radii, r_1 , r_2 , and r_3 [4]. The two intersections of the three sphere surfaces are directly in front and directly behind the point designated intersections in the $z = 0$ plane.

The intersections of the surfaces of three spheres is found by formulating the equations for the three sphere surfaces and then solving the three equations for the three unknowns, x , y , and z . To simplify the calculations, the equations are formulated so that the centers of the spheres are on the $z = 0$ plane. Also, the formulation is such that one center is at the origin, and one other is on the x -axis. It is possible to formulate the equations in this manner since any three non-collinear points lie on a unique plane. After finding the solution, it can be transformed back to the original three dimensional Cartesian coordinate system. Next three equations for the three spheres

$$r_1^2 = x^2 + y^2 \quad (5)$$

$$r_2^2 = (x - d)^2 + y^2 \quad (6)$$

$$r_3^2 = (x - i)^2 + (y - j)^2 + z^2 \quad (7)$$

C. Position estimation using the RSSI and Trilateration

After measuring the intensity of a signal through the RSSI, the distance is calculated by using the Friis formula between the mobile node and the fixed node. Friis formula is as follows.

$$d = \frac{c}{4\pi f} \times 10^{\frac{L}{20}} \quad (8)$$

Distance measurement method by the signal intensity is used to Friis formula above.

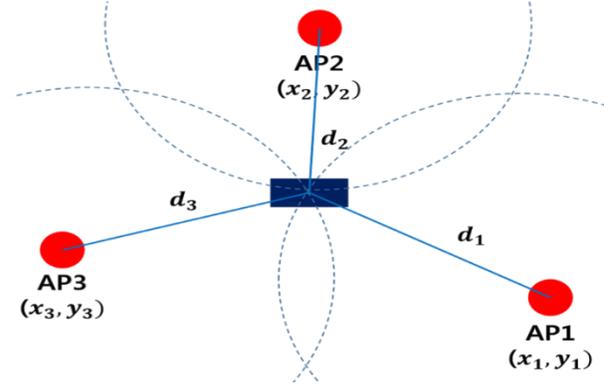


Figure 3. Overview of Position Estimation Method.

c is the Transmitting velocity, f is the frequency of the radio, L is the loss of the received signal. L is obtained by subtracting Intensity of the received signal at the reference point from Received signal intensity Fig. 3. The equations to obtain the distance from each AP as follows. [5]

$$d_1^2 = (x - x_1)^2 + (y - y_1)^2 \quad (9)$$

$$d_2^2 = (x - x_2)^2 + (y - y_2)^2 \quad (10)$$

$$d_3^2 = (x - x_3)^2 + (y - y_3)^2 \quad (11)$$

The using of RSSI values measured at the location with no obstructions is calculated as the actual value. The error between the actual distance and the calculated distance becomes large if the RSSI value is measured from the wrong environment. [6]

III. IMPLEMENTATIONS AND EXPERIMENTS

A. Implementations

Fig. 4 shows a test bed. Test bed was installed in the laboratory. Fig. 4. (a), The starting point in the Test bed is (a) Starting point (0,0) in Fig. 4 (a), which is given coordinate of (0,0). Based on the starting point, 1 o'clock is meant the x-axis and 9 o'clock is meant the y-axis. Fig. 4. (b), Test bed's cell size is 45cm. Fig. 4. (a), In the rest of the sequence, (b) is separated into 6 cells in the x-axis based on the starting point and is given coordinate of (6,0). (c), based on the start point, is located in the 6 cells in the x-axis direction and 7 cells in the y-axis direction and is given coordinate of (6,7). (d), based on the start point, is located in the 7 cells in the y-axis direction and is given coordinate of (0,7). (e) is an Access point 1 and AP channel is 'ISRC_PYD'. (f) is an Access point 2 and AP channel is 'ISRC-5G'. (g) is moving node, and (g)'s coordinate is not fixed. The total size of the test bed is '6x7'. And that is '270cm x 315cm'.

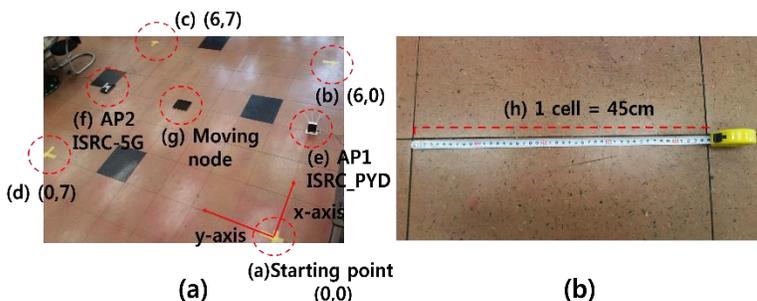


Figure 4. Implementation: (a) test bed, (b) test bed's cell size.

Fig. 5 shows an AR Drone which has a WiFi module to connect another WiFi AP. This AR Drone will carry out the AP3 outside of the laboratory's test bed. In this experiment, the AR Drone will be 3m in height from the ground. And AR Drone's coordinate is not fixed. In addition, the glass wall exists between the AR Drone and moving node. The actual building will exist a cement walls or glass walls between the moving node in the building and AR Drone. This glass wall will interfere communication with the AP and AR Drone. This position minimizes the disturbance factors than the actual environment. And This Drone's AP Channel is 'ardrone2_096244'.

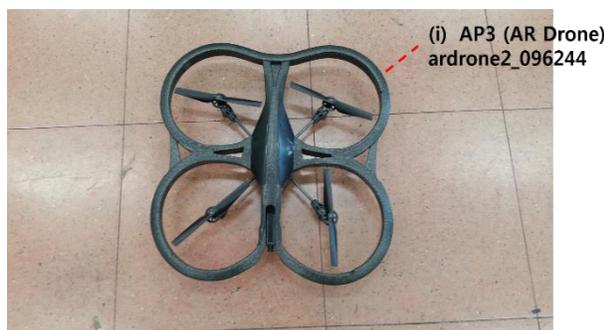


Figure 5. AR Drone, role of AP3.

B. Experiments

Fig. 6 expresses the test bed as a picture. AP1 is located at the starting point (0,0), AP2 is located at (3,7). AP3 is AR Drone, AR Drone exists outdoors a distance 110cm to the minus x-axis direction from the position of the test bed (0,5). And glass wall exists between the AR Drone and moving node. Moving node was defined as a device that can measure the RSSI value for each APs. And this moving node is located at a (3,4). This experiment was performed as follows.

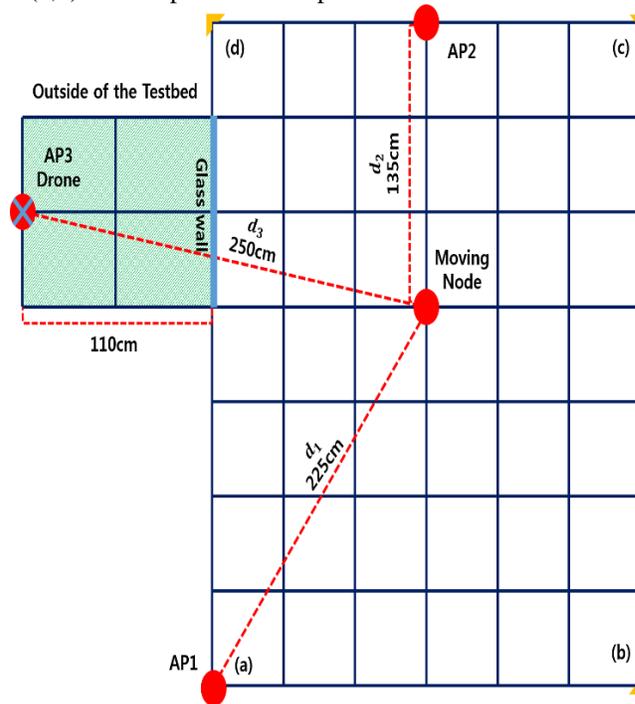


Figure 6. Illustrates of the test bed.

First, each AP is set on the coordinates of the test bed. And then, the actual distance is calculated from between APs and the moving node. As a result,

- Distance between AP1 and moving node is measured 225cm.
- Distance between AP2 and moving node is measured 135cm.
- Distance between AP3 and moving node is measured 250cm.

Then, RSSI values for each of the AP is measured in (3,4) located at the moving node in the test bed. As a result of the measurement,

- AP1 RSSI value is -54 dB
- AP2 RSSI value is -56 dB
- AP3 RSSI value is -62 dB

And, RSSI values for each of the AP is measured in (0,0) located at the moving node in the test bed. As a result of the measurement,

- AP1 RSSI value is -59 dB
- AP2 RSSI value is -57 dB
- AP3 RSSI value is -63 dB

In Section 4, the distance is calculated between the AP and moving node by using the RSSI values. Since then, comparing

the calculated distance with the actual distance is determined by the accuracy of the calculation.

IV. EXPERIMENTAL DATA ANALYSIS

A. Analysis results



Figure 7. AP's RSSI values measured at (3,4)

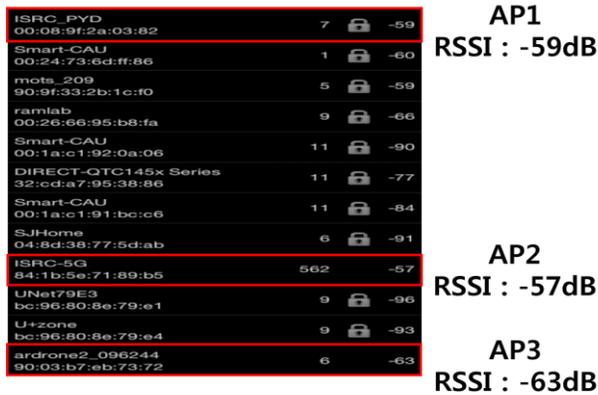


Figure 8. AP's RSSI values measured at starting point (0,0)

Fig. 7 is AP's RSSI values measured at (3,4). Each value is measured that AP1 RSSI value = -54 dB, AP2 RSSI value = -56 dB, AP3 RSSI value = -62 dB. Fig. 8 is AP's RSSI values measured at starting point (0,0). Each value is measured that AP1 RSSI value = -59 dB, AP2 RSSI value = -57 dB, AP3 RSSI value = -63 dB.

Before calculating the distance between the AP and the moving node by the RSSI values using the (8), c is propagation speed, and this value is set to '3 × 10⁸', and the frequency of AP1 is 2.4Ghz, the frequency of AP2 is 5Ghz, the frequency of AP3 is 2.4Ghz. Calculating the (8) have this values, it is as follows.

- $d_1 = \frac{3 \times 10^8}{4 \times \pi \times 2.4 \times 10^9} \times 10^{\frac{(-54+59)}{20}} \cong 176\text{cm}$
- $d_2 = \frac{3 \times 10^8}{4 \times \pi \times 5 \times 10^9} \times 10^{\frac{(-57+56)}{20}} \cong 42.5\text{cm}$
- $d_3 = \frac{3 \times 10^8}{4 \times \pi \times 2.4 \times 10^9} \times 10^{\frac{(-63+62)}{20}} \cong 88\text{cm}$

The actual distance between AP and the moving node in Fig. 6. As follows,

- $d_1 = 225\text{cm}, d_2 = 135\text{cm}, d_3 = 250\text{cm}$

Analysis of the error between the calculated distance with the RSSI value and actual distance is as follows.

- Distance difference (AP1 and Moving node between the actual distance) and (distance calculated as the RSSI value) is (225cm – 176cm = 49cm), and the error rate is 27.8%.
- Distance difference (AP2 and Moving node between the actual distance) and (distance calculated as the RSSI value) is (135cm – 42.5cm = 92.5cm), and the error rate is 217%.
- Distance difference (AP3 and Moving node between the actual distance) and (distance calculated as the RSSI value) is (250cm – 88cm = 162cm), and the error rate is 184%.

B. Error Anlysis

The most significant error was calculated between d_2 and d_2' , the error between d_1 and d_1' was the least. First, an analysis of the largest error that occurred in AP2. The difference between the RSSI value of the Fig. 7 and RSSI value of Fig. 8 hardly occurs because the moving node and AP2 are close. Errors in AP1 showed a relatively appropriate amount to 27.8%. Lastly, an analysis of the error that occurred in AP3. Error has occurred since AP3 is located in the outdoor, so signal is interfered by the wall when the signal transmitted.

The way to reduce the above error is as follows. In order to reduce the error arising from the AP2, the distance between the AP and the moving node has to be ensured over a certain distance. In order to reduce the error that occurred in AP1, RSSI values measured several times, it is possible to reduce the error by using the average value of these values. Lastly, in order to reduce the error that occurred in AP3, the position of the drone moves to less radio interference place [7].

V. CONCLUSIONS AND FUTURE WORKS

In this paper, we confirmed that it is possible to scan a person by a drone that is located outside the building and connected with AP inside of the building. If a mobile device has the appropriate distance between the mobile device and the AP and the little signal interference, we can efficiently estimate the indoor position of the mobile device. If this tracking function is developed, we will quickly find people who need help in a disaster area.

ACKNOWLEDGMENT

This work was supported by the MSIP(Ministry of Science, ICT and Future Planning), Korea, under the ITRC(Information Technology Research Center) support program (IITP-2015-H8501-15-1018) supervised by the IITP(Institute for Information & communications Technology Promotion), Korean Home Operation and Mashup Environment Development and Evaluation (NIA-2015-E1505-1) supported by NIA(National Information Society Agency), and the Human Resources Development

(No.20154030200860) of the Korea Institute of Energy Technology Evaluation and Planning (KETEP) grant funded by the Korea government Ministry of Trade, Industry and Energy.

REFERENCES

- [1] L. Girod and D. Estrin, "Robust Range Estimation Using Acoustic and Multimodal Sensing," Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems, Maui, pp. 1312-1320, 29 October-3 November 2011.
- [2] L. Girod, V. Bychovski, J. Elson and D. Estrin, "Locating Tiny Sensors in Time and Space: A Case Study," In: B. Werner, Ed., Proceedings of the IEEE International Conference on Computer Design: VLSI in Computers and Processors, Freiburg, pp. 214-219, 16-18 September 2002.
- [3] Lee, Jung-Kyu, YoungJoon Kim, and Jong-Ho Lee and Seong-Cheol Kim, "An Efficient Three-Dimensional Localization Scheme Using Trilateration in Wireless Sensor Networks," IEEE Commun. Lett., vol. 18, pp. 1591-1593, 2014.
- [4] Zheng Yang and Yunhao Liu, "Quality of Trilateration: Confidence-Based Iterative Localization," IEEE Trans. Parallel Distrib. Syst., vol. 21, pp. 631-640, 2010.
- [5] P.K. Sahu, E.H.-K. Wu, and J. Sahoo, "DuRT: Dual RSSi Trend Based Localization for Wireless Sensor Networks," IEEE Trans. Parallel Distrib. Syst., vol. 13, pp. 3115-3123, 2013.
- [6] Anindya S. Paul and E.A. Wan, "RSSI-Based Indoor Localization and Tracking Using Sigma-Point Kalman Smoothers," IEEE J. Sel. Top. Signal Process., vol. 3, pp. 860-873, 2009.
- [7] A.H. Wong, M.J. Neve, and K.W. Sowerby, "Antenna selection and deployment strategies for indoor wireless communication systems," IET Communications, vol. 1, pp. 732-738, August 2007.

Edge-Based Technique for Ultra-Fast Gating of Large Array Imagers

Octavian Maciu, Wilfried Uhring, Jean-Pierre Le Normand, Jean-Baptiste Kammerer, Foudil Dadouche

ICube UMR 7357

University of Strasbourg and CNRS

Strasbourg, France

Email: octavian.maciu@icube.unistra.fr

Abstract— Ultra-fast gating of large array imagers can be quite challenging to implement due to the distributed RC (Resistance Capacitance) nature of the metal wires used in all ICs (Integrated Circuits) for electrical connections. For the transmission of a signal across a long path, the metal line reduces the electrical bandwidth and adds a delay. The behavior of these lines has been modeled and a new solution is presented to circumvent these limitations. In this paper, we present an edge-based approach to the gating circuitry that allows sub nanosecond gating with a very low skew across the whole imager. Simulation data shows that our solution is an efficient way of reducing the effect of the distributed RC line delay with a small penalty on surface area and consumption.

Keywords—SPAD, Ultra-fast Gating, Edge-triggered, Fast Pulse, skewless

I. INTRODUCTION

Over the last decade, ultra-fast imaging has been a booming field with many considerable breakthroughs. A key advancement to this technology has been the ability to design image sensors with a sub-nanosecond temporal resolution. These imagers could be configured in either single-shot mode in [4] and [6], or can be repetitive and based on Single-Photon Avalanche Diode (SPAD) arrays rather than classical photodiodes. In this case, the temporal resolution can be even lower. Currently, integrated streak cameras operate at the fastest frequency, therefore also requiring robust acquisition signals. These fast signals are used in order to have a sub-nanosecond shutter speed. Depending on the design, the temporal resolution of an integrated streak camera can vary from a few hundred picoseconds to several nanoseconds. A 1D [5] or 2D [4] approach of integrated streak camera solution can be found in [5], where a delay generator based on the propagation delay of logic gates is used for sub-nanosecond shuttering. The delay could be customized using current starved inverters. In [3] and [4], we see the use of edge-based control signals for fast gating.

Similar to a clock signal in a synchronous sequential circuit, the gating signals are distributed to the entire imager array. Due to the distributed RC (Resistance capacitance) line delay, the integrity of the signal is compromised along the row of pixels. This would degrade the performance as a pixel at the end of the array would not perceive the same signal as another near the beginning of the array. Moreover, it would be completely unusable for ultra-fast gating purposes as the rise time and fall time would be far greater

than the gating time. This is especially true for ultra-fast image sensors as a signal commutes at the order of 100 picoseconds. It behooves us then to ensure that the signal is identical and useable for every single pixel in the array. Therefore, a more reliable approach would be to reduce the dependency on the pulse width and have an edge triggered gating. This solution was implemented in a 0.18 μm process. These approaches have been applied for smaller dimension image sensors and it is therefore interesting to propose a solution that is extendable to full resolution imagers.

II. THEORETICAL APPROACH

As mentioned before, in an imager array, the gating signal is being driven at the beginning of a row. Therefore, the signal perceived at the end of the row will no longer be a clean square pulse. Hence, it is interesting to study what the maximum row length for which the pulse will still be useable. Firstly, we can model this problem by looking at a single distributed RC model. Figure 1 shows the equivalent distributed line representation of a row containing N pixels where R_D and C_D are the resistance and the capacitance per length unit of the metal line respectively. The localized pixel input capacitances C_{pixel} act like a distributed capacitance according to the pixel pitch $C_{D_{\text{pixel}}}$ where,

$$C_{D_{\text{pixel}}} = \frac{C_{\text{pixel}}}{\text{Pixel Pitch}} \quad (1)$$

The total distributed capacitance $C_{D_{\text{total}}}$ is thus given by the sum of the equivalent distributed input pixel capacitance and the line distributed capacitance, $C_{D_{\text{total}}} = C_D + C_{D_{\text{pixel}}}$.

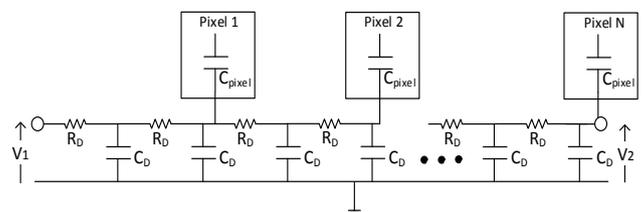


Figure 1. Distributed RC line representation of a sensor row including N Pixels.

The open-circuited Laplace transfer function from the beginning of the line V_1 to the end of the line V_2 can be written as [1][3]:

$$H(s) = \frac{V_2}{V_1} = \frac{1}{\cosh \sqrt{s \cdot R_D C_{D_{total}} \cdot l}}, \quad (2)$$

where s is the Laplace variable, and l the length of the row. The behavior of the line can be approximated by the simplified circuit model depicted in Figure 2[2].

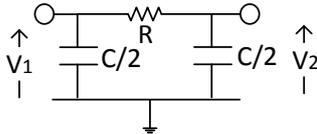


Figure 2. Simplified circuit of entire line

Where $R = R_D \cdot l$ and $C = C_{D_{total}} \cdot l$ are the total resistance and capacitance of the line. The unit step response gives us a clear indication that while not perfect, this approximation is sufficient to illustrate the distributed RC delay problem [2]. The distributed RC effect across the line therefore affects the driver rise time. The fastest rise time of V_2 can thus be obtained by the following expression:

$$T_R = 0.35\pi \cdot R_D \left(C_D + C_{D_{pixel}} \right) \cdot l^2 \quad (3)$$

Equation (3) states that the rising time increases with the square of the row length l and then can dramatically reach a value that makes it impossible to transport a nanosecond pulse across a large array sensor.

A. Parameters extraction

The parameters R_D and C_D can generally be found in the design kit documentation. Otherwise, the unit line and surface capacitance can be obtained by an analog extracted view simulation. Based on variants (square and rectangular) of the diagram in Figure 3, these parameters can be deduced with two different sets of equations (4).

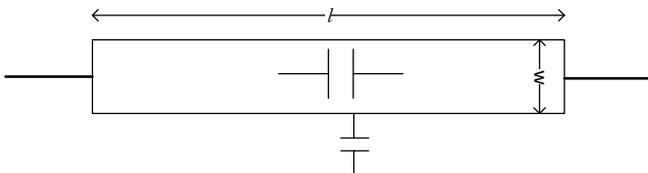


Figure 3. Diagram of layout model variants used (line and rectangle)

$$C_{line} = 2 \cdot C_{edge} (l + w) + C_{surface} \cdot w \cdot l \quad (4)$$

Where w and l are respectively the width and the length of the line, C_{line} is the extracted line capacitance, C_{edge} is the edge capacitance, given in F/m, and $C_{surface}$ the surface capacitance of the line, given in F/m².

Then, the distributed capacitance for a long line can be approximated by:

$$C_D = 2 \cdot C_{edge} + C_{surface} \cdot w, \quad (5)$$

where w is the width of the considered line.

The resistivity can be extracted in the same way or it can be obtained by process data, such as the metal resistivity ρ and thickness T . Therefore, the resistance per length unit R_D of a line of width w is given by:

$$R_D = \frac{\rho}{w \cdot T} \quad (6)$$

III. ULTRA-FAST GATING FOR LARGE ARRAY SENSOR

The Figure 4 shows the shape of a 1 ns pulse propagating in a sensor row at the beginning of the line, where the pulse is applied, in the middle and at the end of the line. If we assume the classic case where the logic of the pixel reshapes the pulse with a threshold voltage of half the power supply (solid line in Figure 4), we clearly see that after a certain length, the conventional gating is no longer adequate.

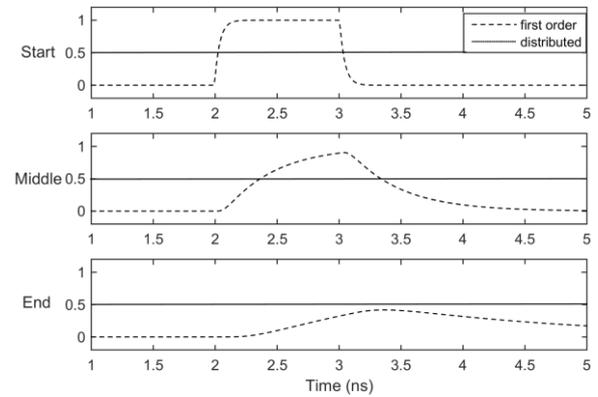


Figure 4. Illustration of lost signal integrity across a row of pixels

Furthermore, it is interesting to create a model in order to anticipate this signal distortion with respect to the length of the sensor. Based on the distributed model and a voltage threshold of half the power supply, Figure 5 compares the Full-Width at Half Maximum (FWHM) ratio of the in-pixel reshaped pulse versus the original pulse according to the original pulse FWHM to the rise time. The effective FWHM inside the pixel decreases as soon as the original FWHM pulse is below the rising time and is reduced to zero for a ratio of 30%.

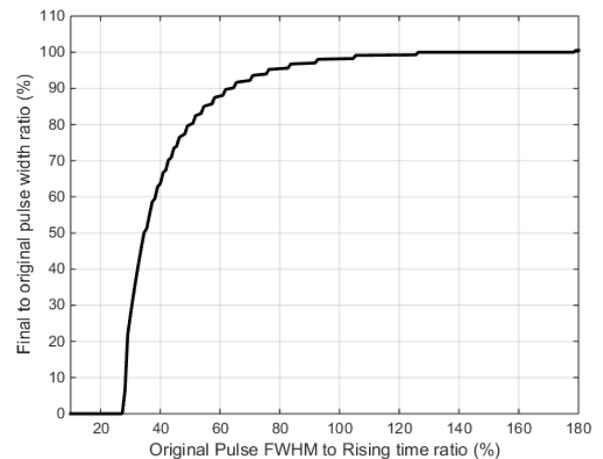


Figure 5. Final to original pulse width ratio versus the original pulse width to rising time ratio.

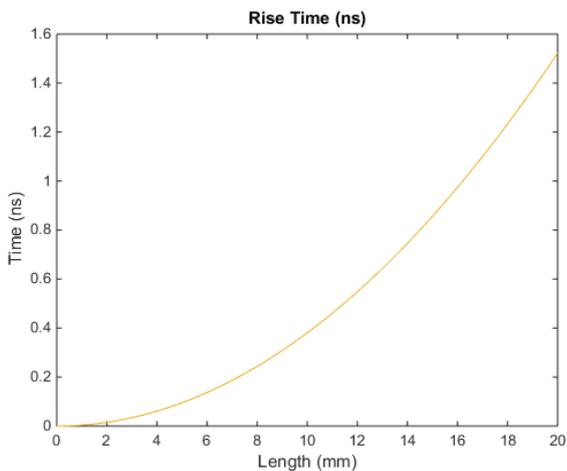


Figure 6. Rise Time of a signal seen across a pixel row through a metal line of up to 20 mm and a pixel pitch of 35 μm with the following parameters: $\rho = 2.65 \cdot 10^{-8} \Omega \cdot \text{m}$, $T = 425 \text{ nm}$, $C_{\text{pixel}} = 2 \text{ fF}$, $C_{\text{surface}} = 0.015 \text{ fF}/\mu\text{m}^2$, $C_{\text{edge}} = 0.032 \text{ fF}/\mu\text{m}$ and $w = 3 \mu\text{m}$.

Using (3), the relationship between the increasing rise time and the length of the array is shown in Figure 6 for a metal line, level 2, of a typical 0.18 μm CMOS (Complementary Metal Oxide Semiconductor) process with a resistivity ρ of $2.65 \cdot 10^{-8} \Omega \cdot \text{m}$, a thickness T of 425 nm, a pixel pitch of 35 μm , an input pixel capacitance of C_{pixel} of 2 fF, a surface capacitance C_{surface} of $0.015 \text{ fF}/\mu\text{m}^2$, an edge capacitance C_{edge} of $0.032 \text{ fF}/\mu\text{m}$ and a width w of 3 μm . Sub-nanosecond rise times are not allowed for sensor dimensions above 16 mm.

A way to enhance the line bandwidth is to increase the line width w because it reduces the resistance of the line. Figure 7 shows the calculated rise time in the middle of the previous line according to the metal width w . Increasing the width of the metal track above 3 μm has minimal impact on rise time and thus becomes irrelevant.

Therefore, generating 200 ps FWHM gating within large sensor arrays over 10 mm is impossible with pulse

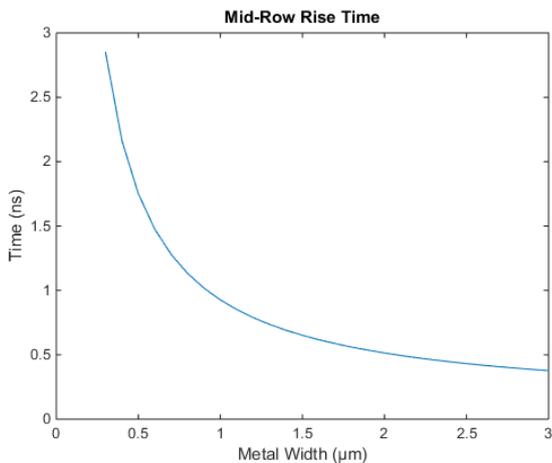


Figure 7. Rise Time with respect to metal track width at midway along a row of pixels

propagation techniques. To ensure signal integrity, the use of edge-sensitive logic is mandatory.

IV. PROPOSED DESIGN OF EDGE-TRIGGERED GATING

A. Edge-based Circuit for Ultra-fast Gating

In SPAD array image sensors, for ultra-fast gating mode there are three critical signals needed for optimal operation: Quench, Reset and Gate. Moreover, SPADs could be gated immediately after the reset to avoid the detection of photons arriving before the investigation time slot. While Quench is a much slower signal, the Reset and Gate signals can be sub-nanosecond pulses.

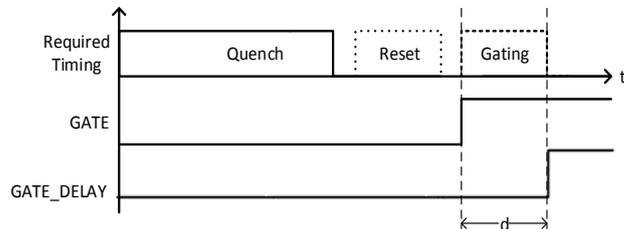


Figure 8. Timing diagram of typical SPAD quench, reset and gating with delayed (d) edge-based signals below.

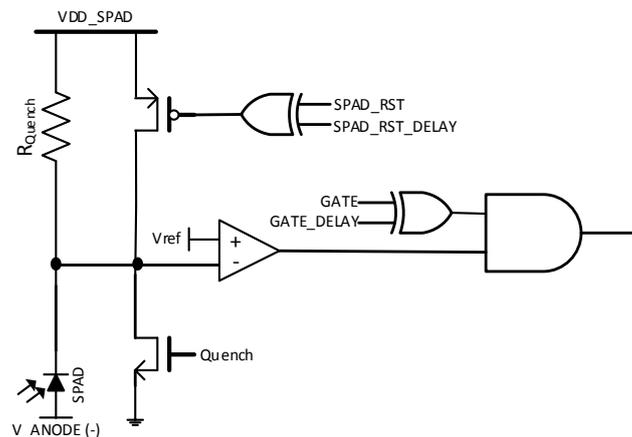


Figure 9. SPAD Edge-triggered circuit with active quench, reset and gating

Having shown clear limitations for ultra-fast gating in large SPAD array image sensors, a different approach is needed. This approach consists of an additional signal GATE_DELAY (Figure 8). GATE_DELAY is delayed thus creating the effective pulse of width d seen above through an XOR gate. It can therefore be extended to the Reset signal of the SPAD as seen in Figure 9. Hence, we are able to achieve gating in the range of 100 ps up to 1 ns while maintaining signal integrity throughout the SPAD array. Quenching time is much longer than the reset and gating signals and do not require an ultrafast signal generation.

B. Edge-based Driver for Ultra-fast Gating

In Figure 9, it can be seen that the GATE signal is delayed to obtain a GATE_DELAY signal. In order to generate the delay, we used a delay generator to drive the signals. The delay generator is constructed using a series of current starved double inverters as seen in [7].

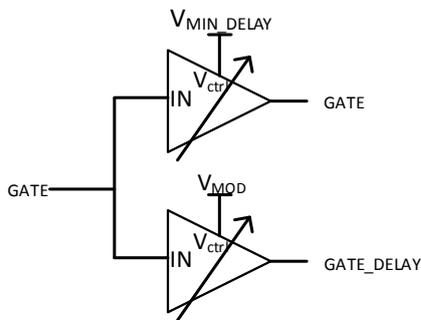


Figure 10. Delay generator block made up of double current starved inverters used to create GATE and GATE_DELAY signals

The delay of both signals (Figure 10) can be determined by a modifiable control voltage (V_{ctrl}). We are able to modify the pulse width by splitting the signal in two branches with one providing minimal propagation delay and the other being delayed with respect to the other. Both signals are then fed into an in-pixel XOR gate to obtain the desired pulse derived from the delay between the two signals.

C. Simulation of Edge-based technique

The presented solution can be demonstrated and validated through a Cadence simulation with two buffered lines of 20 mm long and a width of $3 \mu\text{m}$ based on the model presented in this paper in order to obtain a 300 ps pulse through an XOR gate.

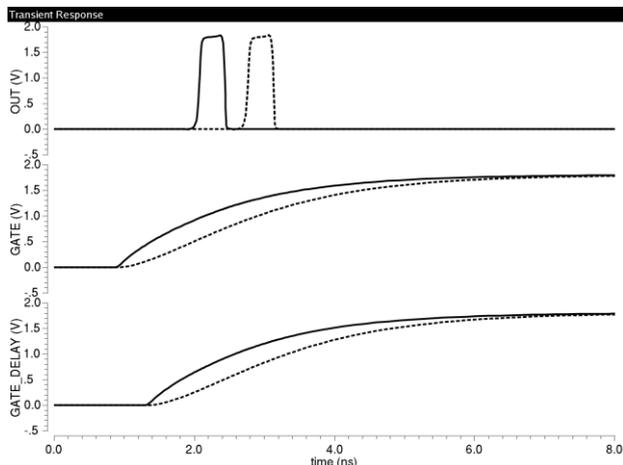


Figure 11. Simulation of a 300 ps pulse created using the edge-based technique (line: 20 mm length, $3 \mu\text{m}$ width). Signals at the beginning of the line (solid line) and at the end (dotted line)

In Figure 11, the FWHM of the pulse is 300 ps and our solution performs well in a Cadence simulation using the model presented in Figure 2 and the defining equations shown previously. However, a non-negligible skew is present.

V. ANALYSIS AND ELIMINATION OF SKEW

For large array ultra-fast image sensors, ensuring signal integrity across the array is only half of the challenge. As can be seen in Figure 6, a 20 mm sensor will introduce a

skew of around 1.5 ns. Hence, the edge based technique is not sufficient for larger dimensions.

In order to present a more robust approach on ultra-fast gating, we proceed to eliminate the large skew present across the array mentioned. This can be done by introducing two skewed inverters on each branch (GATE and GATE_DELAY) before the XOR gate. The threshold voltage of each inverter can be modified by altering the beta ratio effects inside each inverter. Hence, the skew across the sensor array can be compensated.

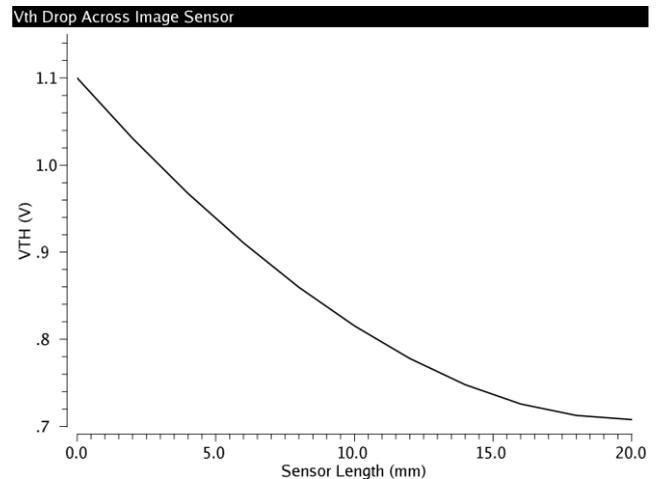


Figure 12. Simulation of needed V_{TH} variation across a 20 mm image sensor row using above mentioned parameters.

Figure 12 shows the simulation of needed V_{TH} variation across a 20 mm image sensor row using above mentioned parameters. The threshold voltage at the beginning of the sensor is arbitrary set at 1.1 V (above $V_{DD}/2$ where $V_{DD} = 1.8$ V) in order to maintain threshold levels across the array well above the $|V_T|$ of each transistor. For a required voltage V_{TH} , the inverter should be adjusted by modifying the geometry of the PMOS and NMOS transistors, thus changing the beta ratio.

This can be obtained by solving the following equation:

$$I_n = I_p, \text{ for } \begin{cases} V_{in} = V_{th} \\ V_{out} = V_{DD}/2 \end{cases} \quad (7)$$

In this case, both the NMOS and the PMOS transistor are in the saturated operation region, thus the currents are given by:

$$I_n = \frac{1}{2} \mu_n C_{ox} \left(\frac{W_n}{L_n} \right) (V_{TH} - V_{T_n})^2 \left(1 + \lambda \frac{V_{DD}}{2} \right) \quad (8)$$

$$I_p = \frac{1}{2} \mu_p C_{ox} \left(\frac{W_p}{L_p} \right) (V_{DD} - V_{TH} - |V_{T_p}|)^2 \left(1 + \lambda \frac{V_{DD}}{2} \right)$$

Replacing I_n and I_p in equation (7) leads to:

$$\mu_n \left(\frac{W_n}{L_n} \right) (V_{TH} - V_T)^2 = \mu_p \left(\frac{W_p}{L_p} \right) (V_{DD} - V_{TH} - |V_{T_p}|)^2 \quad (9)$$

Hence, we obtain the ratio ρ :

$$\rho = \frac{\left(\frac{W_p}{L_p}\right)}{\left(\frac{W_n}{L_n}\right)} = \frac{\mu_n}{\mu_p} \frac{(V_{TH} - V_T)^2}{(V_{DD} - V_{TH} - |V_{Tp}|)^2} \quad (10)$$

Furthermore, in order to keep the best fill factor, the inverter area A given by:

$$A = W_p L_p + L_n W_n = \rho \left(\frac{W_n}{L_n}\right) L_p^2 + \left(\frac{L_n}{W_n}\right) W_n^2 \quad (11)$$

has to be kept as low as possible. For a threshold voltage V_{th} above $V_{DD}/2$, the PMOS transistor has to be more conductive than the NMOS one, thus the ratio ρ should be greater than one. For this, we can consider using $W_n=W_{min}$ and $L_p=L_{min}$ as the best choice. The inverter area becomes:

$$A = \rho \left(\frac{W_n}{L_n}\right) L_{min}^2 + \left(\frac{L_n}{W_n}\right) W_{min}^2 = \rho \alpha L_{min}^2 + \frac{1}{\alpha} W_{min}^2 \quad (12)$$

Minimizing A according to the variable α leads to best sizes of the MOS transistors for the smallest use of surface area:

$$\frac{W_n}{L_n} = \frac{W_{min}}{\sqrt{\rho} \cdot L_{min}} \quad \text{and} \quad \frac{W_p}{L_p} = \frac{\sqrt{\rho} \cdot W_{min}}{L_{min}} \quad (13)$$

In a similar way, the optimal sizes of the transistors for a threshold voltage V_{TH} under $V_{DD}/2$ can be computed as:

$$\frac{W_n}{L_n} = \frac{\sqrt{\rho} \cdot W_{min}}{L_{min}} \quad \text{and} \quad \frac{W_p}{L_p} = \frac{W_{min}}{\sqrt{\rho} \cdot L_{min}} \quad (14)$$

In both case, the area of each threshold adjusted inverter is given by:

$$A = W_{min} L_{min} \sqrt{4\rho} \quad (15)$$

The more ratio ρ is distant from 1, the higher the surface area.

It would be ideal to designate an area of a pixel for the placement of automatically generated skewed inverters based on their position in the sensor row.

By employing this technique, it can be seen in Figure 13

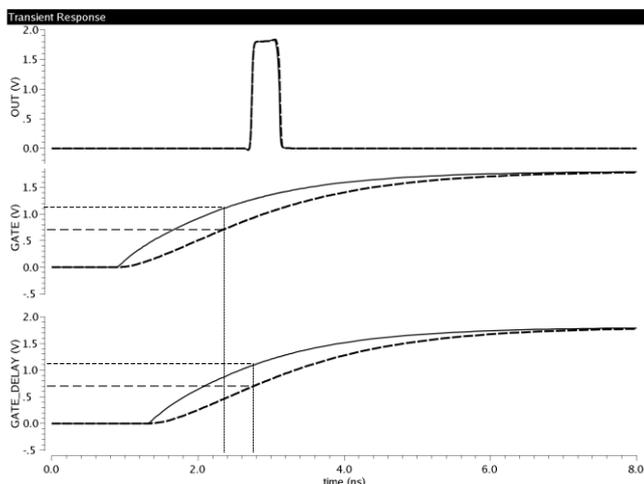


Figure 13. Simulation of skewless pulse. Solid line: signal at the beginning of the sensor, dotted line: signal at the end of a 20 mm sensor.

that the pulse at the beginning of the sensor row and the pulse at the end are synchronized and the skew seen in Figure 11 is eliminated.

In an ideal fabrication process, this technique coupled with the edge-gating approach previously developed and is optimal for sub-nanosecond ultra-fast gating. Moreover, after performing a Monte-Carlo simulation on the possible dispersion existing between the two pulses (beginning and end of line), we obtain a sigma of only 18.42 ps. In order to circumvent the inevitable mismatch, it is conceivable to modify (post-fab) the supply voltages of the inverters added in order to fine tune the additional skew.

VI. CONCLUSION

In the increasingly relevant field of ultra-fast imaging, gating signals play a key role in assuring optimal operation. When designing large resolution sensors, acquisition pulses must be uniformly distributed throughout the entire array. Due to the delay effect of a distributed RC line, this becomes troublesome. For this, a conventional driver at the beginning of the line is no longer adequate and a more robust solution needs to be implemented. This new approach is independent to pulse-based gating and instead uses an edge-triggered gating. Moreover, the designer could place such edge-based drivers on either side of the sensor to obtain a near-perfect gating signal distribution. While this technique is still hindered by signal skew, it can be eliminated using the method presented in this paper.

ACKNOWLEDGMENT

We thank the French agency "Agence Nationale de la Recherche" for financial support under grant ANR-14-CE26-0024-01.

REFERENCES

- [1] M. Ogata, Y. Okabe, and T. Nishi, "Simple RC models of distributed RC lines in consideration with the delay time", Circuits and Systems, 2004. ISCAS '04. Proceedings of the 2004 International Symposium on, vol.4, pp. IV-649-52, 23-26 May 2004.
- [2] V. B. Rao, "Delay Analysis of the Distributed RC Line," Design Automation, 1995. DAC '95. 32nd Conference on., pp. 370-375, 1995.
- [3] Y. Maruyama, J. Blacksberg, and E. Charbon, "A 1024x8 700ps time-gated SPAD line sensor for laser raman spectroscopy and LIBS in space and rover-based planetary exploration," Solid-State Circuits Conference Digest of Technical Papers (ISSCC), 2013 IEEE International , pp. 110-111, 17-21 Feb. 2013.
- [4] F. Morel, J.-P. Le Normand, C.-V. Zint, W. Uhring, Y. Hu, and D. Mathiot, "A new spatiotemporal CMOS imager with analog accumulation capability for nanosecond low-power pulse detections," Sensors Journal, IEEE, vol. 6, no. 5, pp. 1200-1208, 2006.
- [5] M. Zlatanski, W. Uhring, J. Le Normand, C. Zint, and D. Mathiot, "Streak camera in standard (Bi)CMOS (bipolar complementary metaloxide-semiconductor) technology," Measurement Science and Technology, vol. 21, no. 11, p. 115203, 2010.
- [6] M. Zlatanski and W. Uhring, "Streak-mode optical sensor in standard BiCMOS technology," IEEE Sensors, pp. 1604-1607, Limerick, Ireland, 2011.
- [7] Zlatanski, M. Uhring, W. Le Normand, and J.-P. Zint, V., "A new high-resolution Time-to-Digital Converter concept based on a 128 stage 0.35 μm CMOS delay generator," Circuits and Systems and TAISA Conference, 2009. NEWCAS-TAISA '09. Joint IEEE North-East Workshop on , pp.1-4, 2009.

Design Methodology of TDC on Low Cost FPGA Targets

Case Study: Implementation of a 42 ps Resolution TDC on a Cyclone IV FPGA Target.

Foudil Dadouche, Timothé Turko, Wilfried Uhring, Imane Malass, Jérémy Bartringer, Jean-Pierre Le Normand

ICube, UMR 7357, University of Strasbourg and CNRS
Strasbourg, France

Email: wilfried.uhring@unistra.fr

Abstract— This work aims to introduce a design methodology of Time-to-Digital Converters (TDCs) on low cost Field-Programmable Gate Array (FPGA) targets. First, the paper illustrates how to take advantage of the presence of carry chains in elementary logic elements of the FPGA in order to enhance the TDC resolution. Then, it describes how to use the Chip Planner tool to place the partitions composing the system in user specified physical regions. This allows the placement of TDC partitions so that the routing paths are constrained. As a result, the user controls the propagation delay effectively through the connection network. The paper ends by applying the presented methodology to a case study showing the design and implementation of high resolution TDC dedicated to fast imaging systems. The obtained resolution of 42 ps using a low cost FPGA target Cyclone family is very promising and suitable for a large amount of fast applications.

Keywords- Time-to-Digital Converter; FPGA; Chip Planner; Carry Chain Logic.

I. INTRODUCTION

Nowadays, numerous applications require a precise measurement of time duration separating two or several physical events. 3D scanners or 3D console games represent typical application requiring precise time quantification of the interval time to reconstitute a three-dimensional scene. Such systems are generally based on Time Of Flight (TOF) measurement of the light emitted by a laser diode or a light-emitting diode (LED) and detected by a suitable light sensors after reflection by an object. The TOF of the light is proportional to the distance traveled by the latter. The measurement is made independently by several pixels allowing the reconstitution of the 3D scene [1][2].

To measure this duration, we use devices capable of converting extremely low time durations (some tens of ps) into digital values understandable for downstream processing and conditioning chain. These devices are commonly known as Time-to-Digital Converters (TDCs) [3].

In order to design such systems there are several techniques which are proposed in literature. Some of the techniques that can be readily identified are [3]-[5]: Tapped Delay Lines (TDL), Delay Locked Loop (DLL), Vernier Delay Line (VDL), Multilevel TDC, etc. All of these Time-to-Digital conversion techniques are usually designed as Application-Specific Integrated Circuits (ASICs). The latter have the advantage to have high performances but suffer from a higher cost, slow time to market and limited

reconfiguration possibilities. It is also worth noting that the ASIC solutions are not suitable for integration into reconfigurable digital designs mostly described in Hardware Description Languages (HDL). As a result, numerous solutions for implementing TDCs on FPGA circuits have emerged [6]-[11]. However, the most significant limitation of these architectures is the difficulty to predict the placement and routing delays as well as the time delay of logic gates themselves. The consequence of this inevitable hardware restriction is a non-stable resolution of the designed TDC [8].

In this work, we aim to introduce a design methodology for high resolution TDC on low cost FPGA targets. This methodology enables the mastering of the network routing delays as well as the delays of the gates themselves. Therefore, it leads to an optimized TDC design with stable and accurate resolutions.

In order to give some background, the functional principle as well as the structure of the studied TDC is given in the second section. We also point out some associated difficulties encountered while using classical inverters as delay cells. The third section is dedicated to present our approach of implementing a TDC on an FPGA. Firstly, we show how to take advantage of the Carry Chain Logic to enhance and optimize TDC resolution. Secondly, we illustrate how to use the Chip Planner to define the exact physical layout location in the Chip. Therefore, we point out the importance of this operation. Section 4 provides a detailed case study consisting of implementing a 42 ps resolution TDC on Altera Cyclone IV low cost FPGA. The implemented TDC is associated to an FTDI (Future Technology Devices International) USB interface circuit operating in parallel mode with transfer rates reaching up to 40 Mbytes per second. Finally, we end our work by providing some final observations.

II. SETTING IN THE CONTEXT

A. Functional principle of the studied TDC

A TDC is an electronic system that measures the interval time between two occurring events of a given signal. Its main purpose is to convert temporal information to binary sequence understandable for a downstream processing chain.

For an accurate time duration measurement, generally, a TDC is composed of three blocks: two fine measurement blocks and a coarse one. The coarse one counts the number (N) of clock periods between enabling to disabling the

measured interval, and the fine blocks evaluate the uncertainties in both sides that cannot be counted since their duration is shorter than the clock period.

To understand the role of each one of the three blocks, we illustrate by the timing diagram of Figure 1 the functioning principle of a generic TDC.

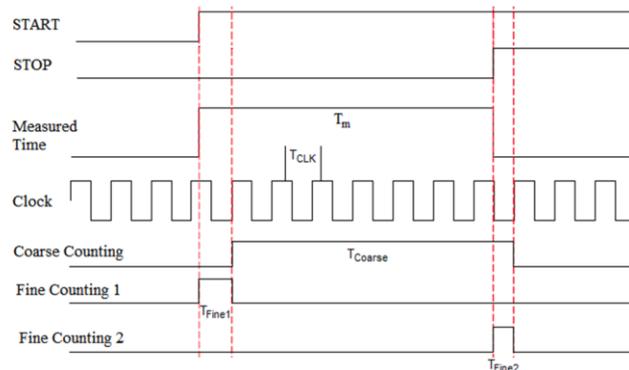


Figure 1. Functioning principle of a generic TDC

As we can see from this timing diagram, the time interval to be measured (T_m) is a combination of three individual durations: (1) T_{Coarse} , which represents the number of clock periods from enabling to disabling coarse measurement, (2) T_{Fine1} , representing the time between the measured signal active edge and the first following rising clock edge, and (3) T_{Fine2} which is the time between the falling edge of the measured interval and the following rising clock edge. Accordingly to this timing diagram, the measured time will be expressed as follows:

$$T_m = T_{Fine1} + T_{Coarse} - T_{Fine2} = T_{Fine1} + N \cdot T_{clk} - T_{Fine2} \quad (1)$$

In practice, TDCs are mostly used in fast imaging systems needing to know the delay separating a photon emission by a laser diode and the detection of that photon by a Single-Photon Avalanche Diode (SPAD). Therefore, the events can be represented by two signals: (1) a START signal, which can be synchronized with the coarse counter clock, and (2) a STOP signal that means that the SPAD has detected a photon. In that case, the whole TDC can be reduced to a coarse counter associated to a fine TDC measuring T_{Fine2} . Consequently, the measured interval time will be given by the expression hereafter:

$$T_m = N \cdot T_{clk} - T_{Fine2} \quad (2)$$

Since the coarse block is a simple counter incremented by the system clock, we will focus in the following section on the implementation of the fine TDC.

B. Structure of the studied TDC

As mentioned previously, there are different techniques of designing TDCs. In this work, we focus on the commonly used Tapped Delay Lines (TDL) architecture depicted in Figure 2.

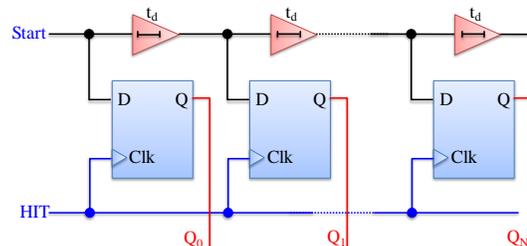


Figure 2. Tapped delay line TDC

A TDL TDC consists of N cascaded delay elements whose inputs are stored in D Flip Flops (DFFs). We would then have as many DFFs as there are delay elements. Therefore, each delay element can be regrouped with its associated DFF to form an elementary cell of the TDC.

The number (N) of these elementary cells depends of the common DFF clock frequency, as well as the propagation time of the delay element (t_d). This is given by the ratio of clock period to propagation time t_d . Since the value of the previous parameter is not provided, it is determined experimentally.

C. Design and validation of the elementary cell

In order to design the elementary cell of the studied TDC, we first used a simple inverter as a delay element associated with a DFF as illustrated by Figure 3.

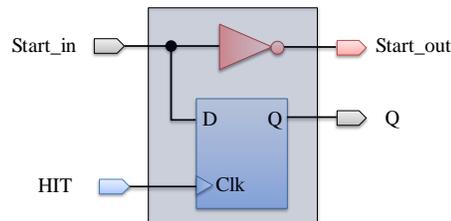


Figure 3. Simple TDC elementary cell

However, implementing a TDC chain on an FPGA by duplication of this cell leads to a simplified circuit entirely different from the desired function. Indeed, if the input signal $Start_in$ and the output signal $Start_out$ are equipotential, the used HDL software (Quartus II) will simplify the logical equation giving the output versus the input so that it saves place and time.

To illustrate this phenomenon, we represent in Figure 4 the RTL view resulting from the implementation of a simple TDC chain composed of four elementary cells. It can readily be seen that, in spite of the presence of inverters, the software has simplified the logical equations. Consequently, all the inverted signals are grouped independently of the non-inverted ones. It is thus evident that this method is not suitable for designing a TDC. Nevertheless, it is worth to notice that, if it is not possible to prevent Quartus II software to optimize data path, it is quite possible to create this path manually by operating directly on the logical resources of the FPGA. Indeed, the Quartus II Chip Planner tool allows physical access to logical resources available on the chip.

Using this tool, we can perform a customized configuration of the logic elements and impose the data path. However, the manual configuration of logic elements is tedious and time consuming in particular for systems with a certain complexity such as TDCs.

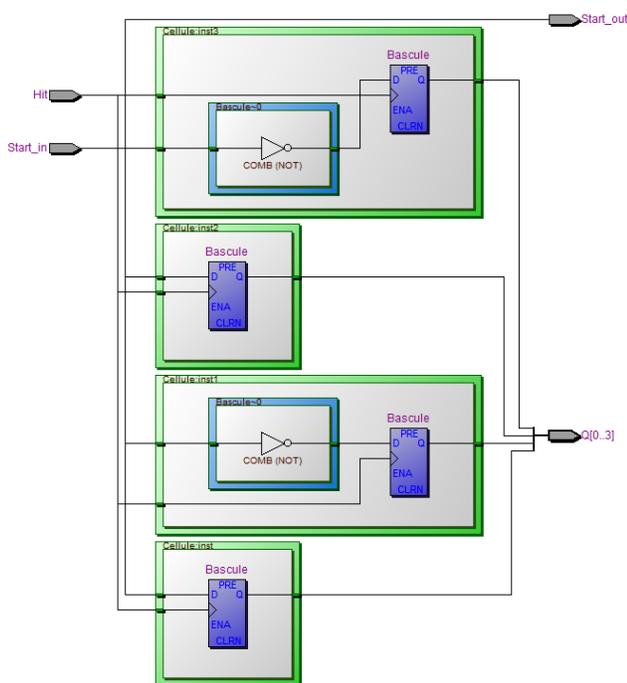


Figure 4. RTL view of the implementation of a simple 4-cells TDC

Even if we can use this technique to implement a TDC on an FPGA, given the large number of logic elements to be configured individually, it is still difficult to set up. Moreover, the TDC chain size can vary from an application to another; it will be therefore preferable to automate the configuration so that the solution will be generic and adaptable. Hence, we propose an appropriate design methodology in the following section.

III. DESIGN METHODOLOGY

In order to provide solutions to the above raised issues, in this section we suggest an alternative approach that can carry out a TDC structure fulfilling the following needs:

- Avoid the software data path simplification
- Increase TDC resolution by reducing the propagation time through delay elements
- Automate the elementary cells set-up process to optimize the design time and make possible the development of generic and adaptable structures
- Use a low cost FPGA target to implement the TDC.

This method is focused on two main areas:

- Using adders as delay elements and utilization of the Carry Chain Logic of the FPGA
- Using the Chip Planner tool.

A. Using Adders and Carry Chain Logic

The implementation of digital circuits on FPGA targets depends on the architecture of the logical resources of the target. In this work, we are aiming to use a low cost FPGA from Altera Cyclone family. The selected target is the Cyclone IV (EP4CE55F23C8) based on the logic element shown by Figure 5 [12].

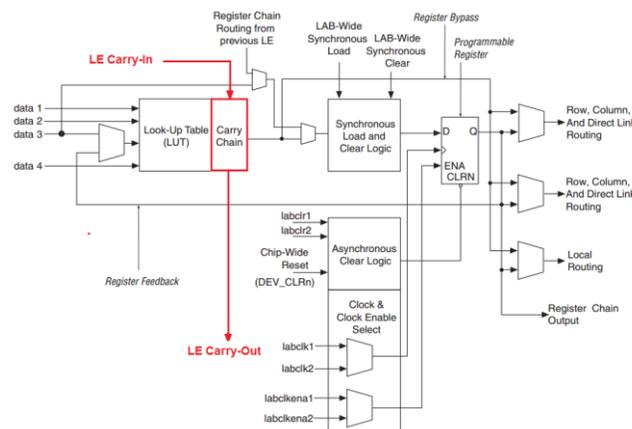


Figure 5. Cyclone IV logic element structure

The Cyclone IV logical element, given by Figure 5, provides a dedicated path for fast carry propagation. The role of this carry chain is to use specific fast paths for carry propagation instead of general-purpose routing network. By doing so, it makes it possible to drastically optimize the propagation time. This is ideal for the enhancement of the TDC resolution. Moreover, it allows to harmonize the delays of the TDC elementary cells.

The problem is that customized handling of carry chains is reserved to high performance FPGAs as such as the Stratix family from Altera whose cost is outstandingly high. However, it is possible to configure the Quartus II synthesis tool to optimize speed. In this case, the synthesis tool uses the carry chain logic automatically when synthesizing an HDL model involving adders.

It is therefore possible to use the carry chain logic to minimize and harmonize propagation delays for components involving adders. It is precisely the idea that is exploited here to design TDC elementary cells based on simple adders. This was done by developing a simple behavioral VHDL model for an adder with a customizable number of elementary cells. The number of the cells depends on the data width modeled by a generic parameter called DATA_WIDTH. The whole model is given by Figure 6.

The fine TDC using adders can be performed by: (1) applying the TDC input signal STOP to the carry input signal (*cin*) of the adder, and (2) choosing values for adder operand inputs (*a* and *b*) so that an output carry is generated (*cout*='1') if input carry is equal to '1'. The output carry is then an exact replication of the input carry delayed by a transmission time through the cell. To do so, all it takes is to set all the bits of the first operand to '1' and the bits of the second operand to '0'. For each bit (*i*) the arithmetic sum

$a(i)+b(i)$ gives '1'. When the input carry is activated ($cin='1'$) by the TDC input signal ($STOP$), the arithmetic sum $a+b+cin$ gives '0' and the carry output moves to '1'.

```

entity signed_adder is
    generic ( DATA_WIDTH : natural := 127 );
    port (
        a      : in signed ((DATA_WIDTH-1) downto 0);
        b      : in signed ((DATA_WIDTH-1) downto 0);
        cin    : in std_logic;
        cout   : out std_logic;
        result : out signed ((DATA_WIDTH-1) downto 0) );
end entity;

architecture rtl of signed_adder is
begin
    PROCESS (cin, a, b)
        VARIABLE s : signed ((DATA_WIDTH) downto 0);
        BEGIN
            s := ('0' & a) + ('0' & b) + ('0' & cin);
            result <= s((DATA_WIDTH-1) downto 0);
            cout <= s(DATA_WIDTH);
        END process;
    end rtl;

```

Figure 6. Adder VHDL model

Figure 7 illustrates the implementation of one elementary cell of a TDC by a logic element of the Cyclone IV target. The adder cell is obtained by the look up table (LUT) and the DFF by the sequential configurable output register.

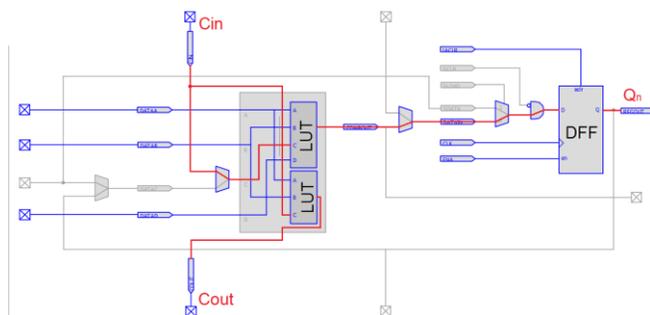


Figure 7. Implementation of a TDC elementary cell by a logic element

Theoretically, to obtain a TDC chain similar to the TDL structure shown by Figure 2, it is sufficient to duplicate the structure of Figure 7 as often as necessary to reach the number of desired cells. However, when implementing such a chain on the FPGA, some DFFs of the TDC elementary cells are dissociated of their corresponding 1-bit adder cells even if the data path is perfectly respected. This phenomenon occurs randomly and leads to the placing of the DFF and the delay element of the same TDC's elementary cell in different logic elements, as shown in Figure 8.

The direct consequence of component misplacing is that the delay time is no longer identical for all cells. This inevitably generates unpredictable artifacts. To ensure a reliable operation, it is necessary to overcome this problem by constraining the placement tool to bring together the

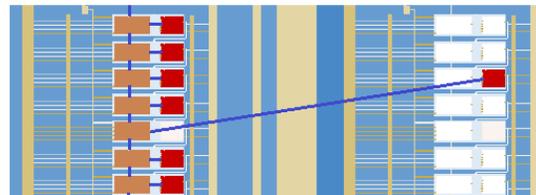


Figure 8. Random placing of DFFs on the chip

components of the same cell in the same logic element. This is the purpose of the next section.

B. Using Chip Planner

Using a TDC in fast imaging systems requires the measurement of very short time durations. It is therefore necessary to master all of the signal propagation delays through the cells as well as the routing network.

As we have seen in the previous section, unconstrained automatic implementation of a TDC on an FPGA usually leads to an inhomogeneous and irreproducible structure. Consequently, the measurement results are tainted by these uncertainties. Therefore, it is necessary to control the exact physical location of TDC cells on the chip.

This could be achieved by using the Chip Planner tool provided by Altera. The latter, according to the user's needs, allows the defining of specified implementation regions on the chip for blocks constituting the whole system. In addition, it supports incremental compilation to preserve the well-implemented parts and reduce the compilation time. This operation takes place in three distinct steps:

- Creating Design Partitions: the first step consists of dividing the design in individual partitions according to system complexity as well as user needs.
- Defining logic regions: after partitioning the design, it is necessary to define logical zones that will be associated to the partitions. This allows individual compiling and optimizing of each region. The tool used to perform this operation is LogicLock Region (LLR) within Chip Planner.
- Physical assignment of logic regions: in order to physically preserve the logic regions defined in the previous step, by means of the LLR tool, physical regions of the chip are assigned to implemented partitions.

The physical delimitation of regions permits to constrain the placing and routing tool to put partitions in their specified regions defined by the user. Doing so, it allows not only avoiding the random placement of certain DFFs away from their associated delay elements, but also implementing the concerned partitions as close as possible to input signal pins (HIT and $STOP$). The purpose of the latter operation is to reduce the propagation delays of input signal before reaching the blocks to which they are intended to be applied.

For illustrative purposes, we represent on Figure 9 the assignments of physical allocations of the partitions defined using LLR and a close-up view of the layout of a 16-cells fine TDC implemented using the method presented above.

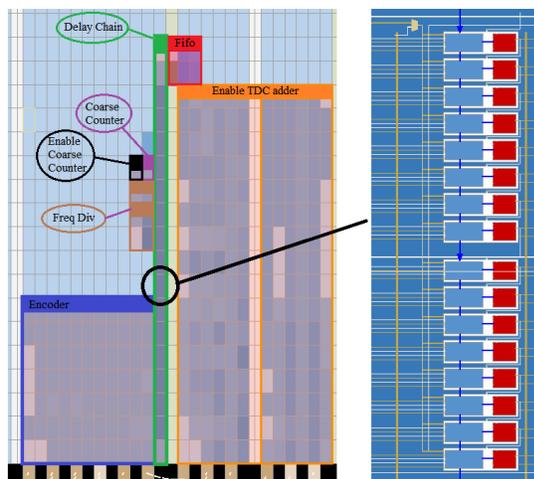


Figure 9. Layout of implemented partitions of a TDC

The TDC fits perfectly within the reserved region that would be assigned to it. Consequently, the DFF and the delay element of each TDC elementary cell are now implemented by the same logic element. The transmission delays are then identical for all cells.

IV. CASE STUDY: IMPLEMENTATION OF A 42 PS TDC ON A CYCLONE IV FPGA

The proposed TDC design has been implemented within the Cyclone IV (EP4CE55F23C8) FPGA target. The coarse counter clock is 200 MHz, i.e., the clock period is 5 ns. The delay line for the fine TDC, based on carry chain adder architecture, comprises 127 cells in order to cover a dynamic of more than 5 ns. The signal that needs to be measured propagates through the delay chain, until the FPGA clock disables the DFFs to block their outputs and then memorizes their states. The value of these DFFs describes the time spent between the signal STOP and Clock. An encoder placed at the output of the TDC prepares the data before saving it into a FIFO (First In First Out) memory. The data is then transmitted to a USB port via an FTDI FT232H operating in parallel mode with transfer rates reaching up to 40 Mbyte per second. To acquire data measurements, we developed a specific application using LabVIEW software. Figure 10 shows the synoptic view of the system.

The TDC has been characterized on its whole dynamic, i.e., from 0 to 640 ns with a step of 5 ps. A Stanford research DG 645 digital delay generator has been used to generate the START and STOP signals. At this range of delay, the jitter of the delay generator is lower than 25 ps rms. Figure 11 shows the detail of the unconstrained and constrained fine TDC measurements between two reference clock edges, i.e., on a range of 5 ns. The unconstrained fine TDC response shows a large discrepancy of the LSB value indicating that some DFFs have been randomly placed. The resulting large steps make the unconstrained fine TDC unusable for sub nanosecond timing. Consequently, the use of the Chip Planner tool as described in section III is mandatory to obtain the behavior of the constrained fine TDC represented by the

green curve. A linear fit is then used to assess the LSB value of the fine TDC which is given by the inverse of the linear fit slope, i.e., 41.5 ps in this study case.

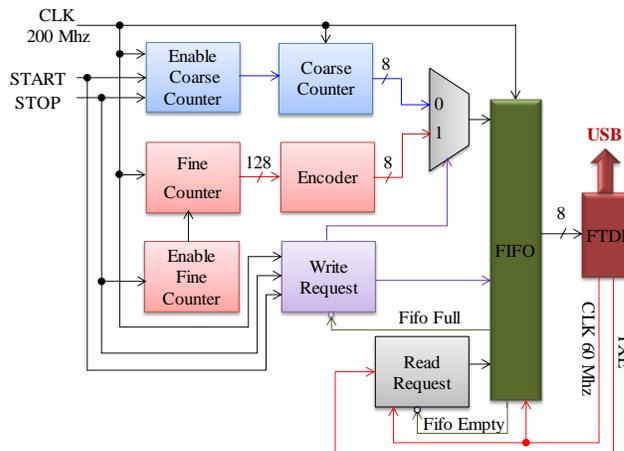


Figure 10. Synoptic view of the implemented TDC system

The noise visible on the fine TDC response is due to the jitter. The latter adds uncertainty on each measurement and it can be evaluated by computing the standard deviation of a set of measurements at a given fixed delay between the START and STOP signals.

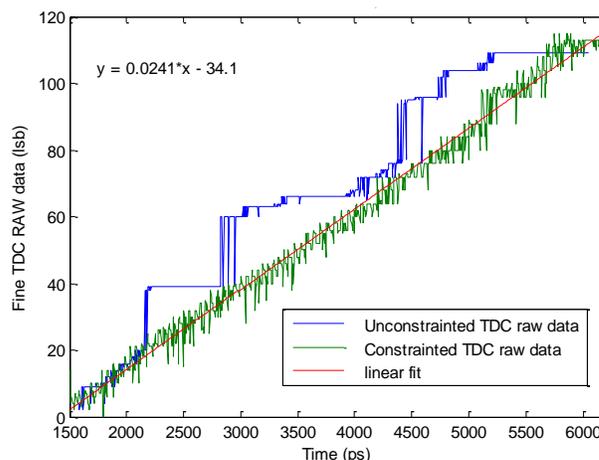


Figure 11. Responses of Fine unconstrained and constrained TDC

The jitter depicted in Figure 12 has been characterized for different delays corresponding to a given signal propagation along the fine TDC line. As each fine TDC elementary cell adds its own jitter [13], the global jitter will then increase as a square root of the number N cells as given by the following expression:

$$\sqrt{\alpha^2 + \beta^2 \cdot N} \tag{3}$$

Where: α is the initial jitter present at the input of the first cell and β the single cell jitter. A curve, following this law is fitted on the jitter profile to underline the jitter's variation relationship in the delay line. The extraction of this

parameters leads to an initial jitter α of 62 ps rms and a single cell jitter β of 5.8 ps rms. The jitter is mainly due to a noise presents on the 1.2 volt FPGA core power supply. The accumulated jitter across the fine TDC delay line leads to a mean jitter of 90 ps rms. Thus, the line length has to be kept as low as possible in order to obtain the best accuracy. This can be done by using the fastest achievable frequency for the coarse counter.

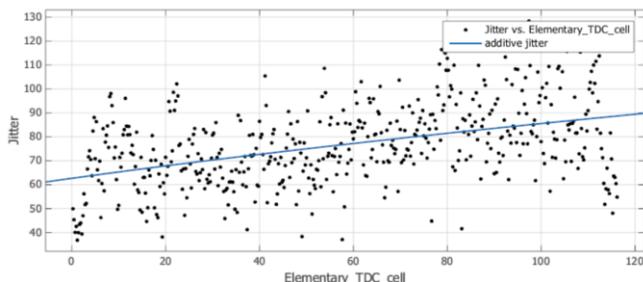


Figure 12. Jitter measurement according to the elementary level, the jitter increases as the signal propagates along the fine TDC cells

The integral non linearity error (INL) and the differential non linearity (DNL) have been measured over the entire range of the TDC. For illustrative purposes, the results from a delay of 0 to 160 ns are represented by the figure hereafter. From Figure 13, it can be seen that, the implemented system shows an INL of 132 ps rms and a DNL of 50 ps rms.

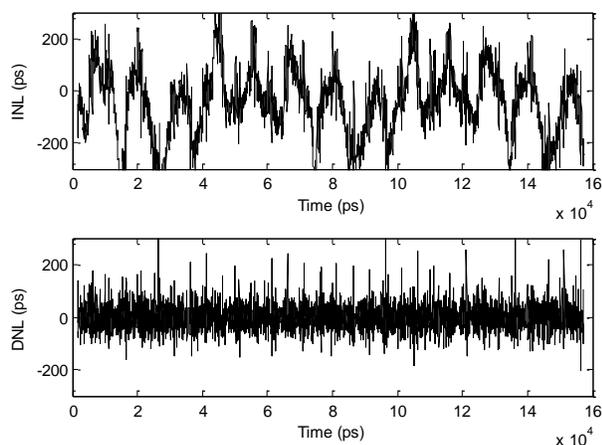


Figure 13. INL and DNL errors of the implemented TDC over a range of 160 ns

V. CONCLUSION

This paper has proposed a global methodology to design and implement Time-to-Digital Converters on low cost FPGA targets. It presents how to use different tools to enhance the TDC resolution by reducing propagation delays through the connection network as well as the logic gates themselves. First, the use of adders as delay elements, to benefit from a dedicated carry chain logic path, is presented. Then we detailed how to take advantage of the chip planner, to constrain the placing and root tool to put the partitions of

the system in user specified physical regions. Doing so, it allowed the mastering of propagation delays and consequently improved the resolution and the stability of the TDC.

The work is ended by a case study that applied this methodology to design a TDC with a resolution of about 42 ps on a Cyclone IV FPGA. The implemented TDC presents a jitter of less than 90 ps rms, and the DNL and the INL has been measured respectively to 50 and 132 ps rms.

The obtained results are very promising, not only because they are suitable for domains requiring high performances, but also because they are achieved by using a low cost FPGA family which opens the door to a broader use in a great amount of fast application fields.

As a perspective in the near future, we plan to integrate the presented TDC in different applications such as image photon counting devices and microfluidic experimentations.

REFERENCES

- [1] L. Li. "Time-of-flight camera – an introduction". Texas Instruments, SLOA190B – Technical White Paper, January 2014, revised May 2014.
- [2] E. Charbon, M. Fishburn, R. Walker, R. K. Henderson, and C. Niclass, SPAD-based sensors TOF Range-Imaging Cameras, Springer-Verlag, F. Remondino and D. Stoppa, Eds. Berlin Heidelberg, 2013, pp. 11–38.
- [3] S. Henzler, Time-to-Digital Converters, Springer Science+Business Media B.V., 2010, DOI: 10.1007/978-90-481-8628-0.
- [4] J. Kalisz, "Review of methods for time interval measurements with picoseconds resolution," Metrologia, vol.41, 2004, pp. 17–32.
- [5] C. S. Hwang, P. Chen, and H. W. Tsao, "A high-precision Time-to-Digital converter using a two-level conversion scheme," IEEE Tr. Nucl. Sci., vol. 51, 2004, pp. 1349–1352.
- [6] J. Kalisz, R. Szplet, J. Pasierbinski, and A. Poniccki, "Field-programmable-gate-array-based time-to-digital converter with 200-ps resolution," IEEE Trans. Instrum. Meas., vol. 46, 1997, pp. 51–55.
- [7] I. Vornicu, R. Carmona-Galán, and Á. Rodríguez-Vázquez, "Wide range 8ps incremental resolution time interval generator based on FPGA technology," 21st Int Conf. Electron. Circuits Syst (ICECS), 2014, pp. 395–398.
- [8] M. Lin, G. Tsai, C. Liu, and S. Chu, "FPGA-Based high area efficient Time-to-Digital IP design," in TENCON 2006, 2006 IEEE Region 10 Conference, 2006, pp. 1–4.
- [9] R. Narasimman, A. Prabhakar, and N. Chandrachoodan, "Implementation of a 30 ps resolution Time-to-Digital Converter in FPGA," Int Conf on EDCAV, Shillong, 2015, pp. 12–17.
- [10] A. Aloisio, P. Branchini, R. Giordano, V. Izzo, and S. Loffredo, "High-precision Time-to-Digital converter in a fpga device," IEEE Nuclear Science Symposium Conference Record, vol. 13, 2009, pp. 290–294.
- [11] S. S. Junnarkar, P. O'Connor, and R. Fontaine, "FPGA based self calibrating 40 picosecond resolution, wide range Time-to-Digital converter," IEEE Nuclear Science Symp. Conf. Rec. NSS '08, 2008, pp. 3434–3439.
- [12] Altera Corporation, Cyclone IV Device Handbook. Volume 1, chapter 2, April 2014.
- [13] M. Zlatanski, W. Uhring, J.-P. Le Normand, and D. Mathiot, "A Fully characterizable asynchronous multiphase delay generator," Nuclear Science, IEEE Transactions on, vol.58, no.2, 2011, pp.418–425.

Temperature Sensor for Hydro Generator Bearings using Thermally Regenerated Fiber Bragg Gratings

Uilian José Dreyer, Kleiton de Morais Sousa,
 Cicero Martelli, Valmir de Oliveira,
 Hypolito J. Kalinowski, Jean Carlos Cardozo da Silva
 Graduate Program in Electrical and Computer Engineering
 Federal University of Technology - Paraná (UTFPR)
 Curitiba, Brazil
 Email: jeanccs@utfpr.edu.br

Erlon Vagner da Silva
 Maintenance Department
 Tractebel Energia, UHE Salto Osório
 Quedas do Iguaçu, Brazil
 Email: erlon@tractebelenergia.com.br

Abstract—This work presents the design and experimental results of a temperature sensor based on quasi-distributed thermally regenerated Fiber Bragg gratings (RFBG) applied to the bearings of hydro generators. The temperature increase in the bearing is generally due to excessive shaft misalignment, low oil level or dirt in the lubricant oil. Expensive and time-consuming damage can be caused to the generator if the temperature sensors are not able to detect the fault in time and accurately. The proposed sensor has three wavelength multiplexed RFBG to monitor the vertical temperature gradient in the lubricant oil. The developed package is designed to avoid the cross sensitivity of the RFBG with mechanical forces. Thermally treated Fiber Bragg Grating (FBG) is heated till about 900 °C far away the temperature range in which the power plant operates, making its use suitable to obtain thermal stability for many years. Temperature stability tests were carried out for over fourteen hours in laboratory. The uncertainties were evaluated considering the sample dispersion and equipment characteristics used during the sensor temperature calibration. The experimental results show that the uncertainty associated to each FBG matches the power plant metrological requirements. The developed package for bearing temperature measurement proved to be robust both mechanically and thermally.

Keywords—*Electrical generator bearing; Optical Regenerated Fiber Bragg Gratings; Temperature measurement.*

I. INTRODUCTION

The Brazilian energy sector is predominantly hydroelectric which represents approximately 62 % of the total energy generated in the country. The installed electrical energy capacity reaches 90.2 GW in a total of 145.5 GW, meaning about five times the capacity of thermal power plants that use natural gas fuel, which is the second largest Brazilian electrical energy source with about 18.1 % of the total [1]. The electric generators used in large power plants have sizes normally reaching hundreds of Mega Watts (MW). They have also heavy structures with dimensions proportional to its power causing the cooling and alignment systems to be an important part of these large machines.

Therefore, hydroelectric power plants must operate uninterrupted for long periods of time before a preventive maintenance of the generation system is required. The electrical generator has a very important role in the generation system because of it is the machine that converts the mechanical

energy coming from the hydraulic turbine into electric energy. However, power generators are susceptible to adverse situations such as unbalanced loads, short-circuits in the transmission line among others. These situations cause, mainly, a higher loss in the windings of the generator (stator and rotor) and therefore a temperature rise above the expected and planned for its operation. When the temperature is above the nominal it causes premature wear of the insulating materials reducing the service life of the generator [2].

In Figure 1, there is a schematic drawing showing the bearings position along the hydro generator structure. The bearing indicated by 1 is the one that supports the vertical force exerted by the weight of the hydraulic turbine (HT) and electrical generator (EG) assembly. The bearings indicated by 2 and 3 are intended to not allowing the misalignment of hydro generator in the horizontal direction. The EG is a three-phase synchronous generator, with a rated voltage of 13.8 kV between phases, and rated current of 7.1 kA. The generator stator has a diameter of 12 m and 2.5 m in height. The structure weighs over 200 tons implying serious safety problems if any alignment issue takes place. With a reliable system to monitor temperature continuously, small temperature increases can be detected before causing further damage.

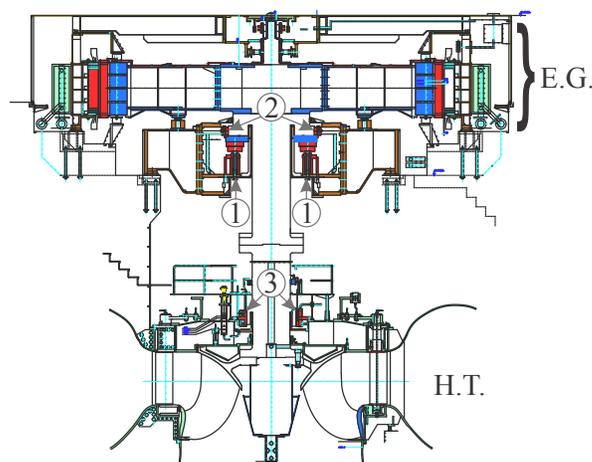


Figure 1. Drawing showing the bearing positions into a 200 MVA hydro generator installed in Salto Osório power plant.

Apart from thermal induced drawbacks, mechanical align-

ment also can lead to faults in the large hydroelectric power generators. The bearings must compensate potential misalignment in the shaft that connects the hydraulic turbine to the generator. The torsional efforts coming from different power cycles that the generator is subject to must also be compensated [3][4]. When the generator operates by varying its power, there is a proportional temperature variation on the bearings, therefore increasing their temperature with the increase in the generated power. There are other factors like lack of lubrication or dirty contaminants in lubricant reservoir, which increase temperature, sometimes leading to the generator shutdown and the necessity to perform corrective maintenance. In Figure 2, it is presented a damaged bearing by metal pieces that fell into the lubricant oil causing the generator shutdown by temperature increase. Another important factor that could affect the bearings temperature is the short-circuit stator winding which induces abnormal electromagnetic flux in the air gap, and, as a result, the excessive vibrational movements in the generator-turbine shaft [5].



Figure 2. Photography of a damaged bearing that is disassembled to perform corrective maintenance.

In this work, we present the development and results of an optical fiber sensor developed to monitor bearings in hydro generator that avoid FBG cross-sensitivity [6]. The proposed design uses three wavelength multiplexed and encapsulated Regenerated Fiber Bragg grating sensors (RFBG) that form a thermal transducer of 1.5 m long and has sensing points distributed along its body. The FBG has some promising features that make it wide applicable in electric machines [7][8]. A FBG sensor [9] consists of an optical fiber passive-device that have shown fast insertion in measuring and control applications. The FBG led to new developments in optical filtering and in the design of sensors. FBG temperature sensors have some advantageous characteristics, such as low mass, immunity to electromagnetic interference, small signal attenuation and the possibility of being multiplexed [10]. Thus, it has the ability to solve measurement issues for which conventional sensing techniques are not appropriate, as in the case of the thermal analysis inside large electric generators working at high voltages.

The next sections provide a more specific approach to the main issues related to the developed optical fiber sensor. A brief review about FBG and RFBG is presented in Section II highlighting this main characteristics. In Section III it is presented the packaging assembly, the temperature calibration process and the method used to determine uncertainties. The experimental results are presented in Section IV. Finally, the conclusions obtained from the temperature tests and future work are presented in Section V.

II. THERMALLY REGENERATED FIBER BRAGG GRATINGS

An FBG is formed by a periodic change in the refractive index caused in a fiber core with an effective refractive index n_{eff} by exposure to an ultraviolet (UV) laser beam [6]. The index change is perpendicular to the longitudinal axis of the fiber with constant periodicity Λ . The index modulation inside the fiber builds a resonant structure to a determinate wavelength as shown in Figure 3. The center peak wavelength is called Bragg wavelength (λ_B) and is given by (1):

$$\lambda_B = 2n_{eff}\Lambda. \quad (1)$$

Figure 3 (a) shows an optical fiber in which three FBGs are inscribed with different modulation period Λ . The different periods lead to different resonances Bragg wavelength λ_B . When the Bragg condition is satisfied (1), the light reflected back by each modulation contributes constructively. Then a central wavelength of the reflected light is defined by the parameters of the grating [13]. The reflected spectrum (S_r) is similar to a very effective band-pass filter and the transmitted spectrum (S_t) resembles a reject-band filter as shown in Figure 3 (b).

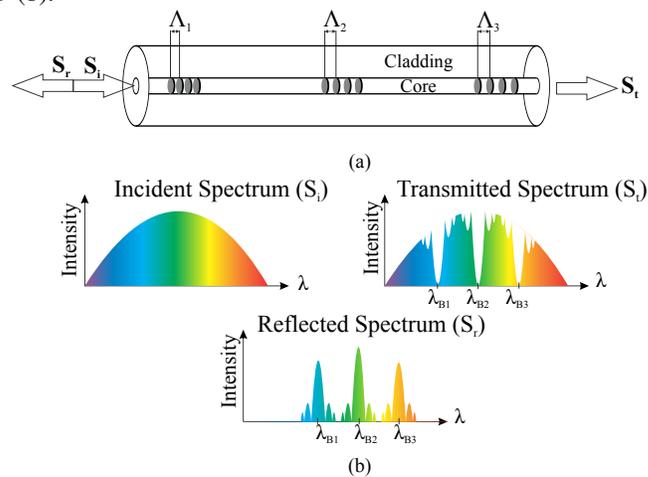


Figure 3. Schematic representations of three wavelength multiplexed FBGs inscribed into an optical fiber (a) and its associated spectra (b).

The Bragg wavelength (λ_B) has dependency on two parameters: effective refractive index and grating periodicity. The change on one of these two parameters from external interference produces a shift in λ_B . The mechanical and thermal environment changes are related to the wavelength shift through the expression:

$$\Delta\lambda_B(l, T) = 2 \left(\Lambda \frac{\partial n_{eff}}{\partial l} + n_{eff} \frac{\partial \Lambda}{\partial l} \right) \Delta l + 2 \left(\Lambda \frac{\partial n_{eff}}{\partial T} + n_{eff} \frac{\partial \Lambda}{\partial T} \right) \Delta T. \quad (2)$$

The first term in (2) represents the strain effect on an optical fiber. It is associated with the shift on λ_B due to the change of the refractive index, called photo-elastic coefficient, and the length of the FBG caused by the optical fiber mechanical deformation. As for the second term in (2), λ_B shift is associated with the thermal expansion coefficient and the thermo-optic coefficient of silica. The influence of the thermo-optic coefficient surpasses by far the thermal expansion coefficient in a silica optical fiber. It is also possible to observe that the FBG has a cross sensitivity between strain and temperature. So, when it is desired to build a temperature sensor it is necessary to develop a package that eliminates the strain effect.

There is a concern to improve the thermal properties of FBG to allow temperature sensing applications with improved the temperature stability of the gratings. This has involved experimenting with various dopants, including Sn and Na₂O [14]. Previous studies have already established that temperature range and stability of FBGs can be improved by several means, including formation of type IIA [15][16][17] and type II [18][19] gratings, including those inscribed using femtosecond IR lasers.

For many applications, the thermal stability of laser induced glass changes determines the limits in which they can operate, for power plant applications, FBGs that can operate at 80 °C for 25 years are required. Thermally regenerated FBG play a special role in temperature measurement inside power plants because of it has great thermal stability within the power plant temperature range [11][12]. In order to make a RFBG, it is inscribed a strong conventional Bragg grating into hydrogen loaded optical fiber called seed grating. Afterwards, the strong seed grating is thermally processed at about 900 °C and the regeneration is observed. Over time as the seed grating disappears completely, a regenerated grating appears [20][21].

The temperature sensing in hydroelectric generators has a great importance as a performance indicator of the machine, and it is frequently measured using RTD or thermocouples as point sensors. Although widely used, such sensors can easily be interchange by new technologies like optical fiber sensors, especially those designed with FBG, as presented in this work. Optical fiber sensing applications in hydro generators were presented in [22][23] demonstrating the monitoring of temperature, vibration, magnetic field among others generator parameters. There is also the use of distributed temperature sensing on the stator surface of 200 MVA generator [24]. Furthermore, the temperature monitoring in bearings using FBG was developed and tested in laboratory during 12 hours [25].

III. PACKAGING AND TEMPERATURE CALIBRATION

The seed FBGs used in this work were produced at the Photonics Laboratory of UTFPR, using ArF Excimer lasers (Coherent Xantos) operating at 193 nm. The gratings were directly written using a phase-mask with typical laser pulse energy of 5 mJ and repetition rate of 250 Hz, with exposition on the order of 5 minutes. Each grating has approximately 3 mm in length.

The fiber samples are hydrogen loaded for several days under 110 atm, 25 °C before the seed grating inscription. The grating was kept at room temperature after inscription to diffuse hydrogen before being subjected to the regeneration process inside a resistive oven. Using a processing procedure identical to [26], strong seed gratings were thermally processed with temperatures between 21 °C to 900 °C (± 15 °C) during approximately 2 h 30 min. At 900 °C the regeneration is observed, and over time as the seed grating is reduced almost completely, the regenerated grating appears. After that, the oven is turned off and the grating remained resting inside to complete the annealing process. Figure 4 shows the spectrum of the seed FBG (dashed line) and RFBG (solid line) after the annealing process.

Due to the harsh environment and the fragility of pristine FBG sensors, the FBG must be packaged with special materials

before installation, to avoid the damage to the fiber sensors. Another feature of the package should be the minimization of the strain-temperature cross sensitive of FBG. In this work the package is formed by one steel rod with an external diameter of 3.18 mm and an internal hole with a diameter of 1.7 mm. Fig 5 (a) and (b) show pictures of the sensor package for bearing sensor.

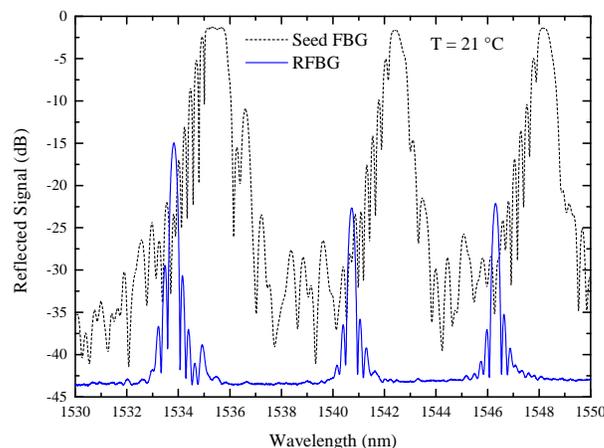


Figure 4. Spectrum of the seed FBG (dashed line) and regenerated FBG (solid line) after the thermal treatment.

In Figure 5 (a), there is a drawing showing the main sensor parts. The sensor presents a splice chamber to keep the optical splices safe and an optical cable exit to connect the RFBGs to the internal optical aqua plant link. The junction applies a force to the collet, it causes a pressure to hold the steel rod on the sensor and inside the thermowell. The bearing sensor has three RFBGs spatially positioned close to it edges (Figure 5 (b)). This distributed sensor configuration was designed because of the three FBG will be immersed in the oil to detect the temperature gradient along the vertical direction.

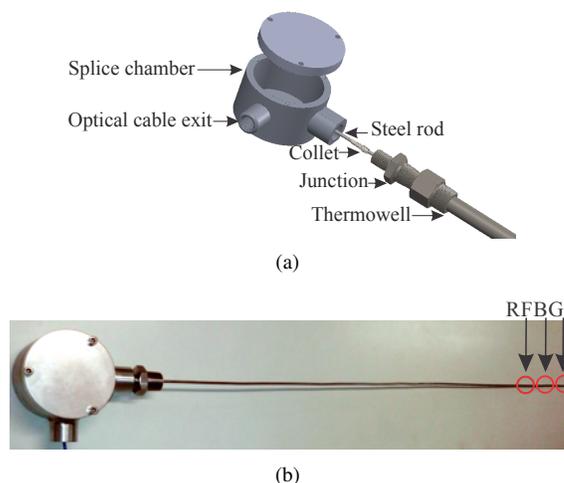


Figure 5. Optical fiber sensor assembly (a) mechanical modeling and (b) photo.

The calibration system and process are based on the diagram presented in Figure 6 (a). The optical fiber sensor (OFS) and the Platinum Resistance thermometer (PT100) sensor are placed inside a copper tube in which water is pumped through it using the thermal bath flow (TB). While the flow is maintained, a digital multimeter (DM) and the optical

interrogator (OI) measure the PT100 and the Bragg gratings, respectively. These are all connected/controlled by a home made software implemented in LabVIEWTM, that automate the calibration process as well as the uncertainty calculations [27].

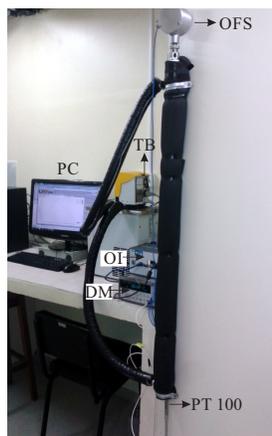
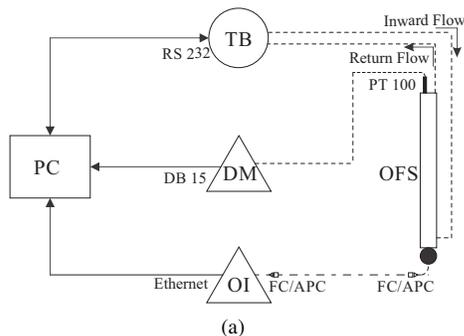


Figure 6. Designed calibration system (a) function diagram and (b) assembly photo.

The measurement uncertainty usually encompasses many requisites. Some can be estimated through a type A evaluation of measurement uncertainty from the statistical distribution values. The values come from several measurements and can be characterized by standard deviations. Other components can be estimated through a type B evaluation of measurement uncertainty. They may also be characterized by standard deviations estimated from probability density functions based on user experience or other information [28][29]. The A and B type uncertainties followed the diagram shown in Figure 7 to this particular sensor.

The calibration process consisted of heating cycles repeated four times, from 10 °C to 80 °C for the optical fiber sensor. The largest uncertainty associated with each sensor is presented in Table I. With the temperature calibration points, a corresponding function is shown to represent the FBG temperature (T) response with its Bragg wavelength as showed in Figure 8 to each RFBG. In this case, a second order polynomial function in (3) is fitted with the α (second order) and β (first order) coefficients presented in the Table I, while the constant coefficient γ , which is characteristic for each grating resonance Bragg wavelength (λ):

$$T(\lambda) = \alpha\lambda^2 + \beta\lambda + \gamma \quad (3)$$

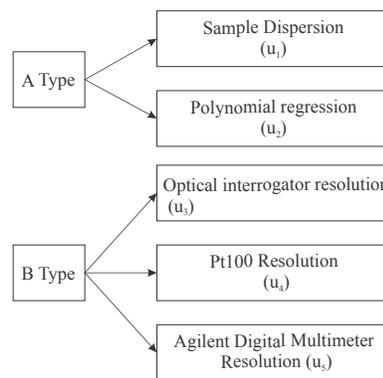


Figure 7. Diagram with parameters involved in uncertainty evaluation for this particular project.

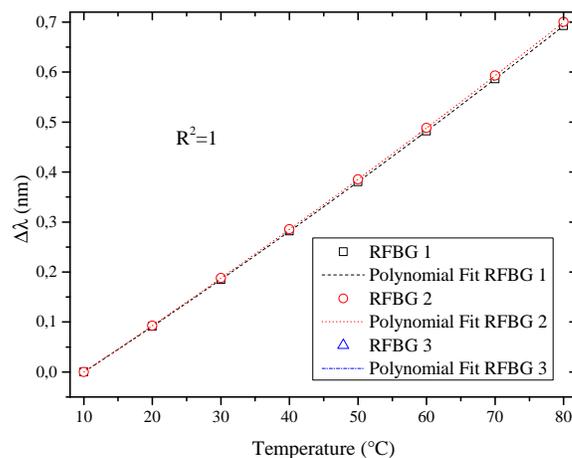


Figure 8. Figure of the polynomial fit for the temperature calibration points using the optical fiber sensor with RFBG.

Sensor	Function fit coefficients	Uncertainty
RFBG 1	$T(\lambda) = 1.3 \times 10^{-5} \lambda^2 + 8.748 \times 10^{-3} \lambda + 1536,951$	0.78 °C
RFBG 2	$T(\lambda) = 1.2 \times 10^{-5} \lambda^2 + 8.916 \times 10^{-3} \lambda + 1540,406$	0.38 °C
RFBG 3	$T(\lambda) = 1.2 \times 10^{-5} \lambda^2 + 8.886 \times 10^{-3} \lambda + 1543,405$	0.36 °C

The fitted functions, having their coefficient values presented in Table I, are approximately the same for every grating of each individual optical fiber sensor.

IV. RESULTS AND DISCUSSIONS

The same calibration system was used for testing the bearing sensor in order to apply temperature step function excitations such as to compare the FBG measurements with a PT100 used in the generator. The thermal cycles ranged from 15 °C to 75 °C, lasting for 15 hours. The thermal cycles had the goal of emulating four loading levels in the hydro generator (Figure 9): 15 °C - turned off, 35 °C - energy compensator, 55 °C - nominal charge and 75 °C - overloaded. The overload occurs at temperatures above 70 °C and the generator is turned off by the temperature system protection. In Figure 9, the 'A' inset indicates a steady state period during the experiment.

It is possible to observe in Figure 9 that the optical fiber Bragg grating sensor temperature measurements follow the

same trend as the PT100 sensor. A small difference is observed in the time response (inset) that can be accounted for the distinct positioning inside of the test chamber of the thermal bath. In four temperature cycles the 3 RFBGs temperature sensors measured the same value as the PT100 sensor and the small differences are smaller than the determined uncertainty for each individual sensor.

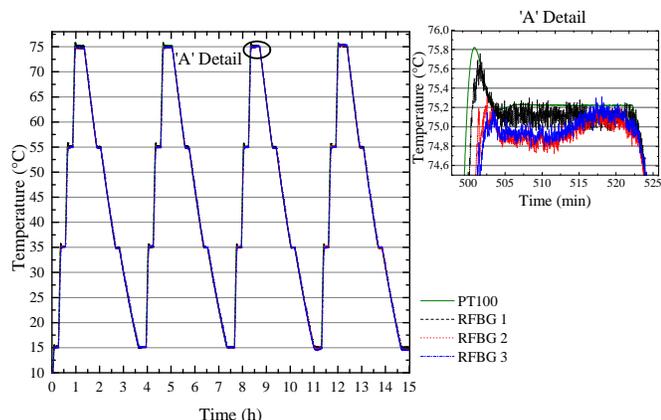


Figure 9. Laboratory essay lasting over 15 h with the 3 RFBG compared to the PT100 sensor.

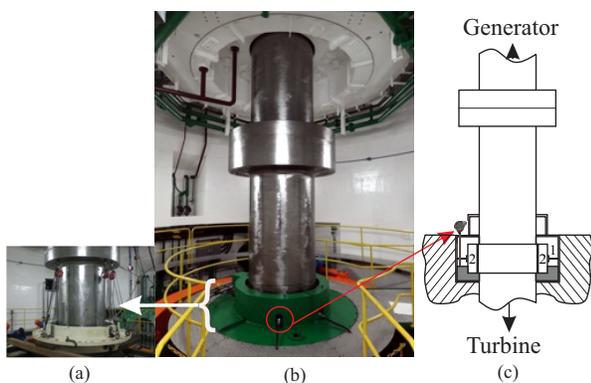


Figure 10. (a) Babbitted shoe metal pulled off from the bearing assembly to perform a maintenance procedure. (b) Bearing photography inside the power plant and (c) the schematic representing the parts that play an important role on the bearing .

The bearing where the RFBG optical fiber sensor will be installed is shown in Figure 10 (b)-(c), and the details (1) – (2) represent the lubricant oil and babbit shoe metal, respectively, being those two points to be monitored. The lubricant oil has the role of reducing the friction between the shaft and the babbit shoe metal, and its temperature is an indicator of the machine operation status. Figure 10 (a) also shows the babbitted shoe metal being pulled off from the bearing assembly to undertake a maintenance procedure.

V. CONCLUSION

The potential of integrating fiber Bragg grating sensors for multiplexed on-line monitoring of temperature and temperature gradients has been experimentally investigated in laboratory and inside a power plant before this presented work [25]. In general, it is noticed that the temperature is fully consistent with the actual operation of the bearing. The presented analysis can indicate the operating condition of the equipment and the results demonstrate that this technique is a potential maintenance tool for predictive purposes with a high efficiency

to continuously monitor the hydro generator facilities. The generator designers could also use these measurements to improve the machine performance and establish their limits of safe operation. Future work include temperature measurements on a bearing to measure the temperature in the turbine bearing of a hydroelectric 180 MW generator inside the power plant.

ACKNOWLEDGMENT

The authors acknowledge the financial support of ANEEL, CNPq, CAPES, FINEP, Fundação Araucária, and Tractebel Energia. This project is developed under Tractebel Energia’s R&D program (PD-0403-0028/2012), regulated by ANEEL.

REFERENCES

- [1] ANEEL, “Brazilian Electricity matrix,”. URL: www.aneel.gov.br [retrieved: 06–2015].
- [2] G. Stone, and J. Kapler, “Stator winding monitoring,” *Industry Applications Magazine*, IEEE, vol 4, 1998, pp. 15–20.
- [3] C. E. J. Bowler, P. G. Brown, and D. N. Walker, “Evaluation of the effect of power circuit breaker reclosing practices on turbine-generators shafts,” *Power Apparatus and Systems, IEEE Transactions on PAS-99*, vol. 5, 1980, pp. 1764–1779.
- [4] D. Stojanovic, D. Petrovic, and N. Mitrovic, “Torsional torques of big turbine-generator shafts due to malsynchronization,” in [Electrotechnical Conference, 2000. MELECON 2000. 10th Mediterranean], vol. 3, 2000, pp. 1051–1054.
- [5] A. Tetreault, “Rotor shape vs. rotor field pole shorted turns: Impact on rotor induced vibrations on hydrogenerators,” in [Condition Monitoring and Diagnosis (CMD), 2012 International Conference], 2012, pp. 133–136.
- [6] A. Othonos and K. Kalli, [Fiber Bragg gratings], vol. 68, 1999, pp. 4309-4341.
- [7] K. M. Sousa, A. A. Hafner, H. J. Kalinowski, and J. C. C. da Silva, “Determination of temperature dynamics and mechanical and stator losses relationships in a three-phase induction motor using fiber bragg grating sensors,” *Sensors Journal, IEEE*, vol. 12, 2012, pp. 3054–3061.
- [8] N. M. Theune, et al. “Multiplexed temperature measurement for power generators,” *Proc. SPIE*, vol. 4074, 2000, pp. 214–221.
- [9] R. Kashyap, [Fiber Bragg Gratings], Electronics & Electrical, Academic Press (1999).
- [10] T. Erdogan, “Fiber grating spectra,” *Journal of Lightwave Technology*, vol. 15, 1997, pp. 1277–1294.
- [11] J. Canning, S. Bandyopadhyay, M. Stevenson, and K. Cook, “Ultra-High Temperature (UHT) Gratings,” vol. no , 2008, pp. 1–3 (2008).
- [12] J. Canning, et al., “Regenerated gratings,” *Journal of the European Optical Society - Rapid publications*, vol. 4, 2009, ISSN 1990-2573.
- [13] K. O. Hill and G. Meltz, “Fiber Bragg grating technology fundamentals and overview,” *Journal of Lightwave Technology*, vol. 15, 1997, pp. 1263–1276.
- [14] G. Brambilla and P. Hua, “Phase separation in highly photosensitive tin-codoped silica optical fibers and fiber preforms exposed to {UV} radiation,” *Journal of Non-Crystalline Solids*, vol. 352, 2006, pp. 2921 – 2924.
- [15] I. Riant and F. Haller, “Study of the photosensitivity at 193 nm and comparison with photosensitivity at 240 nm influence of fiber tension: type IIa aging,” *Journal of Lightwave Technology*, vol. 15, 1997, pp. 1464–1469.
- [16] O. Prakash, et al., “Enhanced temperature 800 °c stability of type-ii-a fbg written by 255 nm beam,” *Photonics Technology Letters, IEEE*, vol. 26, 2014, pp. 93–95.
- [17] E. Lindner, et al., “Thermal regenerated type {IIa} fiber bragg gratings for ultra-high temperature operation,” *Optics Communications*, vol. 284, 2011, pp. 183 – 185.
- [18] H. Y. Liu, H. B. Liu, G. D. Peng, and P. L. Chu, “Observation of type i and type {II} gratings behavior in polymer optical fiber,” *Optics Communications*, vol. 220, 2003, pp. 337 – 343.

- [19] D. Grobnc, S. J. Mihailov, J. Ballato, and P. D. Dragic, "Type i and ii bragg gratings made with infrared femtosecond radiation in high and low alumina content aluminosilicate optical fibers," *Optica*, vol. 2, 2015, pp. 313–322.
- [20] J. Canning, M. Stevenson, S. Bandyopadhyay, and K. Cook, "Extreme Silica Optical Fibre Gratings," *Sensors*, vol.8, 2008, pp. 6448–6452.
- [21] B. Zhang and M. Kahrizi, "High-temperature resistance fiber bragg grating temperature sensor fabrication," *Sensors Journal, IEEE*, vol. 7, 2007, pp. 586–591.
- [22] M. Willsch, "Fiber Optical Sensors in Power Generation," in [Third Asia Pacific Optical Sensors Conference], vol. 8351, 2012, pp. 835137–1–835137–9.
- [23] C. Martelli, et al., "Temperature sensing in a 175MW power generator," in [OFS2012 22nd International Conference on Optical Fiber Sensors], *Proc. SPIE*, vol. 8421, 2012, pp. 84212F–1–84212F–4.
- [24] F. Mezzadri, C. Martelli, E. V. Silva, J. P. Bazzo, and J. C. C. Silva, "175MW Hydroelectric Generator Stator Surface Temperature Monitoring using a DTS System," *IEEE Sensors Journal*, vol. no , 2014, pp. JM5A.64.
- [25] E. V. da Silva, et. al., "Optical fiber instrumentation of a high power generator and turbine," *European*, vol. 8794, 2013, pp. 879446–1–879446–6.
- [26] V. Oliveira and H. J. Kalinowski, "Strongly regenerated Bragg gratings in standard single-mode fibres," *Methodology*, vol. 7653, 2010, pp. 765312–765316.
- [27] S. G. Rabinovich, "Measurement errors and uncertainties," third edition, 2005.
- [28] B. N. Taylor, and C. E. Kuyatt, "Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results," Technical note 1297, 1994.
- [29] A. C. Baratto, et al., "Evaluation of measurement data – Guide to the expression of uncertainty in measurement", 2008.

A Novel Elliptically-Slotted Patch Antenna-based Biosensor Design

Sunday Ekpo, Vijayalakshmi Velusamy and Rupak Kharel

Department of Electrical & Electronic Engineering

Manchester Metropolitan University

Manchester, United Kingdom

Email: r.kharel@mmu.ac.uk

Abstract— Antenna-based biosensors have attracted an increasing research interest in the biomedical and healthcare community due to their reliability, sustainability, portability, scalability and cost-effectiveness. These benefits depend on a holistic functional design and modelling of the antenna subsystem. This paper presents the design and modelling of a novel elliptically-slotted patch antenna (ESPA) based biosensor that yields a total gain of 7.5 dBi. This novel biosensing antenna is label-free, amenable to real-time embedded circuit-emulating implementations and resistant to all solvents and reagents.

Keywords- biomedical sensors; biosensor; antenna based biosensor; patch antenna.

I. INTRODUCTION

New approaches for biological cell detection are being developed to improve the biosensor characteristics such as sensitivity, smaller footprint, selectivity, integration with microfluidic devices for sample handling, reproducibility, non-invasive and rapid real-time detection. Recently, radio frequency (RF) and microwave antenna biosensors have drawn attention in the development of biological cell detection due to their numerous applications including clinical analysis, real-time monitoring (of food, water and environment) and biosecurity [1-3]. Most the intensive research interests have revolved around biological and material science fields. Antenna biosensors enable the real-time detection of the electrical and/or magnetic properties of biological materials (potentially also for bacteria like *E. coli*) at RF/microwave frequency. This approach provides a smart real time detection of biological samples without the need for biosample labelling, material intrusion and chemical transformation.

The detection principle of antenna biosensors is to sense and characterise the changes of the return loss (S_{11}) (or reflection coefficient, Γ) and the dielectric properties (i.e., relative permittivity and loss tangent) in the useable resonant RF/microwave bandwidth caused by the biosample. The dielectric properties of various biosamples at frequency 2.79 GHz such as blood, cortical bone, blood vessels, fat and muscle are given in Table 1. The development focus has been on metamaterials that can provide the needed trade-off between the practical antenna

size and the designed resonant wavelength. We investigated and modelled various novel patch antenna configurations to provide the optimal sensing location for the biosample detection. The elliptically-slotted patch antenna geometry with an eccentricity of $\cos 60^\circ$ (or $\sqrt{0.75}$) yielded the best solution for the designed antenna size and resonant wavelength. RF/microwave antenna have been considered to be more suitable for real-time biosensing due to their unique properties such as non-invasive and non-ionising radiation sensing. These make RF/microwave antenna-based biosensors the more appealing alternatives for specific transducers currently employed in the development of biosensors [1, 4, 5]. The biomedical application domain of the reported novel ESPA system can be extended to include spacecraft-borne space-based wireless biosensing devices [6, 7].

This paper is organised as follows. Section II presents the patch antenna system design including the system design parameters and the elliptically-slotted patch antenna system model. In section III, the simulation parameters, results and discussion are stated. The paper is concluded in section IV.

II. PATCH ANTENNA SYSTEM DESIGN

A. System Design Parameters

The resonant frequency, f_r , of a patch antenna is given by:

$$f_r = \frac{1}{2\pi\sqrt{LC}} \quad (1)$$

where the system capacitance, C , is given by:

$$C = \epsilon_{\text{eff}} \frac{A}{d} \quad (2)$$

and ϵ_{eff} , the effective permittivity (i.e., the average complex permittivity of the ESPA printed circuit board (PCB) and the biosample-under-test (BUT); the inductance, L , is introduced into the feed system through the coaxial feed; A , the effective area (m^2) of the finite substrate relative to the patch and d , the dielectric separation or thickness (m). The operating (or centre-design) frequency of a microstrip (patch) antenna scales with its length, l , according to the following equation thus:

TABLE 1: DIELECTRIC PROPERTIES OF VARIOUS BIO SAMPLES AT 2.79 GHZ

	Density (kg/m ³)	Relative Permittivity ϵ/ϵ_0	Loss Tangent $\tan \delta$	Conductivity σ (S/m)
Blood	1060	57.7	0.31809	2.8581
Cortical Bone	1750	11.184	0.26642	0.46249
Fat	909.4	5.2452	0.14735	0.11996
Muscle	1059.9	52.314	0.244	1.9812
Blood Vessels	1060	42.114	0.2539	1.6596

$$f_c \approx \frac{1}{2l\sqrt{\epsilon_o\mu_o\epsilon_r}} = \frac{c}{2l\sqrt{\epsilon_r}} \quad (3)$$

where ϵ_o is the permittivity in free space (F/m), ϵ_r , the relative permittivity and μ_o , the permeability in free space (H/m).

From (3), the length of the patch antenna should be equal to one-half of a wavelength within the dielectric medium.

The average complex permittivity, ϵ_{eff} , comprises the real and imaginary components given by:

$$\epsilon_{eff} = \epsilon_{eff}' - j\epsilon_{eff}'' \quad (4)$$

$$\epsilon_{eff} = \epsilon_o\epsilon_r(1 - j \tan \delta) \quad (5)$$

$$\epsilon_{eff} = \epsilon_o\epsilon_r - j\frac{\sigma}{\omega} \quad (6)$$

where ϵ_{eff}' is the lossless electromagnetic energy transmittance (EET), ϵ_{eff}'' , the lossy EET, δ , substrate conductivity and ω , the frequency of operation.

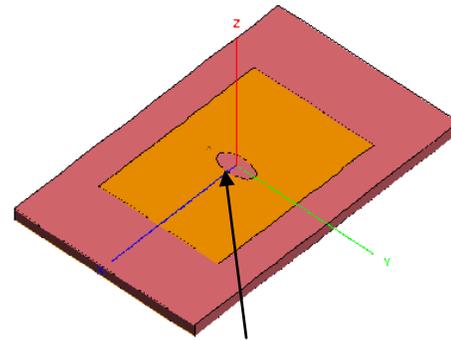
The EM-wave energy loss during transmittance is called the loss tangent, $\tan \delta$, and defined as the ratio of the lossy to the lossless EM-wave energy transmittances. The width of the patch influences the input impedance and hence, the antenna bandwidth. The larger the width of the patch antenna, the broader its bandwidth. Also, an increase in the antenna width results in a decrease in its input impedance (at the expense of a large footprint patch antenna system).

B. ESPA System Model

The CAD model of the ESPA system is created and Figure 1 shows the front view, where the elliptical slot is seen at the centre of the patch.

The substrate is a Rogers RT duroid[®] glass fiber reinforced polytetrafluoroethylene (PTFE) composite designed for microstrip circuit applications. The actual patch component is a substrate-centralised elliptically-slotted copper metal with a thickness of 1–1.6 mm. The ESPA system is excited with a pin, such as a SubMiniature version A (SMA) connector. The ESPA feed assembly is a coaxial cable or probe feed is from underneath through the ground plane. The outer conductor of the coaxial cable is connected

to the finite ground plane and the centre conductor passes through the substrate to the patch antenna.



Elliptical Slot on a Patch

Figure 1. ESPA System (Front View)

An ESPA-based biosensor system model carries the biological sample in the elliptical slot located at the centre of the patch antenna.

III. SIMULATION RESULTS AND DISCUSSIONS

A. Slotless versus Slotted Patch Antenna

A novel ESPA was designed, modelled and simulated for biosensing applications. Table 2 shows the parameters of the ESPA system utilised for simulation. The model was simulated over finite and infinite grounds. The performance of the proposed ESPA system was compared with a conventional slotless patch antenna (CSPA). There is a considerable improvement in performance, weight, size and power consumption requirements of the ESPA design over the traditional patch antenna. Furthermore, the ESPA model allows for an integrated combinational design of metamaterials to be implemented.

TABLE 2. ESPA SYSTEM SIMULATION PARAMETERS

Model Variable	Value
f_{Max}	3.0
f_{Min}	2.6
λ_{Min}	c/f_{max}
c	$3*10^8$
P_d	$0.332\lambda_{Min}$

P_w	$0.468\lambda_{Min}$
S_h	$0.0287\lambda_{Min}$
F_d	$0.089\lambda_{Min}$
F_r	$0.00065\lambda_{Min}$
ϵ_r	2.2
R_{min}	$0.025\lambda_{Min}$
R_{maj}	$0.05\lambda_{Min}$
$\tan \delta$	0.0004

B. ESPA Performance Analysis

Figures 2 and 3 give the reflection coefficients versus frequency for the conventional slotless patch antenna and the ESPA system respectively. The CSPA system resonates at about 2.69 GHz (finite ground) and 2.88 GHz (infinite ground) thereby yielding approximately 6.6 % difference in the resonance frequency (Figure 2). The ESPA architecture resonates at roughly 2.77 GHz (finite ground) and 2.83 GHz (infinite ground) thereby yielding a resonance difference of approximately 2.5 %. The difference in resonance becomes greater if the substrate dimensions are decreased further. In any given EM-based sensor design, a good correlation between simulated and real-life models is required. To ensure that the results of both the finite and infinite substrates improve, the geometrical size of the finite substrate must be increased to better match the infinite approximation. To obviate this requirement for reliable, small footprint and portable antenna systems for biosensing applications, the ESPA system is proposed. The centre-design frequency, f_c , of the patch antenna system is 2.793 GHz and the close-to-real-life ESPA system yields a resonance frequency of $0.993f_c$. Moreover, the simulated ESPA model indicates a good correlation and yields the same reflection coefficient of approximately -14.0 dB at the f_c (Figure 3).

Figure 4 shows the impedance magnitude of the ESPA system for biosensing applications. In any given RF/microwave system application, impedance matching is a critical requirement for system transceiver stability and/or distortionless signal applications. The finite approximation (or near-real-life) model depicts a 50-Ω impedance at 2.77 GHz. This matches the characteristic impedance (typically 50 Ω or 75 Ω) of the input port of RF/microwave transceiver systems for biosensing applications. In Figure 5, the voltage standing wave ratio (VSWR) for the closely-matched finite and infinite models of the ESPA system is approximately 4 dB at the f_c . The VSWR values at 2.77 GHz (finite substrate) and 2.83 GHz (infinite substrate) are roughly 1 dB respectively. Figure 6 shows the mismatch losses of the infinite and finite models for the ESPA system. The maximum loss for each approximation occurs at their respective resonant frequencies and agrees with the scattering-parameters analysis.

The radiation pattern of the ESPA system is illustrated in Figure 7. The fields are linearly polarised and the total gain is 5.0 dB in the vertical direction. Rectangular patch antennas are typically narrowband with a bandwidth of approximately 3 %.

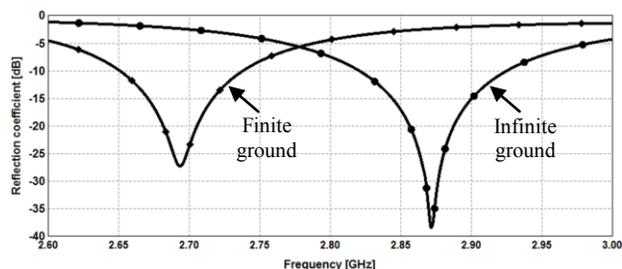


Figure 2. Reflection Coefficient versus Frequency (CSPA)

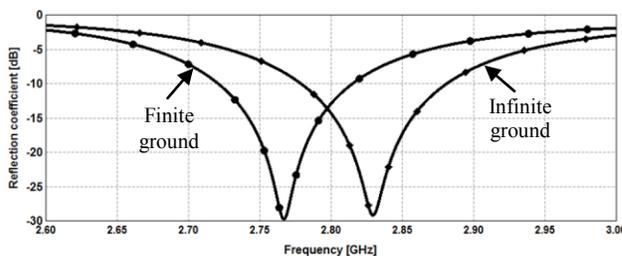


Figure 3. Reflection Coefficient versus Frequency (ESPA)

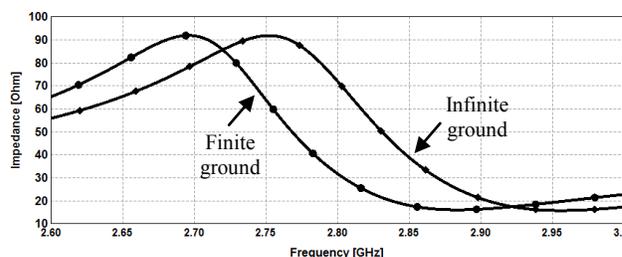


Figure 4. Impedance Magnitude of the ESPA System

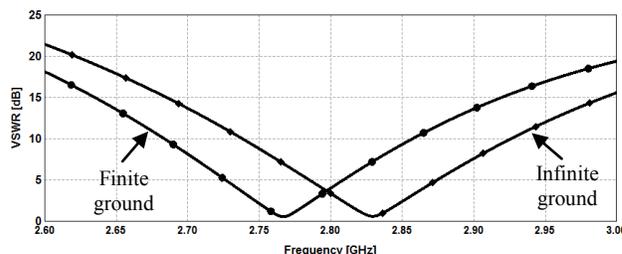


Figure 5. VSWR of the ESPA System

The ESPA was designed to operate at 2.79 GHz but it is resonant at 2.77 GHz. This frequency shift is attributed to the fringing fields around the antenna, which make the ESPA seem longer. The general design principle is to trim the patch antenna by typically 2–4 % to obtain resonance at the desired centre-design frequency. The fringing fields near the surface of the ESPA are responsible for its radiation. Equal but opposite currents flow through the ESPA and cancel the radiation due to current. The fringing fields are more bowed (extending further from patch) by a smaller value of ϵ_r ; the smaller the value of ϵ_r , the better the radiation efficiency. Thus, the proposed novel ESPA is an excellent candidate for biosensing applications involving various microbes.

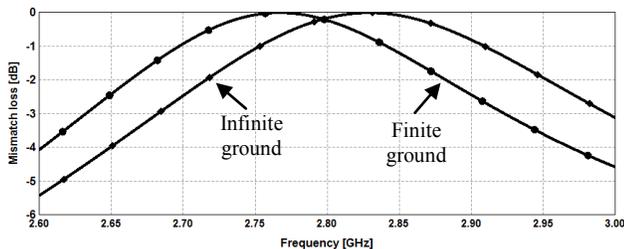


Figure 6. Mismatch Loss of the ESPA System

Figure 8 shows the radiation pattern of the ESPA system with a blood sample at 2.8 GHz with the parameters provided in Table 1 (other samples with their respective parameters can be used as well). The directivity of the antenna changes due to the presence of the sample and yields 7.5 dBi. This helps to differentiate between various biosamples and serves as a real-time biosensor for body tissues and micro-molecules.

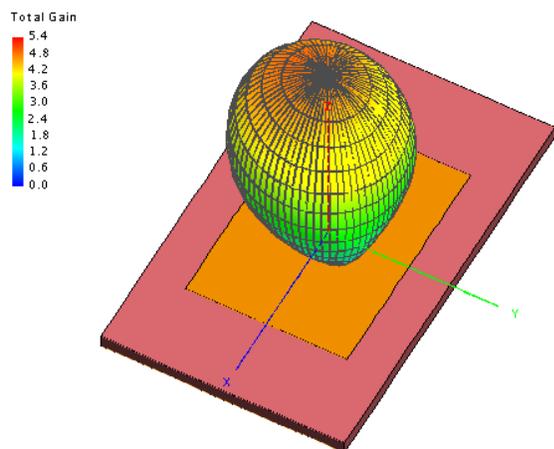


Figure 8. Radiation Pattern of the ESPA with Blood Sample at 2.8 GHz

IV. CONCLUSION

The presented novel elliptically-slotted patch antenna has the lowest electrical loss, low moisture absorption, uniform electrical properties over the operating frequencies and excellent chemical resistance. The ESPA geometry with a semi-major to semi-minor axes ratio of 2:1 and an eccentricity of $\cos 60^\circ$ (or $\sqrt{0.75}$) yielded the best solution for the designed antenna size and resonant wavelength. The ESPA system yields a total gain of 6.0 dBi, 7.5 dBi and 9.0 dBi at 2.6 GHz, 2.8 GHz and 3.0 GHz respectively. The substrate is a glass microfiber reinforced PTFE composite that can be easily integrated into real-time embedded microstrip circuit for low-cost, small footprint, low-profile and easy-to-fabricate biosensor systems applications. This novel antenna is more suitable for biosensing application as it holds a great promise in enhancing the sensitive, selective, real-time detection of low-concentration label-free and “spot-volume” biological samples.

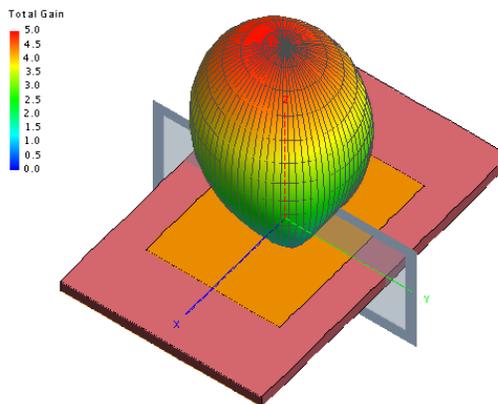


Figure 7. Radiation Pattern of the ESPA System at 2.8 GHz

REFERENCES

- [1] O. Korostynska, A. Mason, and A. Al-Shamma'a, "Microwave sensors for the non-invasive monitoring of industrial and medical applications," *Sensor Review*, vol. 34, pp. 182-191, 2014.
- [2] A. Pal, A. Mehta, M. E. Marhic, K. C. Chan, and K. S. Teng, "Microresonator antenna for biosensing applications," *Micro & Nano Letters, IET*, vol. 6, pp. 665-667, 2011.
- [3] D. M. Elsheikh, H. A. Elsadek, E. A. Abdallah, S. Atteya, and W. N. Elmazny, "Rapid detection of blood entero-viruses using microstrip antenna bio-sensor," in *Microwave Conference (EuMC), 2013 European*, pp. 878-880, 2013.
- [4] H. J. Lee and J. G. Yook, "Recent research trends of radio-frequency biosensors for biomolecular detection," *Biosensors & Bioelectronics*, vol. 61, pp. 448-459, 2014.
- [5] B. Kim, S. Uno, and K. Nakazato, "Miniature on-chip spiral inductor RFID tag antenna fabricated with metal layer of standard CMOS process for biosensor applications," in *Antennas and Propagation in Wireless Communications (APWC), 2011 IEEE-APS Topical Conference on*, pp. 925-928, 2011.
- [6] Ekpo, S. C., Adebisi, B., George, D., Kharel, R. and Uko, M., "A System-level Multicriteria Modelling of Payload Operational Times for Communication Satellite Missions in LEO," *Recent Progress in Space Technology*, Vol. 4, No. 1, pp. 67-77, June 2014.
- [7] Ekpo, S. and George, D., "Impact of Noise Figure on a Satellite Link Performance," *IEEE Communications Letters*, Vol. 15, No. 9, pp. 977-979, June 2011.

Integrated Smart Glove for Hand Motion Monitoring

Brendan O’Flynn, Javier
Torres Sanchez,
Tyndall National Institute
University College Cork,
Cork, Ireland
e-mail:
brendan.oflynn@tyndall.ie

James Connolly, Joan
Condell, Kevin Curran
Computing & Engineering,
Magee College, Ulster
University, Derry, N. Ireland
e-mail:
jp.connolly@ulster.ac.uk

Philip Gardiner
Altnagelvin Hospital,
Western Health and Social
Care Trust,
Derry, N. Ireland
e-mail:
philip.gardiner@westerntrust.h
scni.net

Barry Downes
TSSG,
Waterford Institute of
Technology,
Waterford,
Ireland
e-mail: bdownes@tssg.org

Abstract— Developments in Virtual Reality (VR) technology and its overall market have been occurring since the 1960s when Ivan Sutherland created the world’s first tracked head-mounted display (HMD) – a goggle type head gear. In society today, consumers are expecting a more immersive experience and associated tools to bridge the cyber-physical divide. This paper presents the development of a next generation smart glove microsystem to facilitate Human Computer Interaction through the integration of sensors, processors and wireless technology. The objective of the glove is to measure the range of hand joint movements, in real time and empirically in a quantitative manner. This includes accurate measurement of flexion, extension, adduction and abduction of the metacarpophalangeal (MCP), Proximal interphalangeal (PIP) and Distal interphalangeal (DIP) joints of the fingers and thumb in degrees, together with thumb-index web space movement. This system enables full real-time monitoring of complex hand movements. Commercially available gloves are not fitted with sufficient sensors for full data capture, and require calibration for each glove wearer. Unlike these current state-of-the-art data gloves, the UU / Tyndall Inertial Measurement Unit (IMU) glove uses a combination of novel stretchable substrate material and 9 degree of freedom (DOF) inertial sensors in conjunction with complex data analytics to detect joint movement. Our novel IMU data glove requires minimal calibration and is therefore particularly suited to the multiple application domains such as Human Computer interfacing, Virtual reality, the healthcare environment.

Keywords— Data glove; IMU; Virtual reality, Arthritis, Joint Stiffness, Hand Monitoring

I. INTRODUCTION

Data gloves contain strategically placed sensors controlled by circuitry that communicates finger joint movement to an end device. In recent years data gloves have been evaluated by researchers as an effective replacement for the universal goniometer (UG) [12]–[17]. Results showed comparable repeatability to the UG with the added advantage of simultaneous angular measurement and removal of intra-tester and inter-tester reliability problems associated with the UG. Data gloves however have several drawbacks; they require laborious calibration, are difficult to don and doff; and are designed to fit specific hand sizes and so require small, medium and large gloves to fit all hand variations. The first iteration of our system was developed using a state-of-the-art

5DT Ultra 14 data glove [18]. In this paper, our inertial measurement unit (IMU) Smart Glove is evaluated against this data glove for accuracy and repeatability and further validated using the Vicon motion capture system [19].

Virtual reality (VR) systems can be segmented into one of three experiences: non-immersive, semi-immersive, and fully immersive. Non-immersive systems would be those that can be visualized on a desktop computer. Semi-immersive VR environments incorporate images projected on the walls (e.g., cave automatic virtual environment, better known by the acronym CAVE). For a period of time, the user may superficially succumb to the perception of “being there”, but all the while still be aware of their real world surroundings. Finally, there is fully-immersive technology. In these systems, real-world visual and auditory cues are completely blocked out and the user has a sensory experience of being inside the computer-generated world. The experience is made ever more real through the use of hand-held and/or wearable devices that in some cases deliver haptic feedback which invoke sensations of touch. To enable Human computer interaction in this immersive fashion, high precision data acquisition systems need to be developed which are accurate, require minimal calibration and which provide real-time data streams wirelessly. The development of such a glove based system lends itself to multiple use cases including the Gaming environment and hand healthcare (e.g., Rheumatoid Arthritis (RA) monitoring).

This paper is organized as follows. Section II describes two data glove use case scenarios. Section III describes the glove hardware. Section IV addresses the system implementation. Section V goes into the calibration of the glove. Section VI describes the graphical user interface (GUI). Sections VII and VIII describe the tests and results respectively. Section IX goes into the conclusions. The acknowledgment section closes the paper.

II. MOTION MONITORING GLOVE EXAMPLE USE CASE

Wearable data acquisition systems which provide real time data of high quality are increasingly valuable in a variety of application scenarios. These range from virtual reality, gaming and Human Computer Interaction (HCI) to Connected Health and monitoring of wellness in a clinical context. Two such application spaces are outlined in the following subsections.

A. Virtual Reality (VR)

To be compatible with the Virtual Reality use case, it is important that any glove system developed for Human Computer Interaction adheres to requirements detailed below:

1. *Accuracy & Precision.* Accuracy is the degree of closeness to a quantity's actual true value. Precision is the degree to which repeated measurements give the same quantity. Here, we define accuracy and precision to consist of position and orientation. Different parts of the hand should have priority for accuracy: a) The mapping of the center of the virtual hand is the most important for most VR applications, b) The finger tips are the next most important for accuracy as the joints can be estimated via inverse kinematics and other constraints, c) The skeleton/joints of the hand are the next most important for accuracy.

2. *Consistent recognition of gestures.* Like speech recognition, if a gesture recognition system occasionally misinterprets signals then a break in presence occurs and users can become annoyed. Accidental gestures (known as false positives) are also a problem (e.g., accidentally signaling a command when unconsciously "talking with the hands").

3. *Low latency.* The faster the response of the system, then the more pleasant the user experience and the more easily users can enter a state of flow.

4. *Simulation of button presses.* Some applications will greatly benefit from simulation of button presses that provide a sense of self-haptic feedback (e.g., by touching two fingers together) and to control the game and system.

B. Rheumatoid Arthritis assessment

RA is an auto-immune disease which attacks the synovial tissue lubricating skeletal joints and is characterized by pain, swelling, stiffness and deformity. This systemic condition affects the musculoskeletal system, including bones, joints, muscles and tendons that contribute to loss of function and Range of Motion (ROM). Early identification of RA is important to initiate treatment, reduce disease activity, restrict its progression and ultimately lead to its remission. Clinical manifestations of RA can be confused with similar unrelated musculo-skeletal and muscular disorders. Identifying its tell-tale symptoms for early diagnosis has been the long-term goal of clinicians and researchers. Classifiers such as the Disease Activity Score (DAS) and Health Assessment Questionnaire (HAQ) provide an outcome measurement that reflects a patient's severity of RA disease activity. Such measurements are subjective and can be influenced by other factors such as depression or unrelated non-inflammatory conditions. Traditional objective measurement of RA using the universal goniometer (UG) and visual examination of the hands is labour intensive, open to inter rater and intra-rater reliability problems.

The DAS and HAQ [2] [3] are commonly used to measure disease onset and to assess disease status during clinical assessment [1]. Joint Stiffness is a common condition of RA that affects their ability to perform basic activities and daily functions [4] [5]. Several objective measurement systems have been devised by researchers and assessed in clinical trials for effectiveness as a joint stiffness measurement device [6]–[11].

III. TYNDALL GLOVE HW DESCRIPTION

The objective of the IMU Smart Glove is to measure the range of hand joint movements in a quantitative manner, including flexion, extension, adduction and abduction of the MCP, PIP and DIP joints of the fingers and thumb in degrees, together with thumb-index web space, palmar abduction to assist medical clinicians with the accurate measurement of the common condition of loss of movement in the human hand in patients with arthritis. All the Smart Glove functionality is maintained, controlled and analyzed by our in-house developed software system.

The described glove is a second generation iteration of the system by the authors as described in previous work [20].

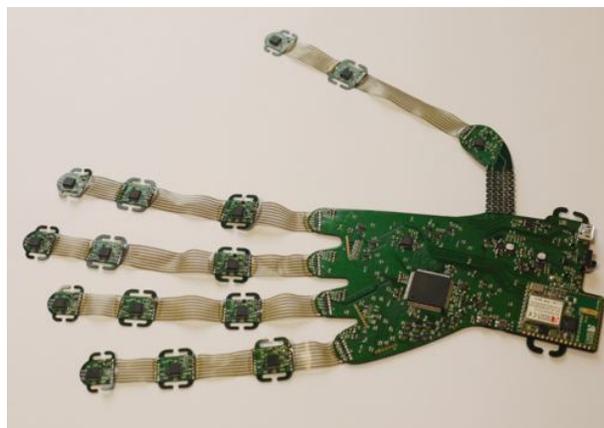


Figure 1. The IMU Smart Glove rev 2

A. HW description

The IMU glove, shown in Figure 1, has been manufactured using a mix of stretchable & flexible technology. Stretchable areas of the device cross each finger joint so they can conform to the human hand.

The glove includes 16 9-axes IMU's (each including 3-axis accelerometer, 3-axis gyroscope and 3-axis magnetometer) strategically placed to account for the degrees of freedom of each finger joint of the hand. IMUs are positioned on the stretchable interconnect and are located on the phalange of each finger segment to measure their orientation and biomechanical parameters.

Each IMU provides 6-degrees of freedom motion information (3 translational + 3 rotational) and 3D orientation information. By placing an IMU at both sides of each finger joint, (that is one per each finger bone and another one on the palm of the hand), the relative orientation of each IMU is calculated and used to generate angular and velocity movement throughout flexion and extension exercise of each finger joint and to calculate splaying of each finger.

B. Microcontroller

The processor selected for use in the system is an AVR32 UC3C 32 Bit Microcontroller. This is a high performance, low power 32-bit AVR microcontroller with built in single precision floating point unit. It was selected to enable complex embedded algorithms focused on motion analysis to be developed for real time low power consumption operation.

C. Wireless Communication

The RS9110-N-11-22 [21] module shown in Figure 2 is a IEEE 802.11b/g/n WLAN device that directly provides a wireless interface to any equipment with a UART or SPI interface for data transfer. It integrates a MAC, baseband processor, RF transceiver with power amplifier, a frequency reference, and an antenna in hardware. It also provides all WLAN protocols and configuration functionality. A networking stack is embedded in the firmware to enable a fully self-contained 802.11n WLAN solution for a variety of applications.

The module incorporates a highly integrated 2.4 GHz transceiver and power amplifier with direct conversion architecture, and an integrated frequency reference antenna. The RS9110-N-11-22 comes with flexible frameworks to enable usage in various application scenarios including high throughput and more network features.

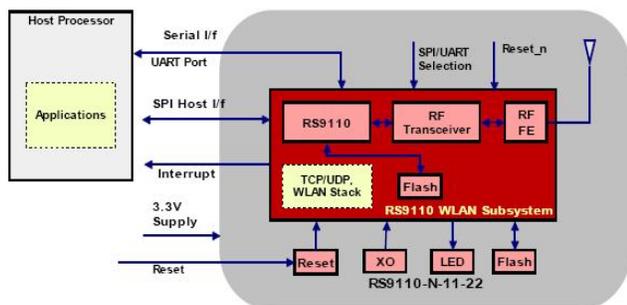


Figure 2. RS9110-N-11-22 System Block Diagram

The system operates according to a low complexity standard 4-wire SPI interface with the capability of operation up to a maximum clock speed of 25MHz. The communications module conforms to IEEE 802.11b/g/n standards and includes hardware accelerated implementation of WEP 64/128-bit and AES in infrastructure and ad-hoc modes. The fact that the module supports multiple security features such as WPA/WPA2-PSK, WEP, TKIP make it compatible with all medical ERP systems.

D. Sensors

The MPU-9150 [22] is a full three axis inertial measurement system incorporating tri-axis angular rate sensor (gyroscope) with sensitivity up to 131 LSBs/dps and a full-scale range of ± 250 , ± 500 , ± 1000 , and ± 2000 dps, tri-axis accelerometer with a programmable full scale range of $\pm 2g$, $\pm 4g$, $\pm 8g$ and $\pm 16g$ and a tri-axis compass with a full scale range of $\pm 1200\mu T$. The module incorporates embedded algorithms for run-time bias and compass calibration, so no user intervention is required. The MPU-9150 features three 16-bit analog-to-digital converters (ADCs) for digitizing gyroscope outputs, three 16-bit ADCs for digitizing accelerometer outputs, and three 13-bit ADCs for digitizing magnetometer outputs. For precision tracking of both fast and slow motions, the module features a user programmable gyroscope full-scale range of ± 250 , ± 500 , ± 1000 , and $\pm 2000^\circ/\text{sec}$ (dps), a user programmable accelerometer full-scale range of $\pm 2g$, $\pm 4g$, $\pm 8g$, and $\pm 16g$, and a magnetometer full-scale range of $\pm 1200\mu T$.

E. Additional Features

To make the system adaptable in operation and compatible with a wide range of use cases outside the immediate application of RA monitoring, the IMU Smart Glove system also incorporates such features as optional storage via a micro SD card, battery monitoring and recharge ability, as well as a USB bootloader, USB communication interface, and 15 Analogue inputs for optional resistive sensors (e.g., bend sensors or force sensors). The analogue front end is a buffered voltage divider to enable additional sensing functionality.

IV. SYSTEM IMPLEMENTATION

All the system embedded code is implemented using the Atmel Studio 6 IDE. Currently the implementation includes full application code that continuously reads the sensor outputs and wirelessly transmits the data through a TCP socket.

The accuracy of IMU-based real time motion tracking algorithms is highly influenced by sensor sampling rate. Therefore a fundamental design requirement of the IMU Smart Glove was high application throughput to facilitate the development of algorithms using suitable PC SW such as MATLAB C# and Unity. In addition, it was envisaged that once the algorithms would have been fully developed and tested, they would be fully implemented on the embedded platform. This eliminates the requirement for a high throughput device and allows for a low power implementation for example using BLE in a third generation of the glove.

To ensure maximum achievable sampling rates and computation time are compatible with the application scenario envisaged as specified in conjunction with clinical partners regarding signal temporal granularity, it was decided not to share the I2C bus between each of the 16 MPU9150's. Instead, dedicated I2C lines are provided to each one of the sensors and are driven in parallel. This provides the added advantage of ensuring synchronization between all IMU sensors.

A. Case 1 Raw data transmission.

The embedded processor enables multiple modes of operation depending on the use case and degree of data granularity required. Having the wireless system transmitting raw data at the highest achievable data rate is desirable for the development of the analytics as it is more practical to develop them using PC based SW (real time or post processing) and then porting them to the embedded system than develop them directly within the embedded system.

B. Case 2: Transmission of Raw data & information

The wireless system transmits raw data and quaternions/rotation matrix (from gyros) at the highest achievable data rate. Quaternions then will be subject to drift/errors and the analytics to correct for this are implemented within the controlling software. At this stage we have a clear idea of the maximum processing time that could be allocated in the embedding to this task and that is taken into consideration when designing these algorithms.

C. Case 3: Transmission of processed data

With the wireless system with full analytics embedded, the internal sampling rate of the sensors should be kept to a

maximum achievable SR, the high wireless data rate might not longer be required.

V. CALIBRATION USING ACCELEROMETRY AND GYROSCOPE

Data glove accuracy and repeatability is affected by the non-linear nature of glove sensor output and any misalignment between the wearers hand and data glove sensor positioning. Data glove sensor calibration improves sensor accuracy and matches the boundaries of each sensor to those of each finger joint. A calibration routine requires the glove wearer to position groups of finger joints such as MCP's and PIP's at specific poses. Each pose places a finger joint group and relevant data glove sensors at their minimum and maximum boundaries. The IMU Smart Glove uses on-board sensors to automatically calibrate each glove sensor, regardless of the wearer's joint flexibility. Each glove accelerometer sensor is sampled when the hand is in a neutral position to calculate finger joint thickness and slope offset, and used during angular calculation. Accelerometers placed on each one of the finger's phalanges provide information with regards to the inclination to gravity of the phalanx. The output response of each sensor provides information on the orientation of the sensor to gravity as shown in Figure 3. The orientation to gravity of each one of the sensors placed on adjacent phalanges can be used to estimate the flexion of the finger.

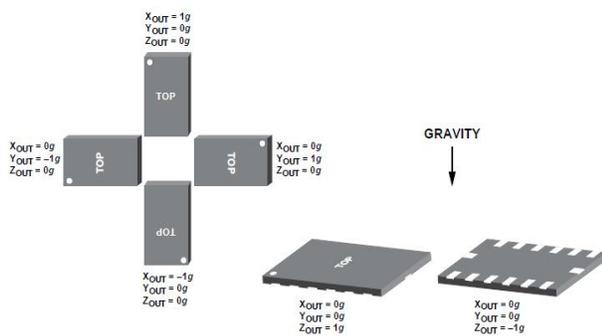


Figure 3. Output response vs. Orientation to gravity

For example, if the measured acceleration for a specific finger from the medial phalanx accelerometer is (Xout, Yout, Zout)= (-1,0,0)g and from the proximal phalanx accelerometer is (Xout, Yout, Zout)= (0,0,1)g, it indicates a flexion of the PIP joint of 90 degrees. The inclination to gravity is determined according to the standard formulas (1), (2) and (3):

$$\theta = \tan^{-1} \left(\frac{A_{X,OUT}}{\sqrt{A_{Y,OUT}^2 + A_{Z,OUT}^2}} \right) \quad (1)$$

$$\psi = \tan^{-1} \left(\frac{A_{Y,OUT}}{\sqrt{A_{X,OUT}^2 + A_{Z,OUT}^2}} \right) \quad (2)$$

$$\phi = \tan^{-1} \left(\frac{\sqrt{A_{X,OUT}^2 + A_{Y,OUT}^2}}{A_{Z,OUT}} \right) \quad (3)$$

Where: θ is the angle between the horizon and the x-axis of the accelerometer, ψ is the angle between the horizon and the y-axis of the accelerometer, and ϕ is the angle between the gravity vector and the z-axis.

VI. GUI/USER INTERFACE

Data is streamed in real-time according to the use cases outlined above and post processed by our controlling software. A pivotal role of this software is its ability to encapsulate movement associated with finger joints in real time. Figure 4 shows an example of the user interface. Algorithms segment recorded data to extract relevant flexion and extension movement information.

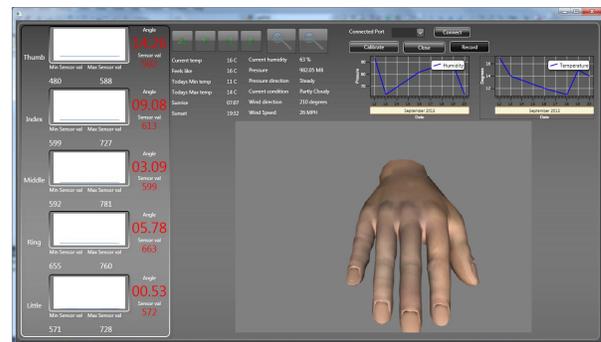


Figure 4. Angular output from the data glove is displayed in 3D

A. Data analytics and Post processing

Each angular calculation is low-pass filtered to remove sensor noise. A complementary filter with error control is implemented to combine accelerometer output with gyroscope rotation angle. Gyroscope rotational angle is initially accurate and drifts over time. Accelerometer angle cannot distinguish between lateral acceleration and rotation. The complementary filter acts as a high-pass and low-pass filter on both signals. It combines estimated gyroscope rotation and accelerometer angle to create an angular output.

VII. TESTING STRATEGIES

Our new data glove was assessed for accuracy and repeatability and was compared with the 5DT state-of-the-art data glove. The Vicon MX motion capture system was used during accuracy testing to independently measure angular values generated at each finger joint. Movement was recorded by Vicon and simultaneously by our in-house developed controlling software whilst each glove was placed on blocks of wood cut to specific angles. Angular readings were assessed using Root Mean Square Error (RMS) to provide an indicator of the variance between each estimated angular repetition value and the expected true value influenced by the angle on each block of wood. RMS error is influenced by both positive and negative errors which are either above or below the expected true value. Therefore RMS output is a measure of the angular error. Repeatability testing examined the ability of each data glove to consistently replicate angular readings when the subjects hand was held in a repeatable position. Testing strategies were originally developed to assess data glove suitability as a replacement for the UG. Although no formal set

of repeatability testing strategies exist, the strategies used by [12] have been adopted by subsequent research groups [13] [16] [23]–[26] and are used in this study to allow comparison between study results.

VIII. RESULTS AND DISCUSSION

The ‘flat hand’ test examines each data glove’s ability to maintain a minimum repeatable value after full stretch of each data glove sensor. The plaster mould test examines the ability of each data glove to reproduce angular readings when positioned in a repeatable position. In all tests, our data glove was not calibrated for the subject. The 5DT data glove was calibrated.

A. ‘Flat hand’ Results

The ‘flat hand’ test results demonstrated in Table I show that the IMU data glove outperformed the 5DT data glove. Mean MCP readings for the IMU glove were near-perfect - 0.38°, with PIP readings of -2.53°. The 5DT produced readings of 4.17° for MCP and 2.27° for PIP. This is particularly impressive since the IMU glove was not calibrated before use.

TABLE I. COMPARISON OF MEAN ANGULAR READINGS RECORDED DURING ‘FLAT HAND’ TESTING

	5DT (Angle / SD)	IMU (Angle / SD)
Index MCP	2.34 (1.59)	-0.59 (1.87)
Index PIP	2.04 (1.05)	-2.74 (0.90)
Middle MCP	5.9 (0.55)	1.32 (2.26)
Middle PIP	3.27 (1.13)	-2.94 (1.25)
Ring MCP	5.14 (0.59)	-2.33 (1.21)
Ring PIP	1.02 (0.52)	-2.7 (1.11)
Little MCP	3.32 (0.88)	0.07 (2.56)
Little PIP	2.76 (1.32)	-1.75 (1.31)
Mean MCP	4.17 (0.90)	-0.38 (1.98)
Mean PIP	2.27 (1.0)	-2.53 (1.14)
Overall mean	3.22 (0.95)	-1.46 (1.56)

B. Plaster mould test results

Table II shows comparison results for plaster mould testing for the 5DT and our IMU data glove. Readings showed the IMU Smart Glove produced better repeatability for MCP and PIP joints and better overall repeatability as indicated by the lower mean range angular reading.

TABLE II. COMPARISON OF MEAN RANGE AND SD READINGS FROM PLASTER MOULD TESTING FOR EACH DATA GLOVE

Glove	MCP		PIP		Mean	
	Range	SD	Range	SD	Range	SD
5DT	8.85	2.13	6.23	2.09	7.54	2.11
IMU	5.99	1.89	5.10	1.58	5.55	1.74

C. Accuracy results

Table III shows comparison of results for the 5DT and our IMU Smart Glove compared with the Vicon motion capture system and the UG. Results showed the goniometer had greatest overall accuracy of 93.23% with overall RMS of 2.76°. This is in agreement with typical findings on goniometric accuracy with 95% of intratester reliability within 5° of measurement and intertester reliability in the range of 7° to 9° [27]–[29]. The Vicon system provided mean accuracy of 89.33% with RMS of 5.19°.

TABLE III. MEAN ACCURACY PERCENTAGE FOR EACH SENSOR INCLUDING MEAN ERROR AND OVERALL ACCURACY PERCENTAGE

Sensor	Vicon	5DT	Goniometer	IMU
Index MCP	93.31	94.20	97.95	89.57
Index PIP	91.23	92.01	90.75	91.47
Middle MCP	91.46	79.66	95.83	82.40
Middle PIP	84.08	74.97	88.96	77.29
Ring MCP	87.20	70.46	97.37	82.02
Ring PIP	86.99	91.99	90.70	89.51
Little MCP	86.14	85.83	91.28	83.38
Little PIP	94.23	74.56	93.03	86.27
Overall accuracy %	89.33	82.96	93.23	85.24
RMS	5.19	7.15	2.76	5.95

This inaccuracy was most likely caused by noise, marker occlusion, and distance of reflective markers from cameras. The IMU data glove provided the best accuracy measurement of all data gloves and demonstrated similar accuracy to the Vicon measurement system. The RMS results obtained show that readings obtained from sensors contained approximately 5.95° of error. Results shown in Table III indicate that all sensors demonstrated accuracy between 82% to 91% except for the Middle PIP sensor that had accuracy of 77.29%.

D. Comparison with previous trials

The results shown in Table IV compare ‘flat hand’ and plaster mould tests for the 5DT and our IMU data glove with previous research studies involving data gloves. The 5DT data glove demonstrated range readings that out-performed data glove findings by [12] [13] and were similar to [26]. The data glove examined by [15] provided better results than all studies including the 5DT and our IMU glove. However this glove contained only 5 sensors that recorded movement of the MCP joints. The IMU glove performed better than all other data glove studies. Readings recorded by earlier studies are averaged for several subjects. This can hide higher inaccurate results for some subjects. For example, [12] recorded range readings from 5 subjects that varied between 2.5° to 6.7°. Results were averaged to 4.4°. Similarly, results from ‘flat hand’ testing from the study by [13] were summarised from a group of 6 male and female participants. Mean male range results went from 2.37° to 5.49° and mean female from 3.90° to 4.75°.

TABLE IV. COMPARISON OF ‘FLAT HAND’ AND PLASTER MOULD TESTS WITH PREVIOUS DATA GLOVE STUDIES

Study	Flat hand test (Range / SD)	Plaster mould test (Range / SD)
Wise et al. [12]	4.4 (2.2)	6.5 (2.6)
Dipietro et al. [13]	3.84 (1.23)	7.47 (2.44)
Simone et al. [15]	1.49 (0.5)	5.22 (1.61)
Gentner and Classen [26]	2.61 (0.86)	6.09 (1.94)
5DT (this study)	2.27 (0.995)	7.54 (2.11)
IMU (this study)	4.86 (1.56)	5.55 (1.74)

IX. CONCLUSIONS

Data gloves have been proven to be a viable replacement for the UG and can offer unbiased finger joint ROM measurement. However their dependence on calibration reduces their usefulness in the many application spaces. The

novel IMU based wireless Smart Glove detailed in this paper removes the requirement for sensor calibration using accelerometers and gyroscopes teamed with intelligent software techniques. Test results showed our IMU data glove had comparable repeatability to the UG with the added advantage of simultaneous angular measurement and removal of intra-tester and inter-tester reliability. Accuracy testing results showed the IMU data glove provided better accuracy and less overall error than the 5DT data glove with which it was compared. Of Note the IMU glove required no calibration before use whilst maintaining results which demonstrated it had similar accuracy to the Vicon system.

ACKNOWLEDGMENT

The support of Science Foundation Ireland (SFI) as well as the National Access Program (NAP) support provided by the Tyndall National Institute is gratefully acknowledged. This work was also supported by Department of Education and Learning (DEL).

REFERENCES

[1] D. M. van der Heijde, van 't H. M. P. L. van Riel, and L. B. van de Putte, "Development of a disease activity score based on judgment in clinical practice by rheumatologists," *J. Rheumatol.*, vol. 20, no. 3, pp. 579–81, 1993.

[2] J. F. Fries, P. Spitz, R. G. Kraines, and H. R. Holman, "Measurement of patient outcome in arthritis," *Arthritis Rheum.*, vol. 23, no. 2, pp. 137–145, 1980.

[3] J. F. Fries, P. W. Spitz, and D. Y. Young, "The dimensions of health outcomes: the health assessment questionnaire, disability and pain scales," *J. Rheumatol.*, vol. 9, no. 5, p. 789–793, 1982.

[4] P. Emery, F. C. Breedveld, M. Dougados, J. R. Kalden, M. H. Schiff, and J. S. Smolen, "Early referral recommendation for newly diagnosed rheumatoid arthritis: evidence based development of a clinical guide.," *Ann. Rheum. Dis.*, vol. 61, no. 4, pp. 290–7, Apr. 2002.

[5] F. C. Arnett, S. M. Edworthy, D. A. Bloch, D. J. McShane, J. F. Fries, et al., "The American Rheumatism Association 1987 revised criteria for the classification of rheumatoid arthritis.," *Arthritis and rheumatism*, vol. 31, no. 3, pp. 315–24, Mar-1988.

[6] J. T. Scott, "Morning stiffness in rheumatoid arthritis.," *Ann. Rheum. Dis.*, vol. 19, pp. 361–8, Dec. 1960.

[7] V. Wright and R. J. Johns, "Quantitative and qualitative analysis of joint stiffness in normal subjects and in patients with connective tissue diseases.," *Ann. Rheum. Dis.*, vol. 20, pp. 36–46, Mar. 1961.

[8] M. L. Ingpen and P. H. Kendall, "A simple apparatus for assessment of stiffness," *Ann. Phys. Med.*, vol. 9, no. 5, pp. 203–5, Feb. 1968.

[9] a Unsworth, P. Yung, and I. Haslock, "Measurement of stiffness in the metacarpophalangeal joint: the arthrograph.," *Clin. Phys. Physiol. Meas.*, vol. 3, no. 4, pp. 273–81, Nov. 1982.

[10] A. Howe, D. Thompson, and V. Wright, "Reference values for metacarpophalangeal joint stiffness in normals.," *Ann. Rheum. Dis.*, vol. 44, no. 7, pp. 469–76, Jul. 1985.

[11] E. Dionysian, J. M. Kabo, F. J. Dorey, and R. a Meals, "Proximal interphalangeal joint stiffness: measurement and analysis.," *J. Hand Surg. Am.*, vol. 30, no. 3, pp. 573–9, May 2005.

[12] S. Wise, W. Gardner, E. Sabelman, E. Valainis, Y. Wong, et al., "Evaluation of a fiber optic glove for semi-automated goniometric measurements," *J. Rehabil. Res. Dev.*, vol. 27, no. 4, p. 411, 1990.

[13] L. Dipietro, A. M. Sabatini, and P. Dario, "Evaluation of an instrumented glove for hand-movement acquisition.," *J. Rehabil. Res. Dev.*, vol. 40, no. 2, pp. 179–89, 2003.

[14] L. K. Simone and D. G. Kamper, "Design considerations for a wearable monitor to measure finger posture.," *J. Neuroeng. Rehabil.*, vol. 2, no. 1, p. 5, Mar. 2005.

[15] L. K. Simone, N. Sundarajan, X. Luo, Y. Jia, and D. G. Kamper, "A low cost instrumented glove for extended monitoring and functional hand assessment.," *J. Neurosci. Methods*, vol. 160, no. 2, pp. 335–48, Mar. 2007.

[16] G. Saggio, S. Bocchetti, C. A. Pinto, G. Orengo, and F. Giannini, "A novel application method for wearable bend sensors," in *2009 2nd International Symposium on Applied Sciences in Biomedical and Communication Technologies*, Bratislava, Slovakia, Nov-2009, pp. 1–3.

[17] K. Li, I.-M. Chen, S. H. Yeo, and C. K. Lim, "Development of finger-motion capturing device based on optical linear encoder," *J. Rehabil. Res. Dev.*, vol. 48, no. 1, p. 69, 2011.

[18] 5DT, "5DT Data Glove 14 Ultra," 2011. [Online]. Available: <http://www.5dt.com/products/pdataglove14.html>. [retrieved: 01-2012].

[19] Vicon Motion Systems, "Vicon," 2013. [Online]. Available: <http://www.vicon.com/>. [retrieved: 01-2012].

[20] B. O'Flynn, J. Sanchez, P. Angrove, J. Connolly, J. Condell, and K. Curran, "Novel smart sensor glove for arthritis rehabilitation" *2013 IEEE International Conference on Body Sensor Networks*, (BSN 2013), May-2013, pp. 1–6.

[21] Redpine Signals, "RS9110-N-11-22: 802.11BGN wireless device server," RS9110-N-11-22 Product brief, 2008. [Online]. Available: <http://www.redpinesignals.com/pdfs/RS9110-N-11-22 Wlan Module.pdf>. [retrieved: 01-2012].

[22] InvenSense, "MPU-9150 Product Specification Revision 4.0," vol. 1, no. 408. InvenSense, California, pp. 1–52, 2012.

[23] G. Kessler, N. Walker, and L. Hodges, "Evaluation of the CyberGlove (TM) as a whole hand input device," *GVU Technical Report; GIT-GVU-95-05*, Georgia Institute of Technology 1995.

[24] M. Mentzel, F. Hofmann, T. Ebinger, B. Jatzold, L. Kinzl, and N. J. Wächter, "Reproducibility of measuring the finger joint angle with a sensory glove," *Handchir Mikrochir Plast Chir*, vol. 33, no. 1, pp. 9–64, 2001.

[25] N. W. Williams, J. M. T. Penrose, C. M. Caddy, E. Barnes, D. R. Hose, and P. Harley, "A goniometric glove for clinical hand assessment," vol. 25, no. 2, pp. 200–207, 2000.

[26] R. Gentner and J. Classen, "Development and evaluation of a low-cost sensor glove for assessment of human finger movements in neurophysiological settings.," *J. Neurosci. Methods*, vol. 178, no. 1, pp. 138–47, Mar. 2009.

[27] A. Hellebrandt, E. Duvall, and M. Moore, "The measurement of joint motion. Part III : Reliability of goniometry," *Phys Ther Rev*, vol. 29, no. 6, pp. 302–307, 1949.

[28] E. Lewis, L. Fors, and W. J. Tharion, "Interrater and intrarater reliability of finger goniometric measurements," *Am. J. Occup. Ther.*, vol. 64, pp. 555–561, 2010.

[29] N. B. Reese and W. D. Bandy, "Measurement of Range of motion and muscle length: background, history and basic principles". *Joint Range of Motion and Muscle Length testing*, 2nd ed. St. Louis, MO: Saunders/Elsevier; 2010. pp. 3–29

Ultra-miniature, computationally efficient diffractive visual-bar-position sensor

Mehjabin Monjur, Leonidas Spinoulas, Patrick R. Gill and David G. Stork

Rambus Labs

1050 Enterprise Way, Suite 700

Sunnyvale, CA 94089 USA

dstork@rambus.com

Abstract—We describe the design and performance of an ultra-miniature lensless computational sensor optimized for estimating the one-dimensional position of visual bars. The sensor consists of a special-purpose wavelength-robust optical binary phase diffraction grating affixed to a CMOS photodetector array. This grating does not produce a traditional high-quality human interpretable image on the photodetectors, but instead yields visual information relevant to the bar-position estimation problem. Computationally efficient algorithms then process this sensed information to yield an accurate estimate of the position of the bar. The optical grating is very small (120 μm diameter), has large angle of view (140°), and extremely large depth of field (0.5 mm to infinity). The design of this sensor demonstrates the power of end-to-end optimization (optics and digital processing) for high accuracy and very low computational cost in a new class of ultra-miniature computational sensors.

Keywords: *Computational sensing, diffractive imager, visual-bar-position sensor, lensless smart sensor*

I. INTRODUCTION

The discipline of computational imaging involves the design of both optics and digital signal processing to achieve a desired end-to-end system performance. Because much of the overall burden of imaging or sensing can be borne by the signal processing, the constraints upon the optical components can be relaxed. [1] As such, optical systems with fewer optical components or smaller form factors can be made. True joint design requires the definition of a global or end-to-end merit function and an explicit functional relationship between this merit function and the optical and signal processing parameters. [2] Under such circumstances, the design process can rely on gradient descent in the end-to-end merit function. Approximations to this method can include taking steps alternately in the optical parameters and then the signal processing parameters and iterating until convergence.

Recently, computational imaging systems have been designed that eschew traditional refracting or reflecting optical elements (lenses or curved mirrors) and which rely instead entirely upon *diffraction*. Such devices have been demonstrated in mobile medical microscopy [3], [4] and far-field imaging. [5]–[13] Simulation studies have shown that such diffractive systems can produce digital images nearly as accurately as do ideal lensed systems, but such lensed systems are very difficult to create at the spatial scales of diffractive imagers ($\sim 100 \mu\text{m}$). [14]

Such computational sensors require more processing than do traditional camera-based systems in order to create a digital

image. If we let n denote the linear size of a square sensor array, image computation by Tikhonov regularization is an $\mathcal{O}(n^4)$ algorithm, which is rather costly for even moderately large sensors. If the optical image is shift invariant (possibly after simple image dewarping to correct for optical barrel distortion), then fast-Fourier-based deconvolution algorithm of complexity $\mathcal{O}(n^2 \ln n)$ can be used. [11] Such convolution algorithms represent a small portion of the overall computational cost of analog-to-digital conversion, data transfer, and so forth.

The above discussion centered on computational *imaging*, and while nearly all high-level computer vision and pattern recognition algorithms (object recognition, bar-code reading, face recognition, ...) operate on image data, many *sensing* and *image estimation* tasks do not require such an image. These sensing tasks are particularly attractive for low-power sensors, including

- overall brightness estimation
- color gamut estimation
- visual flow estimation
- visual orientation estimation
- axial visual flow or “looming” estimation
- image change detection
- visual mark localization and tracking

In computational sensors of the sort described here, the algorithms addressing such sensing tasks can operate on the raw sensor signals, without the need for a traditional image. For this reason, we informally consider the diffractive optical element as a *computational device*—one that performs a signal processing function in parallel, in negligible time (the time it takes light to pass through the thin grating), and with zero electrical power dissipation. Our task in the work reported here is to exploit computational sensing design methodology to simplify the optical component and reduce the computational cost for one such target application: visual-bar-position estimation. Our broad motivation is to create inexpensive, low-power, application-specific sensors for use in mobile and standalone applications in automotive, biomedical, smart architecture, smart cities, and the Internet of Things.

We begin in Section II with a description of the image estimation task at hand, then turn in Section III to our overall design methodology. Next, we discuss in Section IV both our hardware design, in particular our special optical phase grating and the optical signals it captures, as well as the digital signal

processing. We present the overall sensor performance and computational cost in Section VI, and conclude in Section VII with a brief summary and suggestions for future research.

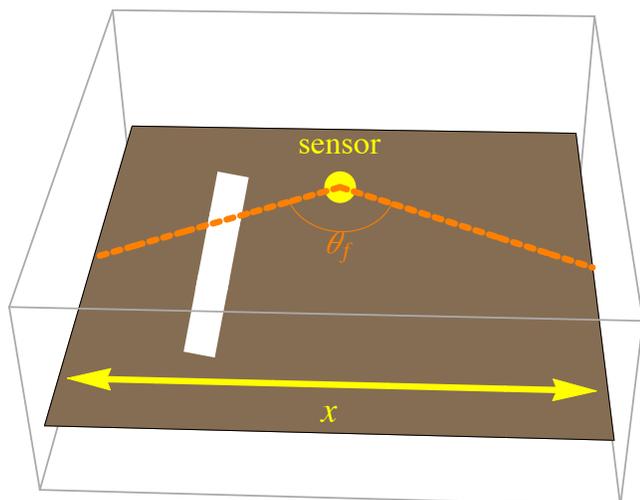


Fig. 1. The image sensing task is to estimate the left-right position, x , of a long vertical visual bar on a dark background in the far field (distance from sensor greater than 2 mm). The full-field angle-of-view of the sensor is $\theta_f = 140^\circ$.

II. IMAGE ESTIMATION TASK

Figure 1 illustrates our image estimation task. The target is a vertical bar on a dark background of unknown horizontal position x within a visual field of $\pm 70^\circ$ or $\theta_f = 140^\circ$ full field. The bar can be of arbitrary spectral composition in the visible range $400\text{ nm} < \lambda < 700\text{ nm}$. Throughout, the field luminance is greater than roughly 500 lux , but performance does not depend significantly upon luminance as long as it is within a few orders of magnitude of this value. Such an image estimation task appears in numerous practical applications from machine inspection (e.g., alignment of moving parts *in situ*), horizon tracking in unmanned aerial vehicles, lane-tracking in autonomous vehicles, estimation of the height of liquids in medical equipment such as test tubes, and others. We report here end-to-end simulation design and testing results, including optical diffraction and digital signal processing.

III. SENSOR DESIGN METHODOLOGY

As mentioned above, computational sensing and imaging relies upon the joint design of both the optics and the signal processing for a desired end-to-end performance of the digital imaging system. The ideal, true joint method is to form a global criterion or merit function, such as predicted image mean-squared error, then determine how the optical and the signal processing parameters affect this merit function, and then iteratively adjust all parameters simultaneously to optimize this global merit function. [2] For complicated optical systems, the dependency of the merit function upon a system parameter, such as the shape parameter governing a diffraction grating, cannot be determined analytically. In such cases, a

number of related or approximate design methodologies can be used instead: one can adjust the optical parameters then the signal processing parameters, then the optical parameters and so on until global convergence of the merit function is reached. Alternatively, one can design a fixed optical element that has desirable application-specific properties and then optimize the signal processing. In our current work, we employed the latter technique.

We model the sensor as a linear system, that is,

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{n}, \quad (1)$$

where \mathbf{x} is a vector of inputs from the scene, \mathbf{y} is the vector of photodetector pixel responses, \mathbf{A} the system matrix describing the linear transformation performed by the two-dimensional optical grating, and \mathbf{n} the additive noise, which describes photodetector noise, Poisson photon statistics, quantization noise, etc. (Other models, such as simple multiplicative noise, could also be assumed.) Here \mathbf{x} is n -dimensional and \mathbf{y} and \mathbf{n} are m -dimensional, and hence \mathbf{A} has dimensions $n \times m$. During design, we computed the system matrix \mathbf{A} once for each candidate diffraction grating, then used it with bars at different positions \mathbf{x} to simulate the signals on the photodetector array. We then implement the signal processing algorithms to estimate the bar position (cf., Sect. V). All simulation steps were implemented in *Matlab*.

IV. SENSOR AND PHASE GRATING

We are not aware of any general theory for deriving analytically the relationship between grating parameters of candidate phase gratings and a final merit function reflecting the accuracy of the final bar position estimate. The physical processes of diffraction, the constraints of grating manufacturability, and so forth, make such an analytical relationship complicated indeed. For that reason, we explored a number of grating designs “by hand,” guided by knowledge of manufacturing and physical constraints. The active portion of the gratings we explored were $120\ \mu\text{m}$ in diameter, commensurate with sensor hardware described elsewhere. [8] Because the sensing problem appears one dimensional, we first tested linear (vertical) gratings. These yielded poor results in part because of geometrical effects due to sources at different depths: the projected images were curved toward the periphery, and that curvature depended upon the optical wavelength of light. Because the intermediate images varied significantly and in highly nonlinear ways for source bars of different spectral composition and spatial depth we could not design simple signal processing methods that reliably estimated the bar position across all such variations.

We then explored several classes of two-dimensional gratings. A traditional Fresnel diffraction grating gave informative images for bars emitting at its single design wavelength, but not for bars emitting elsewhere throughout the visible range. For this reason, basic Fresnel gratings and gratings closely related to it, were unacceptable. A particularly intriguing class of gratings were based on fractals, which were designed

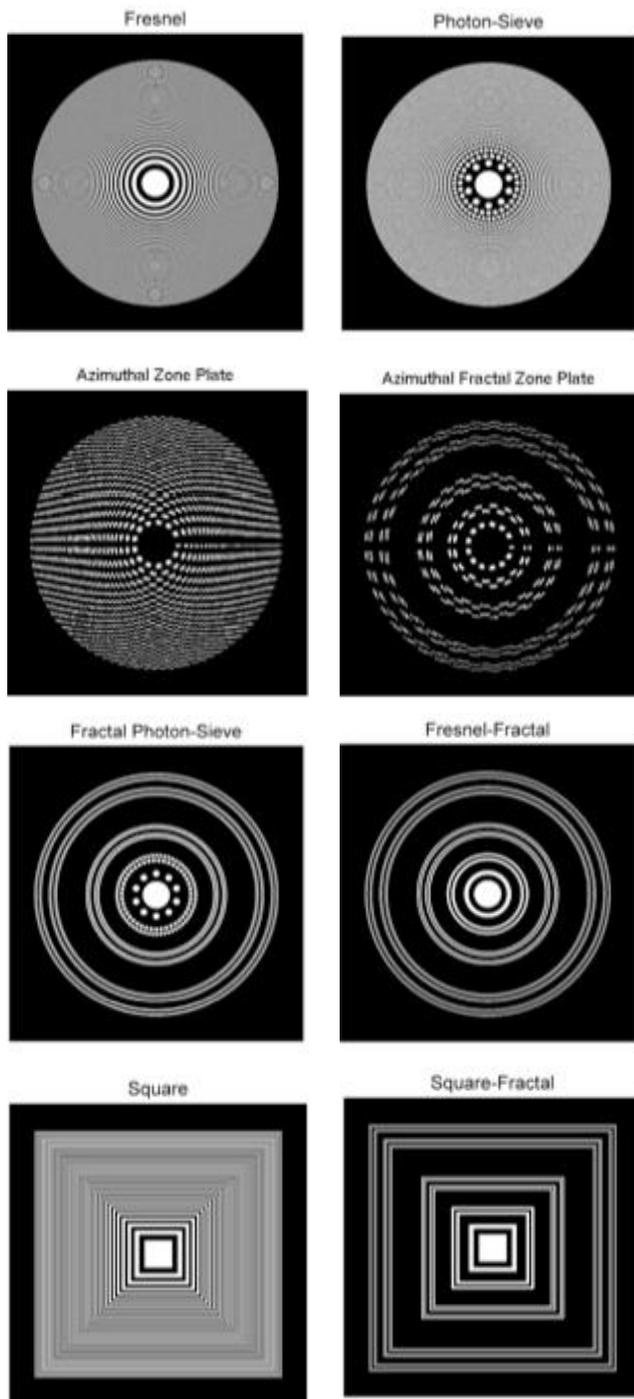


Fig. 2. Representative candidate binary phase gratings tested for one-dimensional bar position estimation.

to be relatively insensitive to wavelength. [15] While these yielded intermediate images that were indeed robust to spectral variations, the spatial shape of the resulting images were somewhat complicated so no simple signal processing algorithms could accurately estimate the bar location. Figure 2 shows representative candidate gratings we explored.

Figure 3 shows the imaging performance of several grating

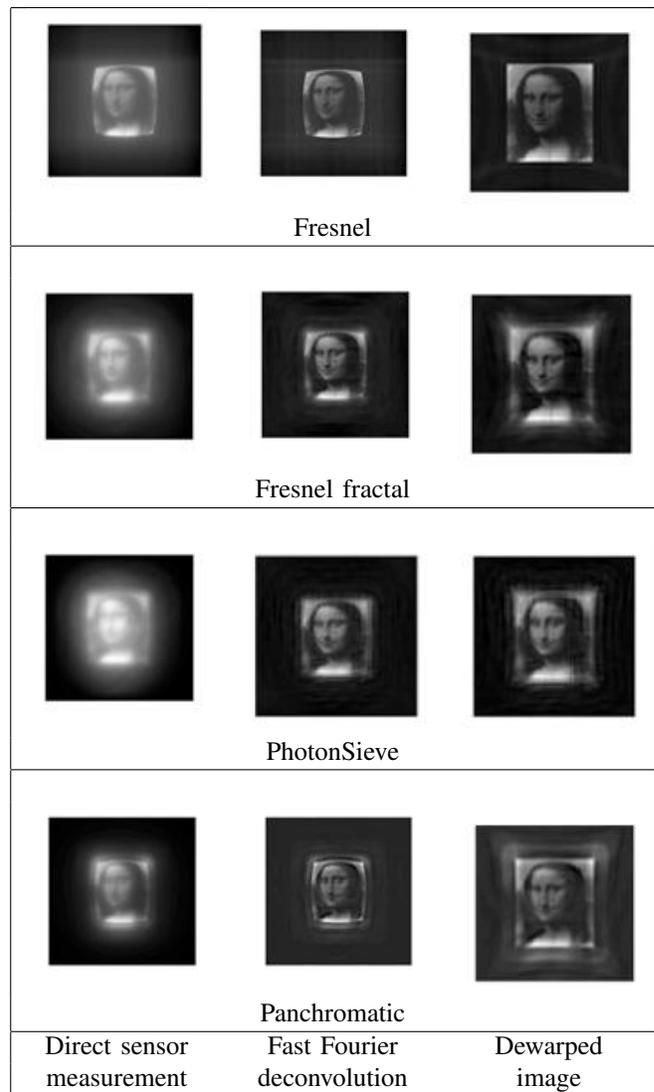


Fig. 3. Images produced by several of the candidate binary phase gratings shown in Fig. 2 for input scene tendered in blue light. These images were computed from the raw sensor signals by fast Fourier deconvolution, described elsewhere. [11] The images in the right column are radially dewarped versions of those in the middle column.

designs (including the panchromatic Fresnel zone plate, see below), designed at the intermediate wavelength $\lambda = 550 \text{ nm}$ but rendered at $\lambda = 470 \text{ nm}$. The image of Leonardo's *Mona Lisa* was computed using fast Fourier deconvolution. [11] In traditional computer vision methodology, the ideal intermediate optical image would be sharp lines independent of wavelength, but such images cannot be achieved using binary phase gratings. Our goal, then, is the create optical images such that signal processing can estimate the visual bar location reliably, and at low computational cost, regardless of wavelength and distance of visual bar.

The grating that provided an easily processed image despite variations in bar spectral composition was the *panchromatic binary Fresnel zone plate*, governed by Eq. 2:

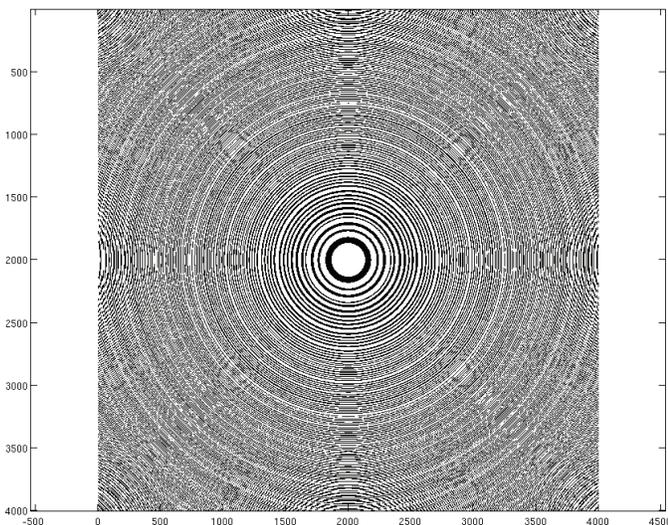


Fig. 4. Panchromatic binary phase Fresnel zone plate generated according to Eq. 2. This circularly symmetric grating yield optical images on the photodetector matrix that are robust to variations in source spectral content, as shown in Fig. 5.

$$PFZP(x, y) = \Theta \left[\sum_{i=1}^3 \left| a + \frac{jk_i}{2\pi z_0} e^{-jk_i(x^2+y^2)} \right|^2 \right], \quad (2)$$

where $\Theta[\cdot]$ is a Heaviside step or threshold function, a a bias constant, z_0 a spatial distance corresponding to the distance of a virtual point source that interferes with normal plane waves, the k_i the wave numbers of the optical wavelengths chosen from the range desired and $j = \sqrt{-1}$ is the unit imaginary number. This grating can be considered the mixture of separate Fresnel zone plates each designed with different wavelengths. Figure 4 shows a typical panchromatic Fresnel zone plate created with three wavelengths $\lambda_i = 470, 550$ and 700 nm , corresponding to blue, green and red ranges of the optical spectrum. This grating yielded intermediate images of visual bars such as shown in Fig. 5.

V. SIGNAL PROCESSING

Every grating we tested produced wavelength-robust, narrow intermediate optical images of visual bars at the center of the visual field ($\theta = 0^\circ$) but somewhat complex multimodal images of bars at large field angles (e.g., $\theta = 40^\circ$). It is likely that a method of spatially varying dictionary learning [16] or complex dewarping followed by Bayesian or other pattern classification method [17] could be used to estimate the bar location from such optical images but such methods are computationally costly. Our goal was to find an algorithm with low computational cost, and thus we sought algorithms that were spatially independent, that is, applied the same algorithmic steps throughout the sensor domain.

The estimation problem is one-dimensional (the left-right position of the visual bar) and our first algorithmic step was to project the sensor signals onto a one-dimensional horizontal

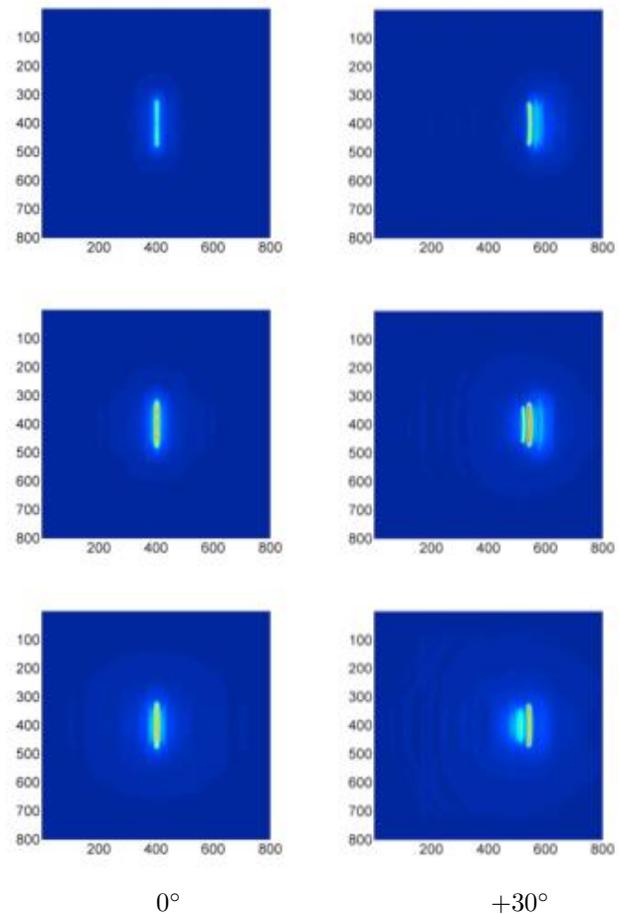


Fig. 5. Raw signals produced by the panchromatic Fresnel zone plate of Fig. 4 for a visual bar at 0° and 30° for $\lambda_i = 470, 550, 700 \text{ nm}$, top to bottom. Notice that on the axis (0°) the intermediate image is sharp and centered in the sensor matrix while off axis (30°) the image is structured, has multiple lobes, and is wavelength dependent. (Through simple signal processing lobes due to different discrete incident wavelengths can be isolated.)

line. This step integrates signals over vertical lines and hence reduces the effects of image noise and variations due to optics. The optical signals were reliable and accurate throughout the full field of view. The next computational step is to estimate some measure of the center or central tendency of the one-dimensional projected signal. We tried several computationally efficient methods to this end:

- peak or mode
- peak or mode after Gaussian smoothing
- mean of a Gaussian fit to the signal
- gradient pre-processing,

three of which are illustrated in Fig. 6.

VI. SENSOR PERFORMANCE

Figure 7 shows the results of five different signal processing algorithms for the most robust diffraction grating, the panchromatic binary Fresnel zone plate. Each subfigure shows three curves representing the estimated bar position (ordinate) versus the actual bar position (abscissa) in long, medium and short

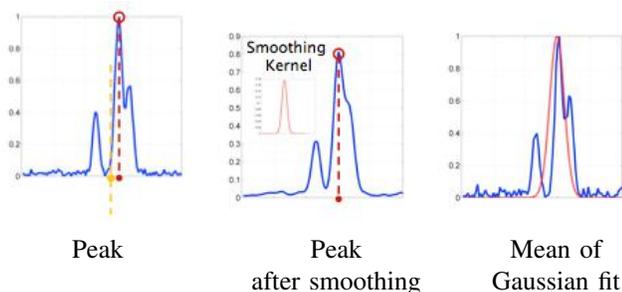


Fig. 6. The blue curves represent the projection of the full two-dimensional sensor signal along the vertical direction (i.e., parallel to the visual bar). The methods of central tendency estimation (for source angle) that were most accurate, reliable and robust to source spectral variation were the peak and the peak after a Gaussian smoothing (see below). This latter estimate did not depend upon the variance of the smoothing Gaussian, throughout a wide range of such variances.

wavelengths—red, green and blue. (In most subfigures, the three component curves overlap significantly and hence appear as one curve.) Each curve has a general sigmoidal or ogive shape due to geometric effects and Snell’s law of light passing through the grating at large incident angles (a version of barrel distortion).

The desiderata for the sensor are:

- all three color curves overlap perfectly throughout the entire angle range, indicating that the location estimate is independent of spectral composition of the visual bar
- all three curves are monotonic (a bijection), so that the unique angle estimate can be computed from the image center estimate by an inverse function or lookup table
- all three curves extend through a large angle of view
- antisymmetry of response with respect to central axis, indicating geometric consistency.

Both first two subfigures—peak and peak after Gaussian smoothing—show excellent performance on all four desiderata and accuracy of roughly 0.2° throughout the field of view, raising to 0.4° at the extremes of the field of view. The difference between the computational costs of these two methods are negligible compared to the full computational costs. Either of these methods, then, would be acceptable in a fielded sensor application.

The overall space computational cost of the estimation algorithm was 1.0 kB without lookup table for inverting the sigmoidal curves (as in Fig. 7), and time cost 6.0 Mflop/sec at video rates, and hence easily implemented in embedded processors or special CMOS. The signal projection step of the algorithm could be parallelized in a SIMD microarchitecture, but such speedups are not needed in practical applications.

VII. CONCLUSIONS AND FUTURE DIRECTIONS

We have designed and tested through extensive simulations an ultra-miniature lensless sensor for estimating the one-dimensional position of a visual bar throughout a large field of view and regardless of the spectral composition of the bar. Our end-to-end design approach led to an optical element (panchromatic binary Fresnel zone plate) that while

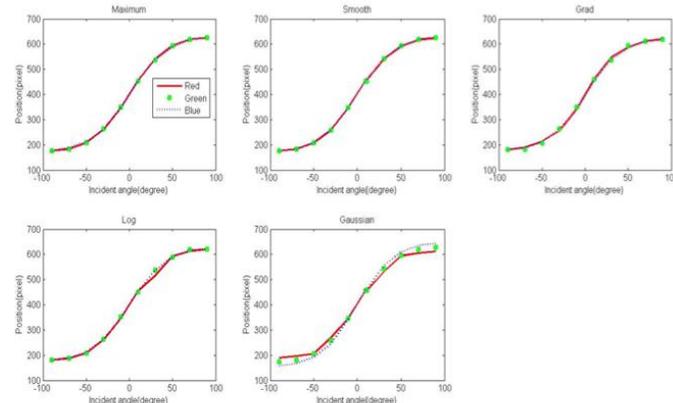


Fig. 7. The estimated bar center (in pixels on the the sensor) versus bar position in the field of view for a panchromatic Fresnel zone plate. Each subfigure contains three curves (for $\lambda = 470, 550$ and 700 nm light), but in several subfigures the curves overlap significantly and hence appear as one curve. The five signal processing algorithms (left-to-right, top-to-bottom) are: maximum or peak; peak after Gaussian smoothing; gradient-based estimation; peak after log processing; mean of a Gaussian fit.

somewhat complicated in design, is simple to manufacture and mount on a CMOS image sensor. The signal processing operates on the raw sensor signal (rather than a reconstructed or computed image) and is very computationally efficient. There are a number of directions for future work in end-to-end optimization based on these results, such as extending these methods to other image sensing functions. Finally, the deep challenge of developing a general-purpose theory for global end-to-end optimization of diffractive electro-optical systems incorporating physical constraints—analogue to the theory developed for traditional reflection and refraction based imagers [2]—remains elusive.

ACKNOWLEDGMENTS

We thank Thomas Vogelsang for helpful comments. M. Monjur and L. Spinoulas of Northwestern University performed this work as 2014 summer interns in Rambus Labs.

REFERENCES

- [1] W. T. Cathey and E. R. Dowski, Jr., “A new paradigm for imaging systems,” *Applied Optics*, vol. 42, no. 29, pp. 6080–6092, 2002.
- [2] D. G. Stork and M. D. Robinson, “Theoretical foundations of joint design of electro-optical imaging systems,” *Applied Optics*, vol. 47, no. 10, pp. B64–75, 2008.
- [3] D. Tseng, O. Mudanyali, C. Oztoprak, S. O. Isikman, I. Sencan, O. Yaglidere, and A. Ozcan, “Lensfree microscopy on a cellphone,” *Lab on a chip*, vol. 14, pp. 1787–1792, 2010.
- [4] A. E. Cetin, A. F. Coskun, B. C. Galarreta, M. Huang, D. Herman, A. Ozcan, and H. Altug, “Handheld high-throughput plasmonic biosensor using computational on-chip imaging,” *Light: Science and Applications*, vol. 3, pp. e122–, 2014.
- [5] P. R. Gill, “Odd-symmetry phase gratings produce optical nulls uniquely insensitive to wavelength and depth,” *Optics Letters*, vol. 38, no. 12, pp. 2074–2076, 2013.
- [6] P. R. Gill and D. G. Stork, “Lensless ultra-miniature imagers using odd-symmetry phase gratings,” in *Proceedings of Computational Optical Sensing and Imaging (COSI)*, Alexandria, VA, 2013.

- [7] —, “Digital camera with odd-symmetry spiral phase gratings supports full-resolution computational refocusing,” in *Advanced Photonics (Optical Society of America Sensors Congress)*, 2013.
- [8] —, “Hardware verification of an ultra-miniature computational diffractive imager,” in *Proceedings of Computational Optical Sensing and Imaging (COSI)*, Kohala Coast, HI, 2014.
- [9] D. G. Stork and P. R. Gill, “Lensless ultra-miniature computational sensors and imagers,” in *SensorComm 2013*, Barcelona, Spain, 2013.
- [10] —, “Optical, mathematical and computational foundations of lensless ultra-miniature diffractive imagers and sensors,” *International Journal on Advances in Systems and Measurements*, vol. 7, no. 3-4, pp. 201–208, 2014.
- [11] P. R. Gill and D. G. Stork, “Computationally efficient Fourier-based image reconstruction in a lensless diffractive imager,” in *Computational Optical Sensing and Imaging (COSI)*, Arlington, VA, 2015 (submitted).
- [12] P. R. Gill, M. Kellam, J. Tringali, T. Vogelsang, E. Erickson, and D. G. Stork, “Computational diffractive imager with low-power image change detection,” in *Proceedings of Computational Optical Sensing and Imaging (COSI)*, Alexandria, VA, 2015 (submitted).
- [13] D. G. Stork, “Computational diffractive sensing and imaging: Using optics for computing and computing for optics,” in *Society for Information Display*, San Jose, CA, 2015, p. (in press).
- [14] L. Spinoulas, O. Cossairt, P. R. Gill, D. G. Stork, and A. K. Katsaggelos, “Performance comparison of ultra-miniature diffraction gratings with lenses and zone plates,” in *Computational Optical Sensing and Imaging (COSI)*, Alexandria, VA, 2015 (submitted).
- [15] W. D. Furlan, G. Saavendra, and J. A. Monsoriu, “White-light imaging with fractal zone plates,” *Optics Letters*, vol. 32, no. 5, pp. 2109–2111, 2007.
- [16] J. Mairal, J. Ponce, G. Sapiro, A. Zisserman, and F. R. Bach, “Supervised dictionary learning,” in *Advances in Neural Information Processing Systems 21*, D. Koller, D. Schuurmans, Y. Bengio, and L. Bottou, Eds., 2009, pp. 1033–1040.
- [17] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York, NY: Wiley, 2001.

A Distributed Scheduling Algorithm to Improve Lifetime in Wireless Sensor Network based on Geometric Placement of Sensors with Coverage and Connectivity Constraints

Diery Ngom ^{+,*}
diery80@yahoo.fr

Pascal Lorenz ⁺
lorenz@ieee.org

Bamba Gueye ^{*}
bamba.gueye@ucad.edu.sn

University of Cheikh Anta Diop ^{*}
Dakar, Senegal

University of Haute Alsace ⁺
Colmar, France

Abstract – Wireless Sensors Network (WSN) are generally used in hostile environment and have several resource constraints such as limited energy, limited memory, limited communication capacities, etc. These limitations can cause changes in topology, isolation of sensors nodes, network partitioning, and loss of communication links. The results are the reduction of network lifetime and the deterioration of the coverage and network connectivity. Therefore placement strategies of sensors nodes and sleep/wake-up policy should be planed in a way to improve their lifetime while maintaining full coverage of the monitored region and optimal network connectivity. This paper proposes a novel algorithm that reduces the energy consumption under coverage and network connectivity constraints. The obtained results show that sensors nodes can save up to 30% of their energy with respect to previous works.

Keywords – Wireless Sensor Network; network lifetime; coverage; network connectivity.

I. INTRODUCTION

WSN has inspired tremendous research in the scientific and industrial communities in recent years. Advances in wireless communication and microelectronic mechanical systems have enabled the development of low cost, low power, multi-functional, tiny sensor node which can sense the environment, perform data processing and communicate each other over short distances. A WSN is an ad hoc network composed of many sensors nodes deployed either randomly or deterministically over a geographical region of interest and communicating via wireless links. These sensors can also collect data from the environment, do local processing and transmit the data to a sink node or base station via multipath routing. A wide range of potential applications have been envisioned using WSN such as environmental conditions monitoring, wildlife habitat monitoring, security surveillance, industrial diagnostic, agricultural of precision, improve health care, etc. Given their small size, low cost and deployment generally in hostile or difficult access areas, the sensors have severe resource constraints such as a limited energy, so a network lifetime limited, limited memory, limited bandwidth, limited computation and communication capacities, etc. These effects generate changes in topology, isolation of nodes, network

partitioning, etc., so that they can reduce the network lifetime and degrade the coverage of the monitored region, and the network connectivity. Many related works are proposed in recent years to address these issues. The most important concern issues for optimizing network lifetime, sensors placement, coverage, and network connectivity. Placement methods, coverage and network connectivity problem are primarily aimed at ensuring full or acceptable coverage and optimal network connectivity. The optimization of the network lifetime methods focalize more on the extension of the lifetime while minimizing the energy consumption in the network; and especially not allow at the same time to ensure a total coverage of the monitoring area and a network connectivity. However, issues for optimizing the network lifetime, guaranteeing full coverage of the monitored region, and maintaining optimal network connectivity are usually conflicting in many applications of WSN such as security surveillance, agricultural of precision, habitat monitoring, etc. These applications require full coverage of the monitored region and optimal network connectivity while maximizing the network lifetime.

We propose in this paper a Distributed Scheduling Medium Access Control (DSMAC) algorithm to optimize the network lifetime while maintaining a full coverage of the monitored region and optimal network connectivity. First, we propose a novel network placement method which minimise the number of require sensors to ensure: the entire coverage of the monitored region and optimal network connectivity. Then, based on the geometric properties of our placement method, we implement a sleep/wake-up policy to improve the network lifetime while saving the full coverage of the monitored region and the optimal network connectivity. Our proposal is based on duty cycle techniques, redundancy of sensing coverage and beacon frames sent by each source node before the transmission of its data packets. Firstly, we present the analytical proof of the full coverage and the network connectivity. Secondly, we validate our proposal by simulations and compare it with the “*TunableMAC*” protocol [12]. The simulation results show that with DSMAC the nodes save about 30% of their energy and the received packets by the Sink node is improved relative to TunableMAC protocol.

The remaining of this paper is organized as follows. Section II surveys the related works. Section III presents our geographic placement method based on grids. Section IV presents the DSMAC algorithm and the analytical

evaluation. Section V presents the validation of the DSMAC algorithm by simulations. Finally, Section VI concludes the paper and discusses future works.

II. RELATED WORKS

Network lifetime, placement methods, coverage, and network connectivity problem are critical issues in WSN. A lot of works have been done in recent years by the researchers for addressing these issues.

Akewar et al. [1] discuss the different deployment strategies such as forces, computational geometry and pattern based deployment. However, they don't address the lifetime issues in their study. With the same goal, Ankur et al. [2] present different placement strategies of sensors nodes in WSN taking into account the lifetime issues. They note that the most objective of placement techniques have focused on increasing the area coverage, obtaining strong network connectivity and extending the network lifetime. A more study of coverage and connectivity issues in WSN are presented in a survey by Khou et al. [3]. In this survey, the authors motivate their study by giving different use cases corresponding to different coverage, connectivity, latency and robustness requirements of the applications considered. They present also a general and detailed analysis of deployment problems in WSN. In their analysis, different deployment algorithms for area coverage, barrier coverage, and coverage of points are studied and classified according to their characteristics and properties. Note that, this survey is good references to have an overall view of coverage and connectivity issues in WSN. However, note that in their survey the network lifetime problem are not addressed while this problem is often in conflict with the coverage and connectivity problems. Zhu et al. [4] address the issues of coverage, connectivity, and lifetime in WSN; and they distinguish two coverage problems: static coverage and dynamic coverage. After the study of coverage problem, they propose a scheduling mechanism for sensors activities in order to reduce the energy consumption in the network and they analyze at the same time the relationship between coverage and network connectivity. Nevertheless, note that placement problem is not study and take account in their proposal. With the same goal, another approach which take account the sensors placement method based on territorial predator scent marking behavior is proposed by Abidin et al. [5]. The main goals of their proposal are: to achieve maximum coverage, to reduce the energy consumed and to guaranty network connectivity. However, note that in their approach the full coverage of the monitored region is not guaranteed. Also in this context, Mulligan et al. [6] present different coverage protocols that try to maximise the number of sensor which put into sleep mode while guaranteeing k-coverage and network connectivity. Singaram et al. [7] present also a recent study in which they propose a self-scheduling algorithm that extends the network lifetime while minimizing the number of active sensors. Note that in these two studies, connectivity issues are also not addressed by these authors. A recent survey for sensors lifetime enhancement techniques in WSN is

presented by Ambekar et al. [8]. Nevertheless, as some previous authors in the related works, coverage and connectivity issues are not addressed by these authors.

With the same goal for optimizing the network lifetime in WSN by scheduling the sensors activities, more energy efficient MAC protocols based on duty cycle are developed. In fact, the duty cycle approach is the main feature of synchronous and asynchronous MAC protocols where any node can alternate between active and sleep states in order to save its energy. In this approach, nodes can only communicate when they are in active state. Several MAC protocols such as "S-MAC", "T-MAC", "B-MAC", and "X-MAC" based on duty cycle approach were proposed respectively by Kaur et al. [9], Kakria et al. [10], and Ullah et al. [11]. In the S-MAC protocol, nodes alternates between active and sleep periods. During active periods, the node radios are turned ON to communicate and during sleep periods the node radios are turned OFF to save energy. However this protocol has some limitations. Firstly, nodes broadcast their schedule to all neighbor nodes using the synchronization packet; so that this mechanism is not efficient in energy consumption. Another limitation of this protocol is that, all the border nodes incorporate the schedules and keep their radios ON during all their active periods. T-MAC extends S-MAC and provides several improvements. In T-MAC, the S-MAC limitations were overcome by including an adaptive duty cycle where the length of the active period is varied according to traffic. Therefore, this protocol significantly increases the network lifetime by downsizing the length of the active periods. However, such as S-MAC, in this protocol nodes broadcast their schedule to all communication neighbors using the synchronization packet. Thus, this mechanism is not efficient in energy consumption and is not suitable in a network with redundancy coverage. B-MAC is a MAC protocol which adopts the LPL (Low Power Listening) technical. In this technical, the nodes periodically switches between active and sleep periods. The active period is usually very short, just allows the node to sense the channel. B-MAC provides good energy efficiency. Nevertheless, since this protocol uses CSMA/CA (Carrier Sense Multiple Access/with Collision Avoidance) for the medium access, it suffers flow problem at the high load due to the collisions and the random backoff periods. Another problem of B-MAC is the over-listening of the preamble by all neighbor nodes because even if the packet is intended only for a particular node (next hop), all other neighbor nodes must still active to listen the preamble. X-MAC is another MAC protocol based on duty cycle approach. This protocol is an improvement of B-MAC to solve the over-listening problem. Instead of transmitting a long preamble, X-MAC divides it into a series of small packets preamble, each of them containing the receiver's address packet to be transmitted. Compared to B-MAC, X-MAC improves energy efficiency and reduces the time using the shortcut preamble. However, X-MAC may choose only one next hop (router) to move the packet to its destination, even if there are multiple paths in the network whose exploitation could make robustness in the

transmission. Another limitation of X-MAC is the low flow problem. Indeed, when the load is high this remains no resolved due to the use of CSMA/CA mechanism for the medium access.

Furthermore, Boulis [12] proposes the TunableMAC protocol based also on the duty cycle approach. As in other MAC protocols note that, in TunableMAC the CSMA/CA mechanism is used for the medium access. It is worth noticing that with his protocol all the nodes are not aligned in their active period, so that each sender transmit an appropriate train of beacon frames to wake up potential receivers before transmitting each data packet. Thus, with respect to this mechanism all neighbour that act as potential receivers of a given sender will be awakened when they received the beacon frame from the sender. Therefore, a lack of efficiency is noted in term of energy consumed. However, TunableMAC is very flexible and can be used to make comparisons with new distributed MAC algorithms developed for WSN.

III. GEOGRAPHIC PLACEMENT OF SENSOR NODES

A. Asymptions and notations

We represent the WSN by a graph $G = (V, E)$ where V represents all vertices (nodes of the network) and $E \subseteq V^2$ represents the set of edges giving all possible communications. There is an ordered pair $(u, v) \in E^2$ if the sensor node u is physically capable to transmit messages to the sensor node v . In this case, sensor node v is located in the communication range of sensor node u . Thus, each node has its key communication range noted R_c that allows it to communicate with others sensor nodes. We assume that all sensor nodes have equal communication ranges. Each sensor node also has a sensing range noted R_s that allows it to sense and capture data from the environment. We also assume that all sensor nodes have the same sensing ranges. All the sensor nodes v located inside the communication range of a given sensor node u are called neighbour nodes of sensor node u and are noted $N(u)$. A bidirectional wireless link exists between a sensor node u and every neighbour $v \in N(u)$ and is represented by the directed edges (u, v) and $(v, u) \in E$. Note that all the neighbour nodes can communicate directly each other. In the following we note $M = \{S_1, S_2, \dots, S_M\}$ the set of sensor nodes in the WSN.

We note also $N = |M|$ the cardinality of the set M that also represents the number of sensor nodes in the WSN. In other hand, we assume in our study that all the sensor nodes transmit their captured data to a Sink node which is the only receiver.

In this part, we will give the modelling of the wireless channel. The performances of a wireless communication system are determined based on the communication channel in which it operates. In a WSN, modelling of communication is very difficult because the sensor nodes communicate in low power, and therefore radio links

nodes are very unreliable. The unit disk model is the simplest deterministic model of communication that illustrates a unidirectional link between two sensor nodes. This model assumes that each sensor node is able to transmit its data to any sensor node being in its communication range. The communication range of each sensor node varies depending on the level of its transmission power. Therefore, we can say that two sensor nodes u and v can communicate each other if and only if the Euclidean distance noted $d(u, v)$ between the two sensor nodes is less than the communication range R_c of these two sensor nodes. Thus, two nodes $u, v \in M$ can communicate if:

$$d(u, v) \leq R_c$$

Regarding the modelling of coverage, we consider in this paper coverage of point and coverage of area (region). Thus, we say that a sensor node S_i covers a point $q \in A$ if and only if:

$$d(S_i, q) \leq R_s$$

A coverage of surface (sensing coverage) means the total surface lying below the range of capture of data at least of a given sensor node. Let $S_i \in M$ a sensor node and note $C(S_i)$ the surface cover by the sensor node S_i , then:

$$C(S_i) = \{q \in A \mid d(S_i, q) \leq R_s\}$$

The surface covered by a subset of sensor nodes $S_c = \{S_1, S_2, \dots, S_c\} \subseteq M$ is then:

$$C(S_c) = \bigcup_{i=1}^{|S_c|} C(S_i)$$

For the coverage of area, we say that a sensor node S_i covers a region A if and only if for each point $q \in A$ then:

$$d(S_i, q) \leq R_s$$

We finish this part for the modelling of network connectivity. Two sensor nodes are said to be connected if and only if they can communicate directly (one-hop connectivity) or indirectly (multi-hop connectivity) [3]. In WSN, the network is considered to be connected if there is at least one path between the sink and each sensor node in the considered area. Thus, connectivity essentially depends on the existence of routes. It's affected by topology changes due to mobility, the failure of nodes, the attacks, etc. The results are the loss of communication links, the isolation of nodes, the network partitioning, thus the coverage of the monitored area can be degrade and/or the network lifetime can be decrease. Therefore, understanding the network connectivity of a WSN allows adapting the communication mechanisms to ensure the smooth running of applications. To study it, we will use the following notations and definitions. The neighbourhood $N(u)$ of a sensor node u represents the set of neighbour nodes that are within the communication range of node u . It is defined by:

$$N(u) = \{(u, v) \in V^2 \mid u \neq v \wedge d(u, v) \leq R_c\}$$

A graph is called k -connected if there are at least k disjoint paths between two nodes of this graph. Coverage is often related to connectivity in WSN. In order to satisfy the conditions of coverage and connectivity, we consider in this paper that the communication range R_c is twice the sensing range R_s .

B. Sensors placement model based on grids

We consider the following placement method.

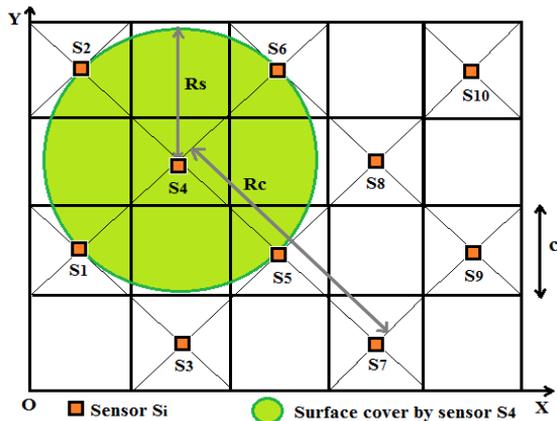


Figure 1. Sensor placement model in the monitoring region

According to the deployment of sensors described in the Figure 1, the geographical area is partitioned into contiguous square grids having the same dimensions that equal to c . The sensor node S_i is placed at a given area of a grid such that the entire area of the monitoring region is covered and the number of necessary sensors is minimized. Our geographic placement of sensors presents the following advantages:

1. The number of sensor nodes required to cover the whole area is minimized;
2. The position and the surface cover by each sensor node are known and can be respectively determined by its coordinates (X, Y) and/or its sensing range R_s .
3. A full coverage and optimal network connectivity are ensured;
4. There are redundancies on the sensing coverage which will be exploited by our proposal for scheduling sensors activities.

Now, the optimal length of c to ensure full coverage and network connectivity of our network placement model can be determined based on the sensing range R_s . This sensing range depends on the communication range R_c based on our assumptions. Furthermore, R_c is determined by the transmission power noted *TX power* and others parameters of the radio communication model. Based on geometric properties of squares and the placement model in Figure 1, the sensing range R_s can be computed using the following equation:

$$R_s^2 = c^2 + c^2 \Rightarrow R_s = c\sqrt{2} \Rightarrow c = \frac{R_s}{\sqrt{2}}$$

Now, let $M = \{S_1, S_2, \dots, S_M\}$ the set of sensor nodes in the WSN and note that each sensor node S_i has (X, Y) coordinate in the coordinate system (O, X, Y) as shown in the Figure 1, where O is the originate of this coordinate system and X, Y represent respectively the X axis and the Y axis of the coordinate system. For instance:

$$S_1\left(\frac{c}{2}, \frac{3c}{2}\right), S_2\left(\frac{c}{2}, \frac{7c}{2}\right), S_3\left(\frac{3c}{2}, \frac{c}{2}\right), \text{ and } S_4\left(\frac{3c}{2}, \frac{5c}{2}\right).$$

We show that, according to our placement method an area may be covered by many sensor nodes at the same time; this is due to overlapping coverage areas of neighbour sensor nodes. To optimize the energy consumption in the network, it is necessary to apply sleep/wakeup strategies by scheduling sensor nodes activities while maintaining the full coverage of the monitored region and optimal network connectivity to ensure robustness in the communication.

IV. PRESENTATION OF DSMAC AND ANALYCAL EVALUATION

A. Overview of DSMAC algorithm

DSMAC algorithm (Algorithm 1) includes the placement method described in Figure 1 and a sleep/wakeup policy based on duty cycle approach. It enables to minimize the energy consumed by the overall network while maintaining a full coverage and network connectivity with respect to all sensors. The DSMAC algorithm exploits the redundancy of sensing coverage due to our geographic placement method. Indeed, according to TunableMAC protocol, each sender should transmit a train of beacons frames in order to wake up its entire neighbourhood before sending any data packets. However, according to our DSMAC deployment of sensor nodes, where we have sensing coverage redundancies due to our placement strategy of sensor nodes, we do not need to wake up all a given sensor's neighbourhood. It is worth noticing that in TunableMAC, the set of sensor nodes has equal sleep interval and equal listening interval. Put simply, DSMAC wakes up only few sensor nodes among a well-chosen sensor's neighbourhood in order to reduce the energy consumed during transmission and reception as well as mitigates the number of collisions between sensor nodes. According to DSMAC, each sensor node uses a neighbourhood's table that contains the *ID* of neighbour's nodes which is determined by the communication range R_c . Also, the sensor nodes have different sleep and listening intervals. DSMAC addresses the following two issues noted in previous studies:

- (i) The set of sender's neighbours that should wake up according to its neighbour's table;
- (ii) The scheduling of sleeping and listening intervals according to the parameters of the duty cycle.

In order to choose the best possible neighbours of a sensor node that enable to minimize the energy consumption during transmission while guaranteeing full coverage and network connectivity, to taking into account the two issues raised above, we consider two types of neighbours for each sensor node: "close neighbours" located at a

maximum distance of $c\sqrt{2}$ from the sender and “remote neighbours” located at a distance strictly greater than $c\sqrt{2}$. For a given sender, its neighbour’s receivers are only its remote neighbours. Therefore, remote neighbours must be woken up and all the remaining sensor nodes within its close neighbourhood must be set in sleeping mode (line 8 to line 14 of Algorithm 1). If they receive other beacons frame, they can decide whether they should wake up again to relay packets. Thus, our algorithm enables the following benefits:

1. Save the energy consumed in the network, so that the network lifetime will be improved;
2. Ensure full coverage and network connectivity;
3. Balance energy consumption in the network;
4. Reduce collisions that may be due to the CSMA/CA mechanism, so that the rate of received packets will be improved.

The Algorithm 1 above describes the DSMAC algorithm in pseudo code.

Inputs:

- “ c ” represents the length c of a given grid
- “ $d(X,Y)$ ” represents the Euclidean distance between X and Y
- “Neighbor_Table” represents the node’s neighbours table
- “ID” represents the ID of a given sensor node
- “ B_i ” represents the beacon frame sent by the source S_i

Output:

A set of active sensor to relay packets and a set of sensor in Sleep mode to save energy

```

1: for each sensor node  $S_i \in M$  do
2:   for each sensor node  $S_j \in M \wedge S_j \neq S_i$  do
3:     if  $d(S_i, S_j) \leq 2c\sqrt{2}$  then
4:       Insert( $ID_{S_j}, Neighbor\_Table[S_i]$ )
5:     end if
6:   end for
7: end for
8: for each sensor  $S_i \in M$  which broadcast a beacon  $B_i$  do
9:   if  $S_j \in M$  receives  $B_i$  and  $ID_{S_j} \in Neighbor\_Table[S_i]$  then
10:    if  $d(S_i, S_j) \leq c\sqrt{2}$  then
11:      Make  $S_j$  in sleep state until it receive a next beacon  $B_i$ 
12:    end if
13:  end if
14: end for
    
```

Algorithm 1. DSMAC algorithm

As shown in the Algorithm 1 that describes the DSMAC algorithm, the lines 1 to 7 of this algorithm enable to compute the neighbour table for each sensor node $S_i \in M$ by inserting all the ID of its sensors neighbour $S_j \in M \wedge S_j \neq S_i$. After this step, each sensor node $S_j \in M$ neighbour of a given sender $S_i \in M$ will decide if it will be switched in Active or Sleep mode based on the beacon frame received by this sender (which precede the

data transmission of the source) and its neighbour table (lines 8 to 14). Therefore, the sensor nodes will alternate in Sleep and Active mode to save their energy; so that the network lifetime will be improved. Note that the full coverage and network connectivity will be preserved during all the network lifetime. We will give in the following part the analytical proof of the full coverage and the network connectivity.

B. Analysis of the full coverage and network connectivity in DSMAC

In this part, we will present respectively the mathematical proof of the full coverage and the network connectivity regarding to DSMAC algorithm.

Consider a sender $S_i(x, y) \in M$. As we mentioned in the previous sections, before this source S_i transmits its data packets, then it broadcast a train of beacon frames noted respectively $B_{i1}, B_{i2}, \dots, B_{ik}$ to wake up all the nodes S_j belonging to its neighbour table and located at a distance strictly greater than $c\sqrt{2}$. These nodes are all the remote neighbours of the sender $S_i(x, y)$. According to the placement method described in the Figure 1 and the illustration of remote/close neighbour described in the Figure 2, then the coordinate of these remote neighbours of $S_i(x, y)$ are the nodes of the following coordinate:

$$(x, y - 2c), (x, y + 2c), (x - 2c, y - 2c), (x - 2c, y), (x - 2c, y + 2c), (x + 2c, y - 2c), (x + 2c, y), (x + 2c, y + 2c)$$

The coordinate of other close neighbour of the sender $S_i(x, y)$ which must be making in sleep mode are the nodes of the following coordinates:

$$(x - c, y - c), (x - c, y + c), (x + c, y - c), (x + c, y + c)$$

Let us consider the sensor node $S_7(x, y)$ shown in Figure 2. Its neighbourhood’s table contains ID of the set of the following sensor nodes:

$$\{S_1, S_2, S_3, S_4, S_5, S_6, S_8, S_9, S_{10}, S_{11}, S_{12}, S_{13}\}$$

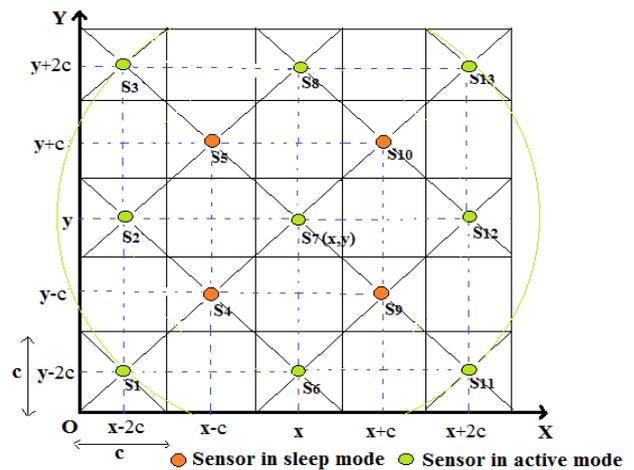


Figure 2. Illustration of close and remote neighbours of $S_7(x, y)$

If the sensor node $S_7(x, y)$ wants to transmit, then the set of sensors located to its neighbourhood table which must wake up after receiving the beacon frames sent by the

sensor node S_7 are: $\{S_1, S_2, S_3, S_6, S_8, S_{11}, S_{12}, S_{13}\}$ and the following sensor nodes $\{S_4, S_5, S_9, S_{10}\}$ should be put in sleep mode. According to the Figure 2, the sensor nodes S_4, S_5, S_9 and S_{10} are in sleep modes at the same time whereas other sensor nodes belonging to S_7 's neighbour table are in active mode and maintain the full coverage of the network. We show that the areas covered by the following sensor nodes S_4, S_5, S_9, S_{10} which are in sleep mode, and the one covered by the four active sensor nodes located at the vicinity of these sleeping sensors are fully covered by the active sensors. Let us consider the sensor node S_4 which is in sleep mode (Figure 2), then according to the definition of the sensing coverage of this sensor node noted $C(S_4)$, we have:

$$C(S_4) = \{q \in A \mid d(S_4, q) \leq c\sqrt{2}\} \quad (1)$$

According to the active sensor nodes S_1, S_2, S_6 and S_7 which are around the sensor node S_4 , the sensing coverage of each of them is:

$$C(S_1) = \{q \in A \mid d(S_1, q) \leq c\sqrt{2}\} \quad (2)$$

$$C(S_2) = \{q \in A \mid d(S_2, q) \leq c\sqrt{2}\} \quad (3)$$

$$C(S_6) = \{q \in A \mid d(S_6, q) \leq c\sqrt{2}\} \quad (4)$$

$$C(S_7) = \{q \in A \mid d(S_7, q) \leq c\sqrt{2}\} \quad (5)$$

Note $S_c = \{S_1, S_2, S_7, S_6\}$. Based on the coverage area described in Section III, we have:

$$C(S_c) = C(S_1) \cup C(S_2) \cup C(S_7) \cup C(S_6) \quad (6)$$

On the other hand, if we compute the Euclidean distance between the sensor node S_4 and the sensor nodes S_1, S_2, S_6 and S_7 , then we have:

$$d(S_4, S_1) = d(S_4, S_2) = d(S_4, S_7) = d(S_4, S_6) = c\sqrt{2} \quad (7)$$

Based on the sensing coverage of the sensor nodes S_4, S_1, S_2, S_6, S_7 described respectively in (1), (2), (3), (4), (5), and based in (6), (7), we have:

$$C(S_4) \subset C(S_1) \cup C(S_2) \cup C(S_7) \cup C(S_6) \quad (8)$$

Hence based to (8), then the sensor nodes S_1, S_2, S_6 and S_7 provide a full coverage with respect to the area covered by the sensor node S_4 . Similarly, we can show that sensors S_2, S_3, S_8 and S_7 (resp. Sensors S_6, S_{11}, S_{12} and S_7) provide a full coverage according to the area covered by the sensor S_5 (resp. sensor S_9). Finally, sensors S_8, S_{13}, S_{12} and S_7 provide a full coverage with respect to the area covered by sensor S_{10} . Since the sensor $S_7(x, y)$ is chosen randomly, we can conclude that the network remains fully covered when running the DSMAC algorithm.

In fact, two sensors S_u and S_v are connected if and only if:

$$d(S_u, S_v) \leq 2c\sqrt{2} \quad (9)$$

In order to prove the network connectivity, it is sufficient to show that all active neighbours of a given sender

$S_i(x, y)$ will be connected to this sender during the execution of the DSMAC algorithm. The remote neighbours of the considered sender $S_i(x, y)$ are:

$$\text{Remote_Neighbor_}S_{x,y} = \{S_{N1}(x, y - 2c), S_{N2}(x, y + 2c), S_{N3}(x - 2c, y - 2c), S_{N4}(x - 2c, y), S_{N5}(x - 2c, y + 2c), S_{N6}(x + 2c, y - 2c), S_{N7}(x + 2c, y), S_{N8}(x + 2c, y + 2c)\}$$

If we compute the Euclidean distance between $S_i(x, y)$ and each of sensor nodes $S_j \in \text{Remote_Neighbor_}S_{x,y}$, we have:

$$d(S_i, S_j) \leq 2c\sqrt{2} \quad (10)$$

For instance:

$$d^2(S_i, S_{N1}) = (x - x)^2 + (y - (y - 2c))^2 = (2c)^2$$

$$\Rightarrow d(S_i, S_{N1}) = \sqrt{(2c)^2} = 2c \leq 2c\sqrt{2}$$

Therefore, from (10) and based to (9) which illustrates the connectivity condition between two sensor nodes, then all sensors $S_j \in \text{Remote_Neighbor_}S_{x,y}$ are connected to the sensor node $S_i(x, y)$. Since the sensor $S_i(x, y)$ is chosen randomly, then all active sensors will be connected during the execution of DSMAC algorithm. In addition, according to the definition of a graph which is k-connected and according to Figure 2, the network is at least 4-connected; therefore, there is an optimum routing topology. However, we will not discuss the routing aspect in this paper.

V. EVALUATION OF DSMAC ALGORITHM

We validated our proposal by extensive simulations done with “*Castalia.3.0*” framework [12]. Castalia is a WSN simulator for Body Area Networks (BAN) and generally networks of low-power embedded devices. It is based on the OMNeT++ platform.

A. Experimental setting

We consider a field of size equal to $(200 \times 200) m^2$. The deployment type is static. We run four simulations scenario with respectively 40, 80, 120, 160, and 200 sensors that send their packets to a given Sink. The simulation time is set to 400 seconds. We used the “CC2420” radio type. The “TX power”, the “communication range”, the “sensing range”, and the “grid length” are respectively set to 0dB, 20m, 10m, and 7m. For the application test, we considered “ThroughputTest” [12] to send constant data payload of 2000 bytes with a rate of 5 packets per second to the Sink. Note that in our simulation, all nodes are the same “initial energy” equal to 18720 Joules (J) corresponding of 2 piles AA.

B. Simulation results

We compared DSMAC and TunableMAC according to different metrics such as the energy consumed, the number received packets by the sink and the application level latency. We performed extensive simulations by considering the same scenarios and the same parameters.

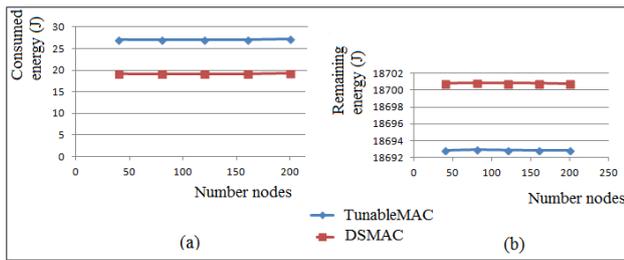


Figure 3. Energy-awareness between DSMAC and TunableMAC

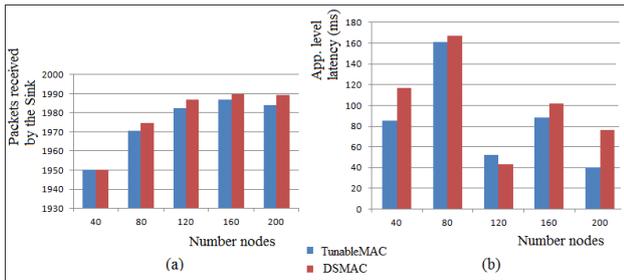


Figure 4. Evaluations of received packets and level latency in DSMAC and TunableMAC

The curves illustrated in Figure 3a and Figure 3b show respectively the average of energy consumed and the average of remaining energy in J for the both algorithms. DSMAC outperforms TunableMAC with respect to the energy consumed. Indeed, with DSMAC only few senders' neighbours are awakened in contrast to TunableMAC where the entire set of a node's neighbours are awakened. Therefore, more active nodes exist and thus the energy consumed is increased. The average of energy consumed in the network is roughly equal to 19.18 J (resp. 27.09 J) for DSMAC (resp. TunableMAC). According to DSMAC sensors can save up to 30% of their energy compared to TunableMAC.

Figure 4a (resp. Figure 4b) show the average packets received by the Sink (resp. the average application level latency in ms). Figure 4a illustrates that DSMAC outperforms TunableMAC according to the number of packets received by the Sink. The main reason is due to the fact that DSMAC algorithm mitigates the number of collisions. Furthermore, Figure 4b shows the average application level latency. As shown in this figure the performance of TunableMAC is slightly upper than DSMAC but the upper level latency in these two algorithms is less than 166.9 ms, thus the level latency is reasonable in DSMAC regarding to the most applications for WSN.

VI. CONCLUSION AND FUTURE WORK

In this paper, we proposed a distributed scheduling algorithm based on a geometric placement model in order to improve the network lifetime while maintaining full coverage and network connectivity. After the implementation of DSMAC, we proved analytically that

coverage and network connectivity are ensured at every time of the network lifetime during the execution of our algorithm.

Simulation results show that DSMAC outperforms the TunableMAC protocol with respect to network lifetime and the average of received packets by the Sink

In future work, we investigate a more realistic modelling of the communication channel, taking into account the path loss and temporal variations. We also intend to make new proposals with respect to routing in WSN based on our sensor placement model.

REFERENCES

- [1] M. C. Akewar and N. V. Thkur, "Study of Wireless Mobile Sensor Network Deployment," in *International Journal of Computer and Wireless Communication*, 2012, vol. 2–No. 4, pp. 533- 541.
- [2] P. Ankur and M. A. Rizvi, "Analysis of Strategies of Placing Nodes in Wireless Sensor Networks," in *International Journal of Advanced Research in Computer Science and Software Engineering*, February 2014, Issue 2, Volume 4, pp. 398-402.
- [3] I. Khou, P. Minet, A. Laouiti, and S. Mahfoudh, "Survey of Deployment Algorithms in Wireless Sensor Networks: Coverage and Connectivity Issues and Challenges," in *International Journal of Autonomous and Adaptive Communications Systems*, 2014, pp. 1- 24.
- [4] C. Zhu, C. Zheng, L. Shu, and G. Han, "A survey on coverage and connectivity issues in wireless sensor networks," in *Journal of Networks and Computer Applications*, 2012, pp. 619-632.
- [5] H. Z. Abidin, N. M. Din, and N. A. M. Radzi, "Deterministic Static Sensor Node Placement in Wireless Sensor Network based on Territorial Predator Scent Marking Behavior," in *International Journal of Computer Network and Information Security*, December 2013, Vol. 5–No. 3, pp. 186-191.
- [6] R. Mulligan, "Coverage in Wireless Sensor Networks: A Survey," in *Networks Protocols and Algorithms*, 2010, Vol. 2– No. 2, pp. 35- 41.
- [7] M. Singaram, D. S. Finney, N. K. Sathish, and V. Chandraprasad, "Energy Efficient Self-Scheduling Algorithm For Wireless Sensor Networks," in *International Journal of Scientific & Technology Research*, Janvier 2014, volume 3, issue 1, pp. 75-78.
- [8] D. V. Ambekar, A. D. Bhoi, and R. D. Kharadkar, "A Survey on Sensors Lifetime Enhancement Techniques in Wireless Sensor Networks," in *International Journal of Computer Applications*, December 2014, Volume 107–No. 19, pp. 22-25.
- [9] J. Kaur, T. C. Aseri, and A. Kakria, "A survey of locally synchronized MAC protocols in Wireless sensor network," in *International Journal of Computer Applications*, April 2014, Volume 91–No. 11, pp. 37- 41.
- [10] A. Kakria and T. C. Aseri, "A Survey on Asynchronous MAC protocols in Wireless Sensor Networks," in *International Journal of Computer Applications*, December 2014, Volume 108 – No. 9, pp. 19- 22.
- [11] A. Ullah, G. Kim, and J. S. Ahn, "Performance Analysis of X-MAC protocol with Collision Avoidance Algorithm," *IEEE INFOCOM*, 2013, pp. 211- 212.
- [12] A. Boulis, "Castalia: a simulator for Wireless Sensor Networks and Body Area Networks", Version 3.2 User's Manual, March 2011.

An Algorithm to Evaluate and Build Schedules for a Distributed Sensor System with Respect to Clock Synchronization

Andreas Puhm, Michael Kramer, and Martin Horauer

University of Applied Sciences Technikum Wien, Vienna, Austria
Email: {andreas.puhm, michael.kramer, martin.horauer}@technikum-wien.at

Abstract—Distributed embedded systems utilizing active sensors that share a common media (e.g., light detection and ranging sensors, LIDAR) have to be coordinated to eliminate or reduce interferences between the different sensor nodes. For LIDAR and similar sensor systems it is possible to apply media access methods, like frequency, code and time division multiple access (FDMA, CDMA or TDMA). FDMA and CDMA methods utilize sensor specific parameters for medium access. This can be a drawback, as it is no longer possible or far more difficult to use error correction or measurement enhancement techniques, like multi-frequency measurements, that depend on these parameters. A TDMA scheme frees these sensor specific parameters, but needs additional infrastructure in the form of clock synchronization. The implementation of the synchronization mechanism constrains the TDMA schedules that can be executed. This paper provides an algorithm to ascertain these constraints, and thus determine the requirements for the synchronization mechanism implementation.

Keywords—Distributed Sensor Network; Clock Synchronization; TDMA; LIDAR.

I. INTRODUCTION

With the advance of distributed sensor systems the need for coordination in such systems arises. Tightly coupled distributed sensor systems need to be coordinated to eliminate measurement errors and correlate results from the individual sensor nodes. For example, if active sensors (e.g., LIDAR or ultrasonic based) are deployed in vicinity, the correct management of the sensor activation and measurement times is a central issue. For applications, like factory automation or indoor positioning, the sensors might be mounted on mobile units. Therefore, simple coordination schemes, like wired triggers, are no longer feasible.

For example, Time-of-Flight (ToF) sensors [1] are available that combine a LIDAR measurement mechanism with a charge-coupled device image sensor. ToF sensors generate image (frame) data similar to a charge-coupled device sensor where each pixel provides distance data in addition to a (gray scale) image. Such a sensor relies on an infrared or laser light source that is activated for each measurement for a certain time. To avoid interference no other light source with the same modulation pattern, e.g., same frequency, should be active during this time. Standard media access schemes, like, FDMA [2], CDMA [3], or phase hopping [4] have been applied to ToF sensor systems. These schemes use different frequencies and/or modulation patterns to eliminate or reduce the interferences between multiple ToF sensors. For a ToF sensor the maximum measurable distance and resolution is determined by the frequency of the light source. Therefore, a

media access scheme that utilizes the light source modulation itself, will constrain the sensor with regard to the measurable distance and resolution.

By employing a TDMA scheme, these sensor parameters can be freely used for measurement enhancement processes and error corrections (e.g., [5] and [6]).

Ultrasonic sensors are another type of active sensor, that can be applied for position measurement in a distributed sensor system, e.g., Constellation [7], BUZZ [8], or Cricket [9]. These systems are based on Time-of-Arrival measurements. A radio signal and an ultrasonic signal are generated at the same time by a sending node. The difference of the reception time of these two signals is measured and used to calculate the distance between sender and receiver. In the above mentioned positioning systems multiple beacons are used so that a mobile device can determine the position by correlating the reception time differences. If transmissions of multiple beacons overlap, the measurement results can be erroneous, as a wrong radio signal to ultrasonic signal match can occur. Therefore, they have to be detected and filtered as explained by [9](p.p. 67) or avoided altogether by coordinating the sensors.

Wired triggers (Synchronous BUZZ), radio signals from a central unit (Constellation) or random signal generation on each sender (Asynchronous BUZZ and Cricket) are typically used to coordinate these sensors. In case of wired triggers, the drawback is the need for additional infrastructure (the trigger wires) and the constrained flexibility of such a system (i.e., this is not feasible for mobile sensors). In case of the randomization, the interferences are reduced but not removed. This reduction can be good enough for a small amount of active transmitters in the system but will fall short for larger systems. Even though the synchronization of the measurements via a radio signal solves these problems, it still shares the same weakness: the interoperability with similar systems that are outside of the control of this scheme.

Similar to the ToF sensors, as explained above, various media access schemes from the field of telecommunication have been evaluated and applied for ultrasonic based sensor systems. For example, the Dolphin system [10] uses a direct sequence code division multiple access (DS/CDMA) scheme for measurement coordination of ultrasonic sensors. [11] compare DS/CDMA and frequency hopping spread spectrum (FHSS) media access schemes for ultrasonic sensors and come to the conclusion that FHSS provides a higher accuracy of the measurement results in comparison to DS/CDMA.

Beside ultrasonic sensors, indoor location measurements are also realized by evaluating the received signal strength

indicators (RSSI). [12] describes a location measurement based on RSSI of IEEE 802.15.4 communication beacon messages. Lau et al. find that interferences from IEEE 802.11 traffic operating in the same frequency band cause a high degradation of the precision of the localization measurement due to lost IEEE 802.15.4 beacon messages. A solution for this problem is the scheduling of exclusive time slots for each network protocol.

Media access schemes are abundantly used in the field of telecommunication but can and are also applied to active sensors. Methods that utilize the signal modulation, e.g., CDMA or FDMA, allow the usage of multiple sensors without need for additional infrastructure. The drawback is that these methods either constrain the measurements (FDMA in case of ToF sensors) or do only reduce the interferences (CDMA or random signal activation, e.g., asynchronous BUZZ). Wired triggers, as central coordination method, limit the application scenarios of the system (e.g., it is harder to support mobile units). Using a centrally generated radio signal as trigger source can be a feasible solution, as long as all sensors are in range of the radio signal. If the measurement triggers cannot be derived directly from a central source a distributed clock synchronization is necessary to execute a TDMA schedule. This can be provided by a higher level clock synchronization protocol (e.g., IEEE 1588 or the network time protocol NTP) or is part of the communication network itself (e.g., EtherCAT or FlexRay). The constraints that are added to the system by the clock synchronization and methods to incorporate these in the schedule planning are explained and proposed in this paper.

A distributed system of four Time-of-Flight sensors is presented in Section III. This section introduces the setup of the sensor system, gives a short overview of the ToF sensor principle and derives an ideal TDMA schedule based on the core parameters of the system. Section IV defines the constraints imposed by the clock synchronization on a TDMA schedule. Section IV-E proposes an algorithm set that can be used to calculate these constraints and build an interference-free TDMA schedule that satisfies these constraints. A TDMA schedule in accordance with the calculated upper bound for the synchronization constraints and a supervisor for the synchronization precision have been implemented in the presented system and are described in Section V.

II. BUILDING A SCHEDULE FOR A DISTRIBUTED SENSOR SYSTEM

In a system that incorporates multiple active sensors it is necessary to control measurements to reduce or eliminate interference. The term active sensors refers to sensors that generate a signal to execute their measurements (e.g., ToF or ultrasonic sensors). These sensors are typically employed to monitor an object or an area of interest in real time. Therefore, data has to be generated by the sensors at a constant rate f_{meas} and delay between measurements. Thus, a static schedule cycle has to be built. Furthermore it is necessary to calculate the worst case execution time for each measurement T_{wct} to allocate a time slot in the schedule. The T_{wct} depends on the sensor type used and is further specified for a Time-of-Flight sensor in the following section. The utilization of the TDMA schedule is the ratio of the allocated time (a sum of all scheduled measurements per time unit) to the available time (e.g., 1 second). In an ideal system it is possible to utilize 100 % of the available time for measurements and to

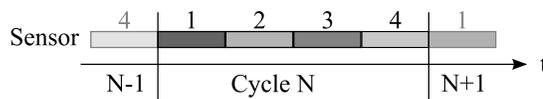


Figure 1: Ideal TDMA Schedule

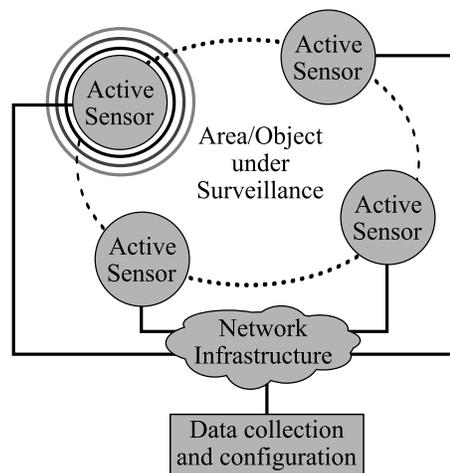


Figure 2: System Overview

execute the different measurements head-to-head as depicted in Figure 1.

III. A DISTRIBUTED SYSTEM OF TOF SENSORS

A system of four ToF sensor nodes of the type Sentis M100 from the company Bluetechnix Group GmbH [13] is used for demonstration as depicted in Figure 2. The sensor measurement configuration is the same for all four sensors. The nodes are connected by 100BASE-TX connections via a switch to a data collector workstation. The workstation is used to calculate the TDMA schedule, configure the trigger modules for each sensor that executes the schedule and to collect measurement data as presented in Section V.

A. Time-of-Flight Sensor Background

A ToF sensor generates a gray scale image of a scene with additional depth information for each pixel. It uses a modulated light source to illuminate the scene. This source is activated for a certain time, also referred to as integration time T_{int} . This typically ranges from $100 \mu s$ to $10 ms$. After this time of activity of the light source and sensor the gathered data has to be read from the sensor, which takes T_{ro} (in case of the Sentis M100, this is $1.35 ms$). During this time another ToF sensor can execute its measurement. However, this is only possible for small integration times ($T_{int} < T_{ro}$). The distance and gray scale data for each pixel can be calculated from a sufficient set of phase measurements N_{phase} (typically 4, [1]). Additional phase measurements can be included for each frame to reduce silicon specific asymmetries of the sensor pixel implementation [1], or to reduce aliasing errors caused by harmonics of the light source [14]. In addition for a worst case measurement time it is necessary to define an upper bound for the integration time, as this parameter can be varied between phase measurements. Therefore, the worst case execution time T_{wct} for one measurement can be calculated by the sum of the

different phase measurement times defined by the upper bound of the integration time T_{int} for each phase i and the read-out time for the sensor T_{ro} as depicted by (1). In case of the example with the Sentis M100 sensor, each phase measurement has the same integration time and read-out time, and thus, can be simplified to a multiplication of the phase measurement time with the amount of phase measurements N_{phase} as depicted by (2).

$$T_{wct} = \sum_{i=1}^{N_{phase}} T_{int_i} + T_{ro} \quad (1)$$

(1) can be simplified in case of the example system:

$$T_{wct} = N_{phase} \cdot (T_{int} + T_{ro}) \quad (2)$$

B. ToF Sensor System Configuration

For a schedulability analysis the total time T_{sys} that has to be allocated can be calculated by the sum of the measurement times for each sensor in the system. If all (N_{sen}) sensors share the same configuration regarding the measurement time T_{wct} and the frame rate f_{meas} this can be done by applying (3).

$$T_{sys} = N_{sen} \cdot T_{wct} \cdot f_{meas} \quad (3)$$

TABLE I: Example system configuration

Parameter	Value	Unit
N_{sen}	4	[1]
f_{meas}	20	[Hz], frames per second
T_{ro}	$1.35 \cdot 10^{-3}$	[s], sensor read-out time
T_{int}	$1.5 \cdot 10^{-3}$	[s], upper bound for the integration time
N_{phase}	4	[1], phase measurements per frame
T_{wct}	$11.4 \cdot 10^{-3}$	[s], (2) per frame measurement
T_{sys}	0.912	[s], (3) time to be allocated
T_{tot}	1	[s], available time

IV. SYNCHRONIZATION MECHANISM CONSTRAINTS

A. TDMA Schedule Problem

To execute a TDMA schedule in a distributed system a common notion of time is necessary. This is provided by a synchronization mechanism (e.g., IEEE 1588). This mechanism can be characterized by two parameters. The synchronization precision Π and the granularity of the global clock Δ_g . These parameters constrain the resource amount that can be allocated for measurements as depicted by (5), (6), and T_{loss} in Figure 3. In a physical system (Π and Δ_g greater 0 s) it is not possible to execute the TDMA schedule given in Figure 1 without interference between the different tasks.

It is common practice to distribute the spare time evenly between the tasks/measurements. For an interference free execution of measurements it is necessary that the spare time per measurement covers the time loss T_{loss} inflicted by the synchronization mechanism. For the example system the time that has to be allocated is $T_{sys} = 912 \text{ ms}$ per second (a utilization of 91.2%).

B. Synchronization Precision

The precision Π_i is the maximum difference of any two clocks of an ensemble at a specific point in time i . The precision of the ensemble (the synchronization precision) Π is the maximum of Π_i over an interval of interest [15, p. 56].

C. Granularity of the Synchronized Global Time

The granularity Δ_g of the global time is the period of the global clock and further the resolution for the triggering of actions (e.g., sensor measurements). Excess time after a scheduled measurement but before the next global clock tick cannot be used for a consecutive measurement.

D. Measurement Time Loss

The synchronization mechanism parameters have to be taken into account for building a TDMA schedule. T_{alloc} (4) depicts the worst case execution time T_{wct} ((1) or (2)) that has to be allocated per measurement with the addition of the synchronization precision to eliminate interference with consecutive measurements. T_{loss} (5) depicts the loss of measurement time due to synchronization precision and clock granularity for a specific measurement duration. The upper bound for the first term of this equation is Δ_g , and therefore, can be simplified to (6), which depicts the maximum loss per scheduled task in the system.

$$T_{alloc} = T_{wct} + \Pi \quad (4)$$

$$T_{loss} = \left(\left\lceil \frac{T_{alloc}}{\Delta_g} \right\rceil \cdot \Delta_g \right) - T_{alloc} + \Pi \quad (5)$$

$$T_{loss}^{max} = \Delta_g + \Pi \quad (6)$$

E. Algorithm Set for the Synchronization Mechanism Constraints

Different synchronization mechanisms and implementation strategies for these mechanisms provide different synchronization precisions and clock granularities. The following algorithm shows how the available slack of a given TDMA schedule can be distributed between these two parameters of the synchronization mechanism.

The time that is not utilized in the TDMA schedule is available as slack S_{tot} (7) and is given by the available time T_{tot} and the necessary time that has to be allocated T_{sys} . By dividing the total slack by the amount of scheduled measurements in the system, the slack budget available per scheduled task S_{task} (8) can be calculated. This can be used to cover the time losses due to the synchronization precision and clock granularity as given by (6).

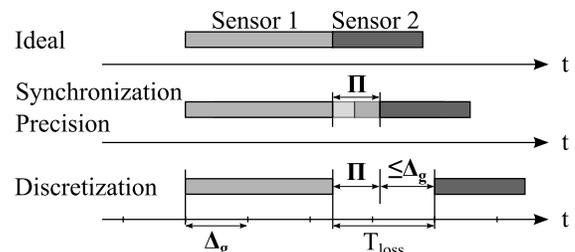


Figure 3: Loss of measurement time due to synchronization precision and clock granularity

$$S_{tot} = T_{tot} - T_{sys} \quad (7)$$

$$S_{task} = \frac{S_{tot}}{\sum_{i=1}^{N_{sen}} f_{meas_i}} \quad (8)$$

$$S_{task} \geq T_{loss}^{max} \quad (6) \quad (9)$$

$$S_{task} \geq \Delta_g + \Pi \quad (10)$$

Therefore, if one of the two parameters is known, the upper bound for the second parameter can be calculated.

F. Applying the Algorithm Set

Starting with the system configuration as given in Table I it is possible to find a set of synchronization parameters that is necessary to execute an interference free TDMA schedule, and therefore, the necessary implementation strategy for the synchronization mechanism. This can be done by setting either the clock granularity or synchronization precision to a fixed value and applying the algorithms defined in Section IV-E to find the corresponding upper bound for the other parameter. The total slack S_{tot} (7) and the available slack per scheduled task S_{task} (8) have to be calculated to solve the inequality (10) for the second synchronization parameter. In case of the example system, (8) can be simplified as all four sensors share the same configuration (f_{meas}).

$$S_{tot} = 1 - 0.912 = 88 \cdot 10^{-3} \text{ s}$$

$$S_{task} = \frac{S_{tot}}{N_{sen} \cdot f_{meas}}$$

$$S_{task} = \frac{88 \cdot 10^{-3}}{4 \cdot 20} = 1.1 \cdot 10^{-3} \text{ s}$$

In case of the example system, the available slack per scheduled task (measurement) is 1.1 ms . This can be distributed between the synchronization precision and clock granularity.

Variant 1: Fixed Synchronization Precision

If a legacy system is extended by a synchronization mechanism the type of implementation for this mechanism might be severely constrained. A range for the achievable synchronization precision is given by the type of implementation and the system characteristic (e.g., processor or network load). If the synchronization precision for a system has been measured or estimated it can be used to determine the necessary clock granularity. A software based IEEE 1588 implementation is capable of a synchronization precision Π below 1 ms .

$$\text{for } \Pi = 1 \cdot 10^{-3} \text{ s}$$

$$\Delta_g \leq S_{task} - \Pi = 1.1 \cdot 10^{-3} - 1 \cdot 10^{-3} =$$

$$= 0.1 \cdot 10^{-3} \text{ s}$$

$$\Delta_g \leq 100 \text{ } \mu\text{s}$$

In this case, the clock granularity Δ_g has to be at most $100 \text{ } \mu\text{s}$ to allow an interference free execution of the TDMA schedule in the example system.

Variant 2: Fixed Clock Granularity

Implementing a fine clock granularity might not be feasible for a system where processor execution time is in high demand. For example, a ToF sensor needs a high amount of execution time for processing the frame data. If a software based implementation of IEEE 1588 is used, a fine clock granularity would add higher processor load to the system. In case of the example system, a coarse granularity of $500 \text{ } \mu\text{s}$ has been chosen to reduce the additional processor load and the corresponding upper bound for the second parameter, i.e. the synchronization precision Π , is calculated.

$$\text{for } \Delta_g = 0.5 \cdot 10^{-3} \text{ s}$$

$$\Pi \leq S_{task} - \Delta_g = 1.1 \cdot 10^{-3} - 0.5 \cdot 10^{-3} =$$

$$= 0.6 \cdot 10^{-3} \text{ s}$$

$$\Pi \leq 600 \text{ } \mu\text{s}$$

A synchronization precision Π of $600 \text{ } \mu\text{s}$ can be reasonably realized by, e.g., a software based IEEE 1588 implementation. Whereas, a typical NTP implementation can achieve a precision down to a few milliseconds, which would not suffice in this case.

G. Choosing or Tuning the Synchronization Mechanism Implementation

With the calculated bounds for the synchronization mechanism parameters it is possible to choose the necessary type and implementation strategy for the system. For example, an IEEE 1588 clock synchronization has two main implementation strategies. A software and a hardware based implementation. In short, a hardware based approach can deliver a synchronization precision down to a few nanoseconds and typically provides a clock granularity of a few nanoseconds. Whereas a software based approach can deliver a synchronization precision in the range of a few microseconds to a few milliseconds. A software implementation of the clock typically provides a clock granularity in the range of $100 \text{ } \mu\text{s}$ to 1 ms . More information regarding these strategies and the achievable synchronization precisions is given in [16].

V. SUPERVISOR FOR THE SYNCHRONIZATION PRECISION

By building a TDMA schedule with respect to the clock synchronization constraints it is possible to supervise the system by comparing the current precision with the upper bound used to build the schedule. In the example system a local supervisor has been implemented on each sensor. A software version of IEEE 1588 has been implemented with a clock granularity of $500 \text{ } \mu\text{s}$. The TDMA schedule for the example system has been built with an upper bound for the synchronization precision of $600 \text{ } \mu\text{s}$ as calculated in variant 2 of Section IV-F.

The synchronization precision Π_i is the highest offset between two sensors in the system at a specific point in time, as defined in Section IV-B. The local supervisors do not have access to values that are generated outside the sensor. Thus, a local supervisor cannot calculate the actual synchronization precision as access to the offsets of the other sensors is not available. The local supervisor has access to the offset from the master (OFM) of the sensor. As shown in Figure 4, the OFM is symmetric around 0. Thus, it is reasonable to assume

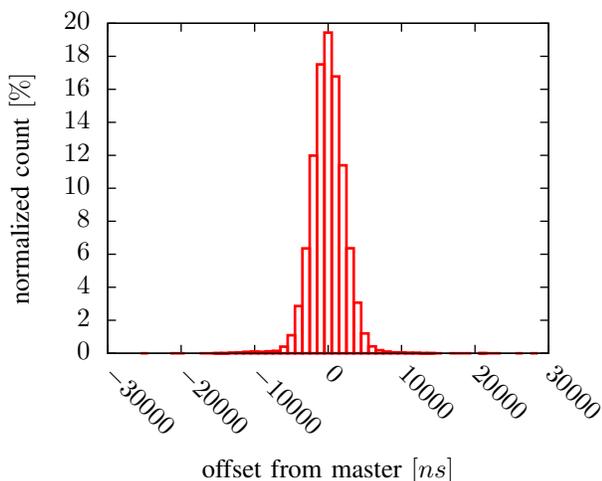


Figure 4: Histogram for the Offset from the Master

that the resulting worst case synchronization precision is two times the worst case OFM and the upper bound for the OFM can be set to $\frac{1}{2}$ of the upper bound for the synchronization precision, i.e. $300 \mu s$. If the OFM rises above this threshold the sensor supervisor disables the measurements for the sensor so that it will not interfere with other sensors.

The synchronization mechanism (in this example IEEE 1588) shares the communication network and the processor execution time with the sensor measurement. Therefore, sensor measurements have an impact on the synchronization mechanism as they need processor time and communication bandwidth.

Three different scenarios have been applied to the system. The internally calculated values for the OFM are transmitted over the communication network and are used offline to calculate the synchronization precision for the system.

Scenario 1: Minimal System Load

In this scenario, no ToF measurements are executed and the communication bandwidth and processor execution time is freely available for the synchronization mechanism. The synchronization precision over the elapsed time (1 hour) is depicted in Figure 5 on the left. A histogram for the synchronization precision is depicted on the right. The average synchronization precision is $3.157 \mu s$ and the worst is $28.763 \mu s$.

Scenario 2: Medium System Load

In this scenario, the ToF measurements are executed as given by the system configuration in Section III and the ToF data of two sensors is transmitted over the communication medium. This results in additional load both on the network and the processor side of the system. The measurement results are depicted in Figure 6. The average synchronization precision is $40 \mu s$ and the worst is $586 \mu s$, which is below the threshold of $600 \mu s$.

Scenario 3: Full System Load

In this scenario, the ToF measurements are executed as given by the system configuration in Section III and the ToF data of all sensors is transmitted over the communication

medium. The measurement results are depicted in Figure 7. The average synchronization precision is $65 \mu s$ and the worst is $618 \mu s$, which is above the threshold of $600 \mu s$.

Extrapolation of Scenario 2 and 3

The network load for scenario 3 is twice that from scenario 2. The average synchronization precision rose by 61% from scenario 2 to 3. Additional sensors, and therefore, additional network load will further deteriorate the synchronization precision. Thus, for a system with more than four sensors, either the clock granularity would have to be decreased, or the system slack increased to allow the execution of an interference free TDMA schedule. For the latter, the measurement parameters, i.e. the integration time T_{int} and the frame rate f_{meas} , would have to be reduced.

A. Comparison and Limitations

A non-TDMA based method, as presented by [17], shows a reduction of the impact of interference down to 1.5 % of the maximum measureable distance for a distributed ToF sensor system. In comparison, the TDMA based approach, taking the synchronization mechanism into account as proposed in Section IV-F, shows a standard deviation of the measurement results of 0.07 % of the maximum measureable distance. This is equal to the standard deviation of the measurement of a single sensor. The TDMA approach without a margin for the synchronization parameters, and therefore, a head-to-head scheduling of the sensors, results in a standard deviation of 0.6 %. The worst case scenario for the sensor system is a measurement of all sensors at the same time. This results in an erratic distribution of the measured distance values between 1 % and 1000 % of the actual distance.

By applying a TDMA schedule and the supervisor to the system it is possible to eliminate the interferences between the sensors, if the following points are accounted for:

- The limits for the precision of the synchronization mechanism. These have to be determined correctly (typically offline) to build a valid TDMA schedule by utilizing the proposed algorithm.
- The accuracy of the online synchronization precision measurement. This value will be used by the supervisor to check against the previously (offline) defined limits for the synchronization precision.

It is also necessary to consider how the supervisor has to interact with the system. It can either mark invalid measurements, or disable measurements if the measured precision deteriorates over a certain threshold. Thus, it has to be considered if a missing measurement might actually cause more harm to a system than a measurement with a certain error due to interference.

VI. CONCLUSION

As it has been shown in Section I, various sensor control methods have been implemented by research groups, ranging from simple central wired or wireless methods to typical media access schemes, like CDMA. The advantages and disadvantages of these methods for active sensors have been discussed. A TDMA scheme provides a control method that is decoupled from the measurement mechanism of active sensors but needs a clock synchronization mechanism in the background and a valid TDMA schedule for operation.

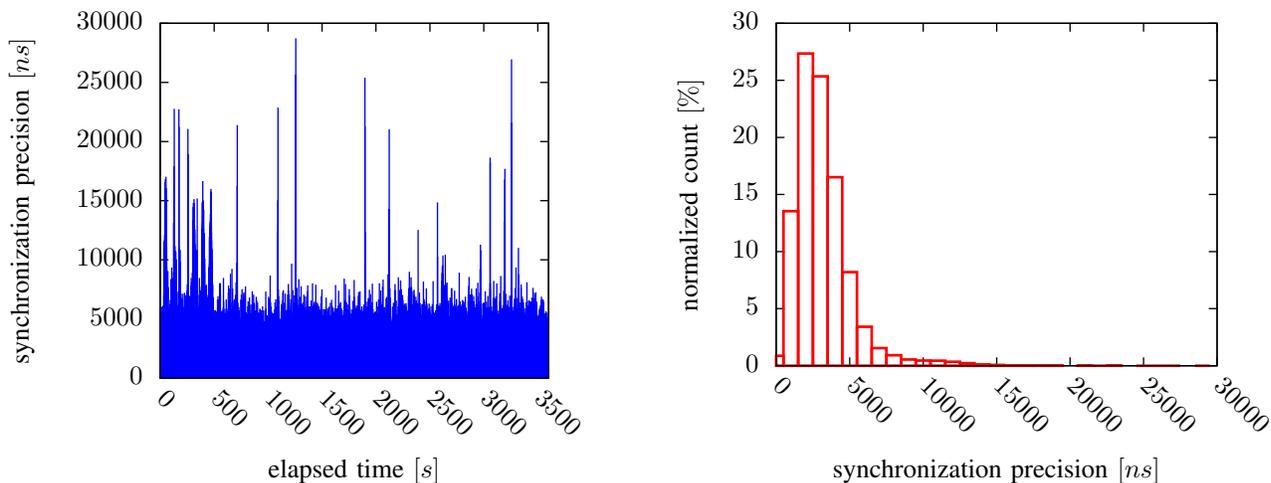


Figure 5: Scenario 1: Minimum System Load

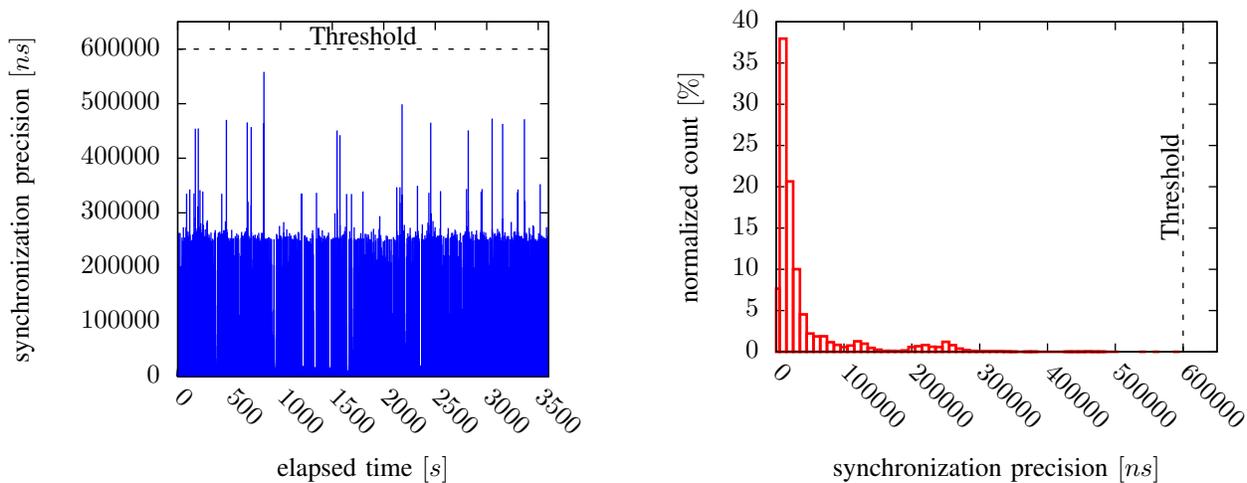


Figure 6: Scenario 2: Medium System Load

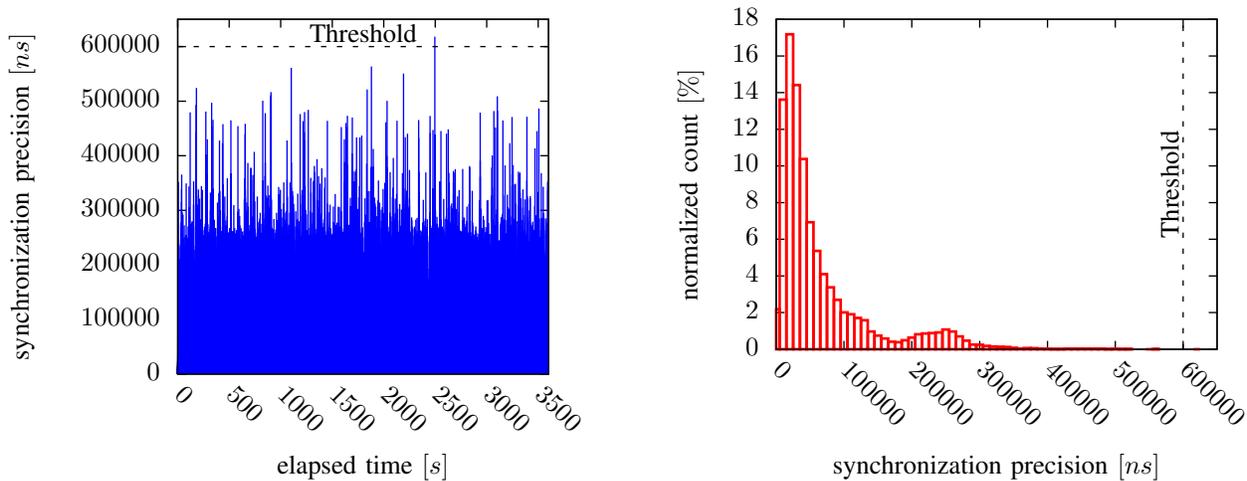


Figure 7: Scenario 3: Full System Load

In Subsection IV-A, it has been shown that not every schedule for a distributed sensor system is feasible if the underlying synchronization mechanism is not taken into account. The imposed synchronization mechanism constraints can be calculated, and thus, incorporated into the schedule planning to deliver a valid and executable TDMA schedule for the system. An algorithm set for this calculation has been proposed in Subsection IV-E.

This algorithm set can be used to set the bounds of the constraints of the synchronization mechanism for the system, especially the synchronization precision. Further, the synchronization precision can be monitored in the system and proper actions can be set, e.g., disabling task executions until the precision is below a specified threshold. This is especially interesting for a system of active sensors, for example, the ToF sensor system described in Section III, where it is necessary to eliminate interference of the scheduled measurements. A simple local supervisor has been implemented in the example system and measurements of the synchronization precision of an IEEE 1588 software implementation for different scenarios have been presented in Section V.

ACKNOWLEDGMENT

This work has been supported by Bluetechnix Group GmbH and by the public funded R&D project Josef Ressel Center for Verification of Embedded Computing Systems (VECS) managed by the CDG.

REFERENCES

[1] R. Kaufmann et al., "A time-of-flight line sensor: development and application," in Proceedings SPIE 5459, Optical Sensing, 2004, pp. 1–8.

[2] C. Bamji. Method and system to avoid inter-system interference for phase-based time-of-flight systems. US Patent 7,405,812. [Online]. Available: <http://www.google.com/patents/US7405812> [retrieved: Jun., 2015]

[3] B. Buettgen, M.-A. El Mechat, F. Lustenberger, and P. Seitz, "Pseudonoise optical modulation for real-time 3-d imaging with minimum interference," in Circuits and Systems I: Regular Papers, IEEE Transactions on (Volume:54 , Issue: 10), 2007, pp. 2109–2119.

[4] T. Oggier, B. Buettgen, and M. Schweizer. System and method for multi tof camera operation using phase hopping. US Patent App. 13/280,154. [Online]. Available: <https://www.google.com/patents/US20120098964> [retrieved: Jun., 2015]

[5] A. Dorrington, J. Godbaz, M. Cree, A. Payne, and L. Streeter, "Separating true range measurements from multi-path and scattering interference in commercial range cameras," in Proceedings SPIE 7864, Three-Dimensional Imaging, Interaction, and Measurement, 2011, pp. 1–10.

[6] A. Payne, A. Jongenelen, A. Dorrington, M. Cree, and D. Carnegie, "Multiple frequency range imaging to remove measurement ambiguity," in In Proceedings of at 9th Conference on Optical 3-D Measurement Techniques, Vienna, Austria, 2009, pp. 139–148.

[7] E. Foxlin, M. Harrington, and G. Pfeifer, "Constellation: A wide-range wireless motion-tracking system for augmented reality and virtual set applications," in Proc. ACM Conf. on Computer Graphics and Interactive Techniques (SIGGRAPH '98), 1998, pp. 371–378.

[8] M. R. McCarthy, "The buzz: Narrowband ultrasonic positioning for wearable computers," Ph.D. dissertation, University of Bristol, 2007.

[9] N. B. Priyantha, "The cricket indoor location system," Ph.D. dissertation, Massachusetts Institute of Technology, 2005.

[10] M. Hazas and A. Hopper, "Broadband ultrasonic location systems for improved indoor positioning," Mobile Computing, IEEE Transactions on, vol. 5, no. 5, May 2006, pp. 536–547.

[11] J. Gonzalez Hernandez and C. Bleakley, "High-precision robust broadband ultrasonic location and orientation estimation," Selected Topics in Signal Processing, IEEE Journal of, vol. 3, no. 5, Oct 2009, pp. 832–844.

[12] S.-Y. Lau, T.-H. Lin, T.-Y. Huang, I.-H. Ng, and P. Huang, "A measurement study of zigbee-based indoor localization systems under rf interference," in Proceedings of the 4th ACM International Workshop on Experimental Evaluation and Characterization, ser. WINTECH '09. New York, NY, USA: ACM, 2009, pp. 35–42. [Online]. Available: <http://doi.acm.org/10.1145/1614293.1614300>

[13] Sentis m100 tof smart sensor. internet. Bluetechnix Group GmbH. <http://www.bluetechnix.com/de/products/depthsensing/>. [Online]. Available: <http://www.bluetechnix.com/de/products/depthsensing/> [retrieved: Jun., 2015]

[14] J. P. Godbaz, M. J. Cree, and A. A. Dorrington, "Understanding and ameliorating non-linear phase and amplitude responses in amcw lidar," Remote Sensing, vol. 4, no. 1, 2011, pp. 21–42.

[15] H. Kopetz, Real-Time Systems, 2nd ed. Springer, 2011.

[16] A. Puhm, M. Kramer, P. Moosbrugger, and M. Horauer, "Problems and solutions for refitting a sensor network with ieee1588 clock synchronisation," in Proceedings of the 19th IEEE International Conference on Emerging Technologies and Factory Automation (ETFA'2014), 2014, pp. 1–7.

[17] B. Buettgen and P. Seitz, "Robust optical time-of-flight range imaging based on smart pixel structures," in Circuits and Systems I: Regular Papers, IEEE Transactions on (Volume:55 , Issue: 6), 2008, pp. 1512–1525.

Self-Stabilizing Structures for Data Gathering in Wireless Sensor Networks

Sandra Beyer*, Stefan Lohs*, Jörg Nolte*, Reinhardt Karnapke[†] and Gerry Siegemund[‡]

*Distributed Systems/Operating Systems Group, BTU Cottbus-Senftenberg

email: sandra.beyer.sb@gmail.com, {slohs, jon}@informatik.tu-cottbus.de

[†]Communication and Operating Systems Group, Technische Universität Berlin

email: karnapke@tu-berlin.de

[‡]Institute of Telematics, Hamburg University of Technology

email: gerry.siegemund@tu-harburg.de

Abstract—Wireless Sensor Networks (WSN) enable a number of applications, with monitoring of habitats, office buildings, or restricted areas most prominent among them. All of these applications have one thing in common: the need to communicate. However, the nature of the wireless medium results in quite a few problems. Lossy communication links with transient faults require acknowledgments, retransmissions, and route repair mechanisms. Tree- or similar structures for data gathering scenarios lead to increased load closer to the sink, with congestion, higher buffer space requirements, and energy drain as results. The second problem is often addressed by aggregation and reduction schemes. These schemes are bound to fail, however, when the underlying structure is compromised due to changes in the connectivity between nodes. Therefore, it is necessary to focus on the structures first of all. We address the problem of transient faults by using the inherent fault tolerance of self-stabilizing algorithms when building and using tree- or tiers (communication-) structures. In this paper we show that self-stabilizing structures are suitable for data gathering scenarios in WSN by comparison of the connectivity achieved by our self-stabilizing tiers algorithm and the tree algorithm from Dolev with that of Collection Tree Protocol (CTP), the standard data-gathering protocol for TinyOS.

Keywords—Wireless Sensor Networks; Self-Stabilization; Routing Structures

I. INTRODUCTION

Wireless Sensor Networks have gained a lot of attention in the research community in the last decade. Application scenarios include for example monitoring of habitats, intrusion detection and house control. This strong interest in sensor networks stems from the fact that they are inexpensive, autonomous systems that have to adapt to ever changing conditions. Due to the fact that sensor networks should be inexpensive, the individual sensor nodes are usually not very powerful but can cooperate to solve complex tasks. In some applications, the nodes need to operate autonomously for many years, because they are deployed in hard to reach areas and human interaction is restricted to the absolute minimum. The restrictions imposed by pricing and form factor result in the need for special protocols, as standard protocols induce too much (computational, communication or memory) overhead.

Sensor nodes are usually powered by batteries, energy harvesting is only rarely possible. Therefore, energy is an important resource and must be conserved as much as possible. In combination with the price factor, nodes should often be as cheap as possible, this leads to the usage of transceivers that require only little energy but can also only transmit over

short distances. This in turn makes the usage of multi-hop communication protocols necessary. Radio communication is always error prone due to the shared medium, collisions and environmental influences. The fact that most sensor nodes only feature a cheap transceiver intensifies this problem. To enable the sensor networks to perform their duties in spite of these errors, repair mechanisms are included. These usually encompass forward error correction, acknowledgments, and retransmissions. However, detecting the errors and storing messages for retransmission increases memory consumption (flash and RAM). The amount required can be arbitrarily large, depending on the number of different kinds of errors that should be detected and the correction mechanisms. Also, the error handling code itself might introduce a new source of errors. To overcome this problem, we propose the usage of self-stabilization, which describes only the 'good' states of the sensor nodes. Any other state is an error state and needs to be changed. This way, there is no need for the programmer to examine all possible error-states. We show that our self-stabilizing protocols are less complex and the memory consumption is reduced when compared to traditional protocols.

The most common applications for sensor networks feature data gathering scenarios, in which all nodes gather their data and transmit it to a central sink from which a user can retrieve it. In a multi-hop environment, the nodes need to forward more messages the closer they are to the sink. This increased communication load leads to congestion, a higher rate of collisions and increased RAM consumption on the nodes close to the sink as they need to store more messages. Also, the need to communicate more results in a change in duty cycle, as those nodes need to stay awake longer, which in turn leads to a higher energy consumption and early node failure.

A common way to deal with this problem is to use aggregation and/or reduction mechanisms on the sensor nodes. Aggregations reduces the number of messages, while reductions reduce their size. This can be useful even for applications with long intervals between message generation where congestion is not a problem, as it also reduces the consumed transmission energy. In order for the aggregation and reduction to be effective, the underlying communication structure needs to operate correctly. Without aggregation and reduction, a lost message results in a single lost value. With aggregation and reduction, this could mean the loss of the data from a whole subnet. In this paper we focus on mechanisms that can be used to build structures that lie underneath the aggregations and

reductions. To build and, most of all, sustain these structures, we rely on self-stabilizing algorithms.

This paper is structured as follows: Related work is presented in Section II, followed by an introduction to self-stabilization in Section III. Our self-stabilizing tiers algorithm (SelfTIER) and the self-stabilizing tree algorithm from Dolev [1] (SelfTREE) are discussed in Section IV and evaluated in comparison with CTP in Section V. We finish with a conclusion in Section VI.

II. AGGREGATION IN WSN

In literature, several approaches for in-network aggregation are presented. All have in common that they first establish a special communication structure and afterward use it for a certain time. The structure is responsible for deciding which node needs to aggregate the data from which other nodes and which path the aggregated data takes on its way to the sink. Also, the times for aggregations are defined. When a node wants to forward the aggregated values, it first needs to wait for all its children to transmit their data. The structure can also include information about deadlines and whether or not all messages from children have arrived. Most aggregation algorithms focus on the error free case and describe the aggregation itself, without offering too much details about the underlying structure and its repair mechanisms. In the following we take a look at three categories of aggregation structures, namely tree based, cluster based, and multi path. Others approaches, which fall into neither or multiple of the categories mentioned above, are also discussed briefly.

A. Tree Based

A tree is the most commonly used routing scheme for aggregation in sensor networks. The data is routed from source nodes to a sink node. Intermediate nodes (parent) are able to collect the data of well defined children and apply a fusion function before forwarding data to the next hop. Using the well organized aggregation structure, each node only forwards one packet per aggregation round, which avoids high network loads and conserves energy. The drawback is that if one packet is lost, the data of the whole sub-tree is lost.

An aggregation scheme for monitoring applications is Tiny AGgregation (TAG) [2]. TAG uses a two phase approach. First, the sink sends a query message to build an aggregation tree. All nodes receiving this message adapt their level (the distance to the sink) and select the sender as parent node. In the second phase, the collection phase, the tree is used to aggregate data and forward it to the sink node. TAG uses a *per-hop aggregation* approach [3], each parent has to wait until all children have sent their data before forwarding the aggregated data. Each round is divided into slots, the number of slots equals the height of the tree. This way, timeouts can be calculated, in case a parent does not receive the messages from all its children. The query message is sent periodically to recover the tree after dynamic link changes.

There are several other approaches that can be used to build aggregation trees, which focus on, e.g., energy consumption, path reliability and/or mobility of nodes [4]-[8].

B. Cluster Based

The second group of schemes are cluster based aggregation schemes. Like tree based algorithms, cluster based algorithms

also organize the network in a hierarchical manner. Each selected cluster head is responsible for aggregating the data of its cluster members and forwarding it to a sink node.

LEACH [9] is one member of this group. LEACH is divided into a setup phase to organize the clusters and a steady-state phase to send the data to the sink. The algorithm is distributed. At first, cluster heads are selected by a probabilistic approach. After that, the cluster head computes a TDMA (Time Division Multiple Access) scheme to avoid collisions in its cluster. Cluster heads send their aggregated data with one single transmission to the sink.

Other members of this group include COUGAR [10] and DRINA [11]. They use different mechanisms to elect the cluster head, to reduce communication overhead or select nodes by different suitable metrics.

C. Multi Path

The idea of multi path approaches is to increase the tolerance against link changes by using multiple paths to the sink node. If a link breaks or one packet is lost because of a collision, the data can be restored using packets that traveled an alternative path.

A member of this group is Synopsis Diffusion [12]. Synopsis Diffusion uses a tier structure as the underlying topology, where each node forwards all messages it receives from nodes on a higher tier. The advantage of this approach is that data is forwarded on multiple paths to the sink node but it also introduces the problem of *duplicate sensitive* aggregation. Nodes may receive duplicates of data which may affect the result of the aggregation. This problem must be solved by the use of suitable aggregation functions. Nevertheless, this approach is more appropriate in harsh environments with a high rate of link breaks and message loss.

D. Other Approaches

Several other approaches exist, whose underlying topologies do not match the previous categories. One Example is the chain based structure of PEGASIS [13], which is focused on conserving energy. Another example is the hybrid approach of Tributaries and Deltas [14], which combines the advantages of tree- and multi path based schemes.

E. Summary

All schemes have in common that they are divided into a setup- and a collection phase. Often the underlying algorithm that is used to establish a structure is exchangeable or can be extended by additional repair mechanisms to increase fault tolerance, periodic rebuilds are the most often suggested approach.

III. SELF-STABILIZATION

The concept of self-stabilizing algorithms was first introduced by Dijkstra in his paper "Self-Stabilizing Systems in Spite of Distributed Control" [15]. He described a network of several processors having a set of registers. Each processor has a so called local view, which consist of the registers of its direct neighbors.

A self-stabilizing algorithm consists of a set of rules in the form *guard* \rightarrow *assignment*. If the guard predicate of a rule is resolved to *true*, the rule is called *enabled* and

```
1 NodeID parent;
2 Integer level;
```

Figure 1. State of the SelfTREE algorithm

the assignment part *may* be executed. Every processor checks (locally) whether any of the rules is enabled, based on its local view. There are two properties necessary for an algorithm to be self-stabilizing: *Convergence* and *Closure*. *Convergence* means that the system will reach a defined (stable) state within finite time, while *Closure* means the execution of any rule will never take the system from a stable state to an unstable one.

In wireless sensor networks, each node represents one of Dijkstra’s processors. The local view of a processor then corresponds to the information from that node and all its neighboring nodes, meaning that the state of a node needs to be communicated. Nodes that receive such a state message may change their own state due to the execution of rules, based on the changes in the state of the neighbor. As state changes propagate through the network, the whole system eventually reaches a stable state, making the system inherently tolerant against transient faults. Exchanging the state messages can make good use of the broadcast character of the medium. Another advantage of the self-stabilizing approach is that the algorithm (eventually) turns the system into a stable state from any given state, meaning that no code for initialization is required.

IV. SELFTREE AND SELFTIER

In this section we describe two self-stabilizing algorithms for a data gathering scenario. Each node of the sensor network has a unique identifier. The goal is to build and maintain a routing topology where all nodes can send their measured data to one distinct node, i.e., the sink. Of course, nodes that are separated from the network will not find a route. The first algorithm builds a minimum spanning tree (SelfTREE) published by Dolev [1] in 2000. The second algorithm is our self-stabilizing tiers algorithm (SelfTIER). Both algorithms consist of only two rules. The first rule of each algorithm is executed by the node which requested the aggregated data (the sink). The second rule is run by all other nodes.

A. SelfTREE

In the SelfTREE algorithm, the state of a node consists of the ID of its parent node and its distance from the sink (Listing 1). Please note that both values can start with any arbitrary value. The state of a node is propagated to its neighbors on a regular basis, and may result in enabled rules on the receiving nodes, which might also lead to state changes.

The first rule of SelfTREE applies only to the sink node. It sets the parent ID to an invalid value and the level as 0. The second rule is applied to all non-sink nodes. Based on the local view, the node checks the distance from the sink of all its neighbors. If the lowest distance from the sink is lower than its own distance minus one, a shorter path to the sink seems to exist. Therefore, the node sets its parent ID to the node with the minimum distance and its own distance as that nodes distance incremented by one. Please note that if there

```
1 Integer level;
```

Figure 2. State of the SelfTIER algorithm

are multiple candidates with the minimum distance, the new parent can be any arbitrary one of them.

A node is in a stable state if it has correct knowledge of the distance to the sink and a correct parent node. A global stable state/ stable system state is reached when all nodes are in a stable state.

When data transmission starts, each node knows its parent node and addresses its message to that parent. Nodes only forward messages in which they are listed as next hop.

B. SelfTIER

In the SelfTIER algorithm, only the distance to the sink is required and included in the state of a node (Listing 2). As in all self-stabilizing algorithms, the state of a node is shared with direct neighbors on a regular basis and may lead to changes on the receiving nodes.

The two rules of the SelfTIER algorithm are fairly similar to those of the SelfTREE algorithm. Rule one is only applied to the sink and sets its level to 0. The second rule is only applied to non-sink nodes. Each node checks the levels of its neighbors and sets its own level to the minimum incremented by one.

In such a tier based structure, only the distance to the sink is relevant for routing decisions. When a node receives a data message, it checks from which tier the message was transmitted. If the tier is higher, i.e., the distance to the sink was higher, the message is forwarded. Otherwise, the message is either from a node on the same level and will be forwarded on the next lower level, or it is from a node that is closer to the sink. Either way, it must not be forwarded. This way, all nodes that are closer to the sink than the transmitting node forward the message, resulting in multiple redundant paths taken by the same message. This increases robustness, but also increases network load and gives rise to a potential problem with duplicates. Depending on the aggregation/reduction scheme that is used on top of the tier structure, this can be ignored (e.g. reduction function: maximum) or must be detected (e.g. reduction function: add).

The big advantage of the tier based algorithm is its redundancy. If the link between a node and its parent breaks in the tree structure, data from this node and all its children will be unavailable until the structure is rebuilt. In the tiers structure, an arbitrary number of links may break without affecting the forwarding of messages, as long as there is still at least one node available in a lower tier for each node.

V. EVALUATION

Both self-stabilizing algorithms have been integrated into the TOLERANCEZONE middleware. It provides a neighborhood discovery protocol and chooses bidirectional neighbors for each node. Also, it takes care of the regular propagation of a nodes state to its neighbors. To reduce the network load, the state and neighborhood messages are transmitted together. The speed with which an algorithm can react to changes in

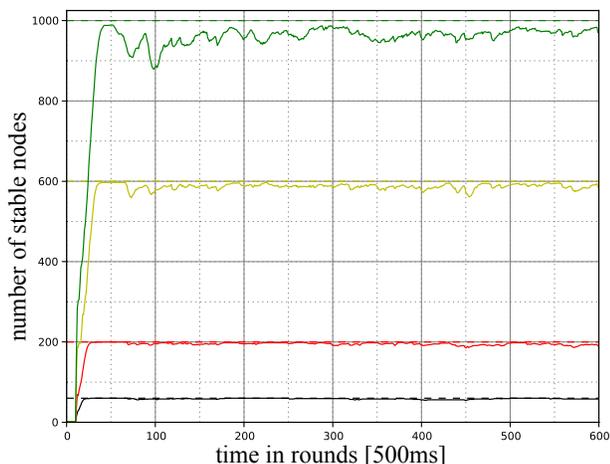


Figure 3. Structure state over time. SelfTIER algorithm with varying number of nodes (60, 200, 600, 1000).

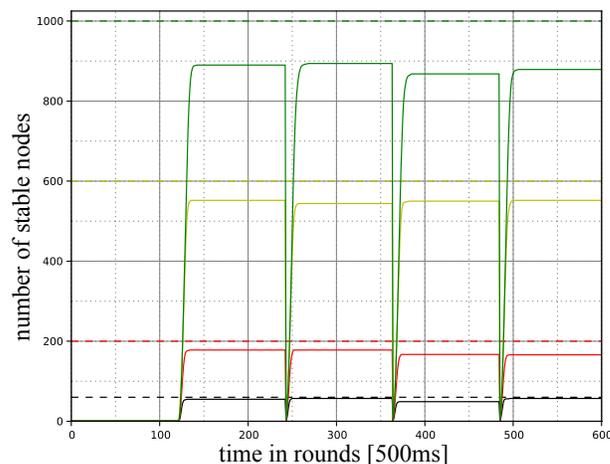


Figure 4. Structure state over time. Ordinary tiers algorithm with varying number of nodes (60, 200, 600, 1000).

a nodes neighborhood depends on the frequency of message transmission. As changes may need to propagate through the whole network, reaching a stable state after a change occurred may take $\mathcal{O}(maxLevel)$ rounds.

The goal of our self-stabilizing algorithms is to increase the tolerance against transient faults. When a transient fault occurs, the system is in a non-stable state and the underlying routing topology must be repaired. In case of self-stabilization, this is done autonomously, in the other cases we have to reconstruct or repair our structure.

Stability

In our first evaluation we compared the SelfTIER algorithm with an ordinary algorithm as proposed in TAG. In the ordinary approach, the sink periodically floods a beacon packet to establish a valid tier structure. Both algorithms were implemented in C++ for the REFLEX operating system [16]. The SelfTIER algorithm is part of our TOLERANCEZONE middleware.

In the first part of the evaluation we measured the stability of the network. In simulations using the discrete event simulator OMNeT++ [17], we ran both algorithms in a grid topology with 60, 200, 600, and 1000 nodes. Each second, a snapshot of the network was taken and the amount of stable nodes was counted. For the SelfTIER algorithm, a node was counted as stable if no rule was enabled. In case of the ordinary tier algorithm, it is counted as stable if it received and forwarded the beacon. To simulate transient faults we introduced link breaks. In case of the network consisting of 1000 nodes, one node lost connections for 30 seconds every second, after this time the connection was reestablished. The probability of a fault at a specific node was the same for all simulated network sizes.

Figure 3 and Figure 4 show the result of our simulations for SelfTIER and the ordinary algorithm respectively. As Figure 3 shows, the SelfTIER algorithm reaches nearly 95% stability for all four different network sizes in spite of the transient faults, while the ordinary algorithm reaches at most 90%.

In the case of the ordinary algorithm, a link break during the construction of the structure results in loss of the construction beacon. Then, the corresponding node is not part of

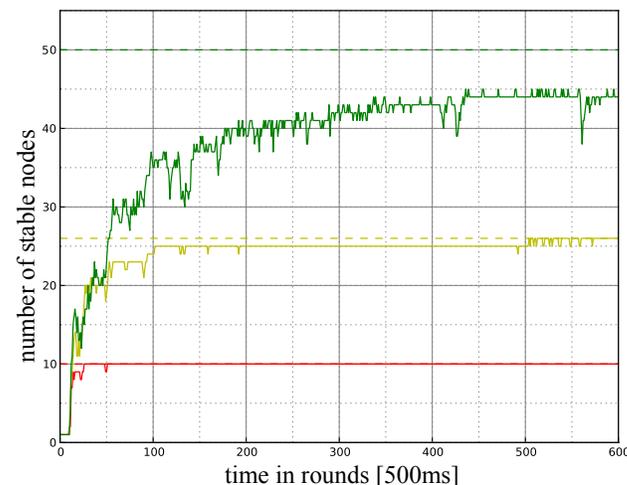


Figure 5. Real world deployment. Structure state over time. SelfTIER algorithm with varying number of nodes (10,26,50).

the structure until it receives the next beacon. In the case of the self-stabilizing algorithm, the node automatically rejoins the tiers structure once the link break is over and the node is connected to its neighbor(s), again.

After the simulations, we also ran experiments with real sensor nodes. We used EZ-Chronos 430 sensor nodes from Texas Instruments and placed them in grids of different sizes on the ground with a distance of 3 meters between nodes. The grids always consisted of a square of nodes plus the sink node, resulting in 10 (3x3 +1), 26 (5x5 +1) and 50 (7x7 +1) nodes used in the experiments. In real world experiments, inducing errors is not easy but also not necessary, as the wireless channel is lossy enough on its own. We measured the stability of the structures generated by both algorithms, using the same nodes and the same placement of nodes to keep the results comparable.

Figures 5 and 6 show the stability of the routing structure measured for 3 minutes after the (re-)start of the network. When the two Figures are compared, it can be seen that the

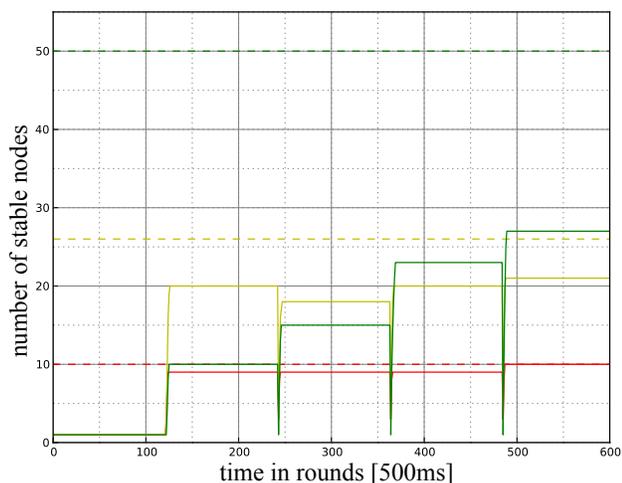


Figure 6. Real world deployment. Structure state over time. Ordinary tier algorithm with varying number of nodes (10,26,50).

size of the network has a strong influence, especially on the ordinary algorithm. Both algorithms perform similar in the network consisting of only ten nodes. SelfTIER can reach a stable structure for all ten nodes almost immediately but the ordinary algorithm reaches nine stable nodes after a rebuild and ten near the end of the three minutes. This could be compensated by allowing for a certain setup time. However, the differences increase when the network size is increased. In the network consisting of 26 nodes, SelfTIER still reaches 95% stability while the ordinary algorithm reaches only about 84%. For the network consisting of 50 nodes it even falls below 60%, while SelfTIER still reaches nearly 90%. As the number of additional nodes that are stable does not rise much for the ordinary algorithm, we expect that there is a certain threshold, above which adding more nodes to the network will not increase the number of stable nodes anymore, for the ordinary algorithm. For the self-stabilizing SelfTIER, however, adding nodes will delay the point in time when most nodes are stable, but eventually this point will be reached, making SelfTIER the algorithm of choice for larger networks.

Connectivity

The second part of our evaluation concerns the quality of the structure. To achieve good aggregation results, as many nodes as possible must know a valid route to the sink. In a data gathering scenario, the data is forwarded hop by hop from all nodes to the single sink. To evaluate the connectivity of the routing structure, we use snapshots of the state of the routing protocol, in addition to the entries in the neighbor tables and the physical layer. In the simulations using OMNeT++, we use the current links between the NIC-modules.

As competitor to the self-stabilizing algorithms, we used a REFLEX implementation of the standard TinyOS data gathering protocol CTP.

CTP uses beacons, which are transmitted by each node regularly, to build a tree structure. After an initial buildup phase the frequency of beacons is reduced. Instead, the quality of connections is measured using acknowledgment messages. To compensate the lack of application messages

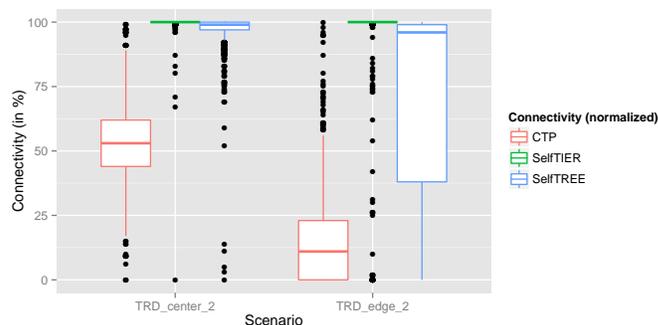


Figure 7. Connectivity of 226 nodes grid. Link disconnect fault injection. Sink node at center (left) and edge (right).

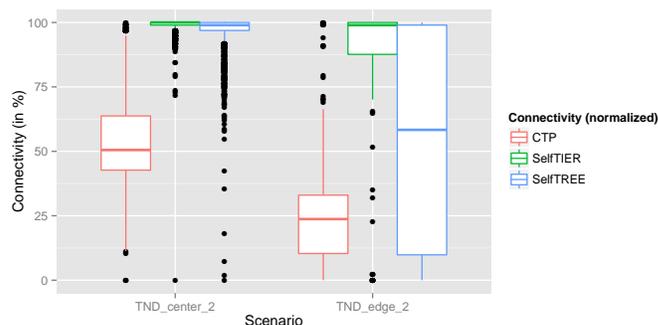


Figure 8. Connectivity of 226 nodes grid. Node disconnect fault injection. Sink node at center (left) and edge (right).

and acknowledgments in our simulations, we kept a constant beacon interval of 500ms, which equals the frequency of status message transmission for the self-stabilizing algorithms. We ran OMNeT simulations with up to 226 (15x15 +1) nodes in a grid topology.

The connectivity achieved by all three algorithms when link failures are injected can be seen in Figure 7. In this scenario, one link is disconnected every second, and remains disconnected for 30 seconds. The Figure is divided according to the placement of the sink, either at the center of the grid (left) or at the corner (right). Connectivity is achieved by a node, if it has a path to the sink (routing table), the next hop is in its neighbor table and the physical connection to that neighbor exists. The Figure shows that a placement of the sink in the middle of the network is better for CTP, which then reaches about 50% connectivity. For this placement, there is no big difference between the SelfTREE algorithm from Dolev and our SelfTIER, both reach nearly 100%. When a sink placement at the corner is evaluated, the performance of CTP drops drastically, to roundabout 12%. The results of SelfTREE are spread a little further, but it still achieves more than 95% connectivity. The results from SelfTIER stay at nearly 100% connectivity.

Figure 8 shows the results the three algorithms achieved for the scenario in which complete node disconnection faults were injected. Every 5 seconds, a node is disconnected from the network for 30 seconds. The Figure shows the normalized result with respect to the physical conditions. A node is counted as connected, if there is a valid entry for a path from

the node to the sink in the routing- and neighbor table on the node, and the used physical links exists. The scenario is further divided according to the placement of the sink node: Either at the center of the grid or at the corner. Placing the sink at the corner naturally doubles path length, compared to the placement at the center. When the sink is placed at the center (Figure 8, left side), the median of reached connectivity for CTP is 50%. Both self-stabilizing algorithms, the SelfTREE algorithm from Dolev and our SelfTIER algorithm, reach nearly 100 % connectivity. When the sink is located at the edge of the network, the performance of CTP and SelfTREE decrease to 25% and 55% respectively, while SelfTIER still offers nearly 100% connectivity. This is due to the fact that SelfTIER used multiple paths to deliver messages. While the failure of one node results in lost connectivity for a whole sub-tree in the tree based algorithms, SelfTIER can usually connect the nodes from lower tiers through different nodes, so only the node on which the failure occurs is 'lost'.

In summary, it can be said that self-stabilizing algorithms are a good choice when building structures for aggregation and/or reduction schemes, with our SelfTIER algorithm outperforming CTP and the SelfTREE algorithm from Dolev when not only links break, but nodes can also fail completely for a certain time.

VI. CONCLUSION

In this paper, we have discussed the prerequisites for a successful aggregation/reduction scheme for data gathering scenarios in wireless sensor networks, namely a robust underlying communication structure. We have argued that self-stabilizing algorithms should be the method of choice, and support our claim with simulations and experiments. In the evaluation we measured stability and connectivity for different structures and network sizes. We compared two different self-stabilizing algorithms with CTP, the standard protocol for data gathering scenarios from TinyOS. The results show that there are performance differences between the two self-stabilizing algorithms that depend on the error scenario and the network diameter, but both protocols always perform better than CTP.

Often, self-stabilizing algorithms are criticized for the higher energy consumption due to the periodic state exchange between nodes. How these exchanges can be reduced without sacrificing too much of the advantages of self-stabilization is currently being evaluated by our group. Sleeping times/duty cycling and other traditional approaches for energy conservation are being discussed, and first evaluations show promising results.

ACKNOWLEDGMENT

This work was partially funded by the Deutsche Forschungsgemeinschaft DFG in the project ToleranceZone (DFG NO 625/6-1).

REFERENCES

[1] S. Dolev, Self-stabilization. MIT press, 2000.

[2] S. Madden, M. J. Franklin, J. M. Hellerstein, and W. Hong, "Tag: A tiny aggregation service for ad-hoc sensor networks," SIGOPS Oper. Syst. Rev., vol. 36, no. SI, Dec. 2002, pp. 131–146. [Online]. Available: <http://doi.acm.org/10.1145/844128.844142>

[3] I. Solis and K. Obraczka, "The impact of timing in data aggregation for sensor networks," in Communications, 2004 IEEE International Conference on, vol. 6, June 2004, pp. 3640–3645 Vol.6.

[4] M. Ding, X. Cheng, and G. Xue, "Aggregation tree construction in sensor networks," in Vehicular Technology Conference, 2003. VTC 2003-Fall. 2003 IEEE 58th, vol. 4, Oct 2003, pp. 2168–2172 Vol.4.

[5] H. S. Kim, T. F. Abdelzaher, and W. H. Kwon, "Minimum-energy asynchronous dissemination to mobile sinks in wireless sensor networks," in Proceedings of the 1st International Conference on Embedded Networked Sensor Systems, ser. SenSys '03. New York, NY, USA: ACM, 2003, pp. 193–204. [Online]. Available: <http://doi.acm.org/10.1145/958491.958515>

[6] T. Banerjee, K. Chowdhury, and D. Agrawal, "Tree based data aggregation in sensor networks using polynomial regression," in Information Fusion, 2005 8th International Conference on, vol. 2, July 2005, pp. 8 pp.–.

[7] O. Gnawali, R. Fonseca, K. Jamieson, D. Moss, and P. Levis, "Collection tree protocol," in Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems, ser. SenSys '09. New York, NY, USA: ACM, 2009, pp. 1–14. [Online]. Available: <http://doi.acm.org/10.1145/1644038.1644040>

[8] H. O. Tan and I. Körpeoğlu, "Power efficient data gathering and aggregation in wireless sensor networks," SIGMOD Rec., vol. 32, no. 4, Dec. 2003, pp. 66–71. [Online]. Available: <http://doi.acm.org/10.1145/959060.959072>

[9] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "An application-specific protocol architecture for wireless microsensor networks," Wireless Communications, IEEE Transactions on, vol. 1, no. 4, Oct 2002, pp. 660–670.

[10] Y. Yao and J. Gehrke, "The cougar approach to in-network query processing in sensor networks," SIGMOD Rec., vol. 31, no. 3, Sep. 2002, pp. 9–18. [Online]. Available: <http://doi.acm.org/10.1145/601858.601861>

[11] L. Villas, A. Boukerche, H. Ramos, H. de Oliveira, R. de Araujo, and A. Loureiro, "Drina: A lightweight and reliable routing approach for in-network aggregation in wireless sensor networks," Computers, IEEE Transactions on, vol. 62, no. 4, April 2013, pp. 676–689.

[12] S. Nath, P. B. Gibbons, S. Seshan, and Z. R. Anderson, "Synopsis diffusion for robust aggregation in sensor networks," in Proceedings of the 2Nd International Conference on Embedded Networked Sensor Systems, ser. SenSys '04. New York, NY, USA: ACM, 2004, pp. 250–262. [Online]. Available: <http://doi.acm.org/10.1145/1031495.1031525>

[13] S. Lindsey, C. Raghavendra, and K. Sivalingam, "Data gathering algorithms in sensor networks using energy metrics," Parallel and Distributed Systems, IEEE Transactions on, vol. 13, no. 9, Sep 2002, pp. 924–935.

[14] A. Manjhi, S. Nath, and P. B. Gibbons, "Tributaries and deltas: Efficient and robust aggregation in sensor network streams," in Proceedings of the 2005 ACM SIGMOD International Conference on Management of Data, ser. SIGMOD '05. New York, NY, USA: ACM, 2005, pp. 287–298. [Online]. Available: <http://doi.acm.org/10.1145/1066157.1066191>

[15] E. W. Dijkstra, "Self-stabilizing systems in spite of distributed control," Commun. ACM, vol. 17, no. 11, Nov. 1974, pp. 643–644. [Online]. Available: <http://doi.acm.org/10.1145/361179.361202>

[16] K. Walther, R. Karnapke, and J. Nolte, "An existing complete house control system based on the reflex operating system: Implementation and experiences over a period of 4 years," in Proceedings of 13th IEEE Conference on Emerging Technologies and Factory Automation, 2008.

[17] A. Varga, "The omnet++ discrete event simulation system," in Proceedings of the European Simulation Multiconference (ESM'2001), Prague, Czech Republic, Jun. 2001.

Wireless Sensor Networks in Structural Health Monitoring: a Modular Approach

Fabio Angeletti, Mario Paoli

Department of Computer, Control and Management Engineering
University of Rome La Sapienza
Email: {angeletti, paoli}@dis.uniroma1.it

Ugo Maria Colesanti, Andrea Vitaletti

Wireless Research Laboratory
Wsense s.r.l.
Email: {ugo.colesanti, andrea.vitaletti}@wsense.it

Abstract—In this paper, we present the Modular Monitoring System (MMS), a low-power wireless architecture dedicated to Structural Health Monitoring (SHM) applications. Our solution features an easily customizable modular architecture, fulfilling the needs of many SHM applications. The MMS supports mesh network topology and offers excellent coverage and reliability, taking advantage of Wireless Sensor Networks (WSN) technology. In this preliminary work we show how the flexibility of our approach offers great advantages with respect to the current state-of-the-art systems dedicated to SHM.

Keywords—Wireless Sensor Networks; Structural Health Monitoring; Modular Architecture; Wireless; Low-Power.

I. INTRODUCTION

Detecting damages to which civil and industrial structures, such as roads, bridges, canals, buildings and aerospace vehicles are subjected, is referred as SHM. It can prevent collapses and breaks, avoiding permanent damages to structures, thus simplifying their maintenance. Depending on application scenarios, SHM requires many different types of sensors, including pressure sensors, vibrating-wire strain gauges, inclinometers, crackmeters, etc. Most monitoring systems dedicated to SHM are wired but, deploying a wired system in a wide area or in a harsh environment, can pose both economical and practical limitations. For such reasons, WSNs were proposed as network infrastructure to support SHM, avoiding prohibitive costs of wired systems and easing the on-field deployment. Nowadays, SHM supported by WSN is an active and well-established research field and some wireless SHM systems are now entering the market. The preliminary work made in this paper introduces the MMS, a wireless, low-power, scalable architecture dedicated to SHM. The main strength point of MMS is its high modularity that allows easy customization of the platform depending on the number and the type of sensors required by the specific application scenario. In addition, MMS fully supports multi-hop wireless communication paradigm and mesh networks. The remainder of this paper is organized as follows: Section II is dedicated to both wired and wireless SHM state-of-the-art systems, in Section III we explain motivations supporting design choices made during the development of the MMS, while in Section IV we introduce the features characterizing our system from an high-level perspective. In Section V, we present the hardware prototype and the achieved results. Finally, in Section VI, we give conclusions and directions for the future.

II. STATE OF THE ART

Nowadays, most SHM systems available on the market, such as the Geomonitor by Solexperts [1], are wired. However, deploying those systems can be cumbersome: besides the installation costs, a detailed deployment plan is required to

face evolving needs of different construction phases. Moreover, cables are subjected to accidental cuts and damages and, in some scenarios, their installation is unfeasible or inappropriate (e.g., historical buildings). Along with wired systems, some standalone data loggers dedicated to SHM are commercially available, among the others: Solexperts SDL [1], Geokon 8002-16-1 [2] and Keller GSM-2 [3]. Those systems are simpler to install, but they do not allow real-time remote monitoring and require frequent in-situ access by qualified personnel to collect data.

The introduction of wireless communications in SHM gives immediate advantages in terms of easier deployments and reduced maintenance and personnel costs. However, when monitoring devices are battery-powered, the employment of wireless communications is among the most energy demanding feature that can significantly limit the devices lifetime. Prominent examples of wireless monitoring systems are: National Instruments Wireless Data Acquisition (WiFi-DAQ) [4] and MicroStrain's wireless sensor network [5]. While the former supports IEEE 802.11 standard, the latter is compliant with IEEE 802.15.4. Both products adopt a conservative approach by limiting wireless communication to 1-hop. Supporting multi-hop wireless communications was investigated in several research papers in the last decade [6][7][8]. Multi-hop networking provides a number of advantages: scalability (larger areas can be covered), reliability (failure and multiple routing paths without single point of failure) and ease of deployments (the presence of multi-hop relay nodes allows to bypass obstacles like walls and metal structures). Recently, National Instruments presented the NI WSN [9]: a multi-hop battery-powered WSN supporting up to 36 nodes configured in a mesh network and up to 3 years node lifetime.

III. MOTIVATIONS

As seen in the previous section, WSNs are slowly entering the market of SHM applications. An attempt to develop a robust solution and to test it in realistic application scenarios was made in the GENESI Project [10]. The main goal of this project was to design and implement a “novel generation of green wireless sensor networks which can be embedded in buildings and infrastructures at the time of construction and be able to provide a monitoring and control intelligence over the whole structure lifetime”. The project involved the monitoring of a bridge construction site in Fribourg [11] and the construction of a tunnel for the B1 underground line in Rome [12], by means of WSNs. The outcomes highlighted the advantages of WSN technology described in Section II compared to old monitoring techniques. According to the application requirements provided by the SHM experts, GENESI nodes can support a number of heterogeneous sensors.

However, a GENESI node can manage only a single sensor per type, while there are some applications in which multiple instances of the same sensor are needed. As an example, in a 3-axis deformation analysis, a single wireless node may need to interface with three instances of a vibrating-wire strain gauge while to monitor a concrete junction the node may need to interface a current-loop inclinometer sensor and a resistive displacement sensor. Other commercial solutions, such as the NI WSN described in Section II, can handle multiple instances of the same sensor but can not support different sensor families simultaneously. In general, we observed that developing a hardware platform flexible enough to handle a high number of sensor combinations required by SHM applications is not practical both in terms of size and costs. In the specific, such solution would require additional hardware and connectors that could remain unused during deployments. This brought us to propose a novel low-power slotted modular system made by a set of modules connected through an internal communication bus. This solution features one wireless-capable master module managing a group of extension modules, each one designed to interface a specific sensor setup. The flexibility of the proposed architecture, named MMS, allows to support a vast number of SHM applications by simply plugging into each node the required extension modules. Another important advantage given by MMS is the possibility to easily swap between wireless technologies by replacing the master module. This characteristic allows to preserve the required level of wireless performance in both indoor and outdoor scenarios, further increasing the flexibility of our system.

IV. SYSTEM ARCHITECTURE

The MMS architecture is based on a master/slave communication abstraction: a *master module*, provided with radio capability and responsible for most of computational tasks, communicates through a *low-power shared bus* with a maximum of four *extension modules* (slaves). All modules are interconnected through a backplane following Figure 1.

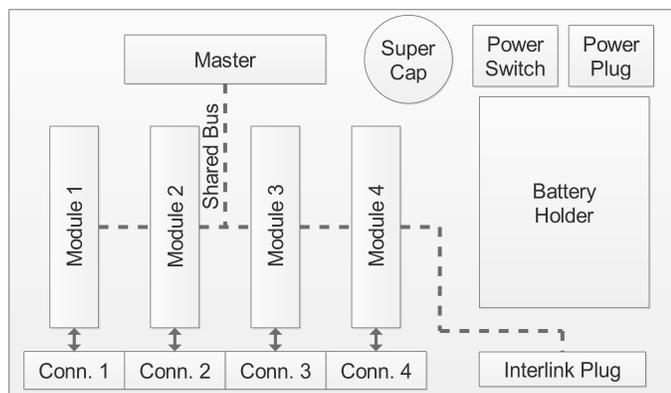


Figure 1. Modular Architecture

The backplane also offers 12 external pins per module to attach external sensors and a battery plug for different kind of power supplies.

A. Low-Power Shared Bus

The low-power shared bus was designed to optimize the energy consumption of the whole system. For this reason, it

uses a communication principle based on Serial Peripheral Interface (SPI) with dedicated chip-selects, interrupts and power lines for each extension module. This enables the master to communicate in an exclusive way with each module, without spurious wake-up of slave modules not involved in the communication (as opposed to other shared bus communication interfaces, i.e., the Two-Wire Interface). The drawback of this approach is a high number of additional lines required, with a subsequent limited number of supported extension modules. Please refer to Section IV-D for a solution to this known limitation.

B. Master Module

The master module is responsible for keeping wireless connectivity with the WSN and manages extension modules. When switched on, the master performs a discovery routine on the low-power shared bus to detect extension modules and their slot locations. For each discovered module, the master retrieves its configuration information and computes the sensing schedule. The master can also update the configuration depending on the user needs. When asking for sensor data, the master has different ways to interact with the extension modules based on power requirements, scheduling and module type. The typical interaction protocol of the master module follows the steps:

- 1) *Optional*: It enables the power line of the extension module
- 2) It pulls down the corresponding chip-select line
- 3) It sends a *Data Request* command with all the required information
- 4) *Optional*: It pulls up the chip-select line and perform other tasks or goes to sleep
- 5) It waits for an interrupt from the module
- 6) *Optional*: It pulls down the chip select line
- 7) It sends a *Data Read* command and retrieves the response
- 8) It pulls up the chip select line
- 9) *Optional*: It disables the power line of the corresponding module

C. Extension Modules

The extension modules provide the hardware interface to external sensors. Each module communicates over the low-power shared bus with the master by means of an SPI-capable microcontroller. They also feature an interrupt pin for chip-select detection and a general-purpose I/O pin to signal the interrupt to the master. In addition, a power pin allows the master to power-down the whole extension module. This power line gives to the system the maximum flexibility in terms of energy management: the module can be switched off or simply kept in a sleep state by the master. This choice varies according to the sampling period, the start-up time and the sleep power consumption of the extension module. The hardware interface provided by each module depends on sensors it supports: it can be fully digital, e.g., to interface RS-232 or RS-485 peripherals, or analog, to connect sensors such as current-loop, vibrating wire strain gauges, resistive etc. Each extension module can support different kinds of sensor or several sensors of the same type.

D. Additional Features

Interlink Bus: As pointed out in Section IV the low-power shared bus cannot interface more than four extension modules, which in some rare cases might represent a limitation. To overcome this issue, we introduced the interlink bus which enables the MMS to support a virtually unlimited number of extension modules. In the specific, it connects the master modules of different MMS creating a tree structure where the parent MMS acts as master for its children.

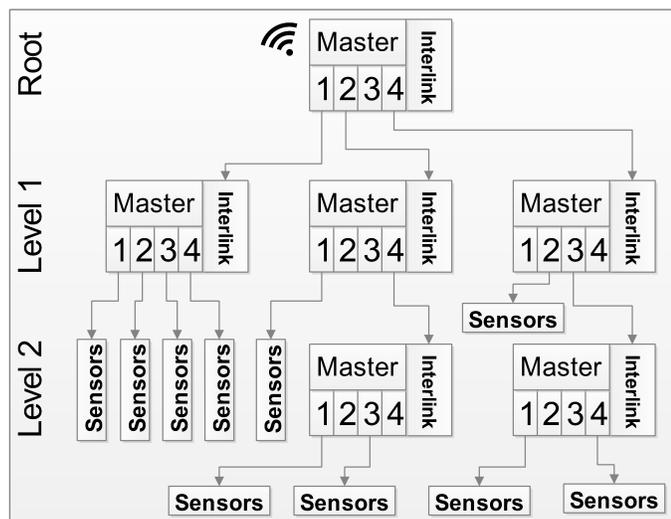


Figure 2. Interlink connections

To do so, the parent MMS uses the extension slots to make a connection through the interlink plugs of its children (see Figure 2). Hence, the parent sees its children as enhanced extension modules. This mechanism can be replicated at each level of the tree to support an unlimited number of sensors. Each leaf node advertises the connected sensors together with its identification number to its parent which aggregates the received information and forwards it to the upper level. This procedure is iterated until the root MMS is reached. The root MMS is the only one which manages wireless communications. From a WSN perspective, this makes the whole tree resulting as a single wireless node.

Extension Module Configuration Advertisement: Thanks to the discovery procedure performed by the master during the boot phase, extension modules attached to the backplane exchange configuration information. Hence, an operator can add an extension module at any time by just power-cycling the MMS, without the need of an in-situ reprogramming. This feature eases the effort required by operators and qualified personnel. Based on our on-field experience, this functionality is often required in evolving environments, such as (but not limited to) SHM scenarios.

V. PRELIMINARY WORK

A. Prototype development

We developed and assembled the first prototype of the MMS at the Department of Computer, Control and Management Engineering of the University of Rome "La Sapienza". The prototype includes a master module, a Resistance Temperature Detector (RTD) extension module and a backplane. The

modules are plugged into the backplane by a standard Peripheral Component Interconnect Express (PCIe) edge connector.

The master module is responsible for radio communication and for managing the extension modules. It is based on the MagoNode OEM [13], a wireless hardware platform specifically designed for WSN applications. The MagoNode is based on the Atmel Atmega128RFA1 (RFA1) System-On-Chip micro-controller equipped with 128KB of ROM (Read Only Memory), 16KB of RAM (Random Access Memory) and an embedded 2.4Ghz radio transceiver fully compliant with the 802.15.4 standard. The radio range is extended through a power efficient RF (Radio Frequency) front-end which enhances radio performance while keeping the power consumption low. The main figure in terms of power consumption are: radio transmission 27.7mA @+10dBm, radio reception 14.5mA and <math><2\mu A</math> in sleep. The master module is equipped with a NOR flash for persistent data storage, three status leds, a micro-usb plug and an RP-SMA (Reverse Polarity SubMiniature version A) antenna connector.

The RTD module is the first extension module developed for the MMS node. It is a hardware interface made for reading up to four PT-100 or PT-1000 RTD temperature sensors. Two accurate 24-bit ADCs (Analog-to-Digital converter), together with a finely tuned filtering circuit, provides sensor readings in 50ms with a maximum error of $\pm 0,1^{\circ}C$. An Atmel 8-bit AVR micro-controller manages the communication with the two ADCs and with the master module over the low-power shared bus. The power consumption is $2\mu A$ in sleep and less than 500nA when powered off.

The backplane hosts the low-power shared bus that connects the master and up to four extension modules through PCIe connectors (see Figure 3). It provides the interlink bus header (see Section IV-D) and 12 pins per extension modules to plug external sensors. The backplane also supplies the system through different battery types. It supports battery holders for 2xAA, 1xC or 1xD thionyl chloride batteries (3.6V) and an external power supply for other power sources (3 to 5.5V).

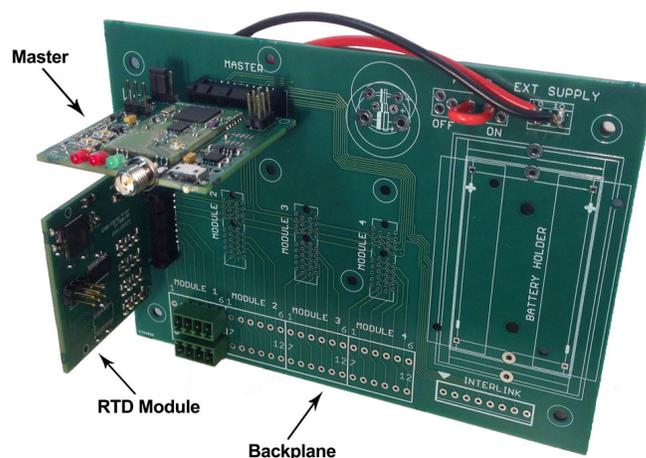


Figure 3. The MMS: master, RTD module, backplane

We developed drivers and firmware for the MMS in compliance with TinyOS [14], an event-driven, open-source operating system (OS) dedicated to Wireless Sensor Networks. TinyOS

is designed to cope with typical constraints imposed by WSNs hardware: low computational capabilities, limited memories and scarce energy resources.

B. Test Results

During our preliminary tests, we evaluated the MMS energy consumption while performing different activities. The MMS hardware configuration used during tests consisted of a master module interconnected through a backplane to an RTD module, which acquires samples from a temperature probe simulated by a 100 Ohm resistor. In order to measure current consumption, the device was powered by a 3.6V external supply and connected to a Rigol DM3068 digital multimeter sampling at a 10KHz rate. The prototype was programmed to wake-up from sleep state, perform a sensor reading (involving ADC initialization, data conversion and retrieval), send acquired data over the wireless medium and go back to sleep. Figure 4 shows the current measurements of the previously cited activities while in Table I significant current values corresponding to different MMS states are summarized.

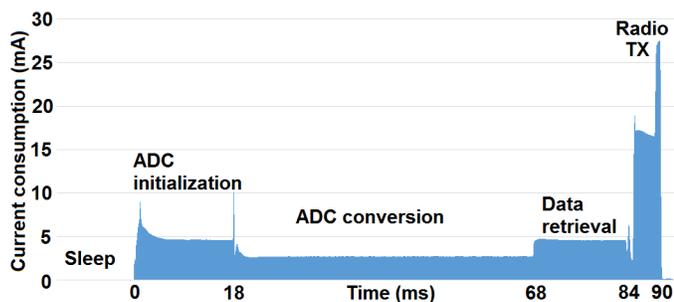


Figure 4. MMS Current Consumption

As expected, Table I shows how the low consumption in sleep state of both master and RTD modules allows the MMS to draw a negligible amount of current ($4.5\mu A$). Such a low value would allow the system to last in this state for more than most batteries self-discharge time. The MMS sleep current can be further reduced by allowing the Master to power off the RTD module. This feature, which is irrelevant in this case, becomes important whenever there is the need to deal with extension modules with a high sleep current consumption.

TABLE I. MMS CURRENT CONSUMPTION

State	Description	Current (mA)
Sleep	Baseline consumption	0.0045
Radio TX	Master radio transmitting	27.7
Radio LISTEN	Master radio idle listening	14.4
Radio RX	Master radio receiving	16.9
Data acquisition	RTD module ADC conversion	2.8

The radio power consumption of the MMS are strictly related to the MagoNode platform, this is why the same values measured in [13] are reported in Table I. Finally, the ADC of the RTD module performs a single conversion in 50ms by generating two 1mA bias current over the RTD sensor that, summed with the $800\mu A$ power consumption of the chip, gives an overall consumption of 2.8mA.

VI. CONCLUSION AND FUTURE WORK

In this work, we presented a novel low-power wireless modular architecture designed for SHM applications. Our platform, the MMS, takes advantage of a low-power shared bus connecting slotted extension modules that interact with a master in a master/slave communication abstraction. The extension modules, which can be combined as needed, allows the MMS to face the continuously evolving needs of most SHM scenarios. Thanks to its peculiar characteristics, the MMS overcomes commercial state-of-the-art WSN solutions for SHM, like the NI WSN, which do not offer enough flexibility to fulfill requirements of many application contexts in a both cost-effective and efficient way.

The preliminary work done toward the implementation of this novel architecture consists of the design and the assembly of a master module, a first RTD extension module and an interconnection backplane. The master module, based on the MagoNode platform, makes the MMS a powerful wireless node fully supporting multi-hop communications. The RTD module gives low-power readings from up to 4 PT100 or PT1000 temperature probes with a precision of $\pm 0,1^{\circ}C$. In the near future, we plan the develop several additional extension modules able to support all kinds of sensors dedicated to SHM, such as strain wire gauges, crack meters, inclinometers, displacement sensors, etc. Furthermore, additional Master modules will be designed to support different wireless frequencies (e.g., 433Mhz, 868Mhz, 915Mhz) and certified industrial wireless protocols like Wireless Hart.

REFERENCES

- [1] "Solexperts website," URL: <http://www.solexperts.com/> [accessed: 2015-03-27].
- [2] "Geokon website," URL: <http://www.geokon.com/8002-16> [accessed: 2015-03-27].
- [3] "Keller website," URL: <http://www.keller-druck.com/> [accessed: 2015-03-27].
- [4] "NI WiFi-DAQ," URL: www.ni.com/data-acquisition/wireless/ [accessed: 2015-03-30].
- [5] "Microstrain WSN," 2015, URL: <http://www.microstrain.com/wireless> [accessed: 2015-03-27].
- [6] N. Xu et al., "A wireless sensor network for structural monitoring," in Proceedings of the 2Nd International Conference on Embedded Networked Sensor Systems, ser. SenSys '04. New York, NY, USA: ACM, 2004, pp. 13–24.
- [7] J. P. Lynch and K. J. Loh, "A summary review of wireless sensors and sensor networks for structural health monitoring," Shock and Vibration Digest, vol. 38, no. 2, 2006, pp. 91–130.
- [8] T. Harms, S. Sedigh, and F. Bastianini, "Structural health monitoring of bridges using wireless sensor networks," Instrumentation & Measurement Magazine, IEEE, vol. 13, no. 6, 2010, pp. 14–18.
- [9] "National Instruments WSN," 2015, URL: <http://www.ni.com/wsn/i/> [accessed: 2015-03-27].
- [10] "GENESI project website," URL: <http://genesi.di.uniroma1.it/> [accessed: 2015-03-30].
- [11] "Structural health monitoring at the pont de la poya, fribourg," in White Paper, GENESI Project, 2013.
- [12] U. M. Colesanti, A. Lo Russo, M. Paoli, C. Petrioli, and A. Vitaletti, "Structural health monitoring in an underground construction site: the roman experience," in Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems. ACM, 2013, p. 46.
- [13] M. Paoli, A. Lo Russo, U. M. Colesanti, and A. Vitaletti, "Magonode: Advantages of rf front-ends in wireless sensor networks," in Real-World Wireless Sensor Networks. Springer, 2014, pp. 125–137.
- [14] P. Levis et al., "Tinyos: An operating system for sensor networks," in Ambient intelligence. Springer, 2005, pp. 115–148.

Building the O-Life Franco-Lebanese Environmental Observatory Using Sensor Web Enablement Framework: Challenges and First Approach

Hicham Hajj Hassan, Anne
Laurent, Nicolas Arnaud, Olivier
Lobry
Univ. Montpellier
Montpellier, France
e-mail: hishamhh@cnrs.edu.lb;
e-mail: laurent@lirmm.fr;
email: Nicolas.Arnaud@univ-montp2.fr;
e-mail: olivier.lobry@umontpellier.fr

Laurent Drapeau
Centre d'Etudes Spatiales de la
Biosphère
Toulouse, France
e-mail: laurent.drapeau@cesbio.ups-tlse.fr

Carla Khater
Remote Sensing Center
CNRS-L, Lebanon
e-mail: ckhater@cnrs.edu.lb

Abstract— This paper introduces an approach how the Sensor Web Enablement (SWE) framework of the Open Geospatial Consortium (OGC) will be used in order to build a franco-lebanese observatory. We present the practical application of SWE services as a source of real-time observation data and the associated technical architecture for making the observations available on the Web to end users near real-time. We discuss the question of crossing sensor data with other data sources, e.g., data provided by human observations. We illustrate our approach with a first use case to monitor the snow weather stations in Lebanese Mountains.

Keywords- *O-Life; Environmental Observatory; Sensor Web Enablement; Lebanon; Snow monitoring*

I. INTRODUCTION

Lebanon is bordering the eastern Mediterranean Sea (See Figure. 1). The Mediterranean basin is a priority area and a

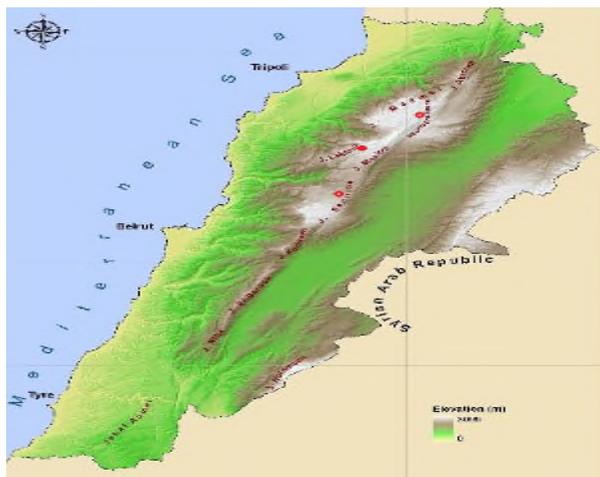


Figure 1. Study area in Lebanon's map of snow weather stations

leading area for the analysis of environmental data and for the extrapolation of trends that will allow a better management of the present and help envisage plausible scenarios for the future. For this reason, actors of the French and Lebanese scientific research community wished to establish a shared observatory between France and Lebanon called "Observatoire Libano-Français de l'Environnement", O-LiFE [1]. The main activity of the observatory is to study the critical zone around the Mediterranean, of water resources, biodiversity, natural hazards, and management of the environment and ultimately the study of land use. To carry out its mission, the O-LiFE observatory aims to construct environmental databases of the critical zone in consideration, create collaborative software tools and organize, share, sustain and enhance environmental data.

Building such an environmental observatory requires defining a technical architecture that must be able to store various types of data and to offer several services. The data and knowledge being handled are mostly spatial and collected from sensors and scientists. For this purpose, we propose an open, standards-based, modular architecture based on the OGC SWE framework. OGC SWE has indeed become currently a very useful monitoring technique at spatial and temporal scale especially after the considerable improvement in the ICT communication infrastructure.

In this paper, we discuss the challenges for building the technical architecture and present our first approach. The next section introduces the framework we propose to use. Section 3 details our context and the architecture we propose. Section 4 illustrates this approach with our first use case dedicated to snow while Section 5 concludes this paper.

II. OGC SENSOR WEB ENABLEMENT

In this section, we present the main ideas lying in the standard OGC that is used in most environmental observatories and that makes it possible to exchange and use data from such observatories. The OGC is an international, non-profit, voluntary consensus standards organization

consisting of more than 500 companies, government agencies and universities [2]. The goal of the OGC is the creation and establishment of standards that enable global infrastructures for delivery and integration of geospatial content and services and to facilitate the adoption of open, spatially enabled reference architectures in enterprise environments worldwide.

Within the last years, the Sensor Web Enablement (SWE) architecture of the Open Geospatial Consortium (OGC) has matured into its second generation [3]. The main goal of this standards framework is the integration of sensors and sensor data into Spatial Data Infrastructures and thus makes it possible to use data measured by sensors in a broad range of applications. Thus, sensor data become an additional source for geospatial information besides conventional data types like maps. In addition, the availability of sensor data offers the possibility to integrate near real-time measurements with conventional geospatial data for visualization purposes.

The SWE framework consists of a set of standards defining data formats for sensor data and metadata, as well as service interfaces to access sensor data, task sensors or send and receive alerts based on sensor measurements [4].

III. PROPOSED IMPLEMENTATION

In this section we discuss the similarities and specificities of the O-Life Observatory and present how the OGC SWE standards are integrated in the O-Life infrastructure so that snow real-time data sources becomes available to end users.

A. Similarities and Specificities of the O-Life Observatory

O-LIFE will be built, as a first step, by combining and sharing data already collected and with direct access to in-situ sensors.

A survey has been conducted for O-LiFE supported research projects showed that there are existing databases and data but most of these data are dispersed, diverse, not updated, unpublished, insecure, not shared, and not accessible to public. Furthermore, in many cases data from many heterogeneous data sources might need to be compared and/or combined and different data sources have to be correlated. Moreover, there are different data sharing policy within the scientific community that prevents access to data. Sharing and leveraging data and research resources can avoid duplication of very expensive and time-consuming efforts, allowing scientists to spend more time in data analysis than in data collection and discovery, and enable more people to benefit from using environmental data. Advanced security concept is necessary so that only authorized users are able to insert/query data into the O-LiFE infrastructure.

Hence, it is necessary to develop a standardized system for collecting data at national level and develop a data sharing policy resulting in a sustainable database, interoperable, shared, and regularly updated. This first step aims subsequently to create an ambitious Circum-Mediterranean observatory network.

B. Crossing Data

O-LIFE aims to collect and exchange data from various sources, including sensor and human observations. As discussed in this paper, sensor data can easily be managed by the Sensor Web infrastructure. However, in some cases it may be difficult to manage human observations as such data do not always have a compatible structure. For instance, the way data are georeferenced, the type of data (e.g., textual data) can make it difficult to integrate the data directly as an input of Sensor Web. For this reason, we envision four possibilities for crossing data that will be explored in our work:

- 1) Transforming non sensor data so as to consider all data sources as inputs of Sensor Web;
- 2) Loading non sensor data directly in relational tables of the PostgreSQL database used to serialize Sensor Web data;
- 3) Crossing data thanks to an application within our infrastructure, thus relying on an application server to deliver the services;
- 4) Crossing data outside our platform, this option being taken by end-users after importing some data from O-LIFE.

C. Proposed Architecture

We choose to use the provided by 52° North German initiative for Geospatial Open Source software, ‘Sensor Observation Service’ (SOS) and ‘Observations and Measurements’ (O&M) [5]. The SOS is a web service to query real-time sensor data and sensor data time series and is part of the Sensor Web framework. A 52° North PostgreSQL database with a predefined schema by 52°North implementation of the SOS will be used to store the sensors data. The PostGIS spatial extension will be included in the PostgreSQL database. Apache Tomcat will be used as Servlet container. The proposed architecture is shown in Figure 2.

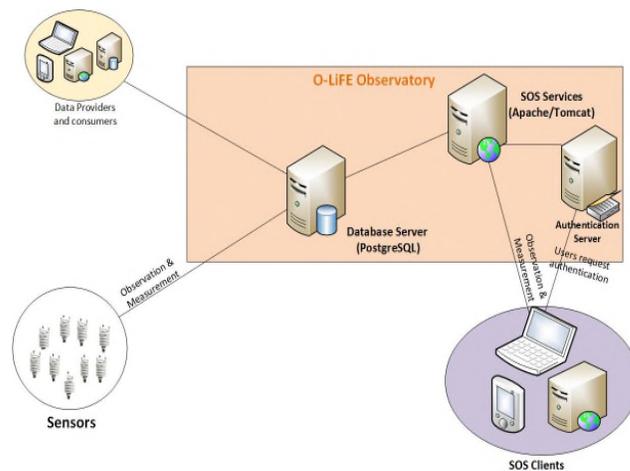


Figure 2. Proposed architecture for integrating the SOS platform into the O-LiFE architecture

The SOS provides an interface for requesting sensor data sets based on temporal and spatial query parameters. The responses of the SOS containing the requested data are then returned using O&M. O&M is a model defining how to link observed values with all relevant properties (i.e., location, time, observed parameter, unit of measurement) that are necessary to interpret them. In the SOS interface, O&M is used in two operations:

- 1) GetObservation: O&M is the default output
- 2) InsertObservation: O&M is used for encoding all values that shall be inserted into a SOS server

The database server contains environmental data from different data providers for different thematic and structures. Sensors data will be hosted on the same server.

IV. CASE STUDY: SNOW WEATHER STATIONS

The observatory system is designed to support real-time monitoring which can collect sensors data. Snow weather stations will be used as our first use case.

A. Presentation

Lebanon is a real water palace, of which the study of snow and dynamics coats is instructive for several reasons: trends, hazards and climate predictability, redistribution of meltwater and impact on the availability of downstream water.

The snow weather stations are installed in Lebanon as part of an agreement for scientific and technical cooperation for research on snow hydrology project and an observatory of joint project of the snow in the Mediterranean between CESBIO, the National Council for Scientific Research in Lebanon and the Saint Joseph University in Beirut [6].

The three snow weather stations are located in three different regions (Mzaar, Cedars, Laqlouq) comprising sensors measuring temperature, snow height, relative humidity, precipitation, air pressure, wind speed and wind direction. The outdoor sensors are powered by solar cells as appears in Figure 3. Sensor observations are sent remotely via a TCP/IP connection through the Internet to the SOS server hosted in the O-LiFE observatory datacenter using



Figure 3. Snow weather stations in Laqlouq region, Lebanon

GPRS modem. Users will be able to query snow sensors data and access animated visualization of spatial-temporal data, including time series presentation. Table I shows a snapshot of data generated by the snow weather stations.

B. Expected Results

The proposed approach for in-situ sensor near real-time data acquisition is supposed to be a first step towards establishing the SWE framework within O-LiFE Observatory. The observatory aims to establish a national Spatial Data Infrastructure (SDI) for Lebanon ensuring acquisition, archiving and management of multi-scale satellite imagery and in-situ datasets for the Lebanese territory and making it accessible by the scientific community and by various public actors involved in environment and management.

Data from in-situ or remote sensing devices form the basis for analyzing gradual processes, such as snow melting, water shortages, or increasing drought. In-situ and satellite measurements are not directly comparable due to their basic configurations. In-situ sensors provide point based measurements at ground level whereas satellites observe the entire atmosphere. The near real time in-situ data could be used to calibrate and validate remote sensing data and models. Later, approaches to consolidate remote sensing data, and in-situ data will be considered to get more accurate results. This is highly prioritized by the O-LiFE observatory to make use and share all available source of data. Nevertheless, a special effort should be made to add additional in-situ sensors in different research areas and to include existing in-situ sensors in the country to the Observatory SDI. Besides, solutions to include mobile human sensors through mobile applications should be considered and developed to have vast amount of incoming georeferenced data and to allocate these data using the SWE standards.

V. CONCLUSION

In this paper, we investigate the possible ways for data management within O-Life observatory. We discuss the specificities of this context and we propose an architecture based on the standards of the Open Geospatial Consortium. This work is being currently implemented in the framework of a collaboration between French and Lebanese organizations. Many perspectives are associated to this first approach. First, we aim at benchmarking our proposal with real data collected in the various data sources. Second, we aim at studying how our architecture can provide valuable services to scientists both for raw data crossing and exchanges and for analytical processing. Finally, we aim at studying how the O-Life data can be shared through Open Data for better enabling interoperability and linked to other open resources [7].

ACKNOWLEDGMENT

This work is part of a PhD thesis and is supported by the CNRS Lebanon PhD scholarship program. This is an O-LiFE contribution number SA 12-2015.

REFERENCES

- [1] <http://www.cnrs.edu.lb/olife> [retrieved:04,2015]
- [2] <http://www.opengeospatial.org> [retrieved:04,2015]
- [3] M. Botts, G. Percivall, C. Reed and J. Davidson, OGC Sensor Web Enablement : Overview and High Level Architecture, OGC Whitepaper, OGC Document Number : 2007, 07-165.
- [4] Bröring, A. et al., New Generation Sensor Web Enablement. MDPI Sensors, 2011, 11, (3), 2652-2699
- [5] <http://52north.org> [retrieved:04,2015]
- [6] http://www.cesbio.ups-tlse.fr/fr/sudmed/sites_ateliers_liban.html [retrieved:04,2015]
- [7] Florian, B. and Martin, K. Florian, B. (R. & Martin, K. (S. W. C., ed. 2012, Linked Open Data, The Essentials - A Quick Start Guide for Decision Makers , edition mono/monochrom, Vienna, Austria.

TABLE I. EXAMPLE OF PART A OF THE DATA GENERATED BY THE SNOW WEATHER STATIONS

TimeStamp	AirTemp (C)	vit_vent (m/s)	Direction vent (degree)	Std direction vent (degree)	Pression (mbar)	Geonor (Hz)	Geonor (mm)	Wind max (m/s)
2015-04-09 08:00:00	8.06	13.49	190.8	12.5	764.1	2090	27.73	24.5
2015-04-09 07:30:00	8.34	13.34	187.1	10.66	764.3	2090	27.74	23.81
2015-04-09 07:00:00	8.97	14.77	189.8	10.99	764.2	2090	27.72	23.86
2015-04-09 06:30:00	9.19	14.24	186.1	11.49	764.2	2090	27.72	24.7
2015-04-09 06:00:00	8.67	11.11	187.2	13.52	765	2090	27.74	20.24
2015-04-09 05:30:00	9.49	11.18	185.6	10.64	764.9	2090	27.74	22.25
2015-04-09 05:00:00	10.21	13.12	183.3	11.19	764.5	2090	27.74	20.78
2015-04-09 04:30:00	10.63	13.9	183	11.6	764.5	2090	27.73	26.17
2015-04-09 04:00:00	11.36	16.05	188.9	12.64	764.4	2089	27.7	32.19

- TimeStamp:date and time of the measured values
- AirTemp (C): Air temperature in Celsius
- vit_vent (m/s): Wind speed in m/s
- Direction vent (degree): Wind direction
- Std direction vent (degree): The standard deviation of wind direction
- Pression (mbar): Atmospheric pressure in mbar
- Geonor (Hz): frequency of vibrating wires in Geonor precipitation gauge
- Geonor (mm): accumulation of precipitation in the bucket
- Wind max (m/s): Maximum Wind speed

A Virtual Force Movement Scheme for Sensor Deployment in Directional Sensor Networks

Chiu-Kuo Liang and Yu-Shu Lo

Dept. of Computer Science and Information Engineering, Chung Hua University

Hsinchu, Taiwan, R.O.C.

Email: {ckliang, e10202025}@chu.edu.tw

Abstract—A directional sensor network is composed of many directional sensor nodes. Unlike conventional omni-directional sensors that always have an omni-angle of sensing range; directional sensors may have a limited angle of sensing range due to technical constraints or cost considerations. Area coverage is still an essential issue in a directional sensor network. In this paper, we study the area coverage problem in directional sensor networks with mobile sensors, which can move to the correct places to get high coverage. We present distributed self-deployment schemes of mobile sensors. After sensors are randomly deployed, each sensor calculates its next new location to move in order to obtain a better coverage than previous one. The locations of sensors are adjusted round by round so that the coverage is gradually improved. Based on the virtual force of the directional sensors, we design a scheme, namely Virtual force scheme. Simulation results show the effectiveness of our scheme in term of the coverage improvement.

Keywords—Directional sensors; mobile sensors; area coverage.

I. INTRODUCTION

In recent years, wireless sensor networks have received a lot of attention due to their wide applications in military and civilian operations, such as fire detection [1], vehicle traffic monitoring [2], ocean monitoring [3], and battlefield surveillance [4]. In wireless sensor networks, target coverage is a fundamental problem and has been studied by many researchers. Most of the past work is always based on the assumption of omni-directional sensors that has an omni-angle of sensing range. However, there are many kinds of directional sensors, such as video sensors [5], ultrasonic sensors [6] and infrared sensors [7]. The omni-directional sensor node has a circular disk of sensing range. The directional sensor node has smaller sensing area (sector-like area) and sensing angle than the omni-directional one. Compared to isotropic sensors, the coverage region of a directional sensor is determined by its location and orientation. This can be illustrated by the example shown in Figure 1.

Area coverage is a fundamental problem in wireless sensor networks. Therefore, sensor nodes must be deployed appropriately to reach an adequate coverage level for the successful completion of the issued sensing tasks [8][9]. However, in many potential working environments, such as remote harsh fields, disaster areas, and toxic urban regions, sensor deployment cannot be performed manually.

Deploying sensors by aircraft may result in the situation that the actual landing positions cannot be controlled. Consequently, the coverage may be inferior to the application requirements no matter how many sensors are dropped. In such a situation, it may need to make the mobile sensors to move to the correct positions for the required coverage.

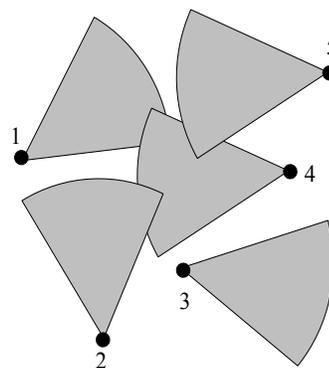


Figure 1. An example of five directional sensors deployed to cover target region

Most previous research efforts on deploying mobile sensors are based on the omni-directional sensor networks. For example, Howard et al. [10] present a distributed, potential-field-based approach to solve the coverage problem. In their approach, sensor nodes are treated as virtual particles that are subject to force, these forces repel the neighboring sensor nodes from each other and from obstacles. Finally, sensor nodes will spread from dense to sparse area. The concept of potential-field was first proposed in the research of mobile robotic route plan and obstacle avoidance by Khatib [11]. In [12], Wang et al. present a set of Voronoi diagram-based schemes to maximize sensing coverage. After discovering a coverage hole locally, the schemes calculate new position for each sensor to move at next round. They use the Voronoi diagram to discover the coverage holes and design three movement-assisted sensor deployment schemes: VECTOR-based (VEC), VORonoi-based (VOR), and Minimax. In [13], Lee et al. designs two movement-assisted schemes: Centroid-based and Dual-Centroid-based. Based on the Voronoi diagram and centroid (geometric center), the proposed schemes can be used to improve the sensing coverage.

In this paper, we study the problem of coverage by directional mobile sensor under the random deployment strategy. We develop a solution that maximizes the sensing coverage while minimizing the computation time in term of rounds. Based on the virtual force of directional sensors, we design a moving algorithm: the Virtual Force scheme. Simulation results show that our distributed algorithm is effective in terms of coverage, deployment time, and movement.

The rest of this paper is organized as follows. In Section 2, we introduce some preliminaries. In Section 3, we state the problem formally and make some assumptions regarding the problem. We present our scheme in Section 4. Section 5 shows some simulation results. Finally, we conclude this paper in Section 6.

II. PRELIMINARIES

A. Directional Sensing Model

Compared to an omni-directional sensor which has a circular disk of sensing range, a directional sensor has smaller sensing area (sector-like area) and sensing angle. This can be best illustrated in Figure 2. As shown in Figure 2, the sensing region of a directional sensor is a sector denoted by a 3-tuple (α, β, R) , and the sensing region is called sensing sector. Here R is the sensing radius, α is the sensing angle, and β is the offset angle.

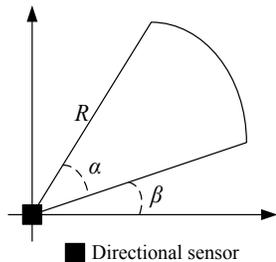


Figure 2. The directional sensing model

B. Virtual Force Point

We assume that there are many virtual force points around the boundary of sensing sector, as shown in Figure 3. Without loss of generality, we assume that there are $3m + 1$ virtual force points of each sensor, in which there are m points on the arc, m points on each of the both straight boundary lines of sensing sector, and sensor point itself. We denoted the virtual force points of s_i , as p_{ik} , where $1 \leq k \leq 3m + 1$. As shown in Figure 3, there are total $3 \cdot 5 + 1$ virtual force points, in which there are 5 points on the arc and on both straight boundary lines respectively. Each virtual force point on sensors can receive a repulsive force from other sensors.

It should be noticed that, as shown in Figure 4, the more virtual force points a directional sensor has, the more repulsive force that exerted on the directional sensor by its neighboring directional sensors. By applying all of the repulsive forces from its neighbors, the directional sensor can be repelled from dense to sparse area.

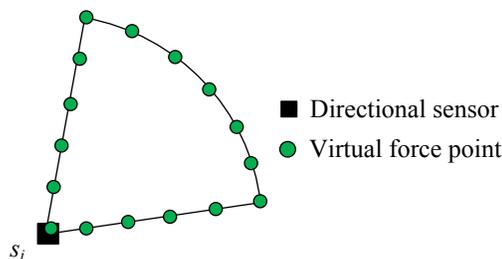


Figure 3. The virtual force points

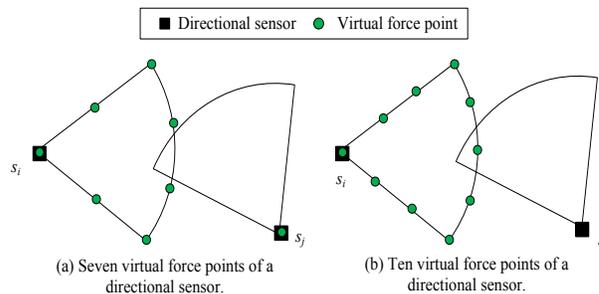


Figure 4. Illustration of two different number of virtual force points

It also should be noticed that, the more virtual force points a sensor has, the more accurate repulsive force can be applied. In Figure 4(a), the directional sensor s_j has the overlapped region with directional sensor s_i , which implies that sensor s_i should be repelled by the sensor s_j . However, sensor s_j will not exert a repulsive force on s_i because that sensor s_j does not cover any virtual force point of sensor s_i . On the contrary, in Figure 4(b), we can see that sensor s_j has covered one virtual force point of s_i . Therefore, sensor s_j would exert a repulsive force on s_i .

III. PROBLEM STATEMENT

Problem: Randomly deploying N mobile directional sensors with sensing range R_s and sensing angle α in a given target sensing region, we are asked to maximize the sensor coverage with less time.

To address the above problem, we need to make the following assumptions:

- All directional sensors have the same sensing range (R_s) and sensing angle α , where $0 < \alpha \leq 2\pi/3$. Directional sensors within $2R_s$ of a sensor are called the sensor's neighboring nodes.
- Directional sensors can move to arbitrary orientation, but its sensing direction is not rotatable.
- Each sensor knows its location information and determines the locations of its neighboring sensors.
- The target region is on a two-dimensional plane with no obstacle. The boundary of target region can be regarded as a wall-like obstacle.

IV. THE PROPOSED MOVING SCHEME

In order to maximize the sensor coverage, we present a moving scheme for directional mobile sensors, namely the Virtual Force scheme.

A. Virtual Force Scheme

The Virtual Force scheme employs the repulsive force between a directional sensor and each of its neighboring sensors as a basis of movement. The virtual force occurred on a directional sensor is basically generated by the repulsive force between each of its virtual force point and each of its neighboring sensor. The main idea of virtual force scheme aims to repelling a sensor node by its neighboring sensors from dense area to sparse area. We assume that C_{ij} is the set of virtual force points of s_i that are covered by its neighboring sensor s_j . When sensors s_i and s_j are overlapped, they will repel each other by the overlapped region. We denote repulsive force occurred on sensor s_i , which is caused by sensor s_j as \vec{F}_{ij} . Then, \vec{F}_{ij} can be obtained as follows. If sensor s_j is located inside the sensing sector of s_i , s_j will act as a repulsive force from the virtual force point p_{ik} to the sensor s_j , in which p_{ik} is the virtual force point of s_i that has the maximal distance to s_j , according to the following equation:

$$\vec{F}_{ij} = \overline{p_{ik}s_j}, \text{ where } p_{ik} \in C_{ij} \text{ s. t. } d(p_{ik}, s_j) \text{ is maximal.}$$

If sensor s_j is outside the sensing sector of s_i , s_j will act as a repulsive force from sensor s_j to the virtual force point p_{ik} , in which p_{ik} has the minimal distance to s_j , according to the following equation:

$$\vec{F}_{ij} = \frac{R_s}{|s_j p_{ik}|} \cdot \overline{s_j p_{ik}} - \overline{s_j p_{ik}}, \text{ where } p_{ik} \in C_{ij} \text{ and } d(p_{ik}, s_j) \text{ is minimal.}$$

Figure 5 illustrates an example of the repulsion model. From Figure 5(a), we can observe that s_j is outside the sensing sector of s_i , and p_{i2} has the minimal distance to s_j and $p_{i2} \in C_{ij}$. So, the repulsive force exerts on s_i is $\vec{F}_{ij} = \frac{R_s}{|s_j p_{i2}|} \cdot \overline{s_j p_{i2}} - \overline{s_j p_{i2}}$. In Figure 5(b), s_j is inside the sensing sector of s_i and p_{i2} has the maximal distance to s_j and $p_{i2} \in C_{ij}$. So, the repulsive force exert on s_i is $\vec{F}_{ij} = \overline{p_{i2}s_j}$.

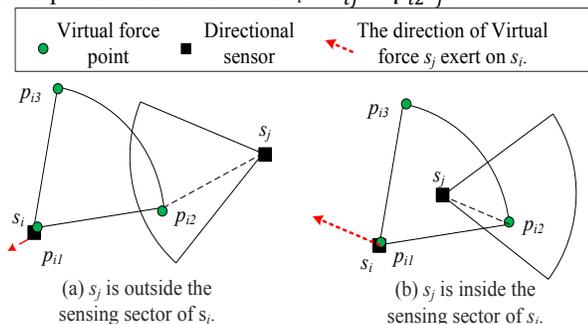


Figure 5. Illustration of repulsion model. s_j exert a repulsive force on s_i .

B. Virtual Force Moving algorithm

According to the repulsion model, we can compute the repulsive forces of each sensor. Then the resultant repulsive force of each sensor is the direction of the new position that sensor should move toward. Figure 6 shows an example that the sensor move toward the direction of the new position with its resultant repulsive force. After repelling by its repulsive force, it really can decrease the overlapped region.

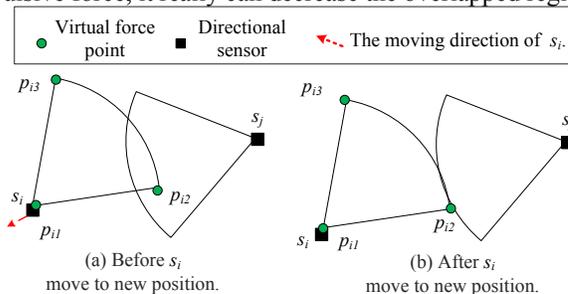
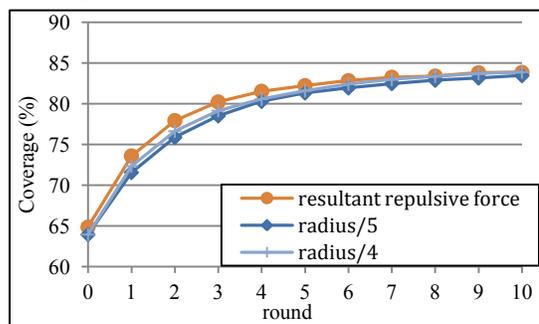
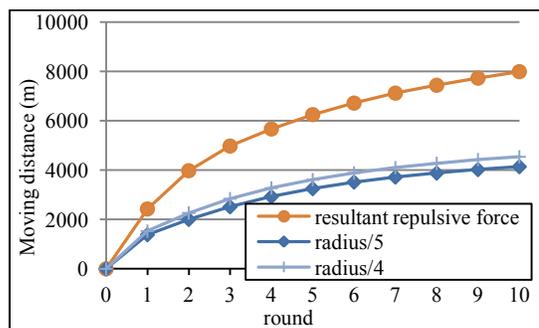


Figure 6. Illustration of moving direction from resultant repulsive forces

In order to save energy, we observe that each sensor does not need to move too far to have better coverage. Thus, in our proposed moving strategy, we take the resultant repulsive force as the moving direction and the moving distance is fixed to radius/5 or radius/4. Figure 7 shows the effect of our moving strategy. As shown in Figure 7(a), the coverage of moving longer distance (i.e. repulsive force) is better than that of moving shorter distance (i.e. radius/5). However, the total moving distance of using radius/5 is significantly better than that of using repulsive force, as shown in Figure 7(b).



(a) Coverage (%)



(b) Moving distance (m)

Figure 7. Illustration of effect of different moving strategies

The movement procedure can be stated as follows: First, the directional sensor determines the direction of movement by the repulsion model. Then the directional sensor checks if the overlapped region with neighboring sensors is decreased by moving to the new destination. If the overlapped region is decreased, the directional sensor will start to move; otherwise, it will stay. The above procedure is called the New-movement-adjustment scheme. Figure 8 illustrates an example of New-movement-adjustment scheme. Then we add an oscillation control on the movement. The purpose of oscillation control is to prevent the sensor from moving back and forth, as shown in Figure 9.

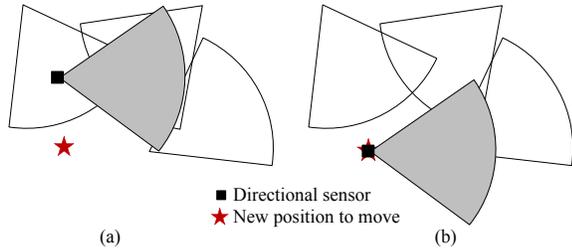


Figure 8. Illustration of the new movement-adjustment scheme. (a) before movement and (b) after movement

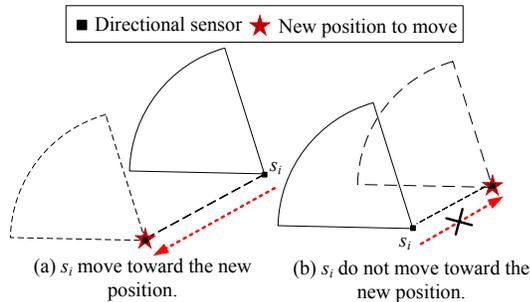


Figure 9. Illustration of the oscillation control

Furthermore, we proposed a move-back scheme to prevent sensors move out of the target region. If the virtual force point of a sensor is out of the target region, sensor should move to the new position until its virtual force point is located on the boundary of the target region as shown in Figure 10. From Figure 10(a), we can see that the virtual points A and B are outside the target region. So, we should move these two virtual points into the target region. This can be done by moving down the sensor with the distance between the virtual point A and the boundary b_1 . After the movement, the virtual point A will be just located on the boundary of the target region as shown in Figure 10(b). Furthermore, the virtual point B is also outside the target region. Thus, the directional sensor will move right with the distance from B to the boundary b_2 . After that, the virtual point B will be also located on the boundary of the target region as shown in Figure 10(c). In Figure 10(d), the sensing sector of the directional sensor will be all inside the target region after moving to the new position. After determining the new position to move, the directional sensor will execute the New-movement-adjustment until new position is reached. Finally, the proposed moving scheme will stop when it

achieves the maximum number of rounds. The complete procedure of Virtual Force Moving algorithm is shown in Figure 11.

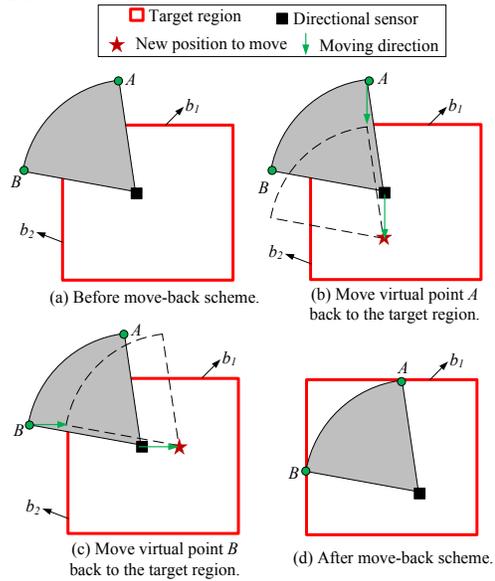
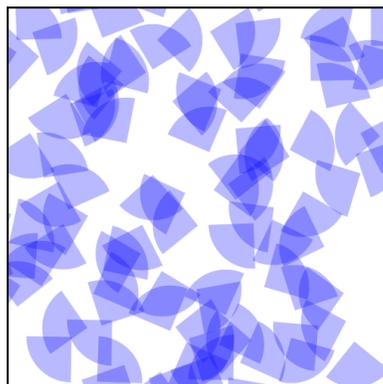


Figure 10. Illustration of the move-back scheme

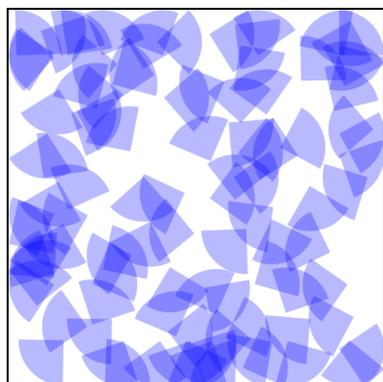
Virtual Force Moving algorithm
<p>Notations : $VPCov_{ij}$, VP_{ik}, $\overrightarrow{S_j P_{ik}}$, FR_{ij} : defined before N_i : the neighbor of sensor S_i \vec{V}_i : moving vector of S_i Max_Round : pre-defined maximum number of round</p> <p>Procedure :</p> <ol style="list-style-type: none"> (1) Enter <i>discovery phase</i> : <ol style="list-style-type: none"> (1.1) set <i>timer</i> to be <i>discovery interval</i> and enter <i>Moving phase</i> upon timeout (1.2) broadcast <i>hello</i> after a random time slot (2) Enter <i>Moving phase</i> : <ol style="list-style-type: none"> (2.1) set <i>timer</i> to be <i>discovery interval</i> and enter <i>discovery phase</i> upon timeout (2.2) Compute the resultant repulsive force <ol style="list-style-type: none"> (2.2.1) $\vec{V}_i = 0$ (2.2.2) for each S_j in N_i If neighbor node is outside the sensing sector and $VPCov_{ij} \neq \emptyset$ and $VP_{ik} \in VPCov_{ij}$ $FR_{ij} = radius - \min(\overrightarrow{S_j P_{ik}})$; $\vec{V}_i = \vec{V}_i + FR_{ij}$ If neighbor node is inside the sensing sector and $VPCov_{ij} \neq \emptyset$ and $VP_{ik} \in VPCov_{ij}$ $FR_{ij} = -\max(\overrightarrow{S_j P_{ik}})$; $\vec{V}_i = \vec{V}_i + FR_{ij}$ (2.2.3) The distance between the new position of S_i and S_j is radius/5 by the moving direction \vec{V}_i (2.3) do oscillation control (2.4) do New-movement-adjustment (2.5) do move-back scheme (2.6) Done when satisfying stop criteria

Figure 11. Procedure of virtual force moving algorithm

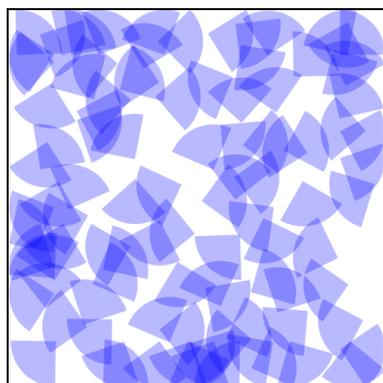
We utilize a case to illustrate the Virtual Force scheme. As shown in Figure 12, we have 100 directional sensors which are randomly deployed in a region of $500 \times 500 m^2$. For each directional sensor, the sensing radius is $60 m$ and the sensing angle is 90° . In Figure 12(a), the initial coverage rate is 61.5712%. After Round 1 (Figure 12(b)) and Round 2 (Figure 12(c)), it can be seen that the coverage ratios are increased to 71.3436% and 76.2856%, respectively.



(a) Initial deployment (61.5712%)



(b) Round 1 (71.3436%)



(c) Round 2 (76.2856%)

Figure 12. Illustration of an example of executing first two rounds of proposed virtual force moving algorithm

V. SIMULATION RESULTS

In this section, we simulate and analyze the performance of Virtual Force scheme from two aspects: coverage and moving distance. Each simulation is executed 10 times then gets the average value. The simulation program is written by C# programming language on .NET platform. We deployed 100 directional sensor of a region of $500 \times 500 m^2$ in our simulation. The sensing radius is $60 m$, the sensing angle is 90° and the number of virtual force point around the boundary of sensing sector is 31. Experimental environment is shown in Table 1.

TABLE I. EXPERIMENTAL PARAMETERS

Network size	$500 \times 500 m^2$
Sensing radius (R_s)	$60 m$
Sensing angle (α)	90°
Number of directional sensors	100
Number of virtual force points	31

The first experiment examines the effect that the number of rounds makes to the performance of coverage rate of our proposed approach on different number of virtual force points. Figure 13 shows the results. In Figure 13, we can see that the more virtual force point a directional sensor has, the more target region can be covered. Thus, we set the number of virtual force points to be 31 in the following experiments.

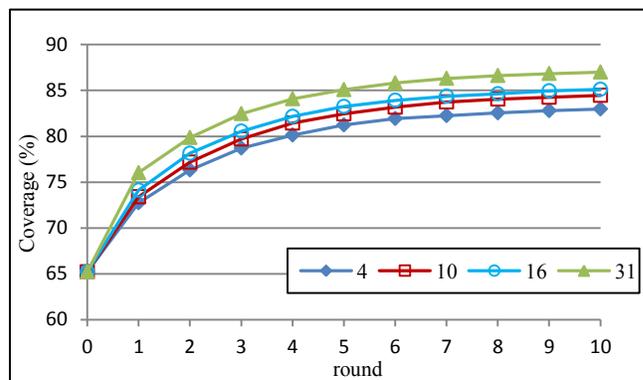


Figure 13. Coverage rate vs. number of rounds

The second experiment evaluates the effect that the number of rounds makes to the performance of accumulated coverage rate of our moving algorithm with 31 virtual force points. Figure 14 shows the results. We can see that our proposed virtual force scheme can increase the coverage rate effectively as rounds increase. This is due to that the directional sensors repel each others from dense to sparse area. Therefore, as rounds increase, the mobile sensors will move to the sparse area. As a result, the coverage holes in the sparse area will be reduced and coverage rate will be increased.

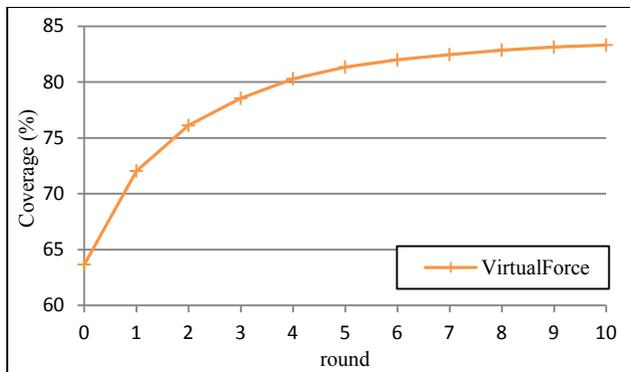


Figure 14. Coverage rate vs. number of rounds – 31 virtual force points

The final experiment examines the effect that the number of rounds makes to the performance of accumulated moving distance of our moving algorithm with 31 virtual force points. Figure 15 shows the results.

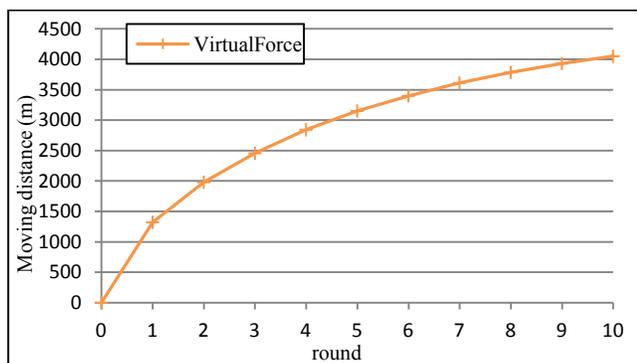


Figure 15. Moving distance vs. number of rounds – 31 virtual force points

We can see that our proposed virtual force scheme will increase the moving distance as rounds increase. This is due to that our approach will move mobile sensors to the sparse area in order to cover the hole area. Therefore, as rounds increase, the moving distance will be increased as well.

VI. CONCLUSIONS

In this paper, we define a new problem regarding how to maximize the area coverage with less moving distance by mobile directional sensors. We propose a scheme, namely the Virtual Force scheme, to improve the coverage. We adopt the virtual force points as the basis of repulsion. Then directional sensor can move toward new position to get

better coverage by this repulsion. Simulation results show that virtual force scheme will increase the area coverage round by round, and the moving distance will be increased when the coverage increases. Specifically, the improvement can be obtained by our proposed Virtual Force scheme up to 30% coverage rate more than initial random deployment.

REFERENCES

- [1] M. Hefeeda and M. Bagheri, "Randomized k-coverage algorithms for dense sensor networks", In: 26th IEEE International Conference on Computer Communications (INFOCOM '07). Anchorage, AK: IEEE Press, 2007, pp. 2376–2380.
- [2] K. Chakrabarty, S. Iyengar, H. Qi, and E. Cho, "Grid coverage for surveillance and target location in distributed sensor networks", IEEE Transactions on Computers, vol.51, no.12, 2002, pp.1448–1453.
- [3] S. Kininmonth, I. Atkinson, S. Bainbridge, G. Woods, G. Gigan, and D. Freitas, "The Great Barrier Reef Sensor Network", Pacem in Maribus XXXI proceedings, Townsville, 2005, pp.361–369.
- [4] I. F. Akyildiz, W. Su, Y. Sankarasubramaniam, and E. Cayirci, "A survey on sensor networks," ACM Trans. on Multimedia Computing, Communications and Applications, Aug. 2002, pp. 102-114.
- [5] M. Rahimi, R. Baer, O. I. Iroezzi, J. C. Garcia, J. Warrior, D. Estrin, and M. Srivastava, "Cyclops: In situ image sensing and interpretation in wireless sensor networks," in ACM Conference on Embedded Networked Sensor Systems(SenSys), 2005.
- [6] J. Djugash, S. Singh, G. Kantor, and W. Zhang, "Range-only slam for robots operating cooperatively with sensor networks," IEEE International Conference on Robotics and Automation, 2006.
- [7] R. Szcweczyk, A. Mainwaring, J. Polastre, J. Anderson, and D. Culler, "An analysis of a large scale habitat monitoring application," ACM Conference on Embedded Networked Sensor Systems (SenSys), 2004, pp. 214-226.
- [8] T. Clouqueur, V. Phipatanasuphorn, P. Ramanathan, and K. K. Saluja, "Sensor Deployment Strategy for Target Detection," WSNA, 2002.
- [9] S. Meguerdichian, F. Koushanfar, M. Potkonjak, and M. B. Srivastava, "Coverage Problems in Wireless Ad-Hoc Sensor Network," Proc. IEEE Infocom, 2001.
- [10] A. Howard, M. J. Mataril, and G. S. Sukhatme, "Mobile sensor network deployment using potential field: A distributed scalable solution to the area coverage problem," Proc. of the 6th Int'l Symp. on Distributed Autonomous Robotics Systems (DARS 2002). 2002.
- [11] O. Khatib, "Real-time obstacle avoidance for manipulators and mobile robots," International Journal of Robotics Research, 5(1), 1986, pp. 90–96.
- [12] G. Wang, G. Cao, and T. L. Porta, "Movement-assisted Sensor Deployment," IEEE Trans. on Mobile Computing, vol. 5, No. 6, June 2006.
- [13] H. Lee, Y. Kim, Y. Han and C. Y. Park, "Centroid-Based Movement Assisted Sensor Deployment Schemes in Wireless Sensor Networks," IEEE 70th Vehicular Technology Conference Fall (VTC2009-Fall), Sept. 2009.

Classification of Human Interactions with Tools Using a Tool-Mounted Wireless Sensor Node to Support Sustainable Manufacturing

Andreas Tilhein, The Duy Nguyen, Jörg Krüger
Institute for Machine Tools and Factory Management,
Technische Universität Berlin
Berlin, Germany
e-mail: theduy.nguyen@iwf.tu-berlin.de

Eduard Wagner, Stephan Benecke, Klaus-Dieter Lang
Research Center for Microperipheric Technologies,
Technische Universität Berlin
Berlin, Germany
e-mail: eduard.wagner@win.tu-berlin.de

Abstract—In this contribution, a customized wireless sensor node with onboard acceleration sensor is applied to support optical motion analysis systems. An integrated signal processing algorithm to classify interactions between user and tool is developed for the specific use case of a cordless screwdriver. Results of the first demonstrator evaluation are discussed with respect to further development of the sensor architecture for smart tools.

Keywords—Human-centered automation; smart tools; wireless sensors; sustainable manufacturing

I. INTRODUCTION

Despite many efforts to automatize manufacturing processes, the human worker is still an essential element in the production line. Humans possess unique skills which recently no robot or machine tool can imitate. Especially when it comes to flexible production, such as in mass customization, employing machines becomes extremely costly due to the need to frequently reconfigure the equipment. On the other hand, qualification and education are essential in order to have flexible workers. Especially industrialized countries with an ageing workforce fear the loss of know-how due to the expected retirements of many experienced workers in the next years. This leaves a lack of staff able to teach the inexperienced beginner. Instead of compensating this loss with simple technological substitution, systems of human centered automation solve the problem by automatically supporting workers in their tasks. These systems enhance the workers' skills, e.g., by amplifying the force or intuitively teach them.

As a core component of these aids, an intelligent sensor system is introduced which is able to provide sufficient information to recognize the actions of the human in order to react appropriately, e.g., by activating actuators or notifying the user. Optical systems suit most of these tasks due to their non-invasive operation principle making bulky equipment to wear obsolete. However, when interacting with objects, e.g., tools, the detection becomes hard due to high variance in appearance and occluded view. Hence, it is proposed to add an additional three-axial acceleration sensor directly at the objects to support the optical systems.

As a dedicated use case for the setup, a teaching system is chosen, which interactively provides the user with information about the currently conducted assembly task.

The sensor information provided here is used to identify the current work step, sequence and proper execution in order to provide the user with feedback. On the basis of the results achieved with the prototypical setup, further developments of wireless sensors are discussed with respect to the defined use-case.

II. RELATED WORK

Past work has focused on gesture recognition or object interactions involving the use of either a complex system of acceleration sensors in combination with RFID sensors [1][2], a microphone [3][4], a gyroscope and a magnetic field [5] or a force sensor [6]. Multi-sensor approaches have been shown to improve recognition performance [7], however, they require more equipment.

Normally accelerometers have been attached to a part of the body or the object whose movement is considered as characteristic for the activity performed [8]. Only the authors in [9] have a tool mounted accelerometer. More commonly used was a wrist worn sensor [1][3][4]. However, a significant disadvantage of attaching the sensors to the body is that, depending on their size and weight, they can noticeably limit the freedom of movement.

As diverse as these sensor combinations, a variety of algorithms have been implemented to detect movement patterns. A variety of approaches to obtain characteristics of the acceleration data for classification have been employed. One option to extract features from the acceleration signal involves their direct derivation from the time domain of the signal. This includes common statistic techniques, such as simple integration methods, mean, standard deviation, skewness, kurtosis, and eccentricity. Furthermore, some researchers have analyzed the signal's frequency spectrum to identify the dominant frequencies. A more recent approach is the wavelet analysis [8]. Here, the original signal is decomposed into a series of coefficients, which contain spectral and temporal information about the original signal. Based on these coefficients, temporal instances with a change in the frequency response of the original signal can be identified [10]. Several researchers have shown that the extraction of features from the time domain of the signal allowed classification performances partly >90% [8][11]. Moreover, since for low sampling frequencies, the detection of time dominant features is superior to the detection of

frequency dominant features [12], we follow a time domain approach.

For the most human activities, hand movements are significant. Yet recognizing the relevant ones in a continuous data stream is difficult. One reason is the hand's high degree of freedom: the same gesture can be performed in different ways. In addition, hands are the most active parts of the body, being constantly in motion even without containing relevant information [3].

Since optical sensors have their problems with classifying fast movements [7] and since they are susceptible to obstruction [4], realizing such a recognition component only with vision based methods is difficult. Especially in a workshop, the distinction between man and tool for the detection of human interactions with tools is problematic.

The implementation of micro system technology (MST) for the described tasks of sensing allows for innovative approaches with respect to the described application. The inclusion of sensor interface, data processing, radio frequency communication and autonomous power supply enables numerous tasks that can be added to the already existing functionalities. Moreover, the high miniaturization potential of MST solutions supports applications that require minimum system size and –weight for least interference with the subordinate technical system and already existing periphery. Small distributed systems are mainly applied for the process monitoring of production equipment [13], but also logistics support with electronic functions beyond RFID identification are a growing sector. Benefits on sustainability from employing an increased number of micro system technology in industrial environments such as manufacturing are currently addressed within the framework of the Collaborative Research Center (CRC) 1026, e.g., [14]. A broader look at the trade-off between benefits and impacts of the additional micro systems as well as the inclusion of teaching tools into sustainability assessment will be part of the ongoing research within CRC1026.

III. CONDUCTION OF THE STUDY

As physical instantiation of the demonstrator, a three-axial acceleration sensor with wireless communication capabilities was attached to a hand-held drilling machine (Fig.1).

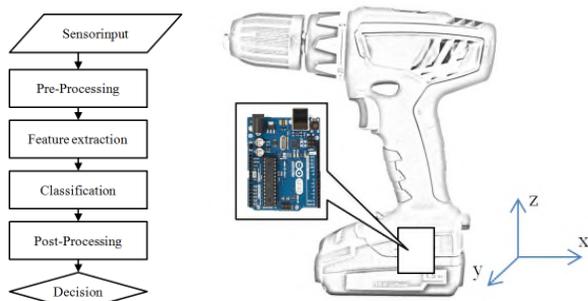


Figure 1. Sensor module attached to drilling machine (right). Pathway of processing the sensor-data till final decision (left).

Five typical drilling activities were examined:

- Picking up the drilling machine from a work table,
- putting down the drilling machine on a work table,
- successively rotating the drilling machine in the sagittal and frontal plane by 90°,
- switching on the drilling machine without contacting the work piece,
- and switching on the drilling machine with contacting the work piece.

In order to develop algorithms that allow the identification of movement/acceleration patterns, the referred activities were recorded five times with short interruptions between each measurement followed by five continuous measurements in a row.

IV. SENSOR HARDWARE AND DATA AQUISITION

The proposed first generation of a sensor node (Tab. II top) is based on open source hard- and software using Arduino UNO R3 board with ATmega328 microcontroller, a customized acceleration sensor layer (ADXL326) and integrated Bluetooth communication interface (BLUETOOTH-SHIELD V2.2). With this setup, wireless communication with a central personal computer is established to transmit all acceleration data required for the classification of human interactions. Through a micro-USB port the Arduino-microcontroller is programmed using customized firmware, covering all data assessment (temperature, 3-axes acceleration), ID of the node and the communication with the Bluetooth interface. Due to the universal functionality of soft- and hardware of the sensor, energy demand is still comparatively high with approx. 110mW in active mode when operating continuously. The applied hardware can therefore be considered as a demonstrator on primary stage for the evaluation of the described scenario of a MST equipped drilling machine. The measurement routine applied is described in the following sections.

V. PRE-PROCESSING

In this section, the pre-processing of the raw data stream is briefly explained. The raw sensor data which is mean adjusted and smoothed. Since the size of the sliding window affects the accuracy and delay of classification, it should be adapted to the characteristic duration of the activities of interest [4]. In order to identify the time intensive activities like picking up or putting down the drilling machine, as well as moments of rest properly, a sliding window with two window sizes of 0.6 s and 1.2 s and a minimal overlap of 1/fs is chosen. With the objective to limit the dynamic range of the acceleration signal to a comparable interval and to implement a state-independent classification, it is essential to align the basic level of acceleration for different machine states on a uniform level. This is realized by a mean adjustment with a signal's outcoming arithmetic mean of zero. The acceleration signal is affected by short frequency signal changes. In order to ensure a reliable detection of

individual activities, signal smoothing is indispensable. For this purpose, a moving average filter has been chosen as it allows efficient computation and quick responses to changes in the amplitude [13]. Despite its simplicity, the moving average filter is optimal for reducing random noise while retaining a sharp step response. On the downside, it allows unacceptable frequency separation [16]. The size of the window in the averaging process has been chosen according to the recommendations in [13]. For fast movements, the recommended time frame ranges from 25 ms to 50 ms, for slow movements, time frames from 100 ms to 200 ms are recommended. Since the activities considered in this work include both – fast and slow movement patterns, an empirical based time frame of 160 ms has been chosen.

VI. FEATURE EXTRACTION

While the selection of features is a critical task for a good recognition performance [4], this paper does not focus on the search for the best possible features. Rather, the goal is to develop an algorithm which leads to an acceptable recognition performance for the underlying problem using low computational effort. One of the main challenges of the pattern recognition task is to distinguish between relevant and non-relevant activities [4]. Characteristic features which allow such discrimination are essential for the classification task. Irrelevant data should be discarded at the same time [17][18]. To do so, the short window activity is quantified by the empirical activity measure (Act) described in [19]. The faster the movement, the greater the change in the acceleration signal is. Act uses these changes of the acceleration vector to detect the sensor's movement represented by a single indicator. While in a static state of approximately zero, the empirical activity of the sensor increases when the sensor is moved. By combining the acceleration of all sensor axes to create one single indicator, movements and non-movements can be distinguished, even if the acceleration changes along one axis. If Act does not meet an empirically determined threshold, the window is classified as non-relevant. Otherwise, the activity is considered relevant and will be examined more in detail.

The goal of implementing a minimum activity criterion is to separate relevant from random, non-relevant information of the acceleration signal. But separation turned out to be difficult, since many diverse movements of the drilling machine exceed the minimum activity threshold, although this information is non-relevant. A comparison of the empirical activity measures of high-frequency and low-frequency ranges of the acceleration signal turned out to be the solution to this problem. It utilizes the circumstance that human movements are characterized by low frequencies [8], fast machine movements by high frequencies. The separation of high and low frequency signal components is carried out by computing the deviation between the smoothed and the original signal.

A. Picking up/ Putting down

If, at the beginning of a machining sequence, the machine is in an upright position, the smoothed vertical component of the acceleration signal is checked for a characteristic pattern

for picking the drilling machine up (Fig. 2). A simple template matching algorithm consisting of a successively scaled sine-function-section turned out to be sufficient for detecting the picking up. As comparison criteria, the maximum and the mean absolute deviation between the smoothed signal and the sine-function are used. For not having to apply the procedure to each data-point within the considered window, the moving standard deviation of the vertical signal component is examined for whether it exceeds a sequence of empirical thresholds. If the picking up of the drilling machine has already been recognized and if the machine is in an upright position, a template matching algorithm is carried out to recognize whether the drilling machine is put down. Except a customized template, this procedure corresponds to that of picking the drilling machine up.

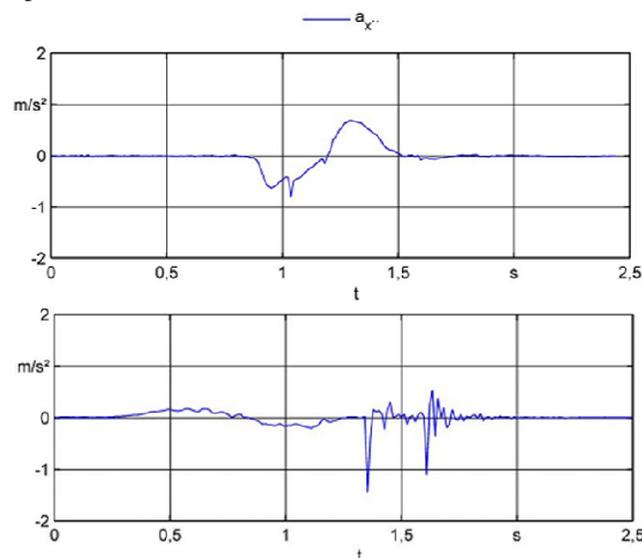


Figure 2. Example signals for activity 'Picking up' (top) and 'Putting down' (bottom).

B. Rotating the drilling machine

To determine whether the drilling machine was rotated in the sagittal or frontal plane, the mean acceleration along all sensor axes in a short time window were calculated and compared with their previous results. If there was an acceleration change of at least ± 0.5 g, activities in the current window were classified as rotating the drilling machine.

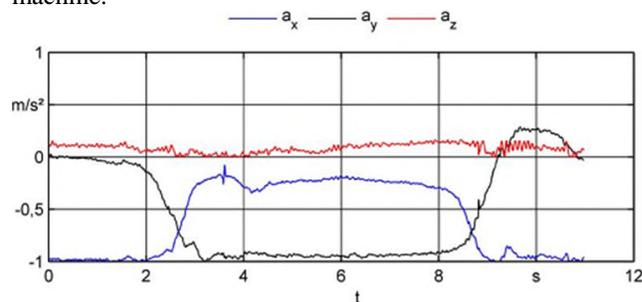


Figure 3. Example signal for activity 'Rotate'.

C. Switching on the drilling machine

If the previously described activities have not been recognized, the signal is analyzed for patterns of drilling movements. The distinction between switching the drilling machine on with and without contact to the work piece is decided via the sensor's horizontal acceleration (Fig. 3). If the drilling machine is switched on with contact to the work piece, low-frequency resonances at high amplitudes along the horizontal axis are observed. This discrepancy is quantified by the empirical activity measure, isolated applied on the horizontal axis.

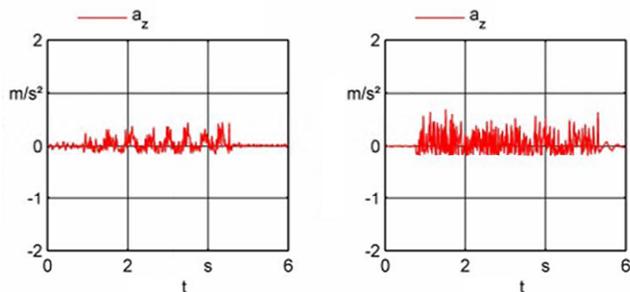


Figure 4. Example signals for activity 'Switch on without contact' (left) and 'Switch on with contact' (right).

VII. CLASSIFICATION AND POST PROCESSING

Although there are a number of classifiers for pattern recognition, to make the classification as understandable and as easily extendable as possible, a simple decision tree has been chosen for classification. A simple decision tree allows a classification of user activities in real time [20] and provides, in terms of movement detection, a good compromise between accuracy and computation time [21][1].

In the first step, the current window of the acceleration signal is analyzed for patterns indicating 'rotation of the drilling machine'. This action can potentially generate an empirical activity higher or less the minimum empirical activity. If this is not the case, there is no relevant information in the observed window. If it exceeds, the drilling machine's position in relation to the vertical axis is identified. If the drilling machine has not yet been picked up from the worktable, the current window is analyzed for typical acceleration patterns for picking up the drilling machine.

In the next step, if the direction of gravitational force is in alignment with the vertical axis and if the drilling machine is already picked up from the working table, the current window is analyzed for patterns of putting the machine down. With sufficient match, the section is assigned accordingly. If there is no match, a comparison of the empirical activity measures of the high- and low-frequency ranges of the acceleration signal is carried out to determine whether there was movement in the window or the drilling machine has been switched on. In a predominance of the high-frequency activity, there is, depending on the empirical activity of the horizontal axis, a switched on drilling machine

with or without contacting the work piece in the observed window.

A subsequent post-processing of the classification results on the basis of empiricism and context knowledge has been indispensable. For this purpose, the classification results of a series of multiple windows are linked adequately to a unitary class.

A. False positive detection of putting down

The false positive detection of putting down the drilling machine is based on the assumption that it indicates the end of a machining process. After that there should be no action until the machine is picked up again. If a signal section has been wrongly assigned to 'putting down', and within a certain time period another empirical minimum activity threshold exceeding acceleration has taken place, the false-positive detection corrects the wrong 'putting down' classification result to 'movement'. The threshold is based on the obtained training data and has been determined by the empirical activity that delimits the hand held from the down put drilling machine. The reason for the wrong assignment is the similarity of the acceleration pattern of 'movement' and 'putting down'.

B. False negative detection of putting down

Depending on the user and surface, there is a large variety of corresponding acceleration patterns for 'putting down' movement of the tool. Since a correct assignment of these patterns is difficult with the developed algorithm, it is necessary to correct the classification result in case the 'putting down' of the drilling machine has not been recognized. Similarly to the false positive detection, the false negative detection is based on the assumption that 'putting down' indicates the end of a machining process, with no subsequent changes in the acceleration signal. Hence, if 'putting down' has not been detected, and if the empirical activity is not exceeding the threshold mentioned before for several time windows, the prior detected activity must have been 'putting down'.

C. Minimum length for drilling

To avoid a false recognition of 'rotating the drilling machine' instead of 'drilling', an empirical based minimum duration for drilling is used.

D. Putting down priority and majority rule

Due to the similarity of the characteristics occurring during 'putting down' with those of moving or switching the drilling machine on, short sequences of nearby windows occur in which an assigned 'putting down' is confused with another activity. In such cases, the putting down priority assigns 'putting down' to the entire window sequence. In sequences of nearby windows with different assignments to 'movement' or 'switching' the drilling machine with or without work piece contact, the majority rule assigns the entire sequence to the activity most frequently encountered in the sequence.

VIII. EVALUATION

The performance of the classification algorithm has been evaluated by determining classification accuracy (percentage of correctly classified test data). Test data from three male and one female operator have been analyzed. Three operators are right-handed. One is left-handed. Each candidate performed a predetermined sequence of following activities:

- picking up the drilling machine from a work table,
- four times of switching the drilling machine on without contacting the workpiece, each in a varied state,
- four times of switching the drilling machine on with contacting the workpiece, each in a varied state,
- putting down the drilling machine on a work table.

Tab. I shows the recognition results in an aggregated confusion matrix. For a sampling frequency of 82 Hz with an overall performance of 96 %, 100 % recognition accuracy has been achieved for ‘picking up’, ‘putting down’ and ‘rotating’. Since all these activities have similar characteristics, the only confusion was at distinguishing ‘switching the drilling machine on’ and ‘with or without contact to the work piece’. However, accuracies of 89% and 91% respectively could still be achieved.

TABLE I. RECONGNITION ACCURACY OF THE IMPLEMENTED ALGORITHM FOR SELECTED MOVEMENTS

Picking up	Rotating	Switching on (no contact)	Switching on (contact)	Putting down	Movement	Total	Class	Accuracy (%)
40	0	0	0	0	0	40	Picking up	100,00
0	400	0	0	0	0	400	Rotating	100,00
0	0	145	7	1	7	160	Switching on (no contact)	90,63
0	0	13	143	0	4	160	Switching on (contact)	89,38
0	0	0	0	40	0	40	Putting down	100,00
40	400	158	150	41	11	800	Total	96,00

Recognition accuracy was then evaluated for different sampling frequencies between 1 Hz to 82 Hz. It could be shown, that for sampling rates below 12 Hz recognition of movement patterns was not successful (Fig. 4).

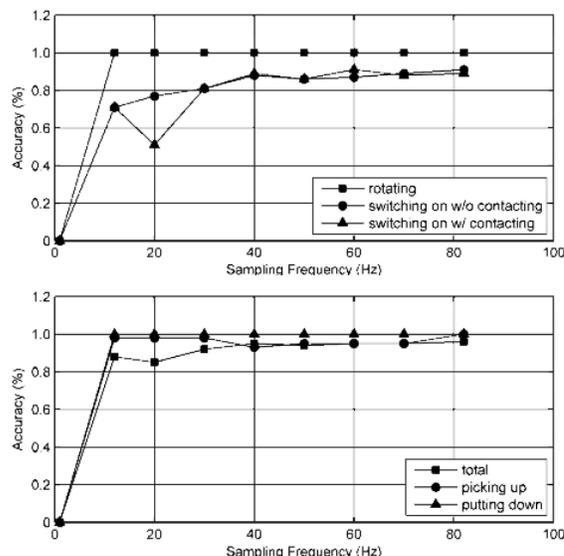


Figure 5. Recognition of movement patterns

However, when sampling rates increase above 12 Hz, movements of ‘putting down’ and ‘rotating’ have been correctly assigned with an accuracy larger than 93 % for ‘picking up’ - Since the high frequency ranges of the acceleration signal gets lost with a decreasing sampling rate [12], the recognition performance for switching the drilling machine on with or without work piece-contact decreases. At lower sampling rates these have been wrongly identified as simple ‘movements’.

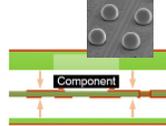
IX. CONCLUSION AND OUTLOOK

In this paper, an algorithm for the identification of acceleration patterns occurring during human interactions with drilling machines has been presented. Five characteristic drilling activities have been examined in a mock up workshop scenario. Data have been gathered from a single tool-mounted three axis acceleration sensor.

With a suitable combination of techniques of preprocessing of the acceleration signal, feature extraction from the time domain of the signal, classification and post-processing, high classification accuracies have been achieved. At a sampling frequency of 82 Hz, the algorithm performs with an overall accuracy of 96 %, with an accuracy of 100 % for picking up; putting down and rotating the drilling machine. For patterns such as ‘putting down’ even further opportunities arise with respect to the realization of low-power sensor systems with reduced sampling rates.

The presented sensor functionalities are currently implemented in advanced MST setups focusing on size reduction and minimized power demand (Tab. II).

TABLE II. FURTHER GENERATIONS OF MST STRUCTRES REDUCING SIZE AND POWER DEMAND

	<p>V1 - Arduino based sensor node with custom layers <i>Purpose:</i> Functional demonstrator for further specification of MT requirements <i>Function:</i> Low resolution temperature sensing, acceleration, ID, communication (Bluetooth) <i>Technology:</i> Commercial sensor platform, customised sensor layers <i>Power consump.:</i> 110mW in active mode; continuous operation</p>
	<p>V2 – Custom sensor node with optimized circuitry <i>Purpose:</i> Evaluation of design approach, debugging, circuitry optimization <i>Function:</i> Precision temperature sensing, 3D acceleration, orientation (compass), identification, optical indication of active status, RF-communication (2.4GHz IEEE 802.15.4) <i>Technology:</i> PCB, 4-Layer, one-sided SMD assembly <i>Power consump.:</i> 58mW in active mode, 30µW in advanced deep sleep mode; Net power consumption 100µW-500µW depending on duty cycle</p>
	<p>V3 – Mini. sensor using advanced packaging tech. <i>Purpose:</i> Demo. of miniaturisation potentials <i>Technology:</i> Size-optimized routing; embedded active and passive devices, bare-Die assembly <i>Function & Power consump.:</i> as V2</p>

ACKNOWLEDGMENTS

This work was funded by the Deutsche Forschungsgemeinschaft (German Research Foundation) within the Collaborative Research Centre (SFB) 1026.

REFERENCES

[1] T. Gu, S. Chen, X. Tao, J. Lu, “An Unsupervised Approach to Activity Recognition and Segmentation based on Object-Use Fingerprints”, *Data and Knowledge Engineering* 69 (2010), pp. 533 - 544.
 [2] P. Lukowicz, A. Timm-Giel, M. Lawo, O. Herzog, “Toward Real-World Industrial Wearable Computing”, *IEEE Pervasive Computing* 6, 4, 2007, pp. 8 - 13.
 [3] J. A. Ward, P. Lukowicz, G. Tröster, “Gesture Spotting Using Wrist Worn Microphone and 3-Axis Accelerometer”, In: *Proceedings of the 2005 Joint Conference on Smart Objects and Ambient Intelligence: Innovative Context-Aware Services: Usages and Technologies*, Grenoble, 12.10. - 14.10.2005.
 [4] J. A. Ward, P. Lukowicz, G. Tröster and T. E. Starner, *Activity Recognition of Assembly Tasks Using Body-Worn Microphones and Accelerometers*. *IEEE Transaction on Pattern Analysis and Machine Intelligence* 28, 10, 2006, pp. 1553 - 1567.
 [5] T. Stiefmeier, et al., “Event-Based Activity Tracking in Work Environments”, In: *IFAWC - 3rd International Forum on Applied Wearable Computing* 2006. Hrsg.: Ottein, H. Berlin, Offenbach: VDE, 2006, pp. 1 - 10.

[6] V. Stanford, “Wearable Computing Goes Live in Industry”, *IEEE Pervasive Computing* 1, 4, 2002, pp. 14 - 19.
 [7] S. Zhou, et al., “2D Human Gesture Tracking and Recognition by the Fusion of MEMS Inertial and Vision Sensor”, *IEEE Sensors Journal* 14 (2014) 4, pp. 1160 - 1170.
 [8] A. Godfrey, R. Conway, D. Meagher, G. ÓLaighin, “Direct Measurement of Human Movement by Accelerometry”, *Medical Engineering and Physics* 30, 2008, pp. 1364 - 1386.
 [9] B. Hartmann, C. Schauer, N. Link, “Worker Behavior Interpretation for Flexible Production. *World Academy of Science*”, *Engineering and Technology* 3, 2009, pp. 494 - 502.
 [10] S. Mallat and W. L. Hwang, “Singularity Detection and Processing with Wavelets”, *IEEE Transactions on Information Theory* 38, 2, 1992, pp. 617 - 643.
 [11] P. Pirttikangas, K. Fujinami, T. Nakajima, “Feature Selection and Activity Recognition from Wearable Sensors”, In: *Ubiquitous Computing Systems - Third International Symposium*. Hrsg.: Youn, H. Y.; Kim, M.; Morikawa, H. Berlin, London: Springer, 2006, pp. 516 - 527.
 [12] A. Krause, et al., “Trading off Prediction Accuracy and Power Consumption for Context-Aware Wearable Computing”, In: *Proceedings of the Ninth IEEE International Symposium on Wearable Computers*, Los Alamitos, 18.10. - 21.10.2005.
 [13] M. Yamaji, Y. Ishii, T. Shimamura, S. Yamamoto, “Wireless Sensor Network for Industrial Automation”, *Ubiquitous Field Computing Research Center Yokogawa Electric Corporation*, Tokyo, 2008.
 [14] K.-D. Lang, et al., “Development of microsystem enhanced machine tool structures for lightweight and accuracy optimised (LEG²O) frames”, *TU-Berlin, Research Center for Microperipheric Technologies*, 2014.
 [15] N. Popovic, “Modellbasierte Erfassung der dreidimensionalen Kinetik der Bewegungen der oberen Extremitäten”, Aachen, Diss., 2012. URL: <http://darwin.bth.rwth-aachen.de/opus3/volltexte/2013/4553/pdf/4553.pdf> (last access: 2015-06-25).
 [16] S. W. Smith, “The Scientist and Engineer’s Guide to Digital Signal Processing”, San Diego: California Technical Publishing, 1997.
 [17] R. O. Duda, P. E. Hart, D. G. Stork, “Pattern Classification”, Weinheim: Wiley, 2001.
 [18] A. Kosmala, “HMM-basierte Online Handschriftenerkennung – Ein integrierter Ansatz zur Text- und Formenerkennung”, Duisburg, Diss., 2000. URL: <http://duepublico.uni-duisburg-essen.de/servlets/DerivateServlet/Derivate-5152/inhalt.htm> (last access: 2014-06-25).
 [19] A. Volmer, “Unterdrückung von Bewegungsartefakten beim Langzeitmonitoring zur Anwendung in Personal-Healthcare-Systemen”, Berlin, Diss., 2011. URL: opus4.kobv.de/opus4-tuberlin/files/2765/volmer_achim.pdf (last access: 2015-06-25).
 [20] Y.-H. Hong, I.-J. Kim, S. C. Ahn, H.-G. Kim, “Mobile Health Monitoring System Based on Activity Recognition Using Accelerometer”, *Simulation Modelling Practice and Theory* 18, 2010, pp. 446 - 455.
 [21] J. Baek, G. Lee, W. Park, B.-J. Yun, “Accelerometer Signal Processing for User Activity Detection”, In: *Knowledge-Based Intelligent Information and Engineering Systems*. 8th International Conference. Hrsg.: Negoita, M. G. et al. Berlin, London: Springer, 2004, pp. 610 - 617.
 [22] U. Maurer, A. Smailagic, D. P. Siewiorek, M. Disher, “Activity Recognition and Monitoring Using Multiple Sensors on Different Body Positions”, In: *Proceedings of the International Workshop on Wearable and Implantable Body Sensor Networks*. IEEE Computer Society, Washington, 03.04. - 05.04.2006

Evidential Network for Multi-Sensor Fusion in an Uncertain Environment

Eric Villeneuve, François Pérès, Cédrik Béler
 Laboratoire Génie de Production
 Université de Toulouse – INPT - ENIT
 Tarbes, France
 e-mail: eric.villeneuve@enit.fr
 francois.perez@enit.fr
 cedrik.beler@enit.fr

Vincente Gonzalez-Prida
 Department of Industrial Management
 School of Engineering, University of Seville
 Seville, Spain
 e-mail: vicente.gonzalezprida@gdels.com

Abstract—Interpreting and quantifying the confidence granted to signals transmitted and received in a sensor network is likely to be called into question by various factors. On an architectural plan, first of all, the nature of the networks or the distance between sensors can induce risk of false alarm or non-detection by misinterpretation of the analyzed signals. External factors related to stresses induced by the environment are also potential sources of measurement errors. Finally, despite the maturity of techniques, internal influence factors related to the accuracy or reliability sensors may also, at a more basic level, impact the confidence placed in the test or the performed diagnosis. A system-embedded intelligence is then necessary to compare the information received for the purpose of decision aiding based on margin of errors converted in confidence intervals. In this paper, we present three complementary approaches to quantify the interpretation of signals exchanged in a network of sensors in the presence of uncertainty.

Keywords: *Sensor networks; Uncertainty; Bayesian techniques; Belief functions; Evidential networks.*

I. INTRODUCTION

Within the field of science and engineering, data imperfection requires the use of tools to define mechanisms for reasoning with partial knowledge and uncertain information. In [1], several types of imperfections are discerned:

- Incompleteness and vagueness are used to qualify the status of a data. It is said to be incomplete if it is impossible for the source to provide information regarding all or part of the aspects of a problem. Vagueness is a form of incompleteness for when the source provides an imprecise data, the resulting information is necessarily incomplete.
- Uncertainty applies when the source is unable to distinguish the veracity of a piece of information (that is to say whether the information is true or false). It therefore characterizes the extent of information compliance compared to reality. It is possible to distinguish two kinds of uncertainty. Random uncertainty is induced by the variability of an entity in a population and is the outcome of random experiments. This type of uncertainty cannot be reduced since it is the result of chance. Epistemic uncertainty is due to the lack of knowledge and therefore relates to the concept of incompleteness.

- Ambiguity represents the fact that a same information can have several interpretations. It is therefore linked to the formalism of information representation which is not always clear and shared by all the stakeholders. This type of imperfection is very common and is often a source of misinterpretation. It can be avoided when formalizing the representation.
- Granularity of information characterizes the difficulties that appear when two very close values have to be distinguished.

A decision requires integrating this imperfection to justify the actions that will be undertaken. In the field of sensor networks, decisions must be taken at every moment with respect to signals individually emitted by each sensor or regarding the signals received and compiled by a centralized basis. Taking into account imperfect data allows the decision maker to legitimate its choices since he will have at its disposal additional knowledge associating the data with an interval of confidence and a margin of error.

In the following are presented two distinct approaches based on Bayesian Networks (Section II) and Belief Functions (Section III) for integrating and propagating imperfect data in a network of sensors. In Section IV, it is shown how both techniques can be combined and reinforced each other within the same tool, namely evidential networks. Finally, in Section V, a comparative view of the different techniques is presented to show their respective conditions of use.

II. BAYESIAN NETWORK APPROACH

Bayesian Networks are graphic models designed to formalize knowledge with the purpose of reasoning about a problem.

A. Principles

Bayes theorem is central in the mechanism of inference in Bayesian Networks. It makes the link between a series of hypotheses, characterized by probabilities of occurrence, and a series of observations representing the actual state of the system. From the Bayes theorem, it is possible to implement two types of reasoning:

- Diagnosis, or reasoning by backward inference, which allows, in the set of assumptions made, the identification of the probable cause of a given result,
- Prognosis, or reasoning by forward inference, which enables the estimation of the occurrence probability of an evidence with respect to the formulated assumptions.

From a mathematical point of view, the space of hypothesis $\theta = \{\theta_1, \dots, \theta_i, \dots\}$ and observations $\Omega = \{\omega_1, \dots, \omega_k, \dots\}$ are defined. To represent the link between observations and assumptions, the probability theory allows the use of a conditional probability distribution based on each verified hypothesis. This distribution can be noted $P^\Omega(\cdot|\theta_i)$.

Whether there is a knowledge regarding the values of the hypothesis $\theta_i \in \theta$, it is then represented in the form of a conditional probability distribution across the observation Ω , noted $P^\Omega(\omega_k|\theta_i)$, which characterizes the likelihood of observation ω_k knowing hypothesis θ_i . Bayes theorem can then be used to provide a reasoning forward inference to determine the most probable cause associated with this evidence by calculating the a posteriori probability distribution $P^\Theta(\cdot|\omega_k)$.

$$P(\theta_i|\omega_k) = \frac{P(\theta_i) \times P(\omega_k|\theta_i)}{\sum_{\theta_i \in \Theta} P(\theta_i) \times P(\omega_k|\theta_i)} \quad (1)$$

When the probability distribution $P(\cdot|\theta_i)$ is known (i.e., the probability of occurrence of an observation given each hypothesis) and a hypothesis $P(\theta_i)$ is assumed, Bayes theorem allows the implementation of reasoning backward inference to estimate the effect of a hypothesis on an evidence $P(\omega_k)$. By calculating the probability distribution on the evidence it is then possible to predict the most likely one:

$$P(\omega_k) = \sum_{\theta_i \in \Theta} P(\theta_i) \times P(\omega_k|\theta_i) \quad (2)$$

Bayesian Networks are used to formalize knowledge in the form of a causal graph associated with a probability space. They are acyclic directed graphs where knowledge is represented by variables. Each node of the graph corresponds to a variable and arcs represent the probabilistic dependencies between these variables. Formally, a Bayesian network is defined by [2]:

- a graph-oriented without circuit, noted $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, with \mathcal{V} , the set of nodes of \mathcal{G} , and \mathcal{E} , the set of arcs of \mathcal{G} ,
- a finite probability space $(\Omega, \mathcal{A}, \mathcal{P})$, where Ω is the universe, i.e., the set of all the elements considered in the problem, \mathcal{A} is a σ -algebra on Ω and \mathcal{P} is a measure on Ω such that $\mathcal{P}(\Omega) = 1$,
- a set of random variables defined on $(\Omega, \mathcal{A}, \mathcal{P})$, corresponding to each node of the graph, such that the set of probabilities associated with these variables defines the distribution of probabilities attached to the network:

$$\mathcal{P}(\mathcal{V}_1, \mathcal{V}_2, \dots, \mathcal{V}_n) = \prod_{i=1}^n \mathcal{P}(\mathcal{V}_i | pa(\mathcal{V}_i)) \quad (3)$$

with $pa(\mathcal{V}_i)$, the parent set (also called predecessors or causes) of \mathcal{V}_i in graph \mathcal{G} . There are two types of probability tables in Bayesian Networks [2]. Tables of prior probabilities (table II.2) characterizes the chances that the variable \mathcal{V}_a without any parent is in state a_i . Tables of conditional probabilities (table II.2) establish the chances that a variable \mathcal{V}_b is in state b_j based on the state of his parents.

Inference in a Bayesian network consists in propagating information in the network. Indeed, a model using this formalism is generally not intended to be a static

representation of knowledge. Beyond the a priori reasoning, evidences may be introduced to update the observed situation and to insert into the model the changes enabling the refinement of the results [3]. This new knowledge, takes the form of a so-called elementary information, denoted J , relative to a particular node. There are two types of basic information. The deterministic information allows instantiating a variable, which is affecting it a precise value, (i.e., $\mathcal{P}(\mathcal{V}_a = a_1|J) = 1$). The imprecise information modifies the distribution of probability of the variable, either by excluding a value of the universe of the variable ($\mathcal{P}(\mathcal{V}_a = a_1|J) = 0$) or, more usually, by changing the law ($\mathcal{P}(\mathcal{V}_a = a_1|J) \neq \mathcal{P}(\mathcal{V}_a = a_1)$).

B. Case Study

The growing need of wiring in avionics, automotive, telecommunications, nuclear plants, buildings, etc., has caused the increase of cable length moving from 200m up to 4km in a modern car. The type of a cable (coaxial, twisted pair, optic fiber, etc.) depends on the nature of the propagating signal (data and energy) into network, the corresponding voltage level and the environment (noise, temperature, vibration, etc.) in which the cable is implemented. One day or another, a cable will show signs of damage involving the appearance of faults (short and open circuit, aging, etc.). These faults can be a consequence of environmental stress (heat, moisture, chafes, etc.) and a cause of dramatic mishaps such as TWA flight 800 in 1996. Therefore, a wiring diagnosis system is needed to detect and locate faults as early as possible. Reflectometry is a suitable diagnosis technique as it requires a single access point to inject a test signal into the cable network. During its propagation, a part of its energy is reflected back to the access point at each impedance discontinuity (fault, junction, etc.). Then, the analysis of the reflected signals, commonly called "Reflectogram", permits to characterize this discontinuity. In the literature, several reflectometry methods are proposed depending on the studied domain and the type of the used test signal [4]. Although standard reflectometry has proven its efficiency in wire fault detection, it suffers from ambiguity problems related to fault location in branched networks. As a solution, a distributed diagnosis strategy is proposed [5].

It consists in using multiple diagnosis systems, called "reflectometer", to make reflectometry measurements at many extremities of the cable network. Here, the major problematic involves the diagnosis system reliability, number and location, signal processing, resource allocation, communication protocol, etc. Based on the uncertainty regarding diagnosis system failure, measurement precision and fault location, the use of Bayesian networks is motivated by the combination of deterministic and stochastic behaviors of such systems of diagnosis [4].

Figure 2 shows the computed reflectogram for the branched network of Figure 1 with an open circuit fault at a distance of 25m from the injection point. Only one reflectometer is placed at the extremity of L1 to diagnose the whole network. The reflectometer and the network are

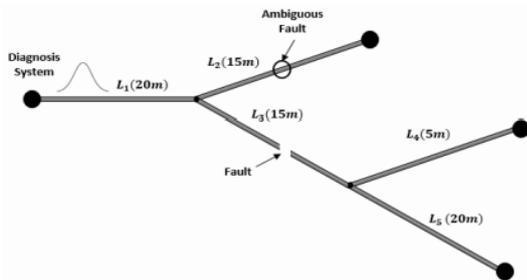


Figure 1: Fault location ambiguity in a branched network

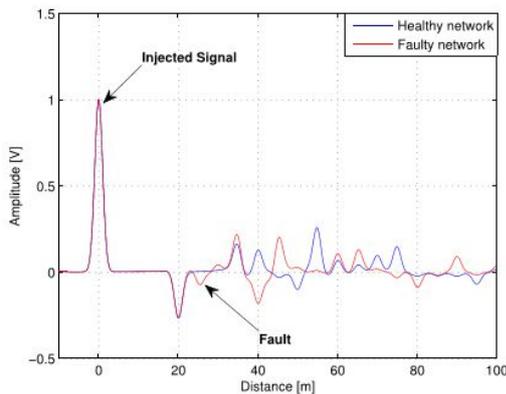


Figure 2 : Obtained Reflectogram using TDR method

considered unmatched, explaining the first positive peak on the reflectogram. The end of lines are also unmatched. Here, the detected fault on L3 cannot be distinguished from the same fault on L2. It is then possible to add another reflectometer at the end of line L. Although, the ambiguity problem is resolved for the fault on branch L, it remains inevitable if another fault appears on the branch L. So, another reflectometer should be added to overcome this ambiguity, which increases the diagnosis cost (number of reflectometers) [6]. As a solution, we proposed to introduce the cable life profile aiming to cancel the fault location ambiguity problem with a low cost. Considering the uncertainty regarding faults location, measurement precision and reflectometer reliability, the use of Bayesian Networks appeared to be appropriate to combine these deterministic and stochastic behaviors [7].

The proposed strategy includes two steps: (1) local diagnosis based on BNs, (2) global diagnosis for the whole network to locate the detected fault(s) [8]. In local diagnosis, each reflectometer introduces the cable life profile to calculate the conditional probability of the presence of the fault on each branch. Then, the obtained results for each reflectometer are integrated into a global BN to locate the fault in the whole network. Simulation results prove the efficiency of the proposed strategy to cancel or mitigate the ambiguity for fault location in a branched network with respect to the reflectometer reliability.

Since this is not the main purpose of the paper, please refer to [6] for more details on the procedures.

III. TRANSFERABLE BELIEF MODEL APPROACH

Based on the work of Dempster [9] generalizing the theory of probabilities by using probability intervals, the belief function theory was developed by Shafer [10] to provide a general framework for the representation of uncertainties.

A. Principles

This combined approach is commonly referred to as the Dempster-Shafer Theory (DST). It is used to represent information based on the belief or the state of knowledge of an entity (person, sensor, etc.) and provides a very rich mathematical framework for:

- the characterization of partial knowledge (including total ignorance)
- the fusion of information from various homogeneous or heterogeneous sources
- the modelling of uncertainty (random and epistemic) related to the state of a system
- the assessment of the degree of confidence associated with a result

The theory of belief is particularly well suited to the representation of different forms of uncertainty. It allows the modelling of problems where the lack of information prevents the reasonable use of probability theory. Belief techniques are used in many fields (decision, data analysis, classification, diagnosis, multi-sensor perception, image processing, etc.) for the very varied tasks such as pattern recognition, likelihood analysis or information merging.

This formalism evolved and saw the emergence of a second approach, developed by Smets in [11]. This approach, called Transferable Belief Model (TBM) is a subjectivist interpretation of the theory of belief functions and allows the analyst to overcome the notion of probability. The TBM approach introduces a decision-making level (also called pignistic level) that irreversibly transforms beliefs (non-probabilistic expression) in a probability-based consistent form to facilitate decision making.

The TBM approach enables the representation of imperfect knowledge from multiple sources [11]. It is based on the assumption that reasoning in uncertainty (credal level) and decision-making (pignistic level) are two cognitive tasks of different kinds:

- The credal level corresponds to the representation and manipulation of the belief statements (without using for example equiprobable distribution).
- The pignistic level makes possible the decision-making by transforming the subjective measures of non-probabilistic beliefs into a measure of probability. This transformation which allows the consideration of risk or bet notions, intervenes only at the time of the decision-making and does not alter the credal level.

Modelling a problem using belief functions, requires to determine the value of the variable ω , which represents the system states. The frame of discernment (FoD) represents all possible n values (or hypothesis) for the variable ω and is denoted by Ω .

$$\Omega = \{\omega_1, \omega_2, \dots, \omega_i, \dots, \omega_n\} = \bigcup_{i=1}^n \{\omega_i\} \quad (4)$$

The definition of the belief mass function, denoted m^Ω , allows the translation of an observation provided by an agent on the power set of Ω (denoted 2^Ω). The power set corresponds to all the subsets that can be formed from the assumptions and the unions of assumptions of Ω . Consequently, the belief mass function is defined as:

$$m^\Omega : 2^\Omega \rightarrow [0, 1] \quad (5)$$

$$A \mapsto m^\Omega(A)$$

with $2^\Omega = \{\emptyset, \{\omega_1\}, \{\omega_2\}, \{\omega_1, \omega_2\}, \{\omega_3\}, \{\omega_1, \omega_3\}, \{\omega_2, \omega_3\}, \dots, \{\omega_1, \dots, \omega_n\}\}$

Notation $m_S^\Omega(A)$ stands for the identification of a source S which provides information on proposal A . The basic belief assignment (bba) is the set of belief masses on the proposition $A \subseteq \Omega$ that satisfies:

$$\sum_{A \subseteq \Omega} m_S^\Omega(A) = 1 \quad (6)$$

A bba can be transformed to highlight information and to improve the dynamic aspects of TBM, including the fusion rules. The credibility measure or belief (denoted $bel^\Omega(A)$) symbolizes the minimal belief in a hypothesis A . It is the part of belief specifically attributed to A (without the part corresponding to the empty set). The plausibility measure (denoted $Pl^\Omega(A)$) represents the maximal belief in a hypothesis A . It is the sum of all the masses for A . Both measures are defined as:

$$bel^\Omega(A) = \sum_{\emptyset \neq B \subseteq A} m^\Omega(B), \forall A \subseteq \Omega \quad (7)$$

$$Pl^\Omega(A) = \sum_{B \cap A \neq \emptyset} m^\Omega(B), \forall A \subseteq \Omega \quad (8)$$

The transformation from credal level to pignistic level is called pignistic transformation [11]. This non-reversible transformation aims to reduce the bba in a probability distribution so as to be compatible with decision theory. Hence, the mass of proposal A is distributed using equiprobability on the singletons of the FoD, Ω , i.e., on the hypotheses related to A . Therefore, this probability distribution, denoted $BetP$, can be obtained as:

$$BetP\{m^\Omega\} : \Omega \rightarrow [0, 1] \quad (9)$$

$$\omega_i \mapsto BetP\{m^\Omega\}(\omega_i)$$

With regard to decision-making itself, it is to select the singleton ω_i of the FoD Ω having the highest pignistic probability to maximize the chances that the hypothesis symbolized by ω_i represents the actual state of the system (or the smallest, if for instance, the goal is to minimize the probability of occurrence of an event). It is therefore to maximize the expected utility in order to rationalize the decision. From a mathematical standpoint, this decision can be expressed by:

$$\omega = \begin{matrix} \text{argmin} \\ \text{argmax} \end{matrix} BetP\{m^\Omega\}(\omega_i) \quad (10)$$

B. Case Study

An example of an illustration of the concept of belief functions applied to sensor networks can be found in [12]. The author of the work deals with the area of Wireless Sensor Networks (WSN) based especially on infrared or ultrasonic techniques. These networks are made up of smaller equipment whose number varies from hundreds to hundreds

of thousands. They can be deployed on one or more locations to observe a particular phenomenon and their mutual collaboration enables the triggering of alerts or the gathering of information on a supervised phenomenon.

The number and the position of the sensors deployed in an area of interest, determine the topology of the network and characterizes its intrinsic properties in terms of coverage, connectivity, cost and life. Therefore, the performance of the WSN depends mainly on the method used in the deployment of sensors. The network architecture is intimately linked to the reliability of the information transmitted by the sensors which determines the quality of network coverage. The belief functions are used in this work to analyze the problem of management of imperfections related to the uncertainty in the data gathering process.

The construction of evidence is based on two states required to specify whether a space point $p \in RoI$ is covered (θ_1) or uncovered (θ_0). Thus, the FoD is the set $\theta = \{\theta_0, \theta_1\}$. Let s be a sensor, R_s be its sensing range and R_u be a distance ($0 \leq R_u \leq R_s$) (Figure 3).

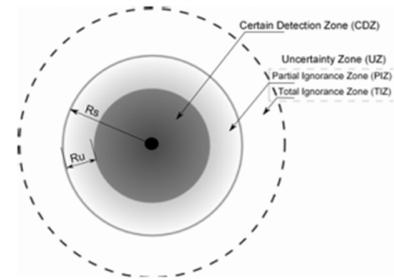


Figure 3: Sensor sensing range representation

Each sensor s provides information on the coverage of a space point $p \in RoI$ (Region of Interest) with a belief $b_{s/p}$. The complementary information $1 - b_{s/p}$ is assigned to the whole FoD because it encodes the sensor ignorance. The output from the sensor s about a space point $p \in RoI$ can thus be represented as a bba $m_{s/p}$ with two focal sets: the singleton $\{\theta_1\}$ and the FoD θ .

$$m_{s/p}(\{\theta_1\}) = b_{s/p}, b_{s/p} \in [0, 1] \quad (11)$$

$$m_{s/p}(\theta) = 1 - b_{s/p}$$

$$m_{s/p}(\emptyset) = 0$$

Relatively to a space point p , a sensor s provides $m_{s/p}$ as a belief function. To decide whether p is covered by s , the pignistic transformation of $m_{s/p}$ (denoted by $BetP_{s/p}$) is constructed. The decision is based on selecting the hypothesis $\hat{\theta}$ with the largest pignistic probability $\hat{\theta} = \underset{i=0,1}{\operatorname{argmax}} BetP_{s/p}(\{\theta_i\})$.

A space point p is covered by a sensor s if:

$$\hat{\theta} = \theta_1 \quad (12)$$

$$BetP_{s/p}(\{\theta_i\}) = Th_p$$

The threshold (Th_p) value is an application-specific user-specified parameter.

One current application of WSNs is target/event detection; the sensors collaborate to arrive at a consensus decision as to whether a target is present in the RoI. A TBM-based approach is considered for reaching this consensus: one of the sensors (called a fusion center: it can be a sink, a cluster-head or any sensor), gathers the evidence from the other sensors, combines the evidence and decides whether an event is present or not.

IV. EVIDENTIAL NETWORKS

In order to take advantage of the two approaches mentioned so far, it is possible to combine TBM and Bayesian Networks.

A. Principles

Basic approaches enable this integration but are limited to binary variables [13]. Evidential networks, introduced in [14] and developed in [15], use conditional belief functions for belief propagation in acyclic oriented graphs. Each node of the evidential network represents a random variable that is associated with a finite number of values. Directed evidential networks were introduced in [16].

In the same way as evidential networks, they use conditional belief functions to propagate the belief in graphs without circuit. This formalism is very close to Bayesian networks but uses conditional beliefs instead of conditional probability functions [17]. Each arc of the graph represents a conditional relationship between two variables, represented by nodes. Each variable is set to a FoD representing the set of values it can take. Parent nodes are characterized by a priori belief functions whereas child nodes are represented by belief functions conditioned by their parent values.

The Generalized Bayes Theorem (GBT) and the Disjunctive Combination Rule (DCR) are used to infer and propagate the knowledge in the network. Figure 4 represents a very simple directed evidential network with only two nodes, Θ and Ω and one arc representing the causal relationship between the nodes. From the perspective of the variables, which are represented by nodes, the edge indicates that node Ω , which represents the observation space characterized by a priori belief mass distribution, noted m_Ω^Ω , is conditioned by node Θ , standing for the hypothesis space characterized by a conditional belief mass distribution $m^\Omega[\Theta]$.

There are several ways to propagate knowledge in a directed evidential network depending on the node receiving the current knowledge expressed in the form of a new belief mass distribution [14].



Figure 4: Basic directed evidential network

The knowledge is spread in the direction of the arc if the Θ node receives a new distribution of masses m^Θ , the Ω node

is then updated taking into account this new information. This type of spread, called forward propagation enables the calculation of the distribution of plausibility Pl^Ω using the equation of the GBT based on an a priori knowledge. This equation uses the DCR rule to determine the plausibility $Pl^\Omega[\theta_i]$ for each subset $\theta_i \in \Theta$ and evaluate Pl^Ω by combining the conditional plausibilities $Pl^\Omega[\theta_i]$ using the Conjunctive Combination Rule (CCR).

The knowledge can also be propagated in the opposite direction of the arc if the Ω node receives a new distribution of masses m^Ω . The information contained in the Θ node is therefore updated to take account of this new information. This type of spread, called backward propagation enables the calculation of the plausibility distribution Pl^Θ by using the following equation:

$$Pl^\Theta(\theta) = \sum_{\omega \in \Omega} m^\Omega(\omega) \times (1 - \prod_{\omega_i \in \omega} (1 - Pl^\Theta[\omega_i](\theta))), \forall \theta \in \Theta \quad (13)$$

B. Case Study

A case study using evidential network applied to sensor networks is being processed (Figure 5). It is related to Advanced Sensor Technologies for Nondestructive Testing (NDT) in the field of aeronautics Structural Health Monitoring (SHM). Sensors based on ultrasonic or Eddy Current techniques are arranged on an aircraft structure to measure the level of damage. Functional recycling of products or just testing the integrity of an aeronautical structure (fuselage or wing of an airplane for example) requires a diagnosis of their current state.

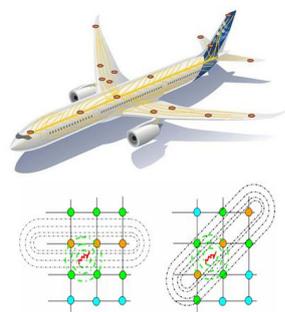


Figure 5: Instrumented aircraft monitored by sensor networks

The main goal is to detect, locate and identify defects, and, subsequently, follow their evolution. The use of NDT techniques allows the decision-maker to call on the compliance of the parts analyzed by quantification of the level of damage. The discussed example aims to develop an informational interface based on an evidential networks superimposed on the NDT sensor network for the interpretation of the signals received and the decision support. The position and the number of sensors lead to a level of uncertainty characterized by belief functions introduced and propagated in the evidential network.

As a possible illustration of the interest of using evidential networks and with respect to the sensor architecture, as well as the emitted and received signals, one can use here this technique both to:

TABLE I. COMPARISON OF FORMALISMS

	Uncertainty kind	Amount of data	Causality modeling	Computational complexity
Bayesian networks	Random uncertainty	High (frequentist approach) Low (Subjectivist approach)	Yes	Low to Mean (depending on the complexity of the network)
Belief functions	Random and epistemic uncertainties	Low	No	Low to mean (depending on the number of singletons of the variables)
Directed evidential networks	Random and epistemic uncertainties	Low	Yes	Very High

- predict the presence of a defect in a forward propagation of the marginal belief functions knowing the signal values
- predict the detection of a defect in a backward propagation of the conditional belief function knowing the existence of a proven deterioration or failure.

We may observe also that upon receiving the information where a particular defect has been detected by any one of several distinct sensors, one can identify the source as belonging to this set. Therefore, one can assign the belief that the signal is present to the set of possibilities. The use of a Bayesian network would have forced to distribute this belief over all the individual nodes (sensors). Evidential reasoning avoids so the need for assumptions on missing data. When beliefs on a sensor is later required they are under constrained resulting from the disjunction and an interval representation is needed to capture the real constraint which enable the explicit representation of uncertain and certain knowledge.

V. COMPARISON OF FORMALISMS

As we mentioned above, each presented formalism aims at modeling and reasoning under uncertainty. However, their conditions of application are different and some are better than others in given circumstances. Table 1 positions the three formalisms (Bayesian Networks, belief functions, evidential Networks) according to four criteria: the kind of uncertainty addressed, the minimum amount of data required, the capacity to model causality and the computational complexity (in relation with the computation time).

VI. CONCLUSION

The purpose of this paper was to present several techniques to introduce the notion of uncertainty in the use of sensor networks. The underlying idea is to be able to assess the measurement error and the corresponding risk to reduce the oversizing of the monitoring architectures and better define the level of confidence placed in the information received from the network. After presenting approaches respectively based on Bayesian Networks and Belief Functions, Evidential Networks have been introduced. The interest of their use lies in their ability to spread non-probabilistic knowledge according to forward or backward modes of propagation to facilitate dynamic decision making.

REFERENCES

[1] D. Dubois, and H. Prade, Formal representation of uncertainty, chapter 3, UK Wiley, 2009.

[2] F. Jensen, and T. Nielsen, Bayesian networks and decision graphs. Springer, 2007.

[3] M. Godichaud, A. P. Tchangani, F. Pérès, and B. Iung, "Sustainable management of end-of-life systems," Production Planning & Control vol. 23 (2-3), 2012, pp. 216-236.

[4] W. Ben Hassen, F. Auzanneau, F. Pérès, and A. P. Tchangani, "Diagnosis sensor fusion for wire fault location in CAN bus systems," In Sensors, 2013 IEEE, 2013, pp. 1-4.

[5] W. Ben Hassen, F. Auzanneau, L. Incarbone, F. Pérès, and A. P. Tchangani, "OMTDR using BER estimation for ambiguities cancellation in ramified networks diagnosis," In IEEE Eighth International Conference on Intelligent Sensors, Sensor Networks and Information Processing, Melbourne, 2013, pp. 414-419.

[6] W. Ben Hassen, F. Auzanneau, L. Incarbone, F. Pérès, and A. P. Tchangani, "Distributed Sensor Fusion for Wire Fault Location Using Sensor Clustering Strategy", International Journal of Distributed Sensor Networks, vol. 2015, 2015, doi: 10.1155/ijdsn.2015.538643.

[7] W. Ben Hassen, F. Auzanneau, L. Incarbone, F. Pérès, and A. P. Tchangani, "Distributed Reflectometry Method for Wire Fault Location using BER in CAN Bus," Transactions on systems, Signals and Devices, 2014.

[8] W. Ben Hassen, F. Auzanneau, F. Pérès, and A. P. Tchangani, "Ambiguity Cancellation for Wire Fault Location based on Cable Life Profile", IFAC World Congress, Cape Town, 2014, pp. 9593-9598.

[9] A. Dempster, "Upper and lower probabilities induced by a multivalued mapping," Annals of Math. Statistics, vol. 38, 1967, pp. 325-329.

[10] G. Shafer, "A Mathematical Theory of Evidence, " Princeton University Press 1976.

[11] P. Smets, and R. Kennes, "The transferable belief model," Artificial Intelligence, vol. 66, 1994, pp. 191-234.

[12] M. R. Senouci, A. Mellouk, L. Oukhellou, and A. Aissani, "Using the belief functions theory to deploy static wireless sensor networks," Belief Functions: Theory and Applications, Springer Berlin Heidelberg, 2012, pp. 425-432.

[13] C. Simon, P. Weber, and E. Levrat, "Bayesian networks and evidence theory to model complex systems reliability," Journal of Computers, vol. 2-1, 2007, pp. 33-43.

[14] P. Smets, "Belief functions: The disjunctive rule of combination and the generalized bayesian theorem," International Journal of Approximate Reasoning, vol. 9, 1993, pp. 1-35.

[15] H. Xu, and P. Smets, "Evidential reasoning with conditional belief functions," Proceedings of the 10th Conference on Uncertainty in Artificial Intelligence, 1994, pp. 598-605.

[16] B. Yaghlane, P. Smets, and K. Mellouli, "Directed Evidential Networks with Conditional Belief Functions," LNCS, vol. 2711, 2004, pp. 291-305.

[17] E. Villeneuve, C. Beler, F. Pérès, and L. Geneste, "Hybridization of Bayesian Networks and belief functions to assess risk. Application to aircraft disassembly," IESM International Conference on Industrial Engineering and Systems Management, Metz, 2011, pp. 1070-1079.

[18] F Pérès, C Martin *Design methods applied to the selection of a rapid prototyping resource* ETFA 1999

[19] AP Tchangani, F Pérès *BOCR framework for decision analysis Large Scale Complex Systems* Theory and Applications 9 (1), 507-513 2010

[20] AP Tchangani, Y Bouzarour-Amokrane, F Pérès *Evaluation model in decision analysis: bipolar approach* Informatica 23, 461

Near Real-Time Oceanographic Data Management through Sensor Web Enablement (SWE) Standards

Latest developments

E. Partescano, A. Giorgetti, A. Brosich

OGS (Istituto Nazionale di Oceanografia e di Geofisica Sperimentale)
Borgo Grotta Gigante-Sgonico (TS), Italy
e-mail: epartescano@ogs.trieste.it

Abstract—Realizing a system to harmonize and disseminate heterogeneous data, collected by different meteo-oceanographic buoys is a challenge. In this paper, we describe a procedure for data validation, conversion in a homogeneous and standard format and dissemination adopting Sensor Web Enablement (SWE) using XML (eXtensible Markup Language) and OGC (Open Geospatial Consortium) standards. To meet the needs of different scientific communities as RITMARE (La Ricerca Italiana per il MARE), Jerico (Towards a joint European research infrastructure network for coastal observatories), MyOcean, ODIP (Ocean Data Interoperability Platform), FixO3 (Fixed-point Open Ocean Observatories), we decided to adopt SWE standard allowing interoperability between data in near real-time, using SensorML (Sensor Model Language) and O&M (Observations and Measurements) standards in a Sensor Observation Service. With last developments, we are able to distribute new collected data in real-time.

Keywords—Sensor Web; Real Time data; Marine Observations.

I. INTRODUCTION

“The OGC’s Sensor Web Enablement (SWE) standards enable developers to make all types of sensors, transducers and sensor data repositories discoverable, accessible and usable via the Web”[1].

A meteo-oceanographic buoy is an array of sensors that monitor environmental parameters, such as temperature, salinity, meteorological conditions, etc.

Measurements made from different sensor systems are acquired using heterogeneous data formats.

The first objective is data sharing; maintaining interoperability and resilience; the second request is to provide new collected data to national and international communities, as RITMARE, JERICO, MyOcean and FixO3. To satisfy both issues and to share open access (near) real-time data and homogeneous data, we used new technologies such as the SWE, using SensorML and O&M standard in a Sensor Observation Service (SOS).

In detail, in this paper, we describe the methodology adopted illustrating each element included into the procedure (Section II) and we synthesized the aims of this paper in Section III.

II. METHODOLOGY

This system developed by OGS (Istituto Nazionale di Oceanografia e di Geofisica Sperimentale), is the answer for the need to find a standard procedure to share (near) real-time data collected by the meteo-oceanographic MAMBO1 buoy (in the North Adriatic Sea) and the observatory site E2M3A (in the South Adriatic Sea). This procedure has seven different elements [2] (Figure 1):

- two different **buoys** that acquire data in (near) real-time: the meteo-marine buoy Monitoraggio Ambientale Operativo (MAMBO1), placed in the Gulf of Trieste, equipped with a meteorological station and two multi-parametric probes, and the E2-M3A, situated in the South Adriatic Sea, hosting the meteo station including a radiometer aimed, to collect air-sea interaction measurements, and a mooring with sensors for physical and biochemical parameters [3];
- **RTLoader** (Real-Time Loader) permits to store in a database real-time heterogeneous data, coming from different kind of instruments and with different formats. This is possible converting XML standard into an object language, using Java Architecture for XML Binding (JAXB). It allows converting heterogeneous ASCII files formats in a unique XML common format. After this, the software “XML2DB” converts XML input into Java objects and then inserts the measurements and the associated metadata into a PostgreSQL database;
- **RTValidator** checks, in the database, the quality of the data, applying some different algorithms developed following the European protocols [4][5] eventually tuned to the regional statistics [6]. The validation process assigns a quality flag without changing or deleting any data points;
- **RTWs** is the RESTful Web Service used to extract data from the database accepting simple request parameterized with temporal range and the output format. It is written using Java and open source libraries like Spring and Jersey;

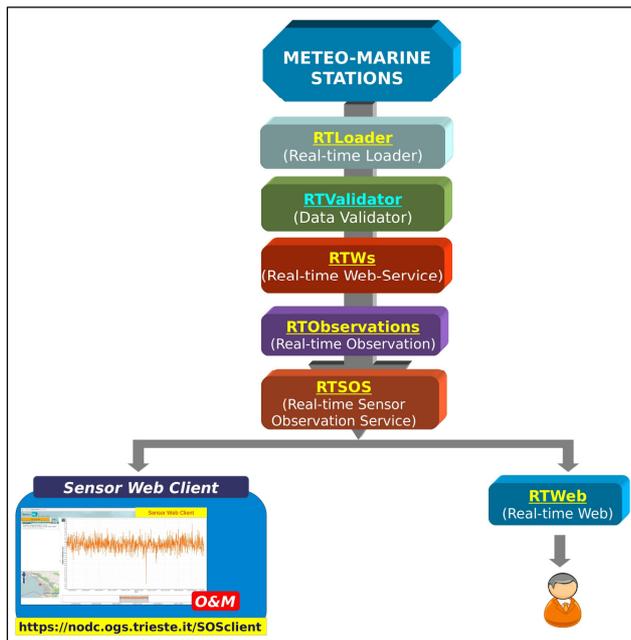


Figure 1. Work flow

- **RTWeb** is the web interface that allows querying the database using the Web Service RTWs. It extracts data into a down-loadable file, satisfying the conditions selected by the users;
- **RTObservations** is a real-time automatic procedure that permits to insert Observations into the Sensor Observation Service. This new element can load in near real-time data into a SOS, using an OGC standard format O&M. This is possible applying a Batch() operation and a POST request using JSON as payload.
- **RTSOS** is an OGC Sensor Observation Service (SOS) that enables to integrate real-time observations of heterogeneous sensors into a Spatial Data Infrastructure. It works using standard requests (e.g., *DescribeSensor()*). The descriptions of sensors and observations are stored in a PostgreSQL/PostGIS database using standard metadata format (respectively SensorML and O&M) and standard requests (*InsertSensor()* and *InsertObservation()*). Then, the data can be obtained by request *GetObservation()* and geo-located by *GetFeatureOfIntersect()* [7]. Also a Web interface was used, in order to visualize observations, the sensors position, their observed properties and long term trends of observations. It has been implemented using JavaScript toolkits (OpenLayers, GeoExt and ExtJS) [8].

III. CONCLUSION

To answer the needs of different scientific communities as RITMARE, Jerico, MyOcean, ODIP and FixO3, we developed an automatic system to store in near real-time data

acquired by two different meteo-oceanographic buoys. Also, we decided to use an innovative technology, adopting OGC standards as SensorML and O&M, using a Sensor Observation Service (SOS), to disseminate these data. A new element is developed, the RTObservations, to store automatically the near real-time data into a Sensor Observation Service, using standard requests. This solution gives us the possibility to share open access and standardized data and to have the opportunity to cooperate, analyze and use new technologies.

ACKNOWLEDGMENT

This work was supported by the Ministry of Education, Universities and Research (MIUR) under RITMARE Project. We also acknowledge funding from the European Program (JERICO, FixO3 and SeaDataNet II).

REFERENCES

- [1] Open Geospatial Consortium: Sensor Web Enablement (SWE) [Online]. Available from: <http://www.opengeospatial.org>.
- [2] E. Partescano, et al., "A (Near) Real-time Validation and Standardization System Tested for MAMBO1 Meteo-marine Fixed Station", SENSORNETS 2014 - 3^o International Conference on Sensor Networks 2014 Lisbon, 7-9 January.
- [3] M. Ravaoli, et al., "The RITMARE Italian Fixed-point Observatory Network (IFON) for marine environmental monitoring: a case study", Journal of Operational Oceanography. Special Issue - "Proceedings of the 3rd Italian GNOC Conference on operational oceanography, innovative technologies and applications". Submitted (2015).
- [4] SeaDataNet: Data Quality Control Procedures, Version 2.0, May 2010, 78 pp. ([http://www.seadatanet.org/content/download/18414/119624/file/SeaDataNet_QC_procedures_V2_\(May_2010\).pdf](http://www.seadatanet.org/content/download/18414/119624/file/SeaDataNet_QC_procedures_V2_(May_2010).pdf)).
- [5] UNESCO: GTSPP Real-Time Quality Control Manual, First Revised Edition. IOC Manuals and Guides No. 22, Revised Edition. (IOC/2010/MG/22Rev.) 2010, 145 pp.
- [6] B. Manca, M. Burca, A. Giorgetti, C. Coatanoan, M. J. Garcia and A. Iona: Physical and biochemical vertical profiles in the Mediterranean regions: an important tool to trace the climatology of water masses and to validate incoming data from operational oceanography. J. Mar. Sys., 48 2004, pp. 83-116.
- [7] RTSOS: OGS's Sensor Observation Service [Online]. Available from: <http://nodc.ogs.trieste.it/SOS4>.
- [8] RTWEB: OGS's Web interface [Online]. Available from: <http://nodc.ogs.trieste.it/SOSclient/>.

Feature Selection and Interpretation of GSR and ECG Sensor Data in Biofeedback Stress Monitoring

Tamás D. Gedeon
 Research School of Computer Science
 Australian National University
 Canberra Australia
 tom@cs.anu.edu.au

Xuanying Zhu, Leana Copeland, Nandita Sharma
 Research School of Computer Science
 Australian National University
 Canberra Australia

Abstract—We have used sensor data previously to detect the stress status of observers of stressful environments. In this work we describe the process of conversion from a passive and post-facto detection of this stress status to a dynamic and real-time or close to real-time process for control of stress using biofeedback. We describe the changes required to feature selection and interpretation of galvanic skin response (GSR) and electrocardiograph (ECG) sensor data for the new setting, using training on data from the post-facto dataset filtered to match, and as far as possible simulate a real-time dataset. We then compare 3 alternatives including a control component.

Keywords- *observer stress; physiological signals; biofeedback; human centred computing.*

I. INTRODUCTION

Our previous work demonstrated that we could construct reliable computational models for observer stress extracted from a range of physiological signals. In that work, we used physiological recordings of the entire experiment to detect the observer's stress level, the experiment consisting of showing a number of short video clips of known stressfulness, as validated by user surveys. This data forms the ANU Stress DB.

We have extended our previous work, and sampled the ANU Stress DB to mimic real-time or close to real-time data collection, to form training data for a biofeedback model. For usefulness in biofeedback the model needs to display results no slower than 1 Hz, based on some preliminary experiments with subjects. Using only the last 2 seconds of data produces a model in 80% in accord with our previous full model using the entire experiment recording.

We then tested this new model in a new experiment, to compare our model with the use of EEG as the biofeedback target, as well as a control curve which displays the stress curve of a randomly chosen prior subject. That is, the stress curve is synchronized with the experiment and is a valid stress curve, but is not the curve of the current subject and so does not reflect any changes made by the current subject.

II. STRESS AND REDUCTION OF STRESS

The term *stress* was coined by Hans Selye in 1936, and defined as “the non-specific response of the body to any demand for change” [1]. There is evidence that too much stress has significant health effects, e.g. [2].

We concentrate on objective measures of stress [3].

A. Observer Stress

We concentrate on a viewer of events, hence the stress encountered is observer stress. In this century, more and more our interactions with the world are virtual or mediated via screens, or for other reasons we have no ability to change what we see. Hence this is a valuable area for stress research. We have done previous work on computational models for stress in a range of environments: abstract (reading) [4], virtual (screens - including video / cctv) [5], and real environments [6].

B. Stress reduction - biofeedback

There are a number of stress management techniques, from meditation [7] to biofeedback. Biofeedback is the process of making unconscious body functions perceptible for individuals so that they can learn and manipulate these physiological activities for the purpose of improving health [8]. Neuro-feedback, which uses EEG data, has been shown to be one of the most effective stress management methods [9].

C. Measures and Sensors

1) Galvanic Skin Response (GSR)

Skin conductance, also known as electro-dermal response or psycho-galvanic reflex, measures the electrical conductance of an individual's skin, which varies due to the amount of sweat on the skin. When the individual is under stress, skin conductance will increase; oppositely, the skin conductance will reduce when the individual encounters less stress [10]. We used the Neulog GSR Logger Sensor [11].

2) Heart Rate Variability

Heart rate variability (HRV) is the variation in the interval between heartbeats. HRV has been shown to be one of the most reliable indicators of stress in nearly a

quarter-century of clinical research [12]. Methods to measure HRV include Electrocardiogram (ECG) and Blood Pressure, but ECG is considered superior as it excludes unnecessary heartbeats and displays a clear waveform. We used the Neulog ECG Logger Sensor [13].

3) Brain Signals

The brain is the key organ of responding to stress, as it perceives what is threatening and thus potentially stressful [14]. Electroencephalography (EEG) is one of the primary methods for brain activity analysis. It re-cords the electrical voltage fluctuations resulting from ionic current flows within the neurons of the brain via multiple electrodes placed on the scalp [15]. EEG equipment usually produces a high degree of intrusion since participants need to wear a head cap attached to specific positions of the scalp with electrically conductive gel. The Emotiv EPOC headset [16] is a less intrusive device with a lighter headset, which is placed on rather than attached to the head. It is still noticeably more intrusive than the GSR and ECG sensors.

D. Extraction of features

Some 59 representative features were derived by calculating the statistics and measures values of two-seconds data (as our preliminary experiments showed that 1 sec data was too little, but 5 secs was too long to still ‘feel’ like real-time). Such summary statistical values included the mean, standard derivation, kurtosis, skewness, interquartile range, minimum and maximum, where measures included the number of peaks for periodic signals.

III. HYBRID GENETIC ALGORITHM AND EXTREME LEARNING MACHINE MODEL

The 59 derived features are too many for good results on this data, hence we must engage in feature selection [17]. The issue is that there is a limited number of subjects for whom we have data, and with many features and a sizeable neural network, the number of free parameters well exceeds the number of data points available.

TABLE I. GA REPRESENTATION AND USE

Vector of best representative features filtered by GA	0	0	1	0	1	...
Vector of derived representative features	.1	.2	.3	.4	.5	...
Vector for best representative features in the current data segment	0	0	.3	0	.5	...

We used a genetic algorithm with feed-forward neural networks as individuals in the population to perform feature selection and training. A simple 59-bit string representation was used to determine whether a feature is used by a particular neural network. Figure 1 shows how the representation was used to construct the train and test sets for each neural network.

A standard 3 layer neural network with two layers of processing elements was used, as shown in Figure 1. There are up to 59 input neurons, some 20 (say) hidden neurons, and one output neuron indicating the degree of stress.

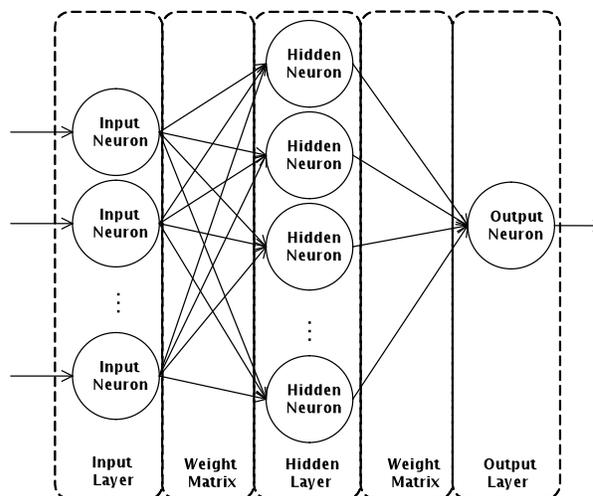


Figure 1. Neural network structure.

We validated our final networks against our previous work in virtual environments (stressful and calm video clips) [5], and using 10 fold cross validation we showed we can achieve 81% reliability.

This is less than our previous results which used the full sensor recording from a participant watching the entire clip, rather than just the preceding 2 seconds. This 81% reliability proved sufficient for our experiment (to be reported below) to be successful.

Our initial approach was to use back-propagation training for our neural network. We observed that it took one week to finish approximately 8% of the whole process, so the overall time for training this classifier was predicted to be more than three months. Instead, we trained our neural networks using the Extreme Learning Machine (ELM) method [18].

The ELM method in this case works as follows: the input weight matrix is assigned random values which are then frozen and not trained (the left ‘weight matrix’ shown in Fig. 1), and only the output weight matrix is trained.

This is possible to do via the delta rule normally used for output layers, but it is even faster to use the Moore-Penrose pseudo-inverse as the input matrix does not change, allowing a much higher efficiency in the process of estimating (training) output layer weights, by computing a ‘best fit’ (least squares) solution.

It is necessary to significantly increase the number of hidden neurons using the ELM approach. In this case, we

increased the number to 400. Then, we were able to complete the training in one week, faster by a multiple of about 15, notwithstanding the 20 fold increase in the number of hidden neurons.

How does it work? Essentially, the random weights provide some random functionality to each hidden neuron, and the training of the output layer selects from these the neurons with useful functionality and then their outputs are combined using the output weights to provide the optimized output value. This explains why we need so many hidden neurons, as we now just select from a menu rather than training individuals.

IV. BIOFEEDBACK EXPERIMENT

Ethics approval to perform the experiment was received from the ANU Human Research Ethics Committee.

Eighteen undergraduate and masters students were recruited for the experiment. The participant cohort was

made up of twelve males (57.1%) and nine females (42.9%) between the ages of 20 and 35 years. The average age was 25.1 years old with a standard deviation of 3.7.

A. Experiment aim

Our goal was to compare effects of our GA-ELM calculated stress curve, EEG curve, and a (plausible) random curve on stress reduction.

B. Experiment Data Collection

Figs 2. and 3. display the experiment setups required for the stress curve / random curve settings, and for the EEG setting, respectively.

Multiple computers were required, as the computational intensiveness of both our own real time stress curve and the Emotiv EEG stress curve were such that a separate data acquisition computer was required. For smooth display of the film clips, we needed yet another computer

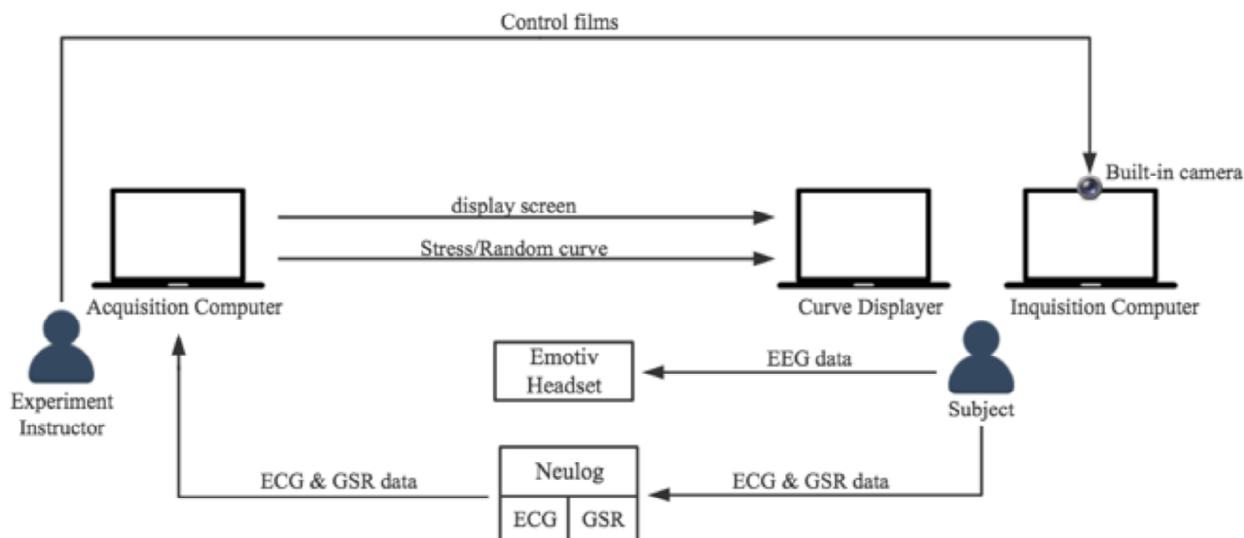


Figure 2. Schematic diagram of equipment setup for stress and random curve group.

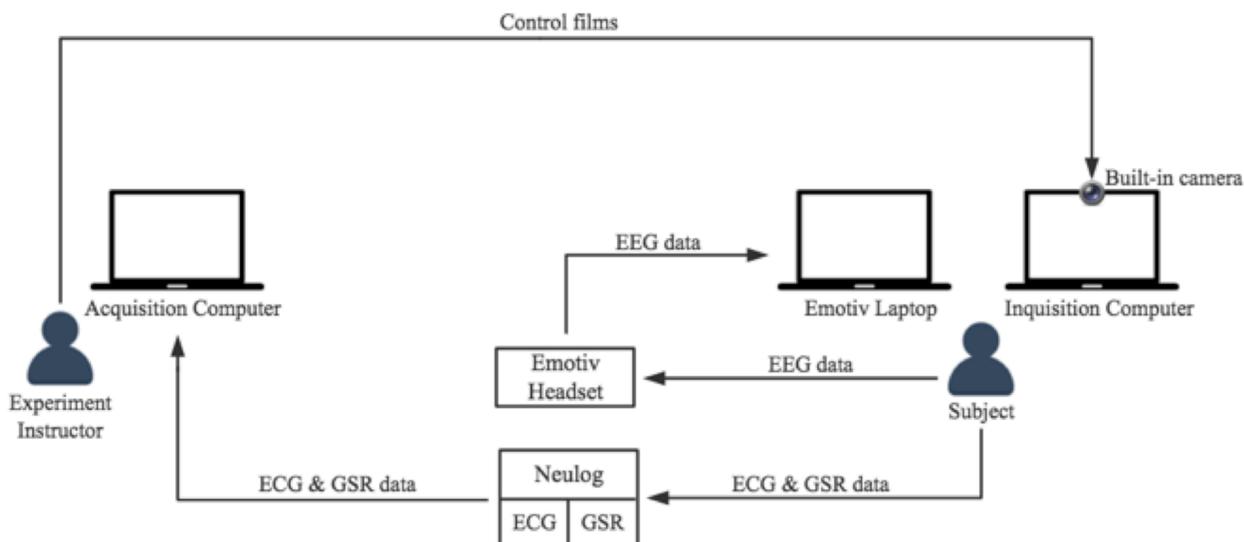


Figure 3. Schematic diagram of equipment setup for EEG curve group.

C. Experiment process

After a countdown display, a blank screen was shown for 15 seconds, which was followed by a sequence of film clips with five-second blank screens in between.

The film clips consisted of stressful and non-stressful film clips and each was approximately one minute in length. These film clips were categorized by the type of environment they create. Some had stressful content in the direction towards distress, fear and tension, see Fig. 4 for a sample screenshot from one of the stressful film clips.



Figure 4. Stressful film clip: Dark Knight.

The non-stressful clips had content that created an illusion of meditation or comfortable environments, see Fig. 5 for a sample screenshot from one of the non-stressful film clips. In total there were three stressful films and three non-stressful film for each experiment session, presented in an order balanced fashion.

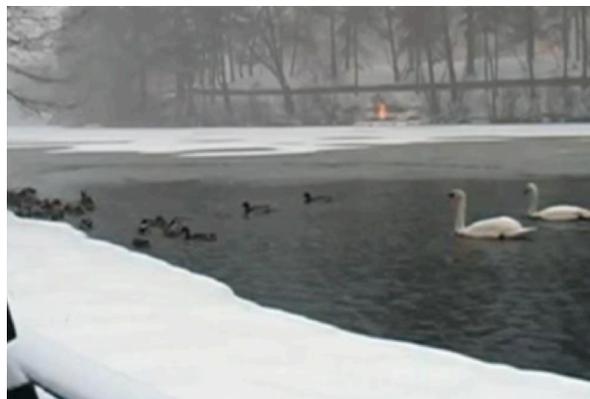


Figure 5. Non-stressful film clip: Ducks on the lake.

One third of the participants were shown their own real-time stress curve as calculated from the their preceding 2 secs of data as described in §1.4, while watching the sequence of film clips.

Another third were shown their EEG stress curve as calculated by the Emotiv Affectiv Suite [19] (see Fig. 4), again while watching the film clips. Since the clips were order balanced, the participants watched them in different orders, to reduce or eliminate any effects from the sequence of presentation of the film clips.

The final third were shown a ‘random’ curve. This curve was based on the ECG and GSR of two subjects from our previous work [5], using the same film clips. Thus, the curves were completely valid stress curves and were in good synchrony with the experiment, they were just not from the participant viewing them.

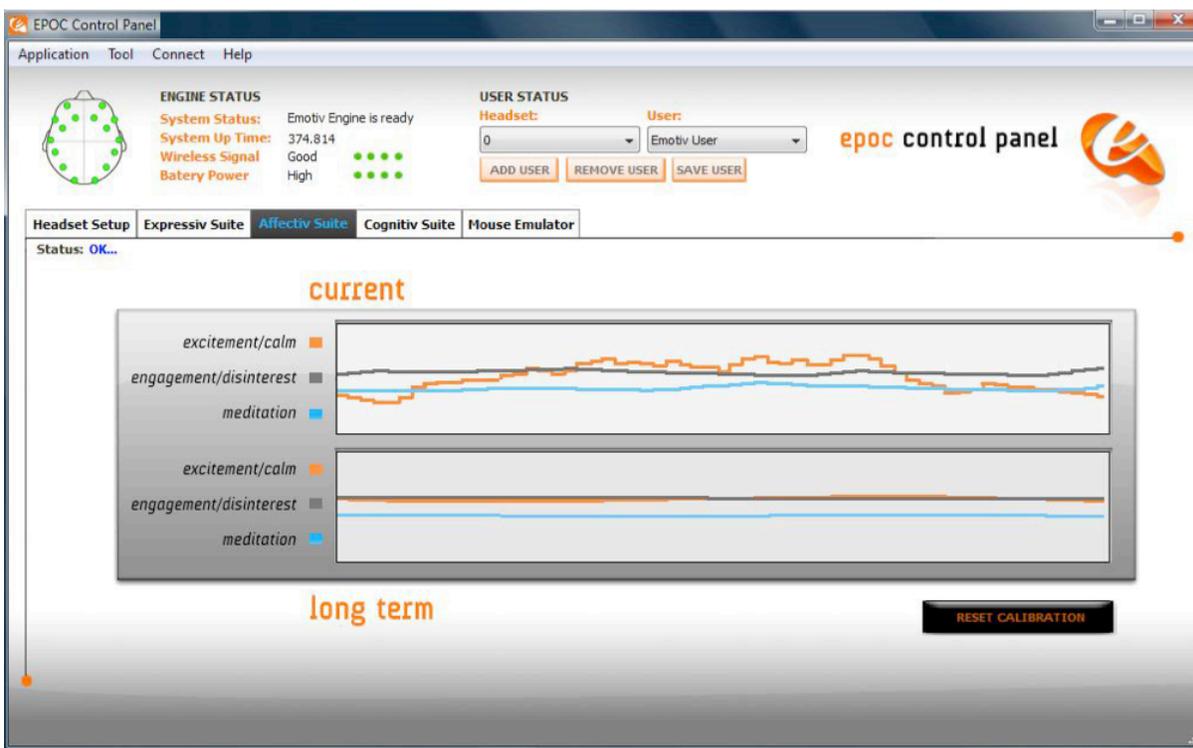


Figure 6. Emotiv Affectiv Suite: the top orange curve is the stress indicator for the EEG curve group.

The point of these ‘random’ curves is to provide a baseline – it is possible the mere intention to reduce stress will have an effect, so for our results to be meaningful we need to show a different (better) result than for this ‘random’ control curve.

V. RESULTS AND DISCUSSION

The data was analysed by visual inspection and clustering analysis. Observation of the stress curve suggested that the stress curve does reflect individuals’ stress in real time. It also revealed the correlation between the effectiveness of physiological biofeedback in stress control. This correlation was confirmed by the use of K-Means clustering. The clustering analysis was conducted on the 2 different stress data sets generated by watching stressful or non-stressful films. The result of clustering on stress data, which were produced when individuals were watching stressful films, showed that biofeedback with our physiological stress curve was effective and it was superior to neurofeedback.

We now provide the results of the clustering, in Table 2. The results are sorted by the curve provided to each participant. Clusters 2 and 3 have a purity of 83%, and correspond to our stress curve, and to the EEG curve, respectively. This is essentially the highest value we could have expected as our pre-experiment estimate for the correctness of our calculated stress curve.

These results show that both our stress curve and the EEG curve have good consistency in terms of their effects on the ability of participants to modify their stress. By observation, the direction of modification is as expected, to reduce their stress.

Cluster 1 has a purity of 50%, for the random curve. The *lack* of high purity in the case of the cluster which represents the random curve indicates that that curve is of some value in reducing the stress of participants,

otherwise this cluster would have high purity due to its consistent uselessness in modifying stress levels.

TABLE II. CLUSTERING RESULTS ON STRESS DATA GENERATED BY WATCHING STRESSED FILMS.

Participant ID	Clustered label	Provided curve
p10	S3	Stress
p12	S2	Stress
p13	S2	Stress
p14	S2	Stress
p15	S2	Stress
p6	S2	Stress
p2	S3	Random
p7	S3	Random
p4	S2	Random
p1	S1	Random
p11	S1	Random
p5	S1	Random
p17	S3	EEG
p18	S3	EEG
p19	S3	EEG
p20	S3	EEG
p21	S3	EEG
p16	S2	EEG

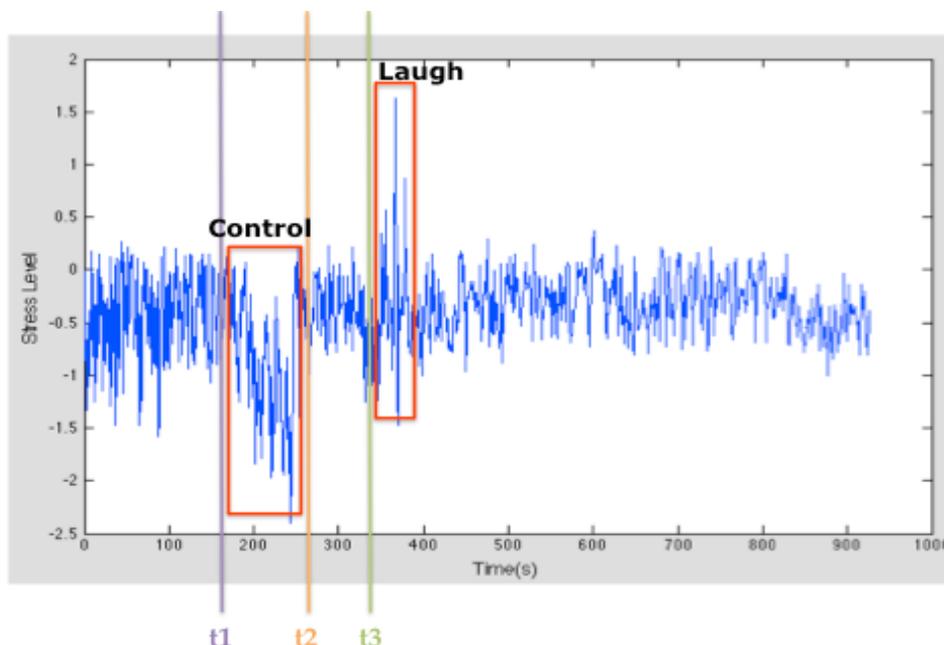


Figure 5. Participant 14: Using our stress curve. The stress axis has arbitrary units, the time axis is in seconds.

An example of the visual inspection analysis is shown in Fig. 5. In the period labeled from t1 to t2, the participant was actively trying to control his stress, as reported during the experiment. At time t2, a new clip started. Unfortunately, at time t3, he laughed and pounded on the table, partially dislodging the sensors.

We performed a statistical analysis of the clustering results to the categorisation of the stress levels of the film clips. The calculation of the p-value derived from the Wilcoxon Statistical Test was: $p < 0.001$. Thus we can conclude that our results are highly statistically significant, as we accept $p < 0.05$ as statistically significant.

CONCLUSIONS, LIMITATIONS AND FUTURE WORK

We have shown via an experiment with 18 participants, and with high statistical significance, that participants could control or modify their stress well with an EEG curve (i.e., by neurofeedback, being an approach known to work well in the literature), which validates our approach.

More significantly, we have shown that with high

statistical significance, that participants could control or modify their stress well with the use of our real time stress curve calculated only from the GSR and ECG. This is a novel contribution, as GSR and ECG is much less intrusive than EEG.

A limitation was that in our experiments, electrodes were attached to the hands of the participants, but in practice this may not be necessary. For example, the use of a GSR enabled mouse could be used (in more active settings than watching film clips). For ECG, a computer wrist rest could be wired to measure this. Alternatively, the wearable revolution in progress right now may soon provide wristwatches or other wrist or arm borne devices, which are both ubiquitous and effective in measuring ECG and GSR.

We modified our initial GA-NN approach for feature selection and training to use ELM training and achieved an approximately 15-fold increase in speed of training.

In future work we will consider the introduction of noise into the training regime [20], as well as the hierarchical structuring of the data [21], especially using fuzzy signatures [22-24].

REFERENCES

- [1] H. Selye, A syndrome produced by diverse noxious agents, *Nature*, 138:32, 1936.
- [2] The-American-Institute-of-Stress, *Stress and Heart Disease*: www.stress.org/stress-and-heart-disease/, 2012.
- [3] N. Sharma and T.D. Gedeon, Objective measures, sensors and computational techniques for stress recognition and classification: A survey, *Computer methods and programs in biomedicine*, 108(3): 1287-1301, 2012. doi: 10.1016/j.cmpb.2012.07.003
- [4] N. Sharma and T.D. Gedeon, Computational Models of Stress in Reading Using Physiological and Physical Sensor Data. In *Advances in Knowledge Discovery and Data Mining* (pp. 111-122). Springer Berlin Heidelberg, 2013.
- [5] N. Sharma and T.D. Gedeon, Modeling stress recognition in typical virtual environments, *Pervasive Computing Technologies for Healthcare (PervasiveHealth)*, 7th International Conference, pp. 17-24, IEEE, 2013.
- [6] N. Sharma and T.D. Gedeon, Modeling Observer Stress for Typical Real Environments, *Expert Systems with Applications*, 41(5):2231-2238, 2014.
- [7] R.R. Michaels, M.J. Huber and D.S. McCann, Evaluation of transcendental meditation as a method of reducing stress, *Science*, 192:1242-1244, 1976.
- [8] H.L. Epstein and E.B. Blanchard, Biofeedback, self-control, and self-management, *Biofeedback and Self-regulation*, 2(2):201-211, 1977.
- [9] E. Hoffmann, *Brain Training Against Stress*, 2005.
- [10] W. Liao, W. Zhang, Z. Zhu and Q. Ji, A real-time human stress monitoring system using dynamic bayesian network, *IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops*, 70, 2005.
- [11] Neulog, *GSR logger sensor NUL-217*. At: neulog.com/gsr/, 2011.
- [12] Z. Ori, G. Monir, J. Weiss, X. Sayhouni and D. H. Singer, Heart rate variability: Frequency domain analysis, *Cardiology clinics*, 10:499-537, 1992.
- [13] Neulog, *Electrocardiogram logger sensor NUL-218*. At: neulog.com/electrocardiogram/, 2011.
- [14] B. S. McEwen, Physiology and neurobiology of stress and adaptation: central role of the brain, *Physiological reviews*, 87:873-904, 2007.
- [15] E. Niedermeyer, F. L. da Silva, *Electroencephalography: basic principles, clinical applications, and related fields*, Lippincott Williams & Wilkins, 2005.
- [16] Emotiv, *Emotiv EPOC Headset*. At: emotiv.com/epoc.php, 2011.
- [17] T.D. Gedeon, Data mining of inputs: analysing magnitude and functional measures, *International Journal of Neural Systems*, 8(02):209-218, 1997.
- [18] G.B. Huang, Q.Y. Zhu and C.K. Siew, Extreme learning machine: a new learning scheme of feedforward neural networks, *IEEE International Joint Conference on Neural Networks, 2004. Proceedings*, 2:985-990, 2004.
- [19] S.F. Verkijkaand L. De Wet, Using a brain-computer interface (BCI) in reducing math anxiety: Evidence from South Africa. *Computers & Education*, 81:113-122, 2015.
- [20] W.M. Brown, T.D. Gedeon and D.I. Groves, Use of noise to augment training data: a neural network method of mineral-potential mapping in regions of limited known deposit examples, *Natural Resources Research*, 12(2):141-152, 2003.
- [21] Gedeon, T. D., and L. T. Koczy. "Hierarchical co-occurrence relations." *Systems, Man, and Cybernetics*, 1998. 1998 IEEE International Conference on. Vol. 3. IEEE, 1998.
- [22] K.W. Wong, T.D. Gedeon and L.T. Koczy, Construction of fuzzy signature from data: an example of SARS pre-clinical diagnosis system, *Proceedings 2004 IEEE International Conference on Fuzzy Systems*, 2004.
- [23] B.S.U. Mendis, T.D. Gedeon, J. Botzheim and L.T. Koczy, Generalised weighted relevance aggregation operators for hierarchical fuzzy signatures, 2006.
- [24] B.S.U. Mendis, T.D. Gedeon and L.T. Koczy, Investigation of aggregation in fuzzy signatures, *Proceedings International Conference on Computational Intelligence for Modelling Control and Automation and International Conference on Intelligent Agents Web Technologies and International Commerce (CIMCA'06)*. Sydney, 2006.

Using Mutual Charge Scheme to Measure Salinity of Ice

Umair N. Mughal & Bhushan Nikumbh

Atmospheric Icing Research Team
 Department of Industrial Engineering
 Narvik University College
 Narvik, Norway
 Email: unm@hin.no

Abstract—For offshore measurements in Cold Regions, salinity of ice is also a critical parameter (together with many other parameters such as icing type, load, icing rate and melting rate) to be identified in order to optimize the performance of anti/de icing systems. Although there are some available sensory solutions in the market to measure real time salinity levels of water, however there are still not many real time techniques or solutions to measure the salinity of ice. In this research task, mutual charge transfer technique is utilized to measure the zero crossover values of different samples of ice and water with varying salt ratios. The aim of this paper is therefore to discuss the testing methodology and testing results.

Keywords—Atmospheric Icing Sensor; Offshore; Salinity; Zero Cross Over; Mutual Charge Transfer.

I. INTRODUCTION

For safer operations in High North, firstly it is important to understand the difficulties that may lead to people and inventory hazards. Continues melting of ice in high north is leading to accelerate the research and development activities towards this part of the world [1]. In order to plan for technological advancements in Cold Regions, it is critical to carefully understand the ice accretion characteristics on vulnerable structures and inventory. Icing on offshore structures have not yet been considered as major hindrance by maritime industry, however it may not be justified in near future. Feedback about icing parameters such as icing rate, type and icing load are considered to be very important for an Intelligent Anti/De Icing system. However, if such a system is utilized during an Offshore Application then real time salinity measurements are also considered to be critically associated with the ice accretion on structures. This leads to define atmospheric icing and sea spray icing. *Atmospheric Icing* is formed during the supercooled transition of water, it freezes as soon as it comes into contact with a cooled surface at temperature far below the freezing point. This ice accretion produces a hard layer of ice on all surfaces and is responsible for different challenges for safe operations specially related with communication, navigation and logistics. Nevertheless, if the offshore vessel is small than this accretion also effects its stability. Also if the sea spray icing also occurs at the same time then it amplifies the risk the accident. *Sea Spray Icing* is in fact more critical risk than the atmospheric icing. As the name suggests, this icing is more saline in nature (freezing point of salt water is $-1.7^{\circ} \rightarrow -1.9^{\circ}C$ till $-30^{\circ}C$ and critically depend upon salinity levels). It may be initiated when the portion of vessel meets the ocean waves or the blowing wind brings transports the water from ocean crests. It is however

more dependent on the relative speed of the vessel and the wind (typically it starts at $8m/s$) and typically limited to $15 \rightarrow 20m$ above sea level. The advantage of this icing is that it supports the melting and cleaning of ice formed on the deck. However, if the scuppers freeze or the railing are covered with ice, then this water is trapped and frozen [2]. It is a general observation that in arctic seas, icing activity continues throughout the year and the ocean’s density is effected more by salinity than by temperature. However, during first two quarters (January till June) atmospheric icing is common whereas during the third quarter (July till October) sea spray icing leads the activity by 50% leaving behind 45% for mixed icing and 9% for atmospheric icing. Generally, during an offshore operation atmospheric icing is not very critical as there it generally occurs on higher parts of ships, which typically get covered with $1 \rightarrow 2cm$ (typically surrounding air temperature is $0^{\circ} \rightarrow -20^{\circ}C$ and wind speed is less than $10m/s$) of thick ice. However, during sea spray icing and mixed icing, ice thickness reaches $1m$ in some cases. Table I reflects some icing intensities and prevailing surrounding parameters [2].

TABLE I. TYPICAL ICING INTENSITIES DURING OFFSHORE ACTIVITY [2]

Intensity Category	Icing Intensity		Temperature	Wind Speed
	ton/hr	cm/hr	$^{\circ}C$	m/s
Slow	< 1.5	< 1	$0 \rightarrow 3$ ≤ -3	Any Speed < 7
Fast	$1.5 \rightarrow 4$	$1 \rightarrow 3$	$-3^{\circ} \rightarrow -8$	$7 \rightarrow 15$
Very Fast	> 4	> 3	≤ -8	> 15

This paper is divided into five sections. As already discussed, Section I is an introduction to form the basis to measure salinity as an important parameter for real time offshore measurements. In the next section II, it is aimed to describe the present salinity measurement techniques, which are presently limited to remote sensing only. Section III will be an overview of mutual charge transfer scheme and in section IV hardware, experimental setup and results will be discussed. In section V, the results will be discussed followed by the conclusions and future work.

II. PRESENT SALINITY MEASUREMENT TECHNIQUES

The present real time salinity measurement techniques are restricted to measure the salinity of water. Salinometer (and Total Dissolved Solids 'TDS' meter) is a conventional instrument, which is designed to measure the salinity of dissolved salt content in a solution.

Technically salinity increases the conductivity of a solution, therefore the measuring sensor normally have electrical conductivity meter calibrated according to the salinity levels. Salinity also changes the specific gravity of the solution therefore it is also measured through hydrometer sometimes. However, in marine aquarium, sometimes a refractometer is used to measure the salinity and specific gravity of the water.

Presently the sea surface salinity variations are detected using remote sensing techniques such as using passive microwave technology (typically in the range of 1 – 3GHz & S bands). The clouds are transparent to such radiation, however sea ice emit such radiation at relatively low energy levels, which are difficult to measure over a smaller area [3][4]. Since 2002, NASA uses Advanced Microwave Scanning Radiometer-Earth Observing System (AMSR-E) sensor installed on Aqua Satellite for measuring the time series of sea ice data to complement the available sea ice data from 1972 [4]. The first global salinity map of Earth's Ocean Surface is revealed by NASA's Aquarius Instrument by recording the data from August 25 till September 2011, see Fig. 1.

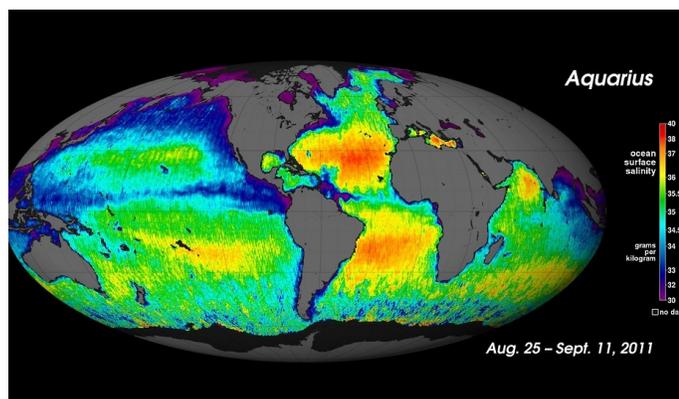


Figure 1. Salt of the Earth by Aquarius, NASA [5]

A review of available atmospheric ice sensing techniques is available in [9] followed by the electromagnetism techniques to detect atmospheric ice [11]. A detailed mathematical description to detect atmospheric icing using dielectric based sensing technique, in particular using standard Debye relation [12](see (1)) and its extended versions is available in [10].

$$\varepsilon = \varepsilon_{\infty} + \frac{\varepsilon_s - \varepsilon_{\infty}}{1 + i\omega\tau_o} - \frac{i\sigma}{\varepsilon_o\omega} \quad (1)$$

where ε is the complex frequency dependent dielectric permittivity; ε_s is the static dielectric permittivity of material; ε_{∞} is the high frequency complex dielectric permittivity of the material; ε_o is the dielectric permittivity of the free space; ω is the angular frequency of the signal; τ_o is the relaxation time of the material and σ is the conductivity of the material.

Similarly using the real part of the Debye dielectric permittivity relation [12](see (1)) at 50MHz a relation to derive the brine volume fraction and bulk salinity in the first year sea have been developed by Backstrom and Eicken [6]. In [6], three different experiments were conducted using the commercially available capacitive probes such as Stevens Water Monitoring Systems Hydraprobe [7][8] in order to find the possibility to measure both the absolute values of the bulk salinity and

changes in bulk salinity. They found good results with large temporal and spacial resolutions, within a reasonable error limit particularly during initial freezing and during ice melt. However, presently there is no available sensor, which can distinguish between different samples of accreted ice with varying salinity levels and deliver the results in a real time.

III. MUTUAL CHARGE TRANSFER

To sense capacitance changes of one order is not difficult; similarly capacitance changes of 10% is although difficult but still not trivial. However, if the capacitance changes are within 1%, in the presence of environmental changes then the measurements can be quite challenging [13]. The charge transfer technique to measure ice carry unique advantage of design flexibility in terms of implementation. The commonly used proximity and capacitive MEMS devices are capable to the influence of dielectric presence. The implementation of this technique has matured over the years and presently they are widely used in industrial automation, human sensory gadgets (human finger capacitance is around 100pF) and automobiles even in freezing domains without much variations. The water rejection techniques (water capacitance approximately 50pF but largely depends upon dimensions of the sensor) are incorporated in most of these devices. Recently ice detection and melting rate measurement results have been reported in [15] followed by some basic experimentation results [16], it is found that this technique have enough potential to detect and measure atmospheric icing parameters by lowering the voltage sensitivity thresholds (see [17]).

The self or mutual capacitance basics can be employed in the design based on charge transfer method. A small printed circuit board aided with the specific electrode configuration can be manufactured as a basic prototyping icing sensor. The benefit of reshaping the electrodes and design to any form with capability to change different dielectric material suitable to harsh environment can be used as a starting point to develop a prototype. This would enable to detect and measure the icing parameters in real time embedded platform with low power consumption, which is ideal for remote installations. The presented design is principally based on the mutual charge transfer (QMatrix) based icing detection, which is advised by Atmel to work better in high humidity or damp environments where the frequency of water droplets and moisture to be collected on the sensor is higher [18]. Furthermore the preference to mutual charge transfer was based due to the following constrains,

a) *Sensitivity*: The self capacitance design are more sensitive in nature. As the electric field is spread outwards through the electrode, in the presence of dielectric, ground loading is provided by the external influencing of the object, in our case will be ice or a water film. But with increase in sensitivity comes the inclusion of the noise in the circuitry, which is un-desirable. The noise in self capacitance might be increased so sensitivity tuning is the vital for design based on this methodology. The other factors, which can effect sensitivity are,

- Electrode dimensions and shape
- Ground loading
- Return path
- Supply voltage
- Charging pulses duration

b) *RF interference rejection*: The mutual capacitance have pairs of electrodes that are always connected to a low impedance circuit, and due to the fact that their electric fields are compact and self-shielded. The pair of electrodes do not respond well to external interference in the environment. In addition, the patent MEMS device *QT60240* used for the experimentation are spread-spectrum technology based, that is highly efficient for suppressing both radiated emissions and external fields susceptibility. Control panels made with this technology routinely pass susceptibility tests exceeding field strengths of $20V/m$.

c) *Dimensional Benefit*: Also mutual capacitance through charge transfer principle is preferable when key area for detection is to be defined. To isolate the areas of detection within a specific limit, this works in efficient manner as compared to the self capacitance. With mutual principle, the PCB tracks for detection circuitry can contain longer routing tracks, which could be the important consideration in the prototype embedded design.

d) *Moisture suppression*: The mutual approach provides two moisture suppression characteristics not present in any other capacitive methodology. First, the presence of a localized water film (due to condensation, mist, or tiny droplets) will induce only a small increase in signal coupling. The contributions to signal coupling caused by moisture will have an influence into the measurement but in the wrong direction of signal of normal readings and hence will cater the false detections. Secondly, the presence of moisture layer, which can conduct charge away are also minimized by the use of narrow *gate timings*. This results in limiting the charge accumulated to a narrow time slot window immediately after the charging pulse edge. As water film can be modeled as a distributed RC network with a time-dependent characteristic, the benefit of narrow gate times (microsecond or less) will greatly suppress the effects of a water films signal potential reduction, further reducing the chances of a false detection. However, this utility of the patented charge transfer based on mutual capacitance have to be modified in ice detection prototype hardware. Hence, the capability of the prototype hardware to distinguish between the thin water film and ice have to be included through PCB design circuitry and charge transfer pulses adjustment.

e) *Other benefits*:

- Temperature stability
- Low power modes

During this research work, experimentation were performed to analyze the measurable response of the icing and liquid ranges with variable salinity levels along with the state transition rates. The initial design was aimed to distinguish the saline ice and saline water layer formation. Later on the design incorporated the saline ice level measurements. The design is kept in the consideration of harsh climatic performance and remote installation mechanism flexible enough to be mounted on various Onshore and Offshore platforms or other sensors as like MuVi-Graphene, see [19]. The necessary condition for a robust flat plate capacitive sensor was to have uniform and sufficient ice deposition on its surface; this has been experimentally demonstrated by Mughal & Virk [20].

IV. EXPERIMENTAL DETAILS

A. Hardware Setup

The hardware implementation of the computing hardware and electrode for icing measurement is shown in Fig. 2. The embedded algorithm of computing hardware was performed using Atmel Studio 6 on the AVR series 8-bit microcontroller with associated peripherals. The in circuit programming was assisted using Atmel STK 500 development kit. The patented charge transfer process is executed by the *QT60240* device and several charge/discharge cycles are executed based on burst of pulses ranging from $500\mu s \rightarrow 2ms$. The micro-controller initializes the communication through *I²C* protocol and determines the change in the zero crossing of detected slope for each charging pulse. Generally, the negative DC voltage ramp for the signal is around $-250mV$, which is achieved by the synchronous switching of capacitors and slope drive cause the zero crossing. The $\tau_{zerocrossover}$ value (differential time measured) is acquired by the μ controller and stored in a buffer array at the same time. The buffered $\tau_{Zero\ Crossover}$ value is converted to *RS232* protocol and sent at the baud rate of $19200bits/s$. The resulting data was displayed in a DAQ factory (data acquisition development software) based user interface for further analysis and experimentation for optimizing sensor configuration.

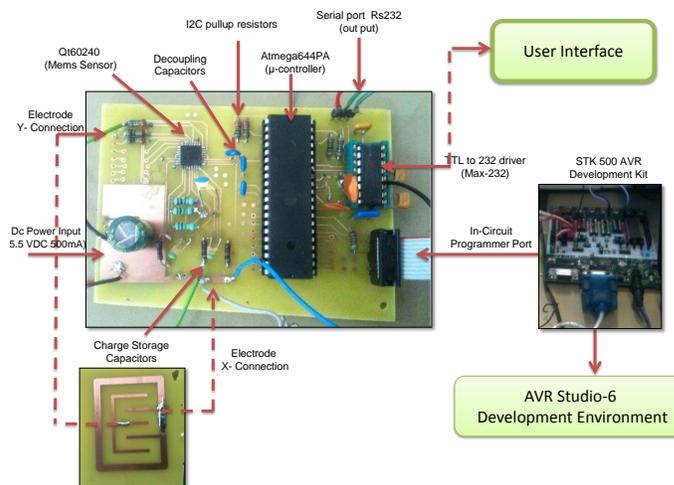


Figure 2. Computing Hardware Module

B. Experimental Setup

A preliminary series of experimentation were performed in Cold Room Chamber of Narvik University College. There are three Cold Rooms where a controlled temperature from $+20 \rightarrow -30^{\circ}C$ can be maintained. A sample of $100ml$ of water was used as a standard liquid content. Different levels of salt $0 \rightarrow 30g$ were mixed in this water sample and tested by pouring these samples on the electrode plate. Similarly the same salt levels were mixed in different water samples of $100ml$ each and allowed to freeze in the local freezer. The electrode panel was made as a non planar structure where the electrodes and rest of the pcb circuitry are lying on different panels (see Fig. 2). This scheme was preferred over the planar structure (Electrodes and computing hardware on same PCB) keeping in consideration of the cold climate environment. The intention was to make a separate robust electrode panel

exposed outside the environment to detect and measure the ice whereas the computing hardware is made on a separate PCB house inside protected covering. The overall dimension of electrodes pair used was $26 \times 36mm$ and the other dimensions are:

- X electrode outer border thickness $2.54mm$
- X electrode finger thickness $1.25mm$
- X-Y inter spacing width $3.8mm$
- Y electrode key thickness $1.25mm$

Common front panel dielectric materials include glass, plexiglass and polycarbonate. However, during this experiment all season packaging tape 60041 was used. This tape was used as it has excellent holding power during extreme hot or cold temperatures and made of pressure-sensitive poly material. The $\tau_{ZeroCrossover}$ of the electrode plate with only front panel dielectric tape 60041 was $438\mu s$.

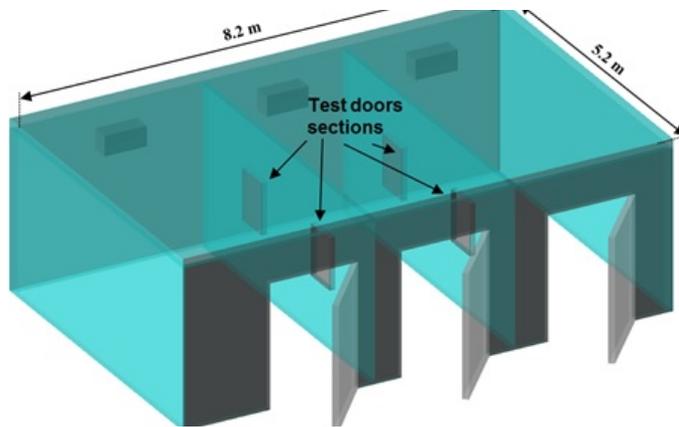


Figure 3. Cold Room Chamber, Narvik University College

C. Results and Discussions

The charge transfer technique outputs $\tau_{ZeroCrossover}$ as a measurement signal. The results of the experiment are tabulated in Table II and plotted in Fig. 4. The results indicate that as the percentage salinity ratio $\%SR_{ice}$ of ice is varying from $0 \rightarrow 23\%$ the $\tau_{ZeroCrossover}$ signal varies from $600 \rightarrow 1250\mu s$. Hence, a linear relationship can be found between $\%SR_{ice}$ and $\tau_{ZeroCrossover}$ with an effective slope of $28\mu s$ (see (2)). However, when the $\%SR_{water}$ of water is plotted against $\tau_{ZeroCrossover}$ signal, then the slope is not effective to demonstrate the linear variation, however this result do indicate the difference between fresh and saline water (see (3)). The Intercept of this curve is intentionally fixed at $2100\mu s$ in order to indicate the linear relationship between the changing salinity levels in water. One of the possible reason of mutual transfer scheme to not indicate the $\%SR_{water}$ can be associated with the water slippage leaving behind small puddles of water on the electrode plate.

$$\%SR_{ice} = 28\tau_{zero crossover} + 559 \quad (2)$$

$$\%SR_{water} = 1.38\tau_{zero crossover} + 2100 \quad (3)$$

TABLE II. SALINE WATER AND ICE TESTING RESULTS

Saline Water			
Volume of Water ml	Mass of Salt g	Salt Water Ratio %	Zero Cross Over μs
100	0	0	600
100	5	4.47	700
100	10	9.11	750
100	20	16.7	1000
100	30	23.1	1250
Saline Ice			
Mass of Ice g	Mass of Salt g	Salt Ice Ratio %	Zero CrossOver Value μs
99.7	0	0	1823
99.7	5	4.47	2089
99.7	10	9.11	2120
99.7	20	16.7	2117
99.7	30	23.1	2137

Real Time Ice Salinity Measurements

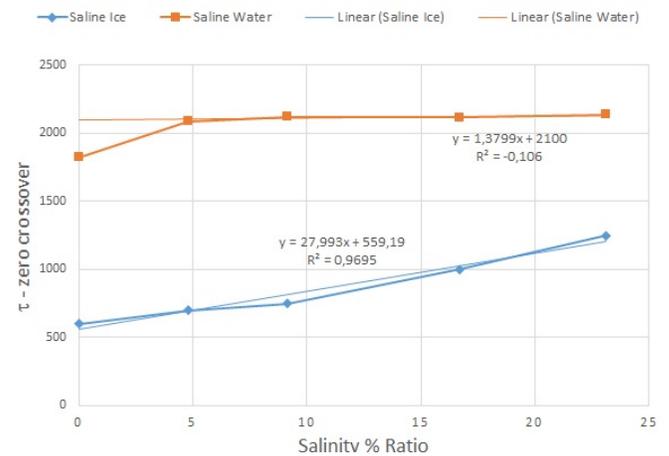


Figure 4. Real Time Ice Salinity Measurements

During the experimentation at room temperature the melting of ice layer/block occurred as a natural process, which could be the case in real environment. This phenomena could also be observed from the $\tau_{zerocrossover}$ measurement as it varies with the melting of ice layer upon the electrode till it slides over the electrode leaving behind the water puddle over the electrode.

D. Limitation

The measurements taken are independently experimented based on the charge transfer technique. The ice samples used were frozen using commercially available freezers.

V. CONCLUSIONS

The icing detection and salinity measurements in harsh cold climate is a demanding challenge. Particularly during an offshore activity, the need is more demanding with the latest developments in the technology and resources. If the parameters such as ice accretion and salinity are measured in real time, then expected damages can be reduced by using them as feedback variables in an intelligent anti/de icing system. The charge transfer technique outputs $\tau_{ZeroCrossover}$ as a measurement signal. This is a real time output to indicate

percentage salinity ratio of ice hence there are minimum delays associated with this technique. This is already proven technology therefore the implementation of this technique have matured even in freezing domains without much variations. The results of this experiment indicate that it is possible to use this technology for offshore platforms to indicate the percentage salinity levels in ice. From present experiment, it is also found that this technique is not useful to indicate the salinity level in water, however it do indicate the presence of salt in fresh water. These results are largely dependent upon the size of electrode plate, plate design and front dielectric panel material and thickness. Hence, optimizing the complete plate design can significantly improve the results. The electrode plate can be covered with different types of dielectric material much more resistant to harsh weather and climatic effects; since there are large number of thin dielectric materials available commercially nowadays. The weather resistant dielectric coating/covering can ensure the protection required to avoid corrosion of the plate. The low voltage requirement mV of the MEMS devices makes the design feasible for battery operated system in a remote location.

It was also quite difficult to delay the melting process of the saline ice sample (as the freezing temperature of saline ice is lower than fresh water ice) during experimentation. This problem was minimized by maintaining a temperature of $-5^{\circ}C$ in the Cold Room Chamber (see Fig. 3) at Narvik University College.

ACKNOWLEDGMENT

The work reported in this paper was partially funded by the Research Council of Norway, project no. 195153 and partially by the consortium of the project ColdTech- Sustainable Cold Climate Technology.

REFERENCES

- [1] "Arctic Regions and its Concerns, Threats and Potential Challenges", Technical Report, 2010.
- [2] "Probability of Icing", Available from http://www.fni.no/insrop/doc_dataprobability_of_icing.html, Last Cited: June 14th, 2015.
- [3] D. S. E. Lagerloef, "Preliminary Assesment of the Scientific and Technical Merits for Salinity Remote Sensing from Satellite", Final Report of the first workshop by Salinity Sea Ice Working Group (SSIWG), La Jolla, CA, USA, 7-8 February 1998.
- [4] "Remote Sensing: Passive Microwave", Available from http://nsidc.org/cryosphere/seaice/study/passive_remote_sensing.html, Last Cited: June 14th, 2015.
- [5] "NASA's Salt of the Earth: Aquarius Reveals First Map", Available from http://www.nasa.gov/mission_pages/aquarius/multimedia/gallery/pia14786.html, Last Cited: June 14th, 2015.
- [6] L. G. E. Backstrom and H. Eicken, "Capacitance probe measurements of brine volume and bulk salinity in the first year sea ice", Journal of Cold Regions Science and Technology, 2006, pp. 167-180.
- [7] J. E. Campbell, "Dielectric Properties of moist soils at rf and microwave frequencies", PhD Thesis, Dartmouth College, Hanover, NH, 1988.
- [8] J. E. Campbell, "Dielectric properties and influence of conductivity in soils at one to fifty megahertz", Soil Science of America Journal, 54(2), 1990, pp. 332-341.
- [9] U. N. Mughal and M. S. Virk, "Atmospheric icing sensors - An insight", The Seventh International Conference on Sensor Technologies and Applications, ISBN 978-161208-296-7, August 25-31, 2013, pp. 191-199.
- [10] U. N. Mughal, M. S. Virk, and M. Y. Mustafa, "Dielectric based sensing of atmospheric ice", AIP Conference Proceedings, 1570(212), 2013.
- [11] U. N. Mughal, M. S. Virk, and M. Y. Mustafa, "Atmospheric Icing Sensors - a conceptual review using matter and icing electromagnetism", International Journal of Multiphysics, 6(4), 2012, pp. 341-353.
- [12] P. Debye, *Polar molecules*, Dover, New York, 1929.
- [13] "Capacitive Sensing Techniques and Considerations", Available from <http://www.cypress.com/?docID=3346>, Last Cited: June 14th, 2015.
- [14] K. B. Shraavan, U. N. Mughal, and M. S. Virk, "Experimental study of relative permittivity of atmospheric ice", International Journal of Energy and Environment, 4(3), 2013, pp. 369-376.
- [15] T. Rashid, U. N. Mughal, and M. S. Virk, "Charge Transfer Scheme for Atmospheric Ice Sensing", Sensors and Transducers Journal, 184(1), 2015, pp. 146-152.
- [16] T. Rashid, U. N. Mughal, and M. S. Virk, "Cognitive Adaptability of Capacitive Sensors for Cold Regions", IEEE Conference on Cognitive Infocommunications, Vietri Sul Mari, Salerno, Italy, 2014, pp. 289-294.
- [17] "Atmel AVR3002: Moisture Tolerant QTouch Design", Available from: <http://www.atmel.com/images/doc42017.pdf>, Last Cited: June 14th, 2015.
- [18] "Proximity Design Guide", Available from: <http://www.atmel.com/Images/doc10760.pdf>, Last Cited: June 14th, 2015.
- [19] U. N. Mughal and M. S. Virk, "Cognitive Decision Making for measuring icing load for Cold Regions", IEEE Conference on Cognitive Infocommunications, Vietri Sul Mari, Salerno, Italy, 2014, pp. 283-288.
- [20] U. N. Mughal, M. S. Virk, K. Kosugi, and S. Mochizuki, "Experimental Validation of Icing Load Using Rotary Physics", Manuscript Submitted in the Journal of Cold Regions Science and Technology, Elsevier, 2014.

Infrared Thermal Signature Evaluation of a Pure Ice Block

Taimur Rashid, Hassan A. Khawaja, K. Edvardsen

Department of Engineering and Safety
The Arctic University of Norway (UIT)
Tromsø, Norway
e-mail: taimur.rashid@uit.no

Umair N. Mughal

Department of Industrial Engineering
Narvik University College (HIN)
Narvik, Norway
e-mail: unum@hin.no

Abstract—Marine operations in cold climates are subjected to heavy ice accretion, which can lead to heavy ice loads causing safety concerns. Atmospheric and sea spray icing can cause severe ice accretion on structures. Ice detection over the relatively larger surface area of marine platforms and vessels could be significant for ice mitigation and its removal in cold climates. A thermal ice signature can be considered a useful property that could be detected using the Infrared (IR) cameras. To study this remote ice detection option, preliminary lab experimentation was performed to detect thermal gradients over two different surface areas of a pure ice sample at room temperature. A complete thermal signature over the surface area was detected and recorded until the meltdown process was completed. The proposed technique can be useful in detecting ice over large areas in cold marine environments than point detection ice sensing. However, challenges remains in terms of the validation of the detection signature and elimination of false detection.

Keywords- icing; infrared; thermal signature; cold region.

I. INTRODUCTION

Ice accretion problems are challenging in the cold climate regions. With a rise in activity in the arctic frontier, icing phenomena will increasingly affect operational activities. The sources of icing may vary according to environmental conditions, for instance atmospheric icing (snow, rain, fog) [1] or sea spray icing generated due to wind or wave-structure collision [2]. In marine icing events, combination of the mentioned causes could also be the case. An example of this is the snow that sticks to a wet surface caused by the sea spray. All of the discussed factors can cause heavy ice accumulation on offshore and mobile platforms operating in the cold climates [3]. Adequate measures are required to minimize the influence of atmospheric and sea spray icing events on structures under different conditions and, therefore, effective anti/de-icing techniques become important [4]. These techniques involve thermal, chemical and mechanical methods to avoid and/or remove the ice [5]. Most often the problem arises in situations where unpredicted or unexpected environmental conditions occur that could cause heavy ice accretion phenomena, as the ice accretion rate is high enough to become a possible threat to the structure. In such cases the time elapsed becomes the critical factor in responding in an effective manner, which also includes ice mitigation and removal. A viable option available is timely ice detection over the structure during heavy icing events, which should support the anti/de-icing systems. Timely detection can be

important to the mitigation of heavy ice loads upon structures.

In this paper a brief insight into the concept of infrared (IR) thermography is given in Section II. Suitable wavelengths to detect cold objects with infrared detectors is also discussed. Section III gives the details about the preliminary experimentation performed to evaluate the IR thermal signature of a pure ice block and the results are discussed in Section IV. Also included is a brief discussion on remote IR detection in marine operations, which will be the applicability of this research project followed by the concluding remarks in Section V.

II. IR THERMOGRAPHY AND THERMAL SIGNATURE DETECTION AT LOW TEMPERATURES

IR light is electromagnetic radiation with a wavelength longer than that of visible light, measured from the nominal edge of visible red light at 0.74μm and extending conventionally up to 300μm. Figure 1 shows the IR spectral band that is generally sub-divided into four sub-bands; near infrared (NIR) ranges between 0.75–1μm, shortwave infrared (SWIR) is from 1-2.5 μm, middle infrared (MWIR) is between 3–5μm, longwave infrared (LWIR) is between 8–14μm.

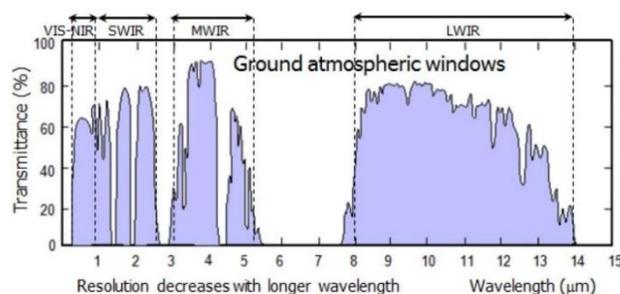


Figure 1. IR Spectrum [6]

For cold objects near 0°C, the prominent radiation wavelength is near 11 μm. Since most commercial IR sensors use wavelengths between 8 to 14 μm, IR thermography can be handful for surface ice studies [7]. Theoretically, each body at any temperature above absolute zero will emit some radiation, but the intensity and frequency distribution of the radiation is different based on the basic structure of the body.

The energy emitted by a true blackbody is the maximum theoretically possible for a given temperature. The radiative power (or number of photons emitted) and its wavelength distribution are given by the Planck’s radiation law (given by (1) and (2)), where λ is the wavelength, T is the temperature, h is Planck’s constant, c is the velocity of light, and k is Boltzmann’s constant.

$$\frac{2\pi hc^2}{\lambda^5} [\exp(\frac{hc}{\lambda KT}) - 1]^{-1} W / (cm^2 \mu m) \quad (1)$$

$$P(\lambda, T) = \frac{2\pi c}{\lambda^4} [\exp(\frac{hc}{\lambda KT}) - 1]^{-1} photons / (scm^2 \mu m) \quad (2)$$

Figure 2 shows a plot of these curves for a number of blackbody temperatures. With a rise in temperature, the energy emission at any given wavelength increases respectively and the wavelength of peak emission decreases, which is specified by Wien’s displacement law. The waveband 1–15 μm in the IR region of the electromagnetic spectrum contains the maximum radiative emission for thermal imaging purposes.

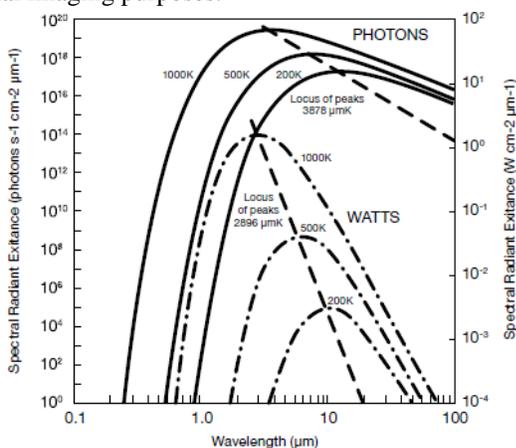


Figure 2. Spectral radiance excitance (Planck’s Law) [8]

The working principle of an IR camera is based on thermography imaging. The major components of the camera are lens, detector, video processing electronics and user interface control. The incident beam of light is focused by the lens upon the detector. The detector contains the IR sensitive elements arranged in the array called focal plane array (FPA). These are IR sensitive elements and miniature in size (micrometers). The resolution of FPA determines the resolution of the IR imagery produced by the camera. Many IR cameras available have user interface to calculate the scene temperature along with imagery recording software. Calibration is often required to read out the correct temperature across the scene that is captured.

With the advancement in the IR camera-technology, significant work has been done to observe ice/snow emissive properties to support remote microwave detection. Snow and clear ice have high emissivity values, which makes them

convenient to thermally image. Snow emissivity has been variously found to be 0.98 for frost [9], above 0.98 for small-grained snow under 1000 μm [10], 0.96 for flat solid ice and water, and as low as 0.8 for old snow. Measurements obtained at particular wavelengths (e.g. 10 μm) have observed emissivity as high as 0.995. Conversely, the thermal reflectance of snow r is less than 2.3% between the 4 to 14 μm wavelengths [11]. This holds good for various snow types including granular, fine, wet, and dry; whereas newly fallen snow reflectance is less than 1% between 4 to 14 μm . For most of the 7.5 to 13 μm spectrum, the emissivity is found higher in comparison. Therefore, IR cameras operating in the range above 7.5 μm can prove significant to observe cold objects.

III. PURE ICE BLOCK THERMAL SIGNATURE

In order to test the potential of IR ice detection, it is necessary to evaluate the thermal signature of icing. In order to observe the IR signature of pure ice, lab experimentation was performed. The sample was prepared from freezing tap water in a commercial freezer. The dimension of the ice block prepared was 14.5 x 14 x 5.5 cm. Although the dimensions were randomly chosen, it was taken into consideration that the ice block should have a large enough surface area to allow observation of a differentiating thermal signature and a wide range of temperature profiles. As a starting point, the viewing angle of the IR camera and the ice sample was kept at 90 degrees. The forward-looking IR camera FLIR A310 was used to observe the thermal signature of the block using the proprietary software of the FLIR device. The observations acquired were produced in the mentioned calibrated software. FLIR A310 series IR camera has a detector operating in the range of MWIR (8-11 μm). The thermal signature was recorded immediately from the time when sample was taken out of the freezer. The observations were recorded until the melt-down phenomena had started at room temperature and surface of the ice block acquired a uniform temperature.

The lab setup performed is shown in the Figure 3. An attempt was made to observe a freely suspended ice block in order to ensure the uniform heat transfer from all directions. To hold the ice block on the stand, a wooden piece was immersed into the cold water during freezing process, so that it could be used as a suspension support assisted by a clamping stand as shown in Figure 3. Minimal thermal conduction from the ice block was expected from the wooden support (since the wood has poor conductivity), apart from fulfilling the suspension support requirement. The ice block was initially frozen to approximate -26°C and suspended on the stand immediately after taking out of the freezer and observed from IR camera. FLIR A310 has an uncooled detector device that sends IR frames processed at the rate of 5Hz. The FLIR calibrated software acquires the frames and saves them as a sequence file. The camera to software communication is performed via an Ethernet protocol. The IR image of the larger surface-1 and relatively smaller surface-2 (Figure 3) was recorded at room temperature to evaluate the surface temperature profiles.

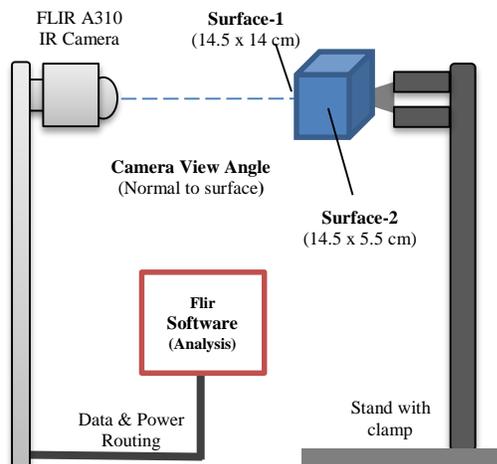


Figure 3. Lab setup of pure ice block thermal signature evaluation

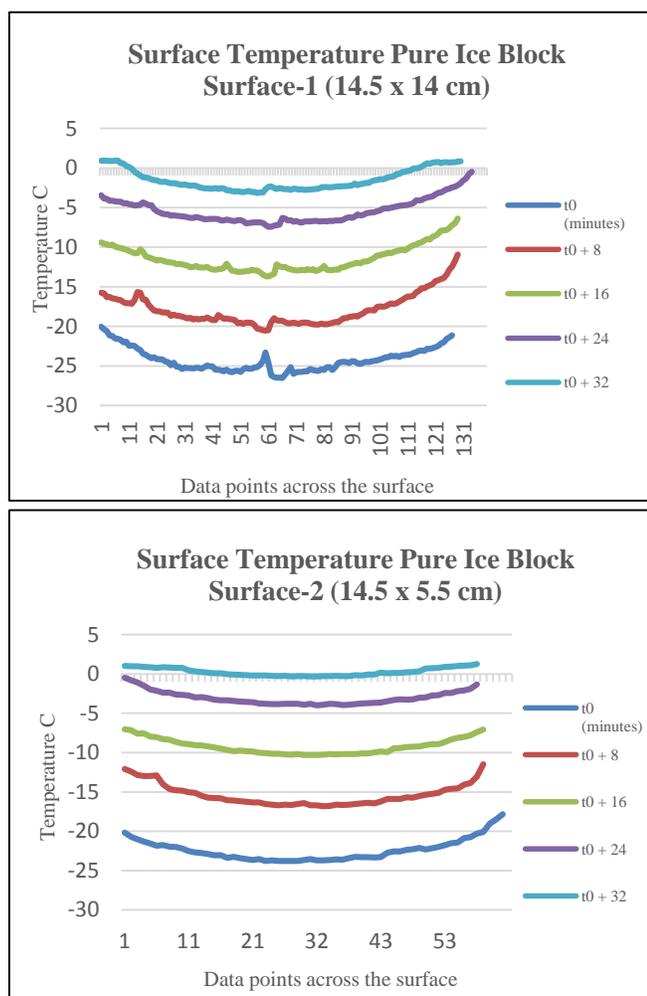


Figure 4. Surface temperature of Ice block recorded from thermal IR signature at different intervals till meltdown starts

The temperature observations were made along a linear path along the ice block's centre line to capture the maximum thermal gradient across the surface. This helped to acquire the maximum and minimum temperature profiles over the surfaces 1, 2 (Figure 4). Minimum temperatures of -26.5°C for surface-1 and -23.7°C for surface-2 were recorded. The gradual rise in temperature for both the surfaces was observed simultaneously as the heat transfer process took place at room temperature. The observations for surface-1 are noted after regular intervals of time t_0 to t_0+32 and also simultaneously for surface-2 as shown in Figure 4. The thermal IR signature available from the camera optics was captured during the course of the observations, which presents the distribution of the thermal profile as shown in Figure 5. The lowest temperature profile was observed almost at the centre of the ice block surfaces whereas the maximum temperature at the boundaries.

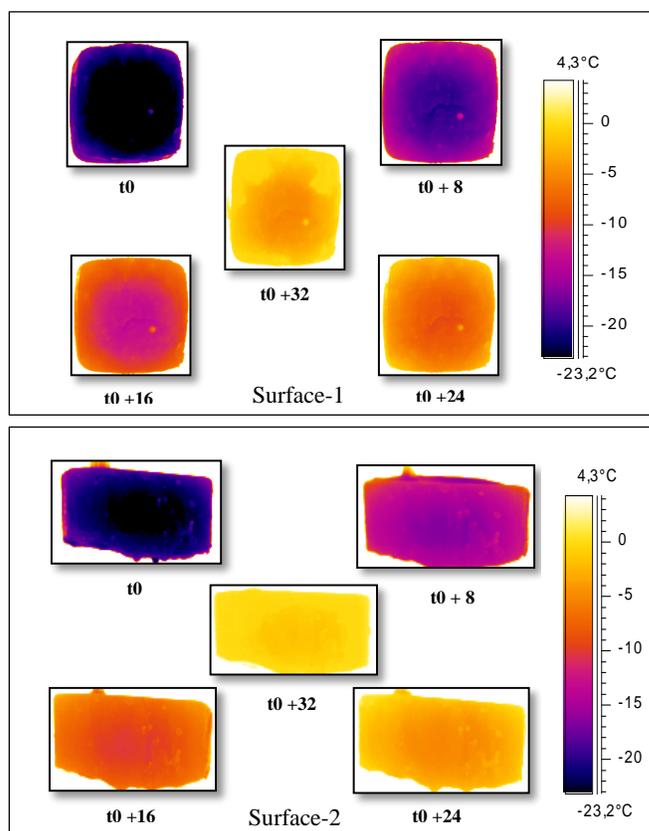


Figure 5 Thermal IR signature distribution across surface

IV. DISCUSSION

Pure ice has a high emissive value and it produces a distinguishable thermal signature shown in Figure 5. The ice block suspended through the wooden stick allowed the ice to conduct heat uniformly. The IR observations show the thermal gradient from the centre of the ice block towards the outside, which follows the heat transfer process. The heat conduction through the boundary layer of the ice can be easily identified from the thermal signatures as shown in

Figure 5. The minimum temperature at surface-1 is -26.5°C at t_0 (Figure 4), which is located in the geometric centre of the surface. A similar behaviour is observed at surface-2. In terms of the maximum temperature profile at the boundary layers, not much difference was observed between both the surfaces. The highest temperature of 0.92°C was observed at the boundary of surface-1 at t_0 as compared to the 1.1°C of surface-2 shown in Figure 4. The maximum-minimum difference in the two surface areas was noted as surface-1 has wider span of temperature range as compared to the surface-2 possibly because of the larger surface area. The maximum-minimum temperature difference at time t_0 to t_0+32 for surface-1 was 6.44°C , 9.61°C , 7.32°C , 6.98°C and 4.07°C respectively. At similar time intervals (t_0 to t_0+32) for surface-2, the difference between maximum and minimum temperature recorded was 5.99°C , 5.24°C , 3.2°C , 3.2°C and 1.59°C . The temperature profiles of the ice block can be correlated with the isothermal images shown in the Figure 5. In the IR image of both the ice surfaces the difference in thermal signature can also be observed at the boundary layer of the ice block, which is at a higher temperature as compared to the rest of the surface block. The temperature profile decreases from corners to the centre of the ice block. Since the room temperature was uniform throughout the experiment, it can be safely said that maximum heat transfer occurred from the boundary through natural convection.

It is to be noted that the wide range of the IR profile of the pure ice block also assists in differentiating it with a relatively hotter environment, though in a cold climate environment this may not be the scenario. IR detection for colder objects seems to work well as long as thermal non-uniformity is present in the scene which is to be observed. Currently, offshore structures and marine ships operating in cold environment use thermal methods for anti/de-icing, apart from other methods. Thermally active heated floors will generate a predictable IR signature. Ice accretion on the floors may result in different thermal signature. This variation may be used for ice detection and may also lead to the detection of the ice accretion rate if used in conjunction with heat transfer theory.

V. CONCLUSIONS

The thermal signature of icing can be applied to detect cold objects especially ice accreted upon the structures. The lab experimentation shows the thermal signature and gradient of two surface areas of a pure ice block. The results shown are preliminary and only reflect the changes in temperature due to warming of ice block under room temperature. LWIR FLIR A310 camera can be used to monitor IR signature of ice. Study of ice block showed that temperature varies over time with maximum value at the boundaries and minimum at the centre. The difference between the boundary and the centre temperatures is higher on a larger surface area. The influence of emissivity values on ice detection is not discussed here and needs to be further investigated.

Considering marine arctic operations, an IR imaging device is to be tested and advanced experimentation is needed to be performed to validate the ice detection and growth.

ACKNOWLEDGEMENTS

The work reported in this paper is funded by the MAROFF, project no. 195153/160 in collaboration with Faroe Petroleum. We would also like to acknowledge the support given by Prof. James Mercer at University of Tromsø.

REFERENCES

1. Fikke, S., Cost 727: atmospheric icing on structures. Measurements and data collection on icing: State of the Art, Publication of MeteoSwiss, 2006. 75(110): p. 1422-1381.
2. Makkonen, L., Salinity and growth rate of ice formed by sea spray. Cold Regions Science and Technology, 1987. 14(2): p. 163-171.
3. Ryerson, C.C., Ice protection of offshore platforms. Cold Regions Science and Technology, 2011. 65(1): p. 97-110.
4. Farzaneh, M. and C.C. Ryerson, Anti-icing and deicing techniques. Cold Regions Science and Technology, 2011. 65(1): p. 1-4.
5. Ryerson, C.C., Icing Management for Coast Guard Assets. 2013, Engineering Research and Development Center, Cold regions research and engineering lab: Hanover NH.
6. Dhar, N.K., R. Dat, and A.K. Sood, Advances in Infrared Detector Array Technology. Optoelectronics - Advanced Materials and Devices. 2013.
7. Rees, W.G., Remote sensing of snow and ice. 2005: CRC Press.
8. Rogalski, A., Infrared detectors: an overview. Infrared Physics & Technology, 2002. 43(3): p. 187-210.
9. Wolfe, W.L. and G.J. Zissis, The infrared handbook. Arlington: Office of Naval Research, Department of the Navy, 1978, edited by Wolfe, William L.; Zissis, George J., 1978. 1.
10. Dozier, J. and S.G. Warren, Effect of viewing angle on the infrared brightness temperature of snow. Water Resources Research, 1982. 18(5): p. 1424-1434.
11. Salisbury, J.W., D.M. D'Aria, and A. Wald, Measurements of thermal infrared spectral reflectance of frost, snow, and ice. Journal of Geophysical Research: Solid Earth (1978–2012), 1994. 99(B12): p. 24235-24240.