

Sergio Saponara
Alessandro De Gloria *Editors*

Applications in Electronics Pervading Industry, Environment and Society

APPLEPIES 2021

Lecture Notes in Electrical Engineering

Volume 866

Series Editors

Leopoldo Angrisani, Department of Electrical and Information Technologies Engineering, University of Napoli Federico II, Naples, Italy

Marco Arteaga, Departament de Control y Robótica, Universidad Nacional Autónoma de México, Coyoacán, Mexico

Bijaya Ketan Panigrahi, Electrical Engineering, Indian Institute of Technology Delhi, New Delhi, Delhi, India

Samarjit Chakraborty, Fakultät für Elektrotechnik und Informationstechnik, TU München, Munich, Germany

Jiming Chen, Zhejiang University, Hangzhou, Zhejiang, China

Shanben Chen, Materials Science and Engineering, Shanghai Jiao Tong University, Shanghai, China

Tan Kay Chen, Department of Electrical and Computer Engineering, National University of Singapore, Singapore, Singapore

Rüdiger Dillmann, Humanoids and Intelligent Systems Laboratory, Karlsruhe Institute for Technology, Karlsruhe, Germany

Haibin Duan, Beijing University of Aeronautics and Astronautics, Beijing, China

Gianluigi Ferrari, Università di Parma, Parma, Italy

Manuel Ferre, Centre for Automation and Robotics CAR (UPM-CSIC), Universidad Politécnica de Madrid, Madrid, Spain

Sandra Hirche, Department of Electrical Engineering and Information Science, Technische Universität München, Munich, Germany

Faryar Jabbari, Department of Mechanical and Aerospace Engineering, University of California, Irvine, CA, USA

Limin Jia, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China

Janusz Kacprzyk, Systems Research Institute, Polish Academy of Sciences, Warsaw, Poland

Alaa Khamis, German University in Egypt El Tagamoa El Khames, New Cairo City, Egypt

Torsten Kroeger, Stanford University, Stanford, CA, USA

Yong Li, Hunan University, Changsha, Hunan, China

Qilian Liang, Department of Electrical Engineering, University of Texas at Arlington, Arlington, TX, USA

Ferran Martín, Departament d'Enginyeria Electrònica, Universitat Autònoma de Barcelona, Bellaterra, Barcelona, Spain

Tan Cher Ming, College of Engineering, Nanyang Technological University, Singapore, Singapore

Wolfgang Minker, Institute of Information Technology, University of Ulm, Ulm, Germany

Pradeep Misra, Department of Electrical Engineering, Wright State University, Dayton, OH, USA

Sebastian Möller, Quality and Usability Laboratory, TU Berlin, Berlin, Germany

Subhas Mukhopadhyay, School of Engineering & Advanced Technology, Massey University, Palmerston North, Manawatu-Wanganui, New Zealand

Cun-Zheng Ning, Electrical Engineering, Arizona State University, Tempe, AZ, USA

Toyoaki Nishida, Graduate School of Informatics, Kyoto University, Kyoto, Japan

Luca Oneto, Dept. of Informatics, Bioengg., Robotics, University of Genova, Genova, Genova, Italy

Federica Pascucci, Dipartimento di Ingegneria, Università degli Studi "Roma Tre", Rome, Italy

Yong Qin, State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University, Beijing, China

Gan Woon Seng, School of Electrical & Electronic Engineering, Nanyang Technological University, Singapore, Singapore

Joachim Speidel, Institute of Telecommunications, Universität Stuttgart, Stuttgart, Germany

Germano Veiga, Campus da FEUP, INESC Porto, Porto, Portugal

Haitao Wu, Academy of Opto-electronics, Chinese Academy of Sciences, Beijing, China

Walter Zamboni, DIEM - Università degli studi di Salerno, Fisciano, Salerno, Italy

Junjie James Zhang, Charlotte, NC, USA

The book series *Lecture Notes in Electrical Engineering* (LNEE) publishes the latest developments in Electrical Engineering - quickly, informally and in high quality. While original research reported in proceedings and monographs has traditionally formed the core of LNEE, we also encourage authors to submit books devoted to supporting student education and professional training in the various fields and applications areas of electrical engineering. The series cover classical and emerging topics concerning:

- Communication Engineering, Information Theory and Networks
- Electronics Engineering and Microelectronics
- Signal, Image and Speech Processing
- Wireless and Mobile Communication
- Circuits and Systems
- Energy Systems, Power Electronics and Electrical Machines
- Electro-optical Engineering
- Instrumentation Engineering
- Avionics Engineering
- Control Systems
- Internet-of-Things and Cybersecurity
- Biomedical Devices, MEMS and NEMS

For general information about this book series, comments or suggestions, please contact leontina.dicecco@springer.com.

To submit a proposal or request further information, please contact the Publishing Editor in your country:

China

Jasmine Dou, Editor (jasmine.dou@springer.com)

India, Japan, Rest of Asia

Swati Meherishi, Editorial Director (Swati.Meherishi@springer.com)

Southeast Asia, Australia, New Zealand

Ramesh Nath Premnath, Editor (ramesh.premnath@springernature.com)

USA, Canada:

Michael Luby, Senior Editor (michael.luby@springer.com)

All other Countries:

Leontina Di Cecco, Senior Editor (leontina.dicecco@springer.com)

**** This series is indexed by EI Compendex and Scopus databases. ****

More information about this series at <https://link.springer.com/bookseries/7818>

Sergio Saponara · Alessandro De Gloria
Editors

Applications in Electronics Pervading Industry, Environment and Society

APPLEPIES 2021

 Springer

Editors

Sergio Saponara
DII
University of Pisa
Pisa, Italy

Alessandro De Gloria
DITEN
University of Genoa
Genoa, Italy

ISSN 1876-1100

ISSN 1876-1119 (electronic)

Lecture Notes in Electrical Engineering

ISBN 978-3-030-95497-0

ISBN 978-3-030-95498-7 (eBook)

<https://doi.org/10.1007/978-3-030-95498-7>

© The Editor(s) (if applicable) and The Author(s), under exclusive license
to Springer Nature Switzerland AG 2022

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

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

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

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

The 2021 edition of the Conference on “Applications in Electronics Pervading Industry, Environment and Society” was held in mixed mode, i.e., in presence and online, during September 21 and 22, 2021.

The sessions in presence were held at the Aula Magna Pacinotti at the School of Engineering, University of Pisa, in Pisa, Italy.

During two days, about 1 hundred of registered participants, from different entities (Universities and industries), discussed electronic applications in several domains, demonstrating how electronics has become pervasive and ever more embedded in everyday objects and processes.

The conference had the technical and/or financial support of University of Pisa (Prof. Sergio Saponara being the general chair), University of Genoa (Prof. Alessandro De Gloria being the honorary chair), SIE (Italian Association for Electronics), and of the H2020 European Processor Initiative.

After a strict blind-review selection process, 19 short presentations and 25 lectures have been accepted and presented in six sessions (four regular sessions, two short sessions) focused on circuits and electronic systems and their relevant applications in the following fields: High performance computing (HPC) and digital continuum, wireless and IoT, health care, vehicles and robots, power electronics and energy storage, cybersecurity, AI and data engineering.

There were also two scientific keynote sessions, focused on the roadmap of EuroHPC and the European Processor Initiative. The keynote “The European Roadmap Towards High Performance Computing: Industrial and Scientific Perspectives” was held by J. P. Panziera from ATOS (worldwide leading industry in the HPC field) and B. Mohr from FZ Juelich, a German research center leader in Europe for HPC.

The keynote “High Performance Computing Continuum: The Italian Industry in the European Processor Initiative and Pilots” was held by F. Magugliani from E4; D. Ghezzi from LEONARDO; F. Ottonelli and G. Venere from SECO.

The articles featured in this book, together with the talks and round tables of the special events, prove that the capabilities of nowadays electronic systems, in terms of computing, storage and networking, are able to support a plethora of application

domains, such as mobility, health care, connectivity, energy management, smart production, ambient intelligence, smart living, safety and security, education, entertainment, tourism and cultural heritage.

In order to exploit such capabilities, multidisciplinary knowledge and expertise are needed to support a virtuous iterative cycle from user needs to the design, prototyping and testing of new products and services that are more and more characterized by a digital core.

The design and testing cycles go through the whole system engineering process, which includes analysis of user requirements, specification definition, verification plan definition, software and hardware co-design, laboratory and user testing and verification, maintenance management and life cycle management of electronics applications.

The design of electronics-enabled systems should be characterized by innovation, high performance, real-time operations and budget compliance (in terms of time, cost, device size, weight, power consumption, etc.). Design methodologies and tools have emerged in order to support teams dealing with such a complexity.

All these challenging aspects call for the importance of the role of Academia as a place where new generations of designers can learn and practice with the cutting-edge technological tools and where new solutions are studied, starting from challenges coming from a variety of application domains. This approach is sustained by industries that understand the role of a high-level educational system and able to nurture new generations of designers and developers.

The Conference on Applications in Electronics Pervading Industry, Environment and Society has reached, in 2021, its edition number 9, confirming its role as a reference point for a growing research community in the field of electronics systems design, with a particular focus on applications.

Contents

An Intelligent Non-cooperative Spectrum Sensing Method Based on Convolutional Auto-encoder (CAE)	1
Qinghe Zheng, Hongjun Wang, Abdussalam Elhanashi, Sergio Saponara, and Deliang Zhang	
Impact of Image Resizing on Deep Learning Detectors for Training Time and Model Performance	10
Sergio Saponara and Abdussalam Elhanashi	
Preliminary Design of a Three-Dimensional Anemometer for Sail Boats	18
Enrico Boni, Luca Pugi, and Alessio Venturi	
Design and Preliminary Testing of an Electrified Directional Drilling Machine	24
Lorenzo Berzi, Francesco Grasso, Luca Pugi, Enrico Boni, and Raffaele Savi	
CRFlex: A Flexible and Configurable Cryptographic Hardware Accelerator for AES Block Cipher Modes	31
Pietro Nannipieri, Luca Baldanzi, Luca Crocetti, Stefano Di Matteo, Francesco Falaschi, Luca Fanucci, and Sergio Saponara	
A M-PSK Timing Recovery Loop Based on Q-Learning	39
Gian Carlo Cardarilli, Luca Di Nunzio, Rocco Fazzolari, Daniele Giardino, Matteo Guadagno, Marco Re, and Sergio Spanò	
Scalable Broadband Switching Matrix for Telecom Payload Based on a Novel SWGs-Based MZI	45
G. Brunetti, G. Marocco, A. Giorgio, M. N. Armenise, and C. Ciminelli	
A Smart Portable Potentiostat for Point-of-Care Testing	53
Marco Bassoli, Valentina Bianchi, Andrea Boni, Simone Fortunati, Marco Giannetto, Maria Careri, and Ilaria De Munari	

Experimental Results of Vectorized Posit-Based DNNs on a Real ARM SVE High Performance Computing Machine	61
Marco Cococcioni, Federico Rossi, Emanuele Ruffaldi, and Sergio Saponara	
An Open-Source Hardware/Software Architecture for Remote Control of SoC-FPGA Based Systems	69
Werner Florian, Bruno Valinoti, Luis G. García, Marcos Cervetto, Edgardo Marchi, Maria Liz Crespo, Sergio Carrato, and Andres Cicuttin	
A Self Referencing Technique for the RC-pLMS Adaptive Beamformer and Its Hardware Implementation	76
Ghattas Akkad, Ali Mansour, Bachar ElHassan, Elie Inaty, and Rafic Ayoubi	
A Data-Driven Method for Reliability Estimation of Auxiliary Power Consumption Prediction in Commercial Electric Vehicles	86
Tommaso Apicella, Edoardo Ragusa, Alessio Canepa, and Paolo Gastaldo	
Compression of NN-Based Pulse-Shape Discriminators in Front-End Electronics for Particle Detection	93
Romina Soledad Molina, Luis Guillermo Garcia, Iván René Morales, Maria Liz Crespo, Giovanni Ramponi, Sergio Carrato, Andres Cicuttin, and Hector Perez	
Assisted Driving for Power Wheelchair: A Segmentation Network for Obstacle Detection on Nvidia Jetson Nano	100
Gianluca Giuffrida, Silvia Panicacci, Massimiliano Donati, and Luca Fanucci	
Analysis of Thermal-Induced Shunt Current Sensor Errors in a Low-Cost Battery Management System	107
Alessandro Verani, Roberto Di Rienzo, Federico Baronti, Roberto Roncella, and Roberto Saletti	
Microaggregation Optimisation Through Random Cluster Shuffling	114
Armando Maya-López, Fran Casino, Agusti Solanas, and Antoni Martínez-Ballesté	
Preliminary Design of a Flexible Test Station for Second-Life Battery Development	120
Andrea Carloni, Stefano Constà, Manlio Pasquali, Federico Baronti, Roberto Di Rienzo, Roberto Roncella, and Roberto Saletti	
Novel Setup to Extend the Temperature Characterization Range of a Sodium-Metal Halide Battery	126
Gianluca Simonte, Roberto Di Rienzo, Ian Biagioni, Federico Baronti, Roberto Roncella, and Roberto Saletti	

An Effective Approach to the Cross-Border Exchange of Digital Evidence Using Blockchain 132
 Pablo López-Aguilar and Agusti Solanas

TinyML Platforms Benchmarking 139
 Anas Osman, Usman Abid, Luca Gemma, Matteo Perotto,
 and Davide Brunelli

Automatic Design Space Exploration of Redundant Architectures 149
 Antonio Tierno, Giuliano Turri, Alessandro Cimatti,
 and Roberto Passerone

Visible Light Communication for Intermittent Computing Battery-Less IoT Devices 155
 Alessandro Torrisi, Federico Baggio, and Davide Brunelli

Resource Optimization in MEC-Based B5G Networks for Indoor Robotics Environment 164
 Tadeus Prastowo, Ayub Shah, Luigi Palopoli, and Roberto Passerone

Signal Alignment Problems on Multi-element X-Ray Fluorescence Detectors 173
 Francesco Guzzi, George Kourousias, Fulvio Billé, Gioia Di Credico,
 Alessandra Gianoncelli, and Sergio Carrato

Low-Level Advanced Design of True Random Number Generators Based on Truly Chaotic Digital Nonlinear Oscillators in FPGAs 180
 Tommaso Addabbo, Ada Fort, Riccardo Moretti, Marco Mugnaini,
 and Valerio Vignoli

Design and Implementation of an FPGA-Based CNN Hardware Accelerator Using Partial Reconfigurability: The CloudScout Case Study 187
 Corrado Comino, Tommaso Pacini, Emilio Rapuano, and Luca Fanucci

Exploring GPS L1 C/A Fast Acquisition with COTS FPGA 194
 Andrea Romani, Franco Bigongiari, and Luca Fanucci

Feasibility Study of a Unified Fast Acquisition Core for Modern GPS Signals 200
 Andrea Romani, Franco Bigongiari, and Luca Fanucci

Evaluating Body Movement and Breathing Signals for Identification of Sleep/Wake States 206
 Maksym Gaiduk, Ralf Seepold, Natividad Martínez Madrid,
 Thomas Penzel, Lucas Weber, Massimo Conti, Simone Orcioni,
 and Juan Antonio Ortega

Comparison of a Medical-Grade and an Open ECG Biosensor Using a Soft Real-Time m-Health Platform	212
Miltos D. Grammatikakis and Michael Androulakis	
FPGA Implementation of a Configurable Vocal Feature Extraction Embedded System for Dysarthric Speech Recognition	221
Iacopo Casalini, Marco Marini, and Luca Fanucci	
Classifying Simulated Driving Scenarios from Automated Cars	229
Marianna Cossu, Jorge Leonardo Quimi Villon, Francesco Bellotti, Alessio Capello, Alessandro De Gloria, Luca Lazzaroni, and Riccardo Berta	
Mismatch Analysis of Parallel Li-Ion Batteries	236
Massimo Conti, Luca Stacchiotti, and Simone Orcioni	
Efficient Training and Hardware Co-design of Machine Learning Models	243
Mohammad Amir Mansoori and Mario R. Casu	
Modeling the Line Interruption Issue in a Railway Network	249
Luca Fronza, Riccardo Berta, Paolo Cesario, Alessandro De Gloria, and Francesco Bellotti	
The SENSIPLUS: A Single-Chip Fully Programmable Sensor Interface	256
Andrea Ria, Mattia Cicalini, Giuseppe Manfredini, Alessandro Catania, Massimo Piotto, and Paolo Bruschi	
DoS Detection on In-Vehicle Networks: Evaluation on an Experimental Embedded System Platform	262
Miltos D. Grammatikakis, Nikos Mouzakis, Lefteris Kypraios, and Nikos Papatheodorou	
Simulation Environment for Mixed AHB-NoC Architectures	273
Massimo Conti	
Convolutional Neural Networks Based Tactile Object Recognition for Tactile Sensing System	280
Ali Ibrahim, Haydar Hajj Ali, Mohamad Hajj Hassan, and Maurizio Valle	
Design of V2X Communications Based on 5G NR: A Physical Layer Perspective	286
Fabiola Sapienza, Vincenzo Lottici, Filippo Giannetti, and Sergio Saponara	
A Low Cost Compact Output Amplifier for Multichannel Muscle Stimulation	293
Massimo Ruo Roch and Maurizio Martina	

The Exploitation of Sustainable Composite Materials for the Manufacturing of High-Efficient Electric Cars 300
Jacopo Agnelli, David Benedetti, Nicholas Fantuzzi, and Sergio Saponara

Developing a Synthetic Dataset for Driving Scenarios 310
Jacopo Motta, Francesco Bellotti, Riccardo Berta, Alessio Capello, Marianna Cossu, Alessandro De Gloria, Luca Lazzaroni, and Stefano Bonora

Smart On-Board Surveillance Module for Safe Autonomous Train Operations 317
G. Mezzina, M. Barbareschi, Salvatore De Simone, Alessandro Di Benedetto, G. Narracci, C. L. Saragaglia, D. Serra, and Daniela De Venuto

Author Index 327



An Intelligent Non-cooperative Spectrum Sensing Method Based on Convolutional Auto-encoder (CAE)

Qinghe Zheng¹, Hongjun Wang¹(✉), Abdussalam Elhanashi², Sergio Saponara²,
and Deliang Zhang¹

¹ School of Information Science and Engineering, Shandong University, Qingdao 266237,
Shandong, China

hjw@sdu.edu.cn

² Department of Information Engineering, University of Pisa, Via G. Caruso 16, 56122 Pisa,
Italy

Abstract. As an opportunistic spectrum utilization technology, cognitive radio can greatly improve the spectrum utilization efficiency and alleviate the scarcity of spectrum resources. Spectrum sensing technique is key premise of realizing legitimate spectrum access in cognitive radio. In this paper, we propose to use a convolutional auto-encoder to solve the instability problem caused by complex environments in the traditional spectrum sensing process. The reconstruction error of deep learning model based on normal spectrum is an effective measure to judge whether the test signals are authorized or not. Moreover, the essential characterization capability of convolutional auto-encoder makes the metric well adapted to different environments and meet practical requirements. Finally, the effectiveness of the proposed method is verified by using a self-built broadcast dataset. Compared with state-of-the-art methods including PCA reconstruction, energy detection, and cyclostationary detection, the CAE based method shows better identification accuracy and robustness for unauthorized radio.

Keywords: Spectrum sensing · Unauthorized broadcasting identification (UBI) · Convolutional auto-encoder (CAE) · Cognitive radio

1 Introduction

As a strategic natural resource, electromagnetic spectrum is the most ideal wireless information transmission medium. With the continuous increase of different networks, services, and access methods in wireless communications, spectrum resources are becoming more and more tight. On the other hand, a large amount of available wireless spectrum is idle or used inefficiently. In this situation, spectrum sensing has become a key task in the cognitive radio system. In recent years, the development of advanced technologies such as machine learning [1], pattern recognition [2], and artificial intelligence [3] has improved the accuracy and efficiency of the perception, reasoning, and prediction of the electromagnetic spectrum, laying a solid foundation for the efficient usage of the

electromagnetic spectrum. Typically, the application of deep learning methods [4] in the intelligent electromagnetic spectrum sensing can help improve the timeliness of information transmission and promote the intelligent decision-making of electromagnetic spectrum usage.

Although deep learning method has made noteworthy achievements in many fields including computer vision [5] and medical analysis [6], its applications in practical spectrum sensing still faces a series of challenges:

- The rapidly increasing use of wireless spectrum challenges the performance of traditional sensing methods, including accuracy and inference speed.
- The variety of authorized signal types brings difficulties to modeling, such as different channel parameters and different levels of channel interference.
- The generalization ability of machine learning models is difficult to meet the requirements of complex electromagnetic environments, such as low signal-to-noise ratio (SNR), uncertain noise, and unknown user prior information.

In this paper, we turn the spectrum sensing problem into the classification problem, that is, to identify unauthorized signals from unknown signals. By training the CAE reconstructor on normal signals, the reconstruction error of the test signal can be used as the criterion to judge whether it is authorized or not. The deep learning model can mine the common implicit characteristics of normal signals, thereby improving the discrimination capability of reconstruction error. Simulation results show that the proposed spectrum sensing method based on CAE has better performance than energy detection, cycle detection, and PCA reconstruction.

The remainder of this paper is organized as follows. Section 2 gives a brief review of related work of spectrum sensing. In Sect. 3, we introduce the CAE model for unauthorized signal identification. Experimental results and analysis are presented in Sect. 4. Finally, we conclude our work and future direction in Sect. 5.

2 Related Work

The spectrum sensing system allows cognitive radio signal to detect unused parts of radio spectrum. Last decades, there has been several activities for investigating the enhancement of detecting the performance of energy by using a dynamic selecting of a threshold [7]. spectrum sensing systems have been deployed in many applications. As explored by the Nyquist theory, a simple method has been proposed to spectrum sensing system, which is to acquire the wideband signal by a high speed analog-to-digital converter. For instance, Quan et al. [8] suggested several band joint algorithms to detect PUs over wideband spectrum by utilizing a high speed analogue to digital converter (ADC) for signal acquisition. In addition to that, Tian and Giannakis [9] suggested a specific method which is the wavelet-based wideband spectrum sensing approach that implements a high-speed ADC. However, this is specific for power limited such as smart phones, and battery-free devices in a wireless power transfer system [10]. Recently, Landau [11] proposed that sampling rate shall be no less than the measured value of occupied portion of the spectrum for steady recreation of multi-band signals.

In addition to that, spectrum sensing indicates a property of sparsity in the domain of frequency as its low utilization in execution [12]. Zhang et al. [13] suggested the dynamic spectrum system with and without moderating for SUs. The author investigates the blocking possibility, the interrupted possibility, the forced termination possibility, the non-completion possibility, and the delayed time of SUs by the Markov method. It is exhibited that the buffering minimizes the blocking possibility with tiny increase in a forced termination possibility. Moreover, Tumuluru et al. [14] proposed two different policies to operate the spectrum duty and handoff for SUs, considering prioritizing the SU traffic. To improve the generalization ability of deep learning models, Zheng et al. [16] introduced a two-stage data augmentation method for expanding sample size without extra CNN training costs. Zhao et al. [17] proposed an asynchronous co-prime sampling technique for reconstructing sparse multi-band signals that occupy a small portion of the given bandwidth. Soni et al. [18] propose an LSTM based spectrum sensing (LSTM-SS) method, which learns implicit features from spectrum data, such as time correlation.

3 Convolutional Auto-encoder (CAE)

3.1 Architecture of CAE

The proposed CAE is consisted of two parts, *i.e.*, an encoder $f(\cdot)$ and a decoder $g(\cdot)$, which combine the convolution and deconvolution operations in convolutional neural networks (CNNs) with conventional AE, as shown in Fig. 1.

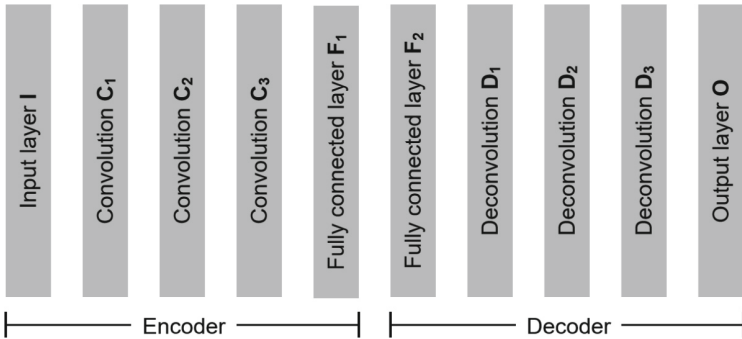


Fig. 1. The specific structure of CAE

The encoder contains an input layer **I**, three convolutional layers **C₁**, **C₂**, and **C₃**, and a fully connected layer **F₁** that maps the signal from input space to the latent code space. In the input layer, I/Q modulated signals s with the size of 1024×2 are fed into the network. Then three convolutional layers including different numbers and sizes of convolution kernels (**C₁**: $256-2 \times 2-1 \times 1-0 \times 0$, **C₂**: $256-2 \times 1-2 \times 1-0 \times 0$, and **C₃**: $128-2 \times 1-1 \times 1-0 \times 0$) extract discriminative features, in which $256-2 \times 2-1 \times 1-0 \times 0$ means the 256 kernels of size 2×2 with a stride of 1×1 and a padding of 0×0 . The fully connected layer containing 256 neurons is used to encode these features. Finally, the

decoder composed of the fully connected layer \mathbf{F}_2 , and three deconvolution layers \mathbf{D}_1 , \mathbf{D}_2 , and \mathbf{D}_3 reconstructs the latent variable back into the input space through the output layer \mathbf{O} . The convolutional layers are used to extract the essential representations of signals, while the deconvolution layers reconstruct signals according to latent variables output by the fully connected layer. Since the model is not very deep, no residual module is added to the structure. The use of dense fully connected layers instead of the global pooling layers in the middle of the structure helps to preserve the complete features of the signal.

In the forward propagation, the output is activated by Swish function, as given by

$$\delta(x) = \frac{x}{1 + e^{-x}} \quad (1)$$

where x is the output and $\delta(\cdot)$ represents the corresponding activation value. Up to the last output layer, the reconstruction loss of the model can be calculated by

$$\mathcal{L} = \frac{1}{M} \sum_{i=1}^M \tau_i = \frac{1}{M} \sum_{i=1}^M \|s_i - f[g(s_i)]\|^2 \quad (2)$$

where τ is the reconstruction error and M is the mini-batch size.

3.2 Training Procedure with RMSProp

Given an initialized CAE model with parameters θ_0 , it can be trained using RMSProp method [15], as given by

$$\theta_t = \theta_{t-1} - \frac{\alpha}{\sqrt{\eta + r_t}} \odot g_t \quad (3)$$

where η is a small constant for numerically stable (usually 10^{-6}) and α represents the learning rate. g and r are gradients and cumulative square gradient respectively and can be calculated according to

$$r_t = \rho r_{t-1} + (1 - \rho) g_t \odot g_t \quad (4)$$

$$g_t = \frac{1}{M} \sum_{i=1}^M \nabla_{\theta_{t-1}} \mathcal{L} \quad (5)$$

where ρ is the decay rate and $r_0 = 0$. The whole training process is iterated until the loss function in Eq. (2) no longer drops. Although many optimization methods have been proposed, such as SGD, mini-batch SGD, and Adam, RMSProp is used to train the model considering its mature hyper-parameter setting.

3.3 Identification Based on Adaptive Threshold

According to the reconstruction error distribution of normal and unauthorized signals in the validation set, the threshold can be adaptively selected as

$$\nu = \frac{1}{2} (\max \min \tau^+ - \min \max \tau^-) \quad (6)$$

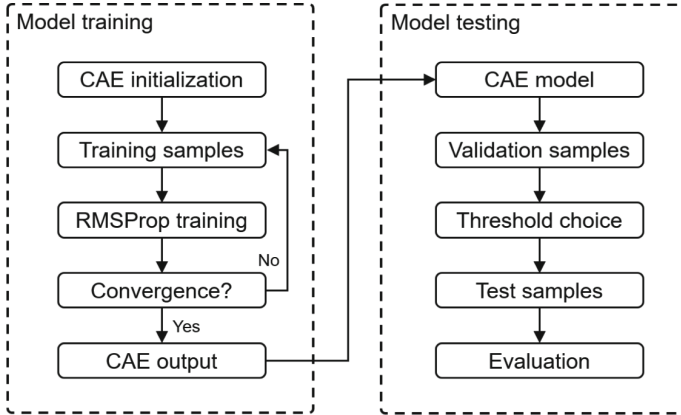


Fig. 2. Training and testing procedures of CAE for unauthorized signal identification.

where τ^+ and τ^- denote reconstruction errors of top ξ normal and unauthorized signals, respectively. The whole training and testing procedures are shown in Fig. 2.

4 Experimental Results and Analysis

In this section, we first introduce the experimental dataset and specific setup during training process of CAE, and then demonstrate the effectiveness of proposed method by comparing a series of algorithms.

4.1 Experimental Dataset and Setup

To verify the spectrum sensing performance of the CAE, a broadcasting signal dataset based on the signal source and spectrum analyzer AV4051D-S is created. A total of 10000 I/Q modulated signals with the size of 2×1024 are collected, of which 2000 normal signals are used for training, 2000 normal signals and 2000 + unauthorized signals are used for validation, and the remaining 2000 normal signals and 2000 unauthorized signals are used for testing. The signal parameters during the acquisition process are summarized in Table 1. Furthermore, the signal waveforms in the time-domain and the corresponding amplitude spectra in the frequency-domain of some example signals are shown in Fig. 3.

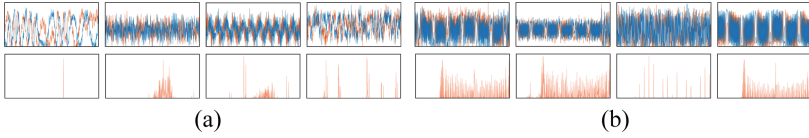
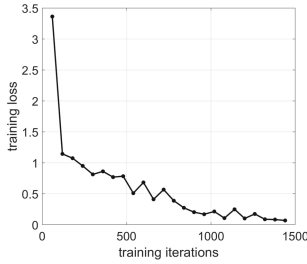
The training and testing process of the CAE is conducted using MATLAB 2019a in the 64-bit Windows 10 operating system, based on a workstation composed of an Intel(R) Core(TM) i9-7900X CPU@3.30 GHz, an NVIDIA TITAN Xp GPU, 2×16 GB RAM, and 2TB HDD. The setup of hyper-parameters before the start of the training is presented in Table 2.

Table 1. Signal parameters during the acquisition process.

Signal parameters	Values
Sampling frequency	0.5 Hz
Frequency step	0.1 MHz
Sampling bandwidth	0.2 MHz
Starting frequency	88 MHz
Termination frequency	108 MHz

Table 2. Hyper-parameters setting during the training process of CAE.

Hyper-parameters	Values
Learning rate	0.01
Batch size	32
Weight decay rate	0.0005
Sparse scale	0.2
Threshold	20%

**Fig. 3.** Examples of authorized and unauthorized broadcasting signals in time (first row) and frequency domains (second row). (a) normal signals, (b) unauthorized signals**Fig. 4.** The descent process of training loss of CAE.

4.2 Unauthorized Broadcasting Identification Performance

In Fig. 4, we present the descent process of training loss of CAE. It can be clearly seen that the model quickly converges to a stable position, *i.e.*, the loss of the model drops

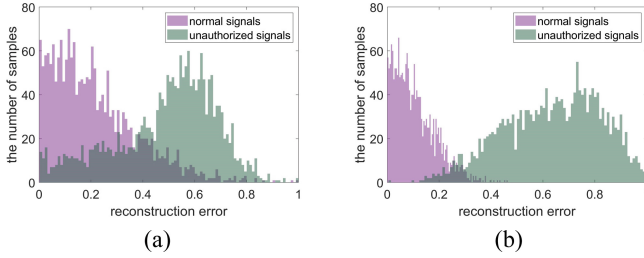


Fig. 5. Reconstruction error distribution of (a) CAE at 100 iterations and (b) convergent CAE on the validation set.

from 3.5 to about 0.2 within 1000 training iterations. In the end, the model can converge at about 0.1 loss.

During RMSProp training iterations, the distribution of normalized reconstruction errors of the validation set gradually become separate due to the discrimination power of the CAE, as shown in Fig. 5. In the beginning, the reconstruction errors of normal and unauthorized signals have a large overlap at 100 iterations. The reconstruction errors of normal and unauthorized signals become more separated at the end. Training on normal samples can the reconstruction errors of normal signals become smaller.

4.3 Comparison with State-Of-The-Art Methods

In this part, we report the identification results of the proposed CAE and compare the accuracy with various state-of-the-art methods, including improved **energy** detection, cyclostationary detection, and PCA reconstruction methods. Their accuracy, precision rate, recall rate, and F1-score are given in Table 3.

Our proposed CAE method achieves the accuracy, precision rate, recall rate, and F1-score of 91.03%, 0.8992, 0.9020, and 0.9006, respectively. The spectrum sensing performance of the energy detection method is easily affected by uncertain noise, and achieves the lowest precision rate. The cyclostationary detection method is difficult to cope with the changing SNRs. The reconstruction method based on PCA shows better performance than the above two methods, but it is not as robust as CAE.

Table 3. Performance comparison with state-of-the-art methods for UBI.

Methods	Accuracy	Precision	Recall	F1-score
Improved energy detection	84.47%	0.8470	0.8530	0.8500
Cyclostationary detection	87.70%	0.8826	0.8600	0.8712
PCA reconstruction	85.66%	0.8507	0.8441	0.8474
Proposed CAE	91.03%	0.8992	0.9020	0.9006

5 Conclusion

In this paper, we proposed an intelligent non-cooperative spectrum sensing method based on CAE, which can improve the performance of spectrum sensing in cognitive radio systems. Through experiments based on self-built dataset, the effectiveness of the proposed method has been verified.

In the future, we plan to test the performance of deep learning models considering temporal information for spectrum sensing. Moreover, some regularization methods to improve the generalization ability of the model will also be considered.

Acknowledgments. This work was supported by National Key R&D Program of China (Grant No. 2018YFF01014304) and Major Basic Research Project of Shandong Provincial Natural Science Foundation (Grant No. ZR2019ZD01).

References

1. Zheng, Q., et al.: Improvement of generalization ability of deep CNN via implicit regularization in two-stage training process. *IEEE Access* **6**, 15844–15869 (2018)
2. Peng, C., et al.: A triple-thresholds pavement crack detection method leveraging random structured forest. *Constr. Build. Mater.* **263**, 120080 (2020)
3. Li, J., et al.: Dynamic hand gesture recognition using multi-direction 3D convolutional neural networks. *Eng. Lett.* **27**(3), 490–500 (2019)
4. Zheng, Q., Zhao, P., Li, Y., Wang, H., Yang, Y.: Spectrum interference-based two-level data augmentation method in deep learning for automatic modulation classification. *Neural Comput. Appl.* **33**(13), 7723–7745 (2020). <https://doi.org/10.1007/s00521-020-05514-1>
5. Zheng, Q., Tian, X., Yang, M., Wu, Y., Su, H.: PAC-bayesian framework based drop-path method for 2D discriminative convolutional network pruning. *Multidimension. Syst. Signal Process.* **31**(3), 793–827 (2020)
6. Ma, X., et al.: Understanding adversarial attacks on deep learning based medical image analysis systems. *Pattern Recogn.* **110**, 107332 (2020)
7. Joshi, D. R., Popescu, D. C., Dobre, A.: Adaptive spectrum sensing with noise variance estimation for dynamic cognitive radio systems. In: *International Conference on Information Sciences and Systems (CISS)*, pp. 1–5, Princeton, USA (2010)
8. Quan, Z., Cui, S., Sayed, A., Poor, H.: Optimal multiband joint detection for spectrum sensing in cognitive radio networks. *IEEE Trans. Signal Process.* **57**(3), 1128–1140 (2009)
9. Tian, Z., Giannakis, G.: A wavelet approach to wideband spectrum sensing for cognitive radios. In: *International Conference on Cognitive Radio Oriented Wireless Network Communications (CROWNCOM)*, pp. 1–5, Mykonos Island, Greece (2006)
10. Zhang, R., Ho, C.: MIMO broadcasting for simultaneous wireless information and power transfer. *IEEE Trans. Commun.* **12**(5), 1989–2001 (2013)
11. Landau, H.: Necessary density conditions for sampling and interpolation of certain entire functions. *Acta Math.* **117**(1), 37–52 (1967)
12. Kolodzy, P., Avoidance I.: Spectrum policy task force report. In: *IEEE Transactions on Information Forensics and Security* (2002)
13. Zhang, Y.: Dynamic Spectrum Access in Cognitive Radio Wireless Networks. In: *IEEE International Conference on Communications (ICC)*, pp. 4927–4932 Beijing, China (2008)
14. Tumuluru, K., Wang, P., Niyato, D., Song, W.: Performance analysis of cognitive radio spectrum access with prioritized traffic. *IEEE Trans. Veh. Technol.* **61**(4), 1895–1906 (2012)

15. Zheng, Q., Tian, X., Jiang, N., Yang, M.: Layer-wise learning based stochastic gradient descent method for the optimization of deep convolutional neural network. *J. Intell. Fuzzy Syst.* **37**(4), 5641–5654 (2019)
16. Zheng, Q., Yang, M., Tian, X., Jiang, N., Wang, D.: A full stage data augmentation method in deep convolutional neural network for natural image classification. *Discrete Dyn. Nat. Soc.* **2020**, 1–11 (2020)
17. Zhao, Y., Xiao, S.: Sparse multiband signal spectrum sensing with asynchronous coprime sampling. *Clust. Comput.* **22**, 4693–4702 (2019)
18. Soni, B., Patel, K., Lopez-Benitez, M.: Long short-term memory based spectrum sensing scheme for cognitive radio using primary activity statistics. *IEEE Access* **8**, 97437–97451 (2020)



Impact of Image Resizing on Deep Learning Detectors for Training Time and Model Performance

Sergio Saponara^(✉) and Abdussalam Elhanashi

Department of Information Engineering, University of Pisa, Via G. Caruso 16, 56122 Pisa, Italy
sergio.saponara@unipi.it

Abstract. Resizing images is a critical pre-processing step in computer vision. Principally, deep learning models train faster on small images. A larger input image requires the neural network to learn from four times as many pixels, and this increase the training time for the architecture. In this work, we presented the evolution of effects of image resizing on model training time and performance. This study is applied on a vehicle dataset. We used You Look Only Once based architectures which include YOLOv2, YOLOv3, YOLOv4, and YOLOv5 with pretrained models to perform object detection. YOLO is designed to detect objects with high accuracy and high speed, which is an advent for real-time applications. Data augmentation method is used in this research to reduce overfitting problems, which approximates the data probability by manipulating the input samples. The experimental results show that if the input image size varies, then it has effects on the training time of the CNN based images classification. Additionally, this research reviewed image resizing and its impacts on the models' performance in terms of accuracy, precision, and recall.

Keywords: Resizing images · Neural network · You Look Only Once · Object detection · Data augmentation

1 Introduction

Convolutional neural network (CNN) is the state-of-the-art approach in computer vision for image classification. CNNs have significant improvement for accuracy, which includes increasing the number of neurons, depth of the neural network, modifying the deep learning models, and tuning the hyper-parameters [1, 2]. Pretrained models such as Resnet [3], Darknet53 [4], and SqueezeNet [5]. These models were designed to obtain the best accuracy for object detection on multiple datasets. All these models have input image layers, which introduce the input images to the neural network architectures. Therefore, every image which has different resolution sizes is rescaled to fit the neural network input's restrain. The size of images differs due to the various input sources. Since deep learning models receive an input of fixed size, all images are required to be rescaled to one size before feeding them to the convolutional neural network architectures [6]. The larger the size of the image, the less shrinking is needed. However, large

images not only occupy additional space in memory size, but also require large neural network architecture, thus, increasing the time complexity and the space in memory. It is necessary that selecting the proper size of images is a tradeoff between accuracy and computational efficiency. Data preprocessing involves using techniques such as data augmentation, standardization, and normalization to rescale input and output values prior to training the neural network architecture. During the training process, learning features are performed with batches, which means a set of data represented as a tensor. These tensors are taken through convolutional layers followed with activation functions to obtain the activation volume, pooling layer for down sampling the images. When classifying an image, all information flows through all these layers in one direction from the input layer to the output layer. All these considerations motivate our work to experiment different sizes of images for vehicle detection with YOLO (You Look Only Once) series and pretrained models. This is to examine the impact of image resizing and its effects on the performance for YOLO detectors. Hereafter, the paper is organized as follows: Sect. 1 and Sect. 2 deal with an introduction and related work. Section 3 presents experimental setup, discussing the global architectures used for vehicle detection. Section 4 shows the results and discussion. Conclusions and other experimental targets are drawn in Sect. 5.

2 Related Work

Considerable work on imaging processing has been carried out in the field of machine learning for object detection. Avidan et al. [7] proposed seam caving which performs retargeting by inserting and removing streams of pixels, which are called seam. It passes through less important features. Pritch et al. [8], presented movement of maps for the rearrangement of pixels. It is composed of graph labelling problems for editing different images in various applications. Lou et al. [9] introduced a method to predict a specific feature. The receptive field focuses on various image regions such as low level, middle level and high level without information left during the classification. Oquab et al. [10] stated that training large neural network architectures will require more time and resources. These models will perform computations for processing the images. Hashemi et al. [11] proposed zero-padding for resizing the images to the same size and compares them to the convolutional methods of scaling up the images by using interpolation. Connor et al. [12] explored that data augmentation can enhance the performance of neural network models and expand limited datasets to take the benefits of utilizing big data. various approaches of regularization the generalization to eliminate overfitting. Mingxing et al. [13] proposed a technique for neural network design architecture for object detection and suggested several ways of optimization to enhance efficiency of neural network models.

Tsung et al. [14] proposed a method which is called pyramidal hierarchy of deep learning to establish feature pyramid with marginal additional cost. Krahenbuhl et al. [15] proposed a method which is called energy map. It consists of many automatic constrains and determined constrains on the key frames. This is to compute a non-uniform pixel warping on video frames. This research focused-on preservation of the local structure and optimizing a warping from the original size to the desired size according to its important regions and permitted regions.

3 Experimental Setup

3.1 Deep Neural Network Models

Object detection is a computer vision technique, which involves the prediction the presence of one or more objects in an image, along with their scores and bounding boxes. YOLO is the state of art of object detection which is targeted for real-time application. There are five versions of the deep learning architectures. The first three yolo versions (YOLO [16], YOLOv2 [17], and YOLOv3 [4]) were released in 2016, 2017, and 2018 respectively by Joseph Redmon. However, in 2020, two major versions were released (YOLOv4 [18], and YOLOv5) which have considerable improvements to the previous versions. YOLO detectors use a single neural network takes an image as input, passes it through a neural network, which has similar convolutional neural network models (CNN), and generates a vector of bounding boxes with classes prediction in the output. The input image is divided into a grid of cell $S \times S$, *see* Fig. 1. Each grid cell is responsible to predict the object in the image. Each grid predicts bounding box B with class probability C . Bounding boxes have five components (x , y , w , h , and confidence). Coordinates x & y represent the center of the bounding box, while the w & h coordinates represent the height and width of the bounding box. These coordinates are normalized to fall between 0 and 1.



Fig. 1. Schematic diagram for YOLO architecture

In this research, we used different detectors of YOLO which involves YOLOv2, YOLOv3, YOLOv4, and YOLOv5 object detectors. We used pretrained models that include SqueezeNet with YOLOv2, and Resnet50 with YOLOv3. YOLOv4 has been utilized with deeper structure. We used backbone in this model which has been evolved from a network that has complete function of CSPDarknet53 structure, which enhances learning capabilities. This is to ensure that performance will not be degraded while the network is deepened. YOLOv5 was used in our experiments. This architecture consists of SPP-NET by modifying the SOTA approach. YOLOv5 is the fastest model which can achieve 140 *fps*. YOLOv5 is only 27 MB, which has the lowest size memory in comparison to the other YOLO detectors. YOLO detectors and pretrained models are implemented in the python library. Kares, sk-learn with TensorFlow. We examined these different detectors to have results and observations based on these models' complexities.

3.2 Datasets and Training

In this paper, we conducted several experiments with different sizes of vehicle images. We used large scale dataset of vehicles, which can provide many different vehicles classes fully annotated with various scenes that have been captured by surveillance camera from various locations and places. Three different images sizes have been used in this research which include 300×300 , 416×416 , and 640×480 . The original image size is 416×416 . We created two different datasets by upscale the original images into 640×480 and downscale them into 300×300 to have three different datasets Each dataset consists of 700 images [19]. The images have been randomly split into 70% for training, 20% for validation and the remaining 10% for testing for each dataset. YOLO detectors have been trained with three different datasets of vehicle images. YOLO detectors were trained with stochastic gradient descent (sdgm) to speed up the gradient's vectors. This will lead faster covering the archicturing during the training phase. The number of epochs has been set at 80 for all YOLO detectors. We set the momentum value at 0.9 for the hyper-parameters. The learning rate has been set at 10^{-3} to control the model change in response to the error.

3.3 Data Augmentation

Data augmentation is an approach which eliminates overfitting problems initiated by limited training data for deep learning models. It manipulates the input samples by utilizing different techniques such as noise disturbance, horizontal flipping, and random cropping in the image. Image acquisition is one of many possible observations, which can be visualized by different spial transformation. In the training stage, training samples in mini-batch set can be expanded in various data augmentation techniques at the time when they are applied to the convolutional neural network, see Eq. (1).

$$Dt = \{(x_i, y_i)\}_{i=1}^M \quad (1)$$

Where Dt is mini-batch samples at t -th training iteration, x_i is i -th sample, y_i is the label of i -th sample, and M is the batch size. Data augmentation is used in this research to enhance YOLO detector performance by randomly transforming the original data in the training process. We used an `augmentedTrainingData` function to add more variety in the training data, see Fig. 2. It generates batches of new images, after preprocessing original images by using operations rotation and flipping. We also used `transform` function to apply custom data augmentations to the training data. We also used `augmentedimageDatastore` and `Image Data augmenter` functions, which are integrated with training the neural network workflow.

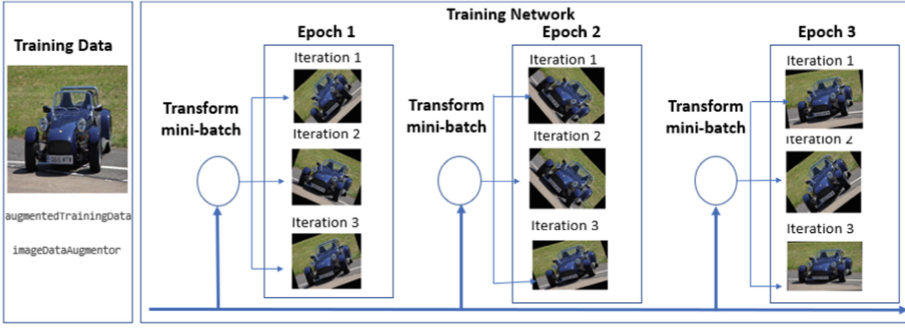


Fig. 2. Workflow for training a network using an augmented image datastore.

4 Results and Discussion

YOLO models were evaluated on the testing dataset for the three different resolutions of vehicle images. Only vehicles are focused as objects to be detected. As the result of this study, vehicle images for all resolution were detected by YOLO. The vehicle class and Intersection over Union (IoU) for each of the detected vehicles are displayed by YOLO detectors with bounding boxes. It is observed that for all YOLO detectors when the image size is varying, the training times are also changing, respectively. The training time decreases when the image resolution is downscaled, while the training time increases as the image resolution is upscaled. Table 1 shows the training computational time with respect to different image sizes for all YOLO detectors. YOLOv4 took a while to train the model due to the size of its model (CSPDarknet53), which is used for this architecture. However, YOLOv5 detector showed the fastest training time for the three different resolutions of vehicle images, which is intuitive to use verses all prior models of YOLO detectors. Further to our experiments, we evaluated all YOLO detectors on the three standard metrics performances: accuracy, precision, and recall, see Eq. (2).

$$Accuracy = \frac{TP + TN}{TP + FN + TN + FP}, Precision = \frac{TP}{TP + FP}, Recall = \frac{TP}{TP + FN} \quad (2)$$

Where TP stands for the number of true positive; TN stands for the number of true negative; FP stands for the number of false positive; FN stands for the number of false negative. The detailed results of each performance score for YOLO models with varying image resolutions are shown in Table 3. The performance metrics decreased when the images are resized to 300×300 and 640×480 pixels verses the performance scores for the original image size 416×416 in all YOLO detectors. This is to note that YOLOv4 showed the best results for performance in all image resolution sizes. This is due to the depth and strength of this architecture across different image sizes. The inference time has been measured for all YOLO detectors. YOLOv4 uses CSPDarknet53 which enabled the model to achieve more accurate detection ability among the other detectors. We examined the execution time of YOLO models for predicting the vehicle classes in the images, and as we can see in Table 2, YOLOv5 is the fastest detector among all different

Table 1. The training time with respect to different image sizes for all YOLO detectors.

Resolution	Model	Training time
300 × 300	YOLOv2	6 h
416 × 416	YOLOv2	7 h, 20 min
640 × 480	YOLOv2	11 h, 40 min
300 × 300	YOLOv3	2 h, 35 min
416 × 416	YOLOv3	3 h, 15 min
640 × 480	YOLOv3	4 h, 22 min
300 × 300	YOLOv4	7 h
416 × 416	YOLOv4	8 h, 30 min
640 × 480	YOLOv4	16 h
300 × 300	YOLOv5	2 h, 20 min
416 × 416	YOLOv5	2 h, 50 min
640 × 480	YOLOv5	4 h

resolutions of vehicle images in comparison to the other YOLO detectors. YOLOv5 is based on ultralytics PyTorch techniques, which is instinctive to and inferences the object detection very fast in the images.

Table 2. The Inference time for YOLO models with respect to different image sizes.

Resolution	Model	Inference time (ms)
300 × 300	YOLOv2	12
416 × 416	YOLOv2	16
640 × 480	YOLOv2	24
300 × 300	YOLOv3	22
416 × 416	YOLOv3	27
640 × 480	YOLOv3	48
300 × 300	YOLOv4	33
416 × 416	YOLOv4	36
640 × 480	YOLOv4	55
300 × 300	YOLOv5	9
416 × 416	YOLOv5	14
640 × 480	YOLOv5	23

Table 3. Impact of imaging resizing on YOLO models performance in terms of (accuracy, precision, and recall).

Model	Resolution	Recall	Precision	Accuracy
Yolov2	300 × 300	83%	86%	88%
	416 × 416	86%	88%	88%
	640 × 480	84%	84.6%	85%
Yolov3	300 × 300	85%	87%	88%
	416 × 416	92%	91.8%	94%
	640 × 480	87%	88.1%	91.2%
Yolov4	300 × 300	94%	93.8%	97%
	416 × 416	96%	94%	98%
	640 × 480	95.3%	95.1%	97.8%
Yolov5	300 × 300	88%	90.1%	90%
	416 × 416	92%	94%	95.1%
	640 × 480	90%	91.5%	92%

5 Conclusion and Other Experimental Target

In this paper, we presented a methodology, applied on the vehicle dataset, for studying the effects of resizing the images on YOLO family detectors. The experimental results showed that when the original image downscaled, it improved the training computation time, while the training time increased when these images upscaled. A drop in performance is observed when the images are resized on YOLO detectors in comparison to the original size of these images. Indeed, it showed that deeper network and YOLOv4 is one of these cases, it tends to learn efficiently in the three different size of image resolution. In the future, we will extend our exploration to experiment this study on multiple datasets, in particular those which contain small objects and complex image information to evaluate the computation time in respect to the deep learning models performance.

Acknowledgments. We thank the Islamic Development Bank for the support of the Ph.D. work of A. Elhanashi and the Crosslab MIUR project.

References

1. Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in Neural Information Processing Systems, pp. 1097–1105 (2012)
2. Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 4700–4708 (2017)

3. He, K., Zhang, X., Ren, S. and Sun, J.: Deep residual learning for image recognition. In: IEEE conference on computer vision and pattern recognition, pp. 770–778 (2016)
4. Redmon, J., Farhadi, A.: YOLOv3: an Incremental Improvement (2018)
5. Iandola, F.N., Han, S., Moskewicz, M.W., Ashraf, K., Dally, W.J., Keutzer, K.: SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size (2016)
6. Hashemi, M.: Web page classification: a survey of perspectives, gaps, and future directions. *Multimedia Tools Appl.* **79**(17–18), 11921–11945 (2020). <https://doi.org/10.1007/s11042-019-08373-8>
7. Avidan, S., Shamir, A.: Seam carving for content-aware image resizing. *ACM Trans. Graph.* **26**(3), 3–10 (2007)
8. Pritch, Y., Kav-Venaki, E., Peleg, S.: Shift-map image editing. In: 2009 IEEE 12th International Conference on Computer Vision, pp. 151–158, September 2009
9. Luo, W., Li, Y., Urtasun, R., Zemel, R.: Understanding the effective receptive field in deep convolutional neural networks. In: *Advances in Neural Information Processing Systems*, pp. 4898–4906 (2016)
10. Oquab, M., Bottou, L., Laptev, I., Sivic, J.: Learning and transferring mid-level image representations using convolutional neural networks. In: *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1717–1724 (2014)
11. Hashemi, M.: Enlarging smaller images before inputting into convolutional neural network: zero-padding vs. interpolation. *J. Big Data* **6**(1), 1–13 (2019). <https://doi.org/10.1186/s40537-019-0263-7>
12. Shorten, C., Khoshgoftaar, T.M.: A survey on image data augmentation for deep learning. *J. Big Data* **6**(1), 1–48 (2019). <https://doi.org/10.1186/s40537-019-0197-0>
13. Tan, M., Pang, R., Le, Q.V.: EfficientDet: Scalable and Efficient Object Detection (2019)
14. Lin, T.Y., Dollár, P., Girshick, R., He, K., Hariharan, B.: Feature Pyramid Networks for Object Detection (2016)
15. Krahenbuhl, P., Lang, M., Hornung, A., Gross, M.: A system for retargeting of streaming video. *ACM Trans. Graph.* **28**(5), 1–10 (2009)
16. Redmon, J.: You only look once: Unifed, real-time object detection. In: *IEEE CVPR*
17. Redmon, J., Farhadi, A.: YOLO9000: Better, Faster, Stronger (2016)
18. Bochkovskiy, A., Wang, C.Y., Liao, H.Y.M.: YOLOv4: optimal speed and accuracy of object detection <https://arxiv.org/abs/2004.10934> (2020)
19. Zenodo.2021. VehiclesDataset_416x416 <https://zenodo.org/record/4106511#.YcNqEGgzaUk>. Accessed 3 Sept 2021



Preliminary Design of a Three-Dimensional Anemometer for Sail Boats

Enrico Boni¹(✉), Luca Pugi², and Alessio Venturi²

¹ Department of Information Engineering, University of Florence, 50139 Florence, Italy
enrico.boni@unifi.it

² Department of Industrial Engineering, University of Florence, 50139 Florence, Italy
luca.pugi@unifi.it, alessio.venturil@stud.unifi.it

Abstract. Control of Autonomous Sail Vehicles involves a precise measurement of amplitude and direction of incoming wind. Ultrasound anemometers are widely used for this purpose. In previous research activities authors have successfully developed, calibrated, and integrated in existing autonomous sailboats a planar ultrasound anemometer. In this work authors propose a further development of this device in which a tridimensional array is proposed. Significant innovations, with respect to existing systems, have been introduced to increase its portability on autonomous mobile systems without losing desirable precision features that are currently granted by more encumbering solutions currently employed in static applications.

1 Introduction

Sail propulsion is encountering an increasing interest in literature both for marine [1] and ground applications [2] as stated by the increasing number of publications that are dedicated to this kind of applications.

Most of the applications proposed in literature are related to autonomous systems or at least to manned systems in which the action of the human driver is strongly assisted by automated systems.

A fundamental feedback needed to properly control a sail vehicle is the measurement of wind direction and intensity.

In previous research activities authors have developed a planar ultrasound anemometer which have been successfully designed [3] assembled and tested, first in a controlled environment represented by the wind tunnel of Florence University [4] and then on an autonomous Sail Vehicle [5] that has been really employed for research activities in a real marine environment.

Performed activities help authors to investigate and better understand some known limits of the proposed planar layout through the adoption of more complex three-dimensional layouts of ultrasonic sensors.

Looking at advanced ultrasound anemometers currently proposed on the market [6] it is possible to understand main differences between planar and three-dimensional layout visible in Fig. 1:

- Planar Anemometer: it is composed by four transducers aligned on two perpendicular planar directions in order to measure corresponding cartesian components of the wind that is traveling between transducers.
- Three-Dimensional Anemometer: transducers able to evaluate wind speed respect to three different independent directions are employed. In commonly adopted configurations sensors are installed as the direction of the three edge of two regular tetrahedrons

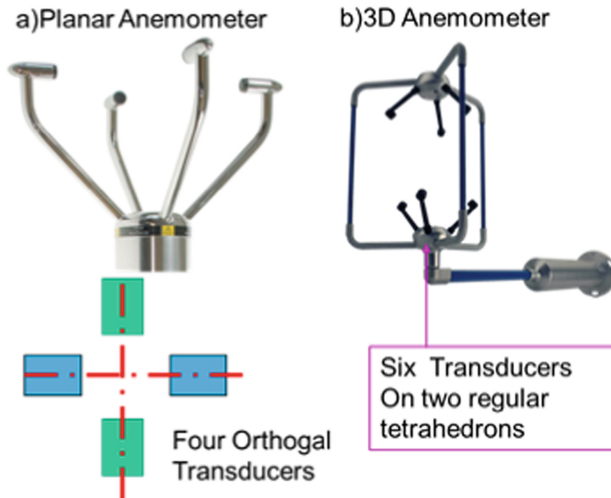


Fig. 1. Comparison between planar (a) and three-dimensional anemometers (b), images are elaborated from commercial examples [6]

Both solutions displayed in Fig. 1 have not been designed for vehicular applications, so weight and encumbrances are largely sub-optimal for this kind of applications.

In this work, authors focused their attention on the optimization of the frame of the three-dimensional solution to properly optimize encumbrances.

The proposed solution is represented in Fig. 2: the proposed sensor must have a diameter which is at least a half respect to a corresponding static solution and a lighter weight. These specifications are constrained by the typical condition of installation in a sail drone. The anemometer is installed at the top of the main mast in a position in which weight and encumbrances are quite critical for flexible behavior and static equilibrium of the system thus justifying the performed optimization.

Positioning on the drone visible in Fig. 2 also explain the reason of the work performed in this paper: anemometer is subjected to important three-dimensional motions since it is constrained to the boat mast. Therefore, a three-dimensional anemometer is a better solution to properly evaluate the real wind direction and intensity. Due to proposed compact structure, the distance between transducers and the aerodynamic interactions between sensor frame and incoming wind represents a critical aspect that must be carefully evaluated, even in much less demanding static applications where these kinds of

disturbances are the object of complex calibration procedures aiming to improve quality of performed measurement [7, 8]. Also, another aspect that is investigated is the comparison between expected behavior of proposed sensor with respect to the previous planar version, which is currently installed in the drone. The reason of this comparison is to understand how the proposed solution should represent an improvement, despite to the drawbacks arising from its very compact construction. This comparison described in the following section is performed using finite element models.

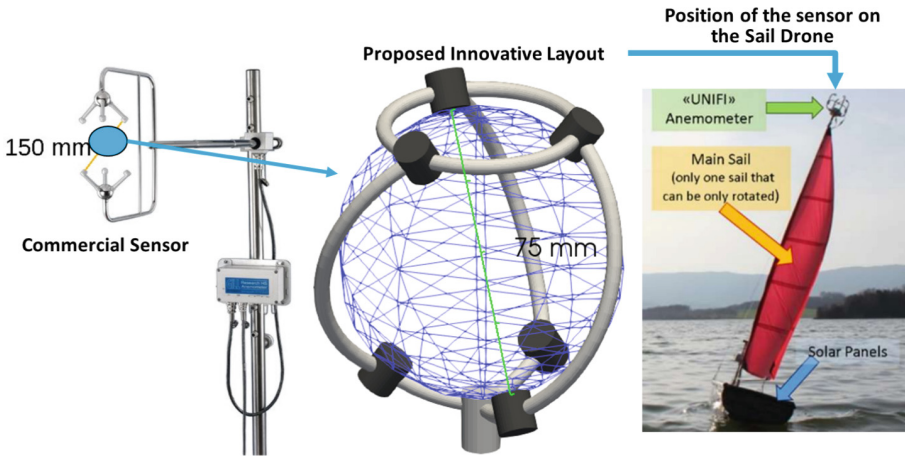


Fig. 2. Reduced encumbrances of proposed solution respect to a commercial one and typical positioning of an anemometer sensor on a sail drone (example is the UNIFI sail drone [5])

2 Finite Element Modelling and Obtained Results

Finite element Modelling is Performed using OPEN FOAM an open tool [9], widely diffused and adopted among various research communities.

For both geometries (planar and 3D anemometer) the speed velocity field is simulated around the sector, considering different speed directions and magnitudes of incoming wind.

Speed measurement performed by the anemometer is substantially related to the spatial average speed in the channel between each couple of transducers [1, 2]. Complex fluidodynamic interactions arising between sensor frame and incoming fluid produce a distortion of the flow field as visible in the examples of Figs. 3 and 4 that are respectively related to planar and 3D layout. This fluid-structure interaction negatively affects the quality of the measurement especially at very low speed where the shading produced by the frame in the fluid is typically more extended.

The aim of the performed analysis was to evaluate expected errors in term of estimated magnitude and direction of the wind, considering different amplitude and direction of the incoming wind.

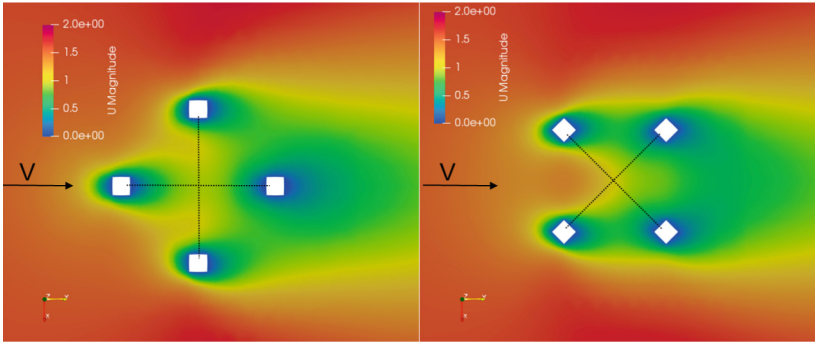


Fig. 3. Example of simulation results for the planar anemometer at low speed (1.6 m/s) considering different directions of incoming wind speed.

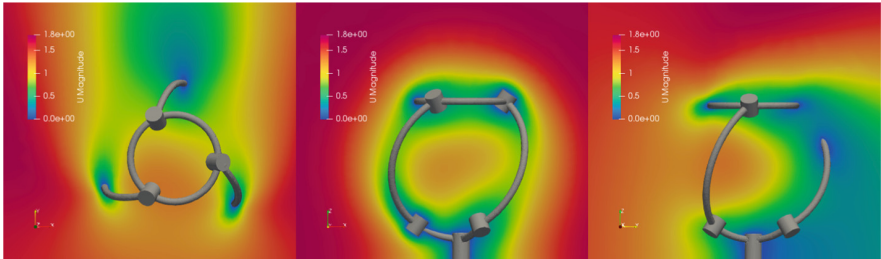


Fig. 4. Example of simulation results for the three-dimensional anemometer at low speed (1.6 m/s) considering different directions of incoming wind speed.

In Fig. 5 some results of performed simulations, considering direction and magnitude of an incoming planar wind, are shown: a planar flow condition is chosen since in this condition both planar and 3D anemometers should be substantially equivalent in terms of expected measurement performances.

Looking at results of Fig. 5 it is interesting to notice that both sensors exhibit a relatively big error in terms of estimated wind speed magnitude. This error is typically compensated using tabulated relations aiming to correct expected errors that are evaluated as functions of magnitude and orientation of wind speed. However, it is clearly noticeable that magnitude errors of the 3D anemometer are better than the ones of the planar one.

Most interesting results regard expected errors on estimated wind direction, where performances of the 3D anemometer are much better. This is a very interesting feature since for a correct guidance of a sailboat estimation of wind direction is a fundamental information for a proper alignment of the sail respect to desired maneuver.

In Fig. 6 some further results are shown: estimated wind speed components are compared with real ones: 3D anemometer can measure properly almost every component, even when the wind direction is not laying on a horizontal plane.

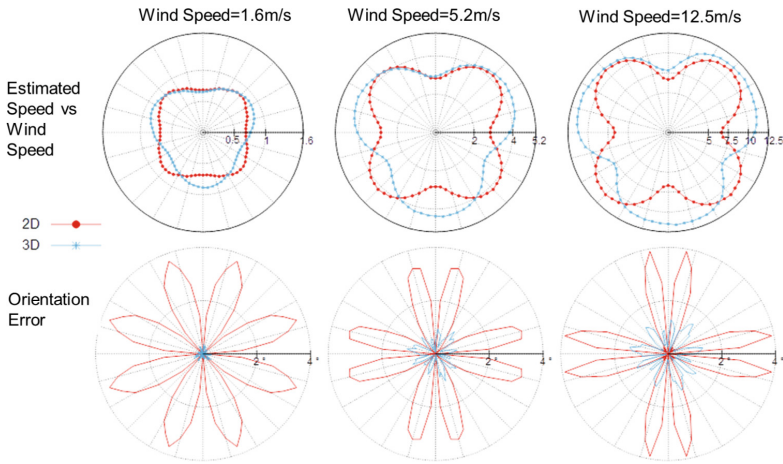


Fig. 5. Polar representation of magnitude and direction errors on performed measurements with respect to intensity and direction of incoming wind.

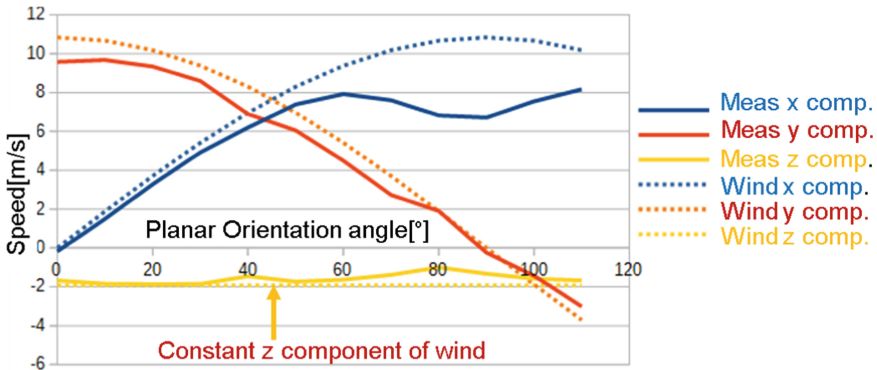


Fig. 6. Comparison wind component and estimated ones for a tridimensional flow.

3 Conclusions and Future Developments

Performed simulations clearly demonstrates, at least at a preliminary level, the advantages arising from the adoption of the proposed layout, justifying the development of the new sensor topology.

The simulation results show that without any compensation the angle estimation error is contained within 1° for the 3D structure, while the absolute speed error can be up to 50% of the full speed. Both measurements are mainly distorted by the systematic error induced by the structure, and thus can be easily compensated as in [3]. The proposed structure performance has comparable performance with existing solutions, like TriSonica-Mini [10], but the proposed spherical layout has the ability of sensing a wider vertical angle.

Currently authors are working to design, assembly and testing of a prototype of the proposed sensor. This activity will be the object of a more extended research publication.

References

1. Zhang, G., Li, J., Wei, Y., Zhang, W.: Event-triggered robust neural control for unmanned sail-assisted vehicles subject to actuator failures. *Ocean Eng.* **216**(15), 107754 (2020). <https://doi.org/10.1016/j.oceaneng.2020.107754>
2. Reina, G., Foglia, M.: Modelling and handling dynamics of a wind-driven vehicle. *Veh. Syst. Dyn.* **57**(5), 697–720 (2019). <https://doi.org/10.1080/00423114.2018.1479529>
3. Allotta, B., Pugi, L., Massai, T., Boni, E., Guidi, F., Montagni, M.: Design and calibration of an innovative ultrasonic, Arduino based anemometer. Conference Proceedings of the 2017 17th IEEE International Conference on Environment and Electrical Engineering and 2017 1st IEEE Industrial and Commercial Power Systems Europe, IEEEIC/I and CPS Europe 2017 (2017). <https://doi.org/10.1109/EEEIC.2017.7977450>
4. Pugi, L., Allotta, B., Boni, E., Guidi, F., Montagni, M., Massai, T.: Integrated design and testing of an anemometer for autonomous sail drones. *J. Dyn. Syst., Measur. Control, Trans. ASME* **140**(5), 055001 (2018). <https://doi.org/10.1115/1.4037840>
5. Boni, E., Montagni, M., Pugi, L.: Autonomous sail surface boats, design and testing results of the MOUNTAINS prototype. In: Saponara, S., De Gloria, A. (eds.) *ApplePies 2018*. LNEE, vol. 573, pp. 453–459. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-11973-7_54
6. Mc Gill Instruments. Technical documentation available at the official site <http://gillinstruments.com/products/anemometer/anemometeritaly.htm>
7. Nosov, V., Lukin, V., Nosov, E., Torgaev, A., Bogushevich, A.: Measurement of atmospheric turbulence characteristics by the ultrasonic anemometers and the calibration processes. *Atmosphere* **10**(8), 460 (2019). <https://doi.org/10.3390/atmos10080460>
8. Van der Molen, M.K., Gash, J.H.C., Elbers, J.A.: Sonic anemometer (co) sine response and flux measurement: II The effect of introducing an angle of attack dependent calibration. *Agri. Forest Meteorol.* **122**(1–2), 95–109 (2004)
9. Open Foam Open CFD Toolbox: tech. documentation available at the official site <https://www.openfoam.com/>
10. Anemoment Trisonica-Mini Documentation: <https://anemoment.com/features/#trisonica-mini>



Design and Preliminary Testing of an Electrified Directional Drilling Machine

Lorenzo Berzi¹, Francesco Grasso², Luca Pugi¹(✉), Enrico Boni², and Raffaele Savi³

¹ Department of Industrial Engineering, University of Florence, 50139 Florence, Italy
luca.pugi@unifi.it

² Department of Information Engineering, University of Florence, 50139 Florence, Italy

³ EGT SRL, Parma, Italy

Abstract. Trenchless technologies represent an innovative opportunity to improve quality, sustainability and overall costs of many construction works dealing with installation of various kind of services and utilities especially in crowded urban environments where interactions of conventional construction yards with surrounding community should produce an additional impact also in terms of costs. In this work it is proposed and investigated the electrification of this kind of machines to improve their sustainability efficiency and productivity. Main topics treated in the work are design and simulation procedures of a machine prototype that has been successfully tested proving to be an innovative market-ready solution.

1 Introduction

There is a wide literature concerning trenchless excavation and potential advantages that should arise from its application to construction and installation of various kind of underground infrastructures [1–4]. This is an important topic especially for further development and upgrade of infrastructure of crowded urban communities, where application of this kind of techniques contribute not only to reduce construction times and costs but also to respect the complicated pre-existing tissue of buildings, infrastructures and on-going activities that are typical of our communities. In this work it is investigated the possibility of a complete electrification of the machine to improve its sustainability and productivity, adding new potentially interesting features and preserving the natural robustness that a construction machine must guarantee to satisfy reliability in harsh environmental conditions.

2 System Specifications

Machine must implement the typical operations that are needed as example for trenchless installation of a pipe:

A small pilot bore is first drilled to produce a pattern that should guide following operations. Trajectory can be curved exploiting flexibility of drilling rods and particular tools in which it is possible to regulate/control a preferential direction of excavation.

Lubricant water is injected at high pressure (tens of bar) by a high-pressure pump, and it also contribute actively to the excavation process. Drilling bore are performed pushing the tool against the ground since no other possible way is practically possible. Drilling rod is composed by modular sections that are assembled as the tool advance in the ground.

After a pilot bore has been drilled, hole/tunnel section can be enlarged using additional tools such as reamers that are more commonly pulled to have a more stable behavior.

Performed excavation can be ensured and stabilized by installing a pipe that is usually pulled by the machine.

To perform these operations, a typical machine is composed by various actuators:

- Rotary: a hydraulic motor is used to provide the drilling torque which is necessary to excavate the ground.
- Mast Motors: push and pull forces exerted by the machine during overcited working phases.
- Pumping Unit: Lubricant pressurized water to assist drilling operation is produced by a dedicated pumping unit.
- Additional Electro-Hydraulic actuators are used to provide additional motion needed to manipulate, assemble, and disassemble rods during the drilling process.
- Traction System: machine is mounted on a tracked vehicle to assure the mobility needed for both transportation and installation of the machine.

With respect to this layout authors focused their attention on a full electrification of the machine aiming to almost substitute hydraulic actuators with electric ones.

Main specification of the machine for what concern rotary and mast actuators are described in Table 1, while some indications concerning traction system are listed in Table 2.

Table 1. Specifications for rotary and mast actuators

Parameter	Value
Approx. Torque load specifications (Rotary)	9000–1500 [Nm] at 30–190 [rpm]
Long. loads (Mast)	120000 [N] at 0.83 [m/s]
Approx. Weight of the machine	15–20 [tons]

Table 2. Specifications for traction system

Parameter	Value
Max speed	5 [kmh]
Max slope	30%
b (track width)	0,3 [m]
l (track length)	7.5 [tons]
rtrack (radius of the track)	0.2 [m]

3 Preliminary Models for Sizing of Actuation System

To properly size the actuation system authors developed a simplified model of the excavation process which is structured in the following way.

Specific cutting energy SE represents the amount of energy E that must be spent to excavate a unitary volume V of material 5. SE is a specific property of excavated material and can be evaluated from literature [6] or from specific experimental tests 7. Also, SE is approximately equal or at least proportional (1) to the compression resistance of the material σ_{compr}

$$SE = \frac{E}{V} \left[\frac{kWh}{m^3} \right] \propto \sigma_{compr} \left[\frac{N}{m^2} \right] \quad (1)$$

According to (2) Drilling Power W_{drill} is proportional to the volumetric flow of excavated material Q and SE energy with an additional factor η_{tool} to take count of the cutting efficiency of the tool respect to rotational speed and advance one; an additional friction coefficient f is introduced to model friction torques generated by rotation inside the bored material (f is proportional to drilled length). Drilling torque T_{drill} is then simply calculated from power by dividing it respect to angular speed ω .

$$W_{drill} = \frac{SEQ}{\eta_{tool}} + f\omega \Rightarrow T_{drill} = \frac{W_{drill}}{\omega}$$

$$\text{where : } \eta_{tool} = \frac{W_{ideal}}{W_{drill}} = \frac{SEAv}{W_{drill}} (\text{tool eff.}); \quad Q = Av (\text{vol. flow of excavated mat.}) \quad (2)$$

Once the pump volume is known pressure delivered by the pump ΔP is roughly proportional to squared value of Q_{pump} and to the length of the excavated bore. In this way, also the power needed for the coolant pump W_{pump} can be calculated (3) where efficiency of the pump η_{pump} is also considered.

$$W_{pump} = \frac{Q_{pump} \Delta P}{\eta_{pump}}; \text{ where } Q_{pump} = k_{pump} Q \quad (3)$$

For what concern axial effort need to cause the advance of the tools axial trust force F (4) is also proportional to specific energy through a specific coefficient α (v, ω) which also depends from both advance and rotation speed of the tools [6, 7]; additional advance resistances R are introduced as proportional to bore length and to additional losses that should be calculated with soft string theory [8, 9]. For simplicity calculation are performed considering straight holes.

$$F = SE\alpha(v, \omega) + R \quad (4)$$

4 Sizing of Rotary and Mast Actuator

Required torque and power profiles calculated by models described in Sect. 3 have been exploited for sizing and design of Rotary and Mast actuators.

For the design of both powertrains the following approach was chosen, mechanical transmissions adopted for the new machine are substantially alike to the ones of the conventional ones, electric motors are sized to work with relatively low speed under 3000–4000 rpm in order to maintain simple robust and reliable transmission systems that are needed by the proposed application. Reliability and robustness issues have also constrained the choice of electrical machines to low-speed induction motors whose torque-speed profiles and adopted reduction ratios are shown in Figs. 1 and 2. For the same reason motor of the rotary is liquid cooled.

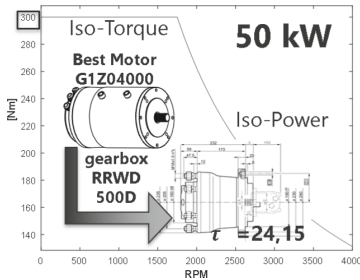


Fig. 1. Torque profile of the ind. motor chosen for the rotary unit (reducer with a reduction ratio of 24.15.)

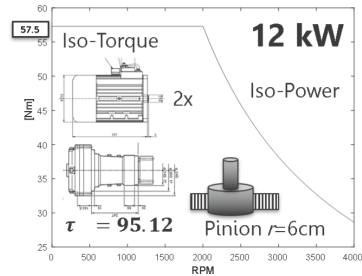


Fig. 2. Torque profiles of the ind. motors chosen to move the mast) speed reducer with reduction ratio of 95.12).

5 Simulation and Sizing of Batteries and Power Management System

A. Simulation of Complete Mission Profiles

To properly size on board accumulators and more generally the power management of the vehicle is performed a full simulation of complete working cycle of the machine. The reason of this choice is relatively simple considering different working phases of the machine in which different actuation and loads are activated. Repeating these simulations on a statistically significant population of different mission profiles authors were able to calculate some statistically reliable ranges of power consumptions respect to different working phase and operational scenarios visible in Table 4: in this way it was established a mean overall power consumption of the machine between 15 and 25 kW.

B. Sizing of Batteries and Energy Management System

For a complete turn of working of eight hours a total capacity of about 200 kWh should be expected. Since working activities should implies the usage of the machine in the same yard for several days, the most practical solution is to integrate within the machine also battery AC charger, also maintaining a direct access to battery trough a DC connection to eventually exploit also external power sources for a fast recharge. This approach is followed to maintain a certain flexibility of use respect to variability

of electrical power sources in the construction yard. It is considered a quite common scenario in which a relatively small AC power is available during the day since yard infrastructure is modest and available power must be shared also with other machines or services. In this case the machine can be connected to the yard grid implementing a “biberonage” strategy in which the machine exploits available power of yard grid even if largely insufficient respect its mean consumptions to reduce load on batteries increasing their autonomy and life. In this way also smaller, cheaper, and lighter battery pack can be employed. For this reason, battery pack is designed to be modular allowing four incremental sizes (50,100,150 and 200 kWh) that should meet different operational scenarios, that are shown in Tables 5: it is calculated machine autonomy considering minimum and maximum mean load and the availability of external power source which appreciable (10 kW) or minimal (3.5 kW). Estimated Machine Autonomy According Battery Size, Available External Power, Intensity of Machine Working Cycle. For batteries it was chosen a LiFePO₄ technology since this technology despite its modest performances in terms of energy density (around 100 Wh/kg) assure a great thermal stability and safety at an affordable cost (Table 3).

Table 3. Autonomy with different levels of floating charge

Battery size	Machine autonomy (hours) Mean working load of 25 kW			Mean working load of 15 kW		
	No floating charge	3.5 kW of floating charge	10 kW of floating charge	No floating charge	3.5 kW of floating charge	10 kW of floating charge
50 [kWh]	2 h	about 2 h 15 ^l	3–4 h	3 h	4 h	10 h
100 [kWh]	4 h	about 4 h 30 ^l	6–7 h	6 h30 ^l	8 h	20 h
150 [kWh]	6 h	7 h	10 h	10 h	12 h	30 h
200 [kWh]	8 h	9 h	14 h	14 h	17 h	40 h

6 Preliminary Testing Activities

Once assembled, machine has been preliminary tested on a site near Massa-Carrara (Italy) on a sandy ground in winter 2020. Drilling tests have been repeated varying some typical process parameters such as rotational and advance speed as visible in Figs. 3 and 4 showing how performed measured on the machine can be very useful even for a simple optimization of the drilling process.

7 Conclusions and Future Developments

The complete electrification of a directional drilling machine has been completed. A working prototype has been successfully assembled and preliminary tested validating

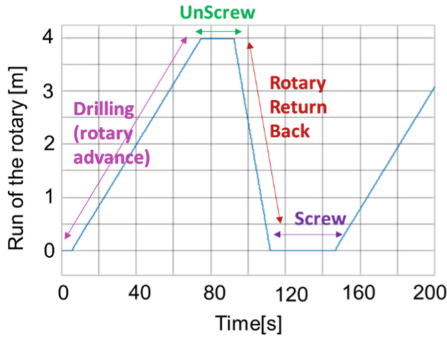


Fig. 3. Behaviour of consumed power when maximum drilling performances are exploited

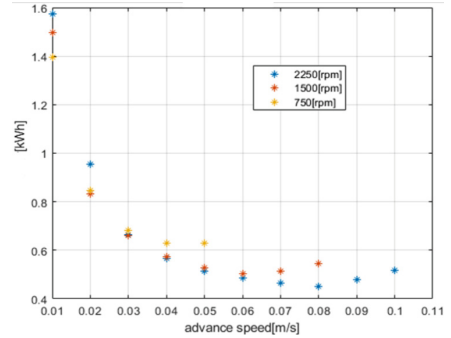


Fig. 4. Energy needed to drill a single rod length as a function of tool advance and rotational speed

proposed design procedure. Future development of the activity should also involve online implementation of model-based identification of ground features such as Specific Energy. Also, a more accurate model of battery behavior based on recent works of Locorotondo [10, 11].

Acknowledgement. Authors wish to thank all the partners of the STIGE (Sviluppo di Trivelle Innovative per attività Geologiche con alimentazione Elettrica) project. Project STIGE was financed by Regione Toscana (Italy) as a part of POR FESR 2014–2020 program.

References

1. Gerasimova, V.: Underground engineering and trenchless technologies at the defense of environment. *Proc. Eng.* **165**, 1395–1401 (2016). <https://doi.org/10.1016/j.triboint.2018.06.014>
2. Marchant, P.: Repair or Replace: Technologies Available for Trenchless Remediation. In: *Water Scarcity and Ways to Reduce the Impact: Management Strategies and Technologies for Zero Liquid Discharge and Future Smart Cities*, p. 197 (2018)
3. Chothe, O.K., Kadam, V.S.: Comparative study of traditional method and innovative method for trenchless technology: a review. *Int. Res. J. Eng. Technol.* **3**(05) (2016)
4. Burden, L.I.: *Synthesis of Trenchless Technologies*, virginia center for transportation innovation and research p.1–25 (2015)
5. Teale, R.: The concept of specific energy in rock drilling. In: *International Journal of Rock Mechanics and Mining Sciences and Geomechanics Abstracts*, vol. 2, no. 1, pp. 57–73. Pergamon March, 1965
6. Pessier, R. C., Fear, M. J.: Quantifying common drilling problems with mechanical specific energy and a bit-specific coefficient of sliding friction. In: *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers January 1992
7. Chen, X., Gao, D., Guo, B., Feng, Y.: Real-time optimization of drilling parameters based on mechanical specific energy for rotating drilling with positive displacement motor in the hard formation. *J. Nat. Gas Sci. Eng.* **35**, 686–694 (2016)
8. Miska, S.Z., et al.: Dynamic soft string model and its practical application. In: *SPE/IADC drilling conference and exhibition*, Society of Petroleum Engineers (2015)

9. Mirhaj, S.A., et al.: Torque and drag modeling; soft-string versus stiff-string models. In: SPE/IADC Middle East Drilling Technology Conference and Exhibition. Society of Petroleum Engineers (2016)
10. Locorotondo, E., Cultrera, V., Pugi, L., Berzi, L., Pierini, M., Lutzemberger, G.: Development of a battery real-time state of health diagnosis based on fast impedance measurements. *J. Energy Storage* **38**, 102566 (2021). <https://doi.org/10.1016/j.est.2021.102566>
11. Locorotondo, E., Pugi, L., Berzi, L., Pierini, M., Pretto, A.: Online state of health estimation of lithium-ion batteries based on improved ampere-count method. In: Proceedings of the 2018 IEEE International Conference on Environment and Electrical Engineering and 2018 IEEE Industrial and Commercial Power Systems Europe, IEEEIC/I and CPS Europe 2018, p.8493825, (2018). <https://doi.org/10.1109/EEEIC.2018.8493825>



CRFlex: A Flexible and Configurable Cryptographic Hardware Accelerator for AES Block Cipher Modes

Pietro Nannipieri^(✉), Luca Baldanzi, Luca Crocetti, Stefano Di Matteo, Francesco Falaschi, Luca Fanucci, and Sergio Saponara

Department of Information Engineering, University of Pisa,
Via G.Caruso, 16, Pisa, Italy
pietro.nannipieri@ing.unipi.it

Abstract. This paper presents a System-on-Chip (SoC) implementation of a cryptographic hardware accelerator supporting multiple AES based block cypher modes, including the more advanced CMAC, CCM, GCM and XTS modes. Furthermore, the proposed design implements in hardware advanced features for AES key secure storage. A flexible interface allows the communication between the hardware accelerator and the chosen processor and makes this implementation suitable to be easily integrated into a generic embedded system. The system has been prototyped and characterized on a Xilinx Zynq 7000 platform. Synthesis results on a 7 nm CMOS Standard-Cell library are proposed too, showing competitive performances and resource usage respect to the State of Art and assessing the portability in different technology libraries of the proposed design. Furthermore, power consumption data are extracted to prove the suitability of the hardware acceleration also in the case of power-constrained devices.

Keywords: Security · Advanced encryption standard (AES) · SoC · Real-time · Hardware accelerator

1 Introduction

Generally, an embedded system is equipped with a microprocessor and a pure software solution can be a valid approach to perform security algorithms. Nevertheless, in specific application cases the performance of the embedded microprocessor could be insufficient to respect the latency and/or throughput requirements, for example in real-time IoT applications [1]. In these cases, a valid solution could be to privilege hardware (HW) domain solutions rather than software (SW) domain solutions, for security algorithms. Furthermore, a HW-based acceleration of cybersecurity functionalities may be needed to minimize power consumption [2–4]. Symmetric key algorithms are preferred over asymmetric ones due to their performance advantage, and the Advanced Encryption Standard

(AES) [5] is the most popular of the symmetric key algorithms. HW implementations of the AES have demonstrated to have better performance than SW ones [6], to the point that pure SW implementations have become uncommon in performance and power critical environments. This work presents an implementation of a configurable AES-based hardware accelerator whose differentiating aspects are reported herein. Multiple block cypher modes have been implemented, ensuring not only confidentiality but also integrity and Authenticated Encryption with Associated Data (AEAD). The proposed Intellectual Property-core (IP) has been named CRFlex core, which stands for Crypto-Flexible core. The flexibility refers to the possibility to select only part of the accelerator architecture at synthesis time, including only the modes that are strictly required for the application, hence minimizing the logic resources. The internal interconnection and the interface to interact with the SW layer are flexible as well and automatically generated depending on the selected features. The solution aims to be complete and easy to integrate in a SoC, where fast and context-aware security approaches are required. Moreover, also advanced policies for secure keys storage are implemented in hardware. The accelerator has been synthesized on 28 nm FPGA device and on a 7 nm Standard-Cell technology. To the best of the authors knowledge, this work presents the first public data of AES implementation on extremely scaled technology (i.e. 7 nm).

2 AES Acceleration: Hardware Architecture

In literature, multiple approaches to achieve an acceleration of cryptographic algorithms can be found, ranging from dedicated ASIC solutions to customized central processing units, passing through a myriad of possible HW/SW hybrid implementations. In [7] a brief comparative analysis between different approaches and algorithms has been presented. Some notable HW/SW acceleration methods are Instruction Set Extension (ISE) [8], flexible dedicated crypto processors [9] and hardware accelerators. Among these methods, the most suitable to be generally employed is the hardware accelerator solution. In [10] the authors propose an AES IP core based on Altera Avalon bus interfacing with a NIOS II softcore, able to support ECB, CBC, OFB, CFB and CTR modes. A similar approach is followed in [11] where the authors have achieved a lower gate count at the cost of a lower speed by using a Xilinx Microblaze processor coupled with a custom accelerator using the Processor Local Bus (PLB). In the mentioned cases, the stand-alone AES modules are capable to maintain very high throughputs but suffer a considerable decrease in performance when connected to the processor, typically of some orders of magnitude. This is due to the bus interconnect that represents the system bottleneck and limits the communication speed between CPU and hardware accelerator. Our proposed design is a SoC based hardware accelerator that performs the basic block cypher modes such as ECB, CBC, OFB, CFB and CTR. Then, it offers advanced block cypher modes to guarantee integrity and authenticity along with confidentiality, such as CMAC, GCM and CCM modes. Furthermore, CRFlex performs also the XTS block cypher

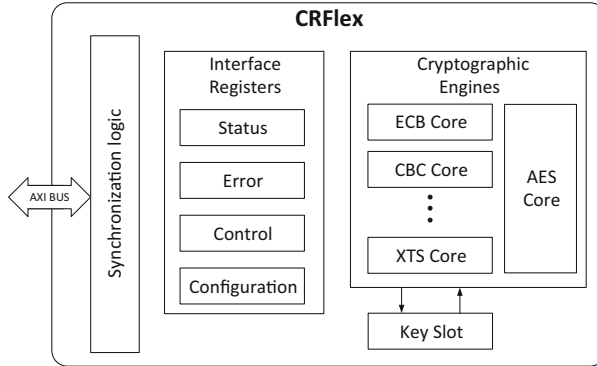


Fig. 1. CRFlex hardware accelerator architecture.

mode for disk encryption support. Our solution has been conceived to give the user the maximum flexibility in the selection of the minimum subset of AES-based cypher block modes to be included in the finally implemented hardware accelerator, making it suitable for different applications while minimizing the area overhead. The architecture of the CRFlex accelerator is shown in Fig. 1. The main system components and features include: a **Memory mapped interface** to connect the hardware module with the CPU via AXI bus, including a dedicate synchronization logic to manage clock domain crossing; a **Interface registers** to handle the interpretation of the input commands (OPcodes); an **Internal key slot register** to perform a secure key storage and reduce system latency in case of multiple AES operations with the same cipher key; a **Multiple AES cipher block** modules, conditionally instantiated and interconnected; a **Globally shared AES Core**. The operations are controlled by the set of interface registers, that receives as input the 32-bit control register and generates the control signals that drive the underlying sub-modules. The micro-controller shall read each output before providing to it a new input.

AES Core

Our proposed architecture instantiates a single AES core, to be shared by all the others implemented block cipher modes accelerator. The drawback is that a single mode at a time can be driven from the processor; the advantage is that the resource utilization is optimized. During the design phase, we carried out a trade-off between performance and complexity to properly implement the AES module. We implemented hardware functions belonging to one round, which are then used iteratively for several rounds (10 to 14, depending on the key size). To further save logic resources, we decided to support only 128 and 256-bit key lengths. It is important to observe that the AES-256 algorithm is considered secure against post-quantum attacks [12], which makes this solution also applicable to future designs. The AES core is based on the architecture presented in [13].

ECB, CBC, OFB, CFB, CTR Cores

All the basic AES modes of operations described in the National Institute of Standards and Technology NIST special publication 800–38A [14] have been implemented. These modes are similar to each other and the convenience to use one instead of the other depends on the application. Among CBC, CFB, OFB and CTR, only the CTR mode does not present any dependencies among encryption/decryption results of current and previous data blocks. These block cypher modes (ECB, CBC, OFB, CFB and CTR) share a single AES core.

CMAC Core

The Cipher-based MAC (CMAC) mode from NIST special publication 800–38B [15] is included in the CRFlex to provide a mode capable to guarantee the integrity of information. It performs customized processing of the input message using a cypher block chaining technique and returns a bit string called Message Authentication Code (MAC), also known as Tag.

CCM and GCM Cores

The NIST special publications 800–38C and 800–38D, [16] and [17], describe the CCM mode and the GCM mode respectively, two advanced cryptosystems to achieve simultaneously confidentiality, integrity and authentication over sensible data, or rather the Authenticated Encryption (AE).

XTS Core

The XTS mode, described in NIST special publication 800–38E [18], is included inside the CRFlex design. With the support of XTS mode, the CRFlex is a suitable solution also in the case of applications connected with external storage devices (e.g. hard disk). XTS-AES provides confidentiality of data, with the security strength of the AES algorithm, for block-oriented storage devices.

3 Implementation on FPGA and Standard-Cell Technology

All data presented in this section refer to the folded architecture of AES core inside CRFlex design, i.e. instancing only one stage of AES core. Our design can instantiate an unfolded architecture with multiple stages of the internal AES (from 2 up to 14 in the case of AES-256), using synthesis parameters. In this work, we report only the single-stage case because our focus is a compact and low power design. The CRFlex module was synthesized on the Xilinx Zynq-7000 xc7z045ffg900-2 FPGA using Xilinx Vivado. All data herein reported referring to post-implementation (i.e. place and route) results. The HW frequency chosen for the synthesis is 166.66 MHz. In order to assess the portability of the proposed system, the CRFlex module was synthesized also on two different standard-cell

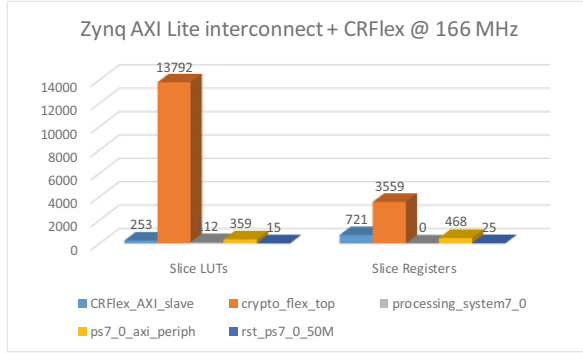


Fig. 2. System complexity: Slice LUTs and Slice Registers for 1 signal AES core with all modes of operation. (Color figure online)

technologies (i.e. 45 nm and 7 nm). On 45 nm, the CRFlex core can reach 620 MHz of clock frequency and a maximum throughput of 7.93 Gbps. Table 1 shows the results in terms of logic resources usage, expressed in kilo Gate Equivalent (kGE), and power consumption for two different configurations of the CRFlex, taken as representative examples of the multiple configuration of supported by our IP core: the AES-ECB mode and the AES-GCM mode. The synthesis was performed under *typical* conditions: 1.1 V of supply voltage and 25 °C for environmental temperature. The frequency chosen for the power characterization on standard-cell is 100 MHz, because this frequency is a good trade-off between throughput and power consumption. The power analysis was performed also considering the switched activity of the hardware modules, extracted by means of dedicated testbenches. The Electronic Design Automation (EDA) tools used for this characterization are: *Design Compiler* by Synopsys for netlist extraction, *Questa* by Mentor Graphics for post synthesis simulation and *Prime Time* by Synopsys for power extraction.

Table 1. Synthesis on 45 nm Standard-Cell technology: cases of CRFlex only instancing AES-ECB and AES-GCM respectively (AES single stage architecture).

Cipher mode	Logic usage	Leakage power	Dynamic power @ 100 MHz
AES-ECB	14.02 kGE	0.27 mW	7.33 mW
AES-GCM	61.01 kGE	0.96 mW	25.1 mW

The power reports show that the CRFlex hardware accelerator is a valid and suitable solution also for power-constrained applications (e.g. mobile, IoT, etc.).

Synthesis on 7 nm Technology On 7 nm technology the CRFlex core can reach 3.0 GHz of clock frequency and a maximum throughput of 27.4 Gbps

Table 2. Synthesis on 7 nm standard-cell technology: case of CRFlex only instancing AES-ECB (AES single stage architecture).

AES-ECB-256		
# Stage(s)	Logic usage	Throughput
1 stage	28 kGE	27.4 Gbps
2 stages	55.7 kGE	55 Gbps
7 stages	195 kGE	192 Gbps
14 stages	370 kGE	384 Gbps

with a single AES stage instantiated. Table 2 shows the results in terms of logic resources usage and throughput for all the possible number of cascaded stages for the AES-256 algorithm. The synthesis was performed with the maximum frequency reachable by the CRFlex. The modifications to increase the throughput would be suitable for contexts where system speed is more important than flexibility and chip size. On 7 nm technology, the CRFlex is a valid solution also for high speed applications, such as hardware acceleration in the new generation of High Performance Computing (HPC) systems.

4 Comparison to the State of the Art

In this section we compare our CRFlex against other AES hardware accelerators implemented on 45 nm Standard-Cell technology. To our knowledge, no results of AES hardware implementations on 7 nm technology can be found in other academic and commercial works. The comparison with other architectures is very complex due to the differences in supported functionalities among the different implementations and the lack of sufficient performance data in the literature. We compared the performance respect only to the CBC mode, instantiating only one AES core inside the CRFlex.

Table 3 shows the comparison among different works in terms of supported modes of operation, area and throughput. Area is measured in Gate Equivalent (GE) and throughput refers only to CBC mode. The work in [19] is a commercial AES hardware IP that supports different modes of operation. The circuit area is around 49–53 kGE and the information about throughput are not published. Work in [20] is an AES hardware accelerator that support only ECB and CBC modes. It is aimed to be used for content protection for high-performance applications. Work in [9] is a flexible crypto-processor that supports different symmetric key algorithms. Respect to the other works, our CRFlex is the more complete in terms of supported modes of operation, with performance in line with the State of the Art.

Table 3. Comparison among different implementations of AES

	Our work	[19]	[20]	[9]
Key lenght	128–256	128–256	128–256	128
Supported modes of operation	ECB,CBC,OFB CFB,CTR,CCM GCM,CMAC,XTS	ECB,CBC CTR,GCM	ECB,CBC	ECB,CBC OFB,CFB
Technology	7 nm/45 nm	45 nm	45 nm	45 nm
Area [kGE]	75.41/73.26	49–53	187.67	7925.31
Throughput CBC [Gbps]	29.45/7.93	–	16.74	6.40

5 Conclusions

This work presented a SoC implementation of a co-processor for AES-based block cypher modes, suitable to be employed in embedded applications. The accelerator was implemented in a 28 nm FPGA device and 7 nm standard-cell technologies. A complete system with an ARM[®] Cortex A9 and the proposed accelerator has been prototyped on a Zynq 7000 SoC. Complexity and performance have been thoroughly characterized for all the AES modes of operations in all the above mentioned technologies. To the best of the author’s knowledge, this work is the first presenting in public literature AES implementation results on 7nm technology. In addition, its flexibility allows for a selection of the minimum resources needed for the chosen application, thus minimizing the cost of the resources, thanks to the shared AES engine. Even in the case in which only a single block cypher mode is selected the overhead of the flexibility is negligible since the crypto-arbiter module slice LUT occupation is 0.03% of the AES one (the registers composing the memory map interface is just 30 32-bits registers, thus again negligible); moreover, the hardware accelerator interface can be shared with other accelerators on the bus. The system interface allows easy integration in embedded systems that require high-performance cryptographic acceleration. The CRFlex interface can be easily modified to match the specific bus used. The module can be accessed either as a common memory-mapped device or as using the DMA engine, depending on the required throughput.

References

1. Rahman, F., Farmani, M., Tehranipoor, M., Jin, Y.: Hardware-assisted cybersecurity for IoT devices. In: IEEE 18th International Workshop on Microprocessor and SOC Test and Verification (2017)
2. Nannipieri, P., et al.: Sha2 and sha-3 accelerator design in a 7 nm technology within the european processor initiative. *Microprocessors and Microsystems* (2020)
3. Nannipieri, P., et al.: True random number generator based on fibonacci-galois ring oscillators for fpga. *Appl. Sci. (Switzerland)*, **11**(8) (2021)
4. Stefano, D., et al.: Secure elliptic curve crypto-processor for real-time iot applications. *Energies*, **14**(15) (2021)

5. NIST. FIPS 197: Advanced Encryption Standard (AES). Federal Information Processing Standards Publication, **197**(441), 03110 (2001)
6. Baldanzi, L., Crocetti, L., Di Matteo, S., Fanucci, L., Saponara, S., Hameau, P.: Crypto accelerators for power-efficient and real-time on-chip implementation of secure algorithms. In: 2019 26th IEEE International Conference on Electronics, Circuits and Systems (ICECS) (2019)
7. Rashid, M., Imran, M., Jafri, A.R., Al-Somani, T.F.: Comparative analysis of flexible cryptographic implementations. In: Reconfigurable Communication-centric Systems-on-Chip (ReCoSoC), 2016 11th International Symposium, pp. 1–6. IEEE (2016)
8. Ben Hadjy Youssef, N., El Hadj Youssef, W., Machhout, M., Tourki, R., Toriki, K.: Instruction set extensions of AES algorithms for 32-bit processors. In: 2014 International Carnahan Conference on Security Technology (ICCST) (2014)
9. Sayilar, G., Chiou, D.: Cryptoraptor: high throughput reconfigurable cryptographic processor. In: Proceedings of the 2014 IEEE/ACM International Conference on Computer-Aided Design, pp. 154–161. IEEE Press (2014)
10. Tao, X.C., Zhang, D.L., Song, Y.K.: An implementation of configurable and Small-Area AES IP Core oriented Avalon Bus (2015)
11. Chang, K., Chen, Y., Hsieh, C., Huang, C., Chang, C.: Embedded a low area 32-bit AES for image encryption/decryption application. In: Circuits and Systems, 2009. ISCAS 2009. IEEE International Symposium on, pp. 1922–1925. IEEE (2009)
12. Mavroeidis, V., Vishi, K., Zych, M. D., Josang, A.: The impact of quantum computing on present cryptography. (IJACSA) Int. J. Adv.Comput. Sci. Appl. **9**(3) (2018)
13. Ueno, R., Morioka, S., Homma, N., Aoki, T.: A high throughput/gate aes hardware architecture by compressing encryption and decryption datapaths - toward efficient cbc-mode implementa. Cryptol. ePrint Archive, Report 2016/595 (2016)
14. Dworkin, M.: NIST Special Publication 800-38A. Technical report (2001)
15. Dworkin, M.: NIST Special Publication 800-38B. US Department of Commerce, Technology Administration, National Institute of Standards and Technology (2005)
16. Dworkin, M.: NIST Special Publication 800-38C. US Department of Commerce, Technology Administration, National Institute of Standards and Technology (2004)
17. Dworkin, M.: NIST Special Publication 800-38D. US Department of Commerce, Technology Administration, National Institute of Standards and Technology (2007)
18. Dworkin, M.: NIST Special Publication 800-38E. US Department of Commerce, Technology Administration, National Institute of Standards and Technology (2008)
19. Crypt-ip-120 aes crypto, rambus. www.rambus.com/security/crypto-accelerator-hardware-cores/basic-crypto-blocks/crypt-ip-120/. Accessed 6 Apr (2021)
20. Mathew, S.K., et al.: 53 gbps native $gf(2^4)^2$ composite-field aes-encrypt/decrypt accelerator for content-protection in 45 nm high-performance microprocessors. IEEE J. Solid-State Circuits, **46**(4), 767–776 (2011)



A M-PSK Timing Recovery Loop Based on Q-Learning

Gian Carlo Cardarilli, Luca Di Nunzio, Rocco Fazzolari, Daniele Giardino, Matteo Guadagno, Marco Re, and Sergio Spanò^(✉)

Department of Electronic Engineering, University of Rome “Tor Vergata”, Via del Politecnico 1, 00133 Rome, Italy

{cardarilli, di.nunzio, fazzolari, giardino, re, spano}@ing.uniroma2.it, matteo.guadagno@students.uniroma2.eu

Abstract. In this work we propose a digital symbol synchronizer for M-PSK modulations based on the Q-Learning algorithm. Through Reinforcement Learning, the system is able to autonomously adapt to environment changes, learning the correct Timing Recovery Loop behavior. The proposed synchronizer has been tested considering a white gaussian noisy channel. We analyzed the modulation error rate and the signal to noise ratio. The obtained results show improved timing recovery capabilities exhibiting a lower locking time.

1 Introduction

In a digital telecommunication system, regardless of the used modulation, it is mandatory the synchronization between receiver and transmitter [1]. A typical synchronizer must perform two processes: the carrier recovery and the symbol synchronization. The latter is achieved by Timing Recovery Loops (TRL) [2, 3].

Among the most used TRLs, there are the Early-Late loop [4], the Muller & Müller loop [5], and the Gardner loop [6]. Recently, novel TRLs have been proposed [7], which includes also the carrier recovery, and loops based on Machine Learning (ML) techniques [8, 9].

These last systems are based on Reinforcement Learning (RL), a subset of ML used to train an agent to achieve a certain task [10]. In RL, the agent performs some actions in its environment and, via a trial&error approach, it would learn an optimal action-selection policy [11]. An interpreter is in charge to evaluate the state of the environment and to give the agent a positive or negative reward (reinforcement) for its past action. These features make RL particularly suitable for time-variant optimization problems.

Q-Learning [12] is one of the most used RL algorithms. It is based on a quality matrix (Q-Matrix) which size is $N \times Z$, where N is the possible number of states for the environment and Z is the possible number of actions for the agent. The learning process consists in the update of the Q-Matrix according to the following formula:

$$Q_{new}(s_t, a_t) = Q(s_t, a_t) + \alpha \left(r_t + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t) \right) \quad (1)$$

The variables in Eq. (1) are defined below.

- s_t and s_{t+1} : present and future state of the environment, respectively.
- a_t and a_{t+1} : present and future action performed by the agent, respectively.
- γ discount factor, $\gamma \in [0, 1]$: it defines how much future rewards must be taken into account instead of immediate ones.
- α learning rate, $\alpha \in [0, 1]$: it defines the pace of the whole training process (like other ML algorithms).
- r_t : present value of the reward.

In the last years, a hardware implementation for the Q-Learning algorithm has been proposed [13]. This architecture will be also used to design the TRL in this work.

2 Proposed Architecture

In this work we propose a symbol synchronizer based on the Q-Learning algorithm. The system is based on the *state evaluator* described in [9], and this design has been conveniently modified to work just with M-PSK modulations. Moreover, the *reward evaluator* is based on the one from [7].

The proposed system is able to retrieve the correct sampling time with raised-cosine shaped signals. The input $x[n]$ is obtained via down-sampling the modulated signal of a factor L . This is the same approach of [5, 7, 9]. In our design, L is equal to the number of samples M that represents the symbol of the input signal.

The size of the Q-Matrix in the proposed architecture is $M \times 3$, so it has 3 actions and a number of states equal to the number of the samples of the symbols. The actions are:

- $a_0 = 0$: no change in the sampling, so the synchronization is correct;
- $a_1 = +1$: the sampling time is moved earlier by one sample;
- $a_2 = -1$: the sampling time is moved later by one sample.

The proposed architecture is shown in Fig. 1.

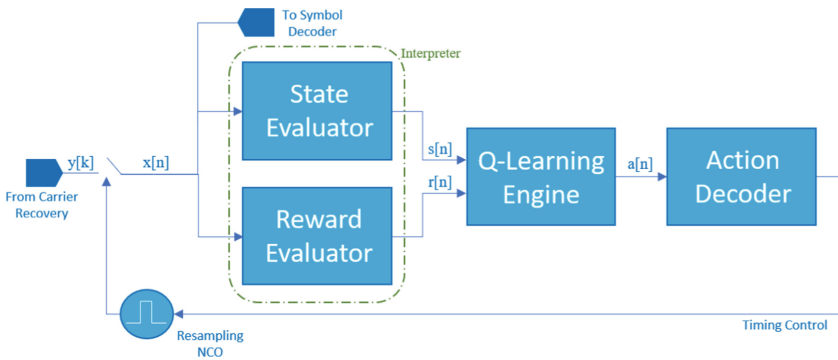


Fig. 1. Top-level architecture of the proposed system. (Color figure online)

The main blocks of the top-level design are the following.

- *State evaluator*: it evaluates the environment state, being part of the interpreter.
- *Reward evaluator*: it evaluates the reward for the previous taken action, it is also part of the interpreter.
- *Q-Learning engine*: it updates the Q-Matrix [13] and implements the action selection policy [11].
- *Action decoder*: it converts the generated actions to a suitable signal for the sampler.
- *Numerically Controlled Oscillator*: it samples the input signal according to the control given by the action decoder.

The state estimation is performed with Eq. (2) and implemented by the architecture in Fig. 2.

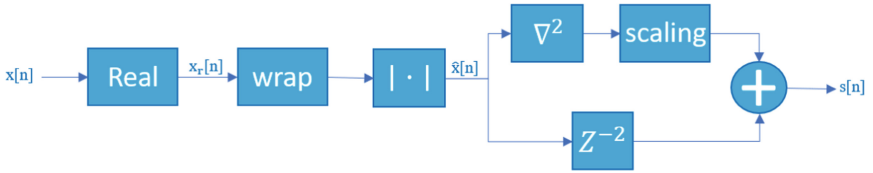


Fig. 2. Architecture of the state estimator. (Color figure online)

$$s[n] = N/2 \cdot \left(\nabla^2 \hat{x}[n] + 1 \right) + \hat{x}[n-2] \quad (2)$$

$$\hat{x}[n] = |x_r - \text{round}(x_r)| \quad (3)$$

$$\nabla^2 \hat{x}[n] = \hat{x}[n] - 2\hat{x}[n-1] + \hat{x}[n-2] \quad (4)$$

N is the number of states, while the “Real”, “Wrap” and “Abs” blocks in Fig. 2 represent the operations in Eq. (3). The “scaling” block performs the first part of Eq. (2).

As stated before, the Q-Learning engine has been implemented according to [13] and policy generator according to [11]. The latter has been set to work using an ϵ -greedy policy.

The reward evaluator architecture (Fig. 3) is based on the one in [7] by simply multiplying for a negative constant λ the output of the *error block* in the referenced work.

The variables in the previous architecture are defined in the following equations.

$$R[n] = |I| + |Q| \quad (5)$$

$$\bar{R}[n] = \sum_{i=0}^{N-1} \frac{R[n-1]}{N} \quad (6)$$

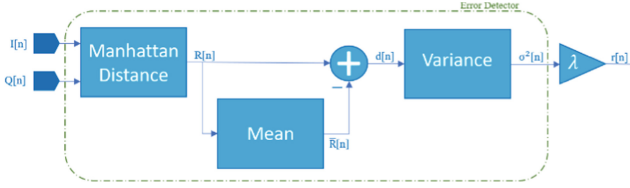


Fig. 3. Reward evaluator architecture (Color figure online)

$$d[n] = R[n] - \bar{R}[n] \quad (7)$$

$$\sigma^2[n] = \sum_{i=0}^{N-1} \frac{d^2[n-1]}{N} \quad (8)$$

$$r[n] = \lambda \cdot \sigma^2[n], \quad \lambda < 0 \quad (9)$$

3 Experimental Results

The experimental setup is shown in Fig. 4.

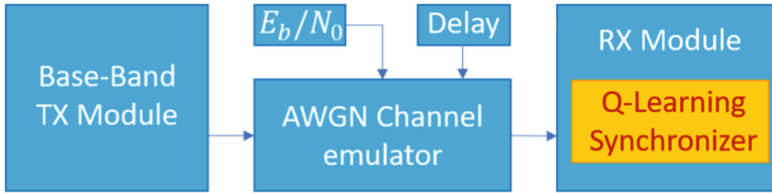


Fig. 4. Experimental setup (Color figure online)

We tested the proposed architecture considering a carrier already recovered, therefore we can use a Base-Band Transmitter (TX) module for the generation of the M-PSK symbols. The same module applies a root raised cosine shape to the signals. The Average White Gaussian Noise (AWGN) channel emulator has a controllable bit energy over spectral power density ratio (E_b/N_0), which adds the noise, and also a controllable delay. The Receiver module (RX) contains the proposed Q-Learning based synchronizer described in the previous section. It also applies a root raised cosine filtering.

In the simulations, we set the following parameters.

- An ϵ -greedy policy with $\epsilon = 10^{-5}$ to allow the system to explore the telecommunication environment.
- A leaning rate $\alpha = 0.1$ to achieve a tradeoff between convergence speed and the avoidance of local maxima in the reward function.
- A discount factor $\gamma = 0.01$ to speed-up the synchronization process.

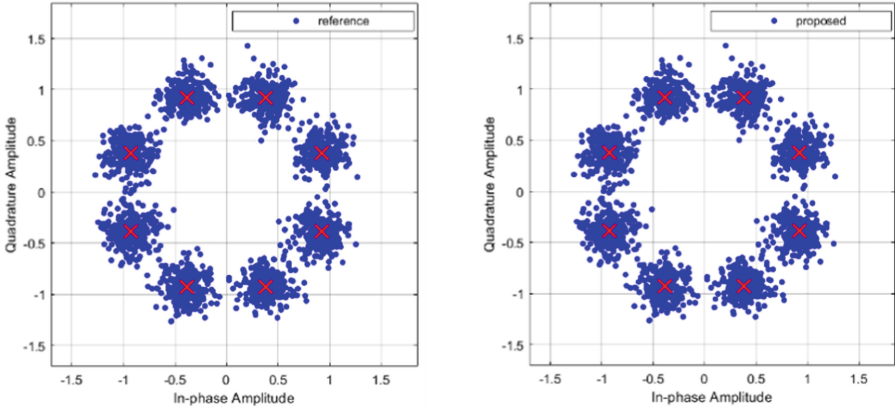


Fig. 5. Comparison of the constellation diagrams for a 8-PSK modulation. (Color figure online)

We compared the synchronization Modulation Error Rate (MER) between the system in [7] and the proposed approach obtaining always the same values but with a consistent speed-up of the convergence.

As example, we show the results for a 8-PSK modulation where $M = 16$, the delay introduced by the channel is 8 samples, and $E_b/N_0 = 15$ dB. The constellation diagrams are shown in Fig. 5, in both cases we obtained a MER of 14.9 dB.

Regarding the convergence time, we show in Fig. 6 how the proposed method is able to retrieve the correct delay of 8 samples faster than the reference method.

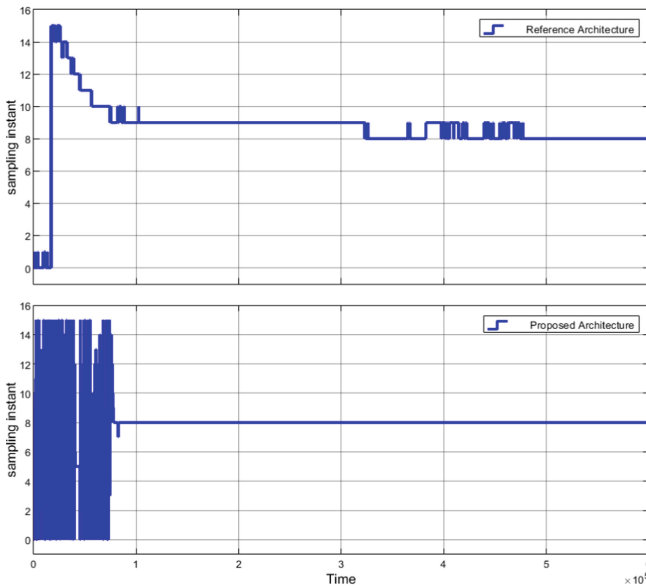


Fig. 6. Comparison of the locking time. (Color figure online)

4 Conclusion

We proposed a novel timing recovery loop based on the Reinforcement Learning algorithm called Q-Learning. From the experimental results, we obtained a system able to synchronize every M-PSK symbol with the same modulation error rate of the reference algorithm.

The proposed approach allows for a speed-up in the synchronization process and can adapt itself to every modification to the telecommunication channel in terms of noise and delay of the samples.

References

1. Haykin, S.: *Communication Systems*, vol. 103, no 6, 4th edn. Wiley, New York (2000). (Simon Haykin With Solutions Manual.pdf. Cell)
2. Ling, F., Proakis, J.: *Synchronization in Digital Communication Systems*. Cambridge University Press, Cambridge (2017). <https://doi.org/10.1017/9781316335444>
3. Naidoo, G.M.: Digital communication: information Communication Technology (ICT) usage for teaching and learning. In: Montebello, M. (ed.) *Handbook of Research on Digital Learning*, pp. 1–19. IGI Global (2020). <https://doi.org/10.4018/978-1-5225-9304-1.ch001>
4. Sklar, B.: *Digital Communications: Fundamentals and Applications*. Signals, 2nd edn. Communications Engineering Services, Tarzana (2001)
5. Mueller, K., Muller, M.: Timing recovery in digital synchronous data receivers. *IEEE Trans. Commun.* **24**(5), 516–531 (1976). <https://doi.org/10.1109/TCOM.1976.1093326>
6. Gardner, F.: A BPSK/QPSK timing-error detector for sampled receivers. *IEEE Trans. Commun.* **34**(5), 423–429 (1986). <https://doi.org/10.1109/TCOM.1986.1096561>
7. Giardino, D., et al.: M-PSK demodulator with joint carrier and timing recovery. *IEEE Trans. Circuits Syst. II Express Briefs* (2020). <https://doi.org/10.1109/TCSII.2020.3041342>
8. Cardarilli, G.C., et al.: A Q-learning based PSK symbol synchronizer (2019). <https://doi.org/10.1109/ISSCS.2019.8801727>
9. Matta, M., et al.: A reinforcement learning-based QAM/PSK symbol synchronizer. *IEEE Access* **7**, 124147–124157 (2019). <https://doi.org/10.1109/ACCESS.2019.2938390>
10. Sutton, R.S., Barto, A.G.: Reinforcement learning: an introduction. *IEEE Trans. Neural Networks* **9**(5), 1054–1054 (1998). <https://doi.org/10.1109/TNN.1998.712192>
11. Cardarilli, G.C., et al.: An action-selection policy generator for reinforcement learning hardware accelerators. In: Saponara, S., De Gloria, A. (eds.) *ApplePies 2020*. LNEE, vol. 738, pp. 267–272. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-66729-0_32
12. Watkins, C.J.C.H., Dayan, P.: Q-learning. *Mach. Learn.* **8**(3–4), 279–292 (1992). <https://doi.org/10.1007/BF00992698>
13. Spanò, S., et al.: An efficient hardware implementation of reinforcement learning: the q-learning algorithm. *IEEE Access* **7**, 186340–186351 (2019). <https://doi.org/10.1109/ACCESS.2019.2961174>



Scalable Broadband Switching Matrix for Telecom Payload Based on a Novel SWGs-Based MZI

G. Brunetti, G. Marocco, A. Giorgio, M. N. Armenise, and C. Ciminelli^(✉)

Optoelectronics Laboratory, Politecnico di Bari, Via E. Orabona 6, 70125 Bari, Italy
caterina.ciminelli@poliba.it

Abstract. Photonics is a disruptive technology also for enabling telecommunication payloads to achieve high performance and potentially low costs. In this paper, we propose the design of a scalable switching cell that consists of a Mach-Zehnder interferometer (MZI), based on a subwavelength grating (SWG) coupler and a thermo-optical phase shifter. The scalability of the device was demonstrated by evaluating the performance of a 4×4 dilated strictly non-blocking Banyan network. Worst-case insertion loss (IL) of 9 dB, crosstalk (XT) of -34 dB and extinction ratio (ER) of 42 dB have been theoretically proved, as well as power consumption of 46 mW and time response of $9.2 \mu\text{s}$, within a footprint of $1080 \mu\text{m} \times 288 \mu\text{m}$.

1 Introduction

In recent decades, the growing demand for data transmission has increased the attention towards telecommunication satellite payloads. In this context, photonics enables the development of systems that guarantee high performance at potential low costs [1]. To meet the increase in data traffic, an improvement of switching technologies, in terms of high-speed, broadband operation, low-power consumption, is strongly required [2, 3]. For example, the European Space Agency guidelines for developing next optical switching technologies prescribe more than two input/output ports, a few μs response time, low insertion loss, and -20 dB as maximum crosstalk [4].

This paper describes the design of a SOI scalable broadband switching matrix, whose key element is a 2×2 SWG-based MZI, aiming at achieving large bandwidth and low crosstalk. First, we present the study of a 2×2 MZI discussing its operational behavior and performance. To demonstrate the feasibility of the proposed device as a key building block of a strictly non-blocking scalable matrix, a 4×4 dilated Banyan topology was designed. The proposed device features a wide operating bandwidth (>150 nm), fast response ($9.2 \mu\text{s}$), low crosstalk (< -34 dB), insertion loss of 9 dB, ultra-compact footprint, and total power consumption less than 50 mW, making it highly promising for space applications.

2 State-Of-The-Art

Photonic switches have been recognized as key components that have significantly transformed the optical transport networks in the last decades, delivering functions such as protection, signal monitoring, cross-connection and add-drop multiplexing [5, 6] and, then, supporting a continuous evolution and increased performance in terms of capacity, flexibility and coverage, so that they are currently seen as promising for being used in data centres, cloud servers and high-throughput computing [6, 7].

The literature on photonic switching technologies is huge and can be declined in terms of studies on both system applications and basic technologies for the single switching cell, categorized on the basis of operating principles, physical effects used and/or geometric configurations. Deeply investigated technologies for photonic switching cells [5] include LiNbO_3 , III-V, liquid crystal, all-optical configurations, alongside with micro-electromechanical systems (MEMS) or optical semiconductor amplifiers. Several books and review papers can be used to get a general overview also on specific applications and related target performance (see as examples [5, 6, 9]).

An $N \times N$ photonic switching matrix [10] with a large input/output count number can be composed of a set of identical smaller modules (starting from 1×2 or 2×2 switching cells) that are joined in a suitable topology. Their specific choice depends on the matrix target characteristics, such as blocking or not condition, and the consequent number of stages and modules, together with the overall matrix performance as extinction ratio, crosstalk insertion loss, switching time, power consumption and signal-to-noise ratio.

One of the most promising approaches for large switching matrices are the MEMS, thanks to clear benefits such as low optical loss and power consumption, monolithic integration, and high scalability. This approach has also already provided commercial devices.

Our research focuses on integrated photonic switches, among which the silicon-based ones are largely investigated due to their reduced volume, low power consumption, ease of integration with CMOS electronics and other photonic components, potential high density capability, and stability and robustness of the full circuits. Silicon-based 2×2 switching cells are usually made by micro ring resonators (MRR) and MZIs [5]. MRR switches are resonant devices offering a wavelength selective filtering functionality. They have relatively low switching power consumption and, when compared to MZI-based switches, are also more compact building blocks, which is good for high-density, large port count matrices. However, the intrinsic wavelength selective behaviour is in contrast with a broadband operation, requested for some applications, unless to use configurations of cascaded MRRs [11], able to engineer the spectrum and to improve the crosstalk between adjacent channels. The technology is, in any case, sensitive to the fabrication tolerances and the spectrum drift due to the temperature.

MZIs convert very efficiently a phase change into an intensity modulation. These devices are typically broadband because not limited by any channel spacing as the resonant ones. A 2×2 MZI switch consists of two 3 dB couplers and two waveguide arms between them [12]. By controlling the phase difference between the two arms, and, then, the interference between the light paths, one can force the output from the cross port to the bar one. Both thermo-optic and free-carrier plasma dispersion effect can be exploited in silicon to locally change the refractive index, obtaining the required phase

shift. Several switching fabrics based on some topologies [13] have been proposed by combining a number of 2×2 MZIs as basic cells.

To obtain wavelength-insensitive operation over a large bandwidth, in addition to the configurations reported in the literature, including straight/bent directional couplers and multi-mode interferometers (MMIs) [14–16], we propose a MZI including two directional coupler, at input and output, based on SWGs.

3 2×2 MZI Switch

Figure 1(a) sketches the proposed device, which consists of a balanced MZI made of two 2×2 3 dB directional couplers. Each coupler consists of two coupled SWGs, aiming at obtaining a wavelength insensitive behavior. The π shift between the optical beams travelling in the two MZI decoupled arms is obtained using a thermo-optical phase shifter placed on one of the MZI arms. Figure 1(b) and Fig. 1(c) show waveguide cross section and longitudinal view of the subwavelength grating, respectively.

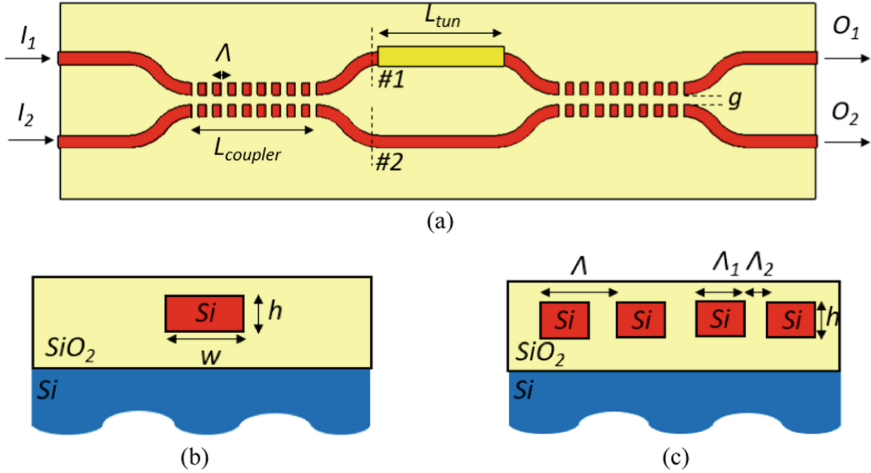


Fig. 1. (a) Proposed SOI SWG-based MZI. (b) Cross-section of the bare waveguide. (c) Longitudinal view of the SWG. ($I_{1,2}$ = input signals, $O_{1,2}$: output signals, #1,2: output signals of the 3 dB 2×2 coupler, h : waveguide thickness, w : waveguide width, Λ : SWG period, $\Lambda_{1,2}$: length of the SWG portions, $L_{coupler}$: 3 dB 2×2 coupler length, g : gap between the 2×2 3 dB coupler arms, L_{tun} : length of the phase shifter). (Color figure online)

The SWG regime has been exploited to achieve a broadband operation with high extinction ratio and low insertion loss. At subwavelength scale, the diffraction effects within the grating are negligible, and the SWG can be approximated to a homogeneous medium [17], with an effective refractive index $n_{eff} = DC \cdot n_1 + (1 - DC) \cdot n_2$, which depends on the duty cycle (DC), expressed as the ratio Λ_1/Λ , and the refractive indices n_1 , n_2 corresponding to the SWG portions. This assumption leads to a significant decrease in the length, $L_{coupler}$, of the coupler, according to the supermode theory [18], and mitigates

the dispersion effect. The proposed SWG consists of a top grating, made by a thickness modulation of a silicon (Si) strip waveguide, with width w and unperturbed thickness h , fully embedded in silicon dioxide (SiO₂). A propagation loss of 0.4 dB/cm has been assumed for the bare waveguide [19].

Although a SWG is, in theory, a lossless structure, we deemed the excess losses to take into account the loss contributions due to evanescent fields and backreflections of the grating, and, mainly, losses caused by any variation or irregularity in the grating characteristics induced by the fabrication process. In particular, a value of 4 dB/cm for the SWG was assumed in our simulations to take into account all the coupler loss contributions, achieving SWG-coupler excess losses compliant to some results reported in the literature [20]. Furthermore, to ensure the feasibility and repeatability of the manufacturing process, a gap g of 200 nm has been chosen.

The $L_{coupler}$ dependence on the operating wavelength should be attenuated to obtain broadband output spectra. The SWG has been designed using the Coupled Mode Theory (CMT) approach [21], aiming at maximizing the insensitivity of $L_{coupler}$ from the wavelength λ by engineering the period Λ , the modulation depth, and the duty cycle DC . A period $\Lambda = 300$ nm was considered a good compromise between enhancing the SWG effect and the manufacturing constraints in the C band.

Since an increase of the modulation depth leads to a decrease of the n_{eff} , and then of $L_{coupler}$, with a remarkable suppression of its dispersion, a modulation depth of 200 nm was considered (see Fig. 1(c)). Furthermore, as expected, the $L_{coupler}$ dispersion decreases when DC also decreases, due to the reduction of n_{eff} at the expense of a reduced efficacy of the SWG effect. Therefore, we assumed a $DC = 50\%$. Then, the designed parameters for the 3 dB coupler can be summarized as $\Lambda = 300$ nm, $\Lambda_1 = 150$ nm, $\Lambda_2 = 150$ nm, $DC = 50\%$, $g = 200$ nm and $L_{coupler} = 2.47$ μm .

The transmission spectrum of the 2×2 SWG-based coupler has been simulated through the 2D Finite Difference Time Domain (FDTD) approach, assuming the port I_1 (see Fig. 1(a)) as input. The device includes also bend waveguides, with a curvature radius of 2 μm , a total length of 8 μm and optical losses of 0.2 dB. In the simulation, the Sellmeier equations for Si and SiO₂ have been considered to take into account the material dispersion.

Figure 2 shows the power transfer efficiency η at the 3 dB output ports #1 and #2, expressed as the relative percentage power ratio between the two output ports. The SWG-based 2×2 coupler shows a power efficiency of about 50% at both output ports #1 and #2 over a bandwidth of 150 nm, from 1500 nm to 1650 nm, with a maximum imbalance of $\pm 13\%$.

For evaluating the achieved performance, the SWG-based coupler has been compared to a conventional directional coupler, in the same material platform with $g = 200$ nm and $L_{coupler} = 7$ μm . As reported in [22], the conventional coupler shows a $\pm 15\%$ power imbalance over 100 nm bandwidth. The power imbalance worsens as the bandwidth increases. Therefore, introducing a SWG within a directional coupler leads to a broadband operation with also a significant footprint saving.

A thermo-optical phase shifter was designed to impose a π shift to the propagating wave and, then, to get switching between the cross and bar state. It consists of a 100 nm thick and 0.8 μm wide TiN layer, placed on top of the silicon core waveguide at a

distance of $0.4 \mu\text{m}$. These design parameters were determined to optimize both power consumption and time response [23]. The optical and electrical performances of the phase shifter were simulated using a Finite Element Method (FEM)-based solver, assuming a temperature equal to 300 K at the simulation domain boundaries.

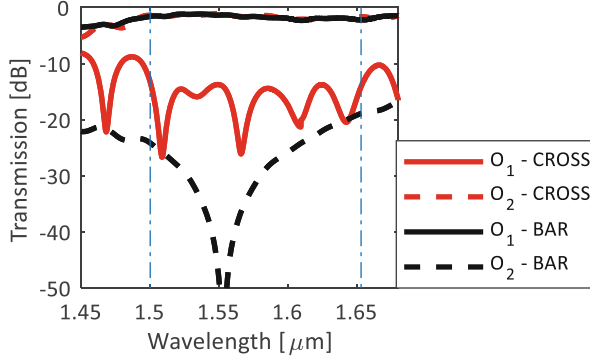


Fig. 2. Transmission spectra at O_1 (solid) and O_2 (dotted) for cross (red) and bar (black) state. (Color figure online)

With $L_{shifter} = 200 \mu\text{m}$, distance from the waveguide of $0.4 \mu\text{m}$ and width of $0.8 \mu\text{m}$, it was also possible to achieve a good device compactness, together with a time response of $9.2 \mu\text{s}$ and a power consumption of 11.5 mW .

To estimate the performance of the whole 2×2 switching cell, the transfer function of the SWG-based MZI, shown in Fig. 1(a), was derived by using a matrix approach [24]. The overall S-matrix was derived as the matrix product of cascade S-matrices related to the MZI key building blocks, including 2×2 couplers and phase shifter (see Fig. 1(a)), whose S parameters were calculated by the FDTD approach. The transmission spectra of the designed 2×2 MZI switch in the bar and cross state at the output ports O_1 and O_2 (see Fig. 1(a)) are shown in Fig. 2, where the spectrum for the bar state has the typical MZI trend, while the ripple in the cross-state spectrum is related to the oscillations of the coupler spectrum shown in Fig. 2. In both switching states, the switch can operate within a wavelength range from 1500 nm to 1653 nm , with the worst extinction ratio of 13 dB and 19 dB for the cross and bar state, respectively. Furthermore, both switching states show $\text{IL} < 2 \text{ dB}$, with $\text{XT} = -16 \text{ dB}$ and -12 dB for the bar and cross state, respectively. The total footprint of the switching cell is $240 \mu\text{m} \times 9 \mu\text{m}$.

A mature manufacturing technique is available to fabricate the proposed device. Starting from a 220 nm thick SOI wafer, the waveguide can be patterned through e-beam lithography and etching processes. The SWG regions can be realized by further lithography and etching processes. Then, 400 nm thick silica upper cladding can be deposited using the plasma enhanced chemical vapor deposition technology. Finally, 100 nm thick and 800 nm wide titanium heaters can be fabricated on top of the MZI arms using a metal deposition technique.

4 4×4 Dilated Banyan Network

To demonstrate the feasibility of the proposed 2×2 switching cell as a key building block of a larger switching matrix, an $N \times N$ dilated Banyan topology has been designed with $2N^2 - 2N$ switches. When the dilated Banyan topology is fully loaded with N connected lightpaths, $2\log_2(N)$ cells are in the connected lightpaths ($2\log_2(N)$ cells/lightpath), and the remaining cells are idle and can be in any state. A dilated Banyan topology allows suppressing first order crosstalk with a strictly non-blocking behaviour since each switching cell routes just one signal, avoiding any interference. In particular, each input-output connection has a dedicated pathway, where only 1×2 and 2×1 switches are used, preserving the signal-to-noise ratio performance from first-order crosstalk contributions. For the whole matrix, with $N = 2, 4, 8$, XT of -25 dB, -35 dB, and -38 dB, has been predicted, according to [25], with IL of 4 dB, 9 dB, and 15 dB, respectively. The main design criterion was to balance XT and IL.

Routing of the signals In_1 , In_2 , In_3 , and In_4 to the output port Out_3 , Out_1 , Out_2 , and Out_4 , respectively, with a non-blocking behavior, was investigated, as in Fig. 3. We optimized the waveguide crossings in the matrix to achieve low insertion loss and high crosstalk. Using the FDTD approach, for a 90° degree crossing with an adiabatic trend of the waveguiding sections [26] resulted $IL < 0.1$ dB and $XT < -36$ dB.



Fig. 3. Schematic of a proposed example 4×4 dilated Banyan network. (Color figure online)

Figure 4 shows the simulated transmissions at the four different output ports for all the connections that can be activated in the matrix. Transmission values that lead to the worst XT at each port have been considered. The worst value $XT \approx -34$ dB was obtained with $ER \approx 42$ dB and $IL \approx 9$ dB. The continuous improvement of the manufacturing technology could lead to the reduction of SWG-coupler excess loss and then the network IL.

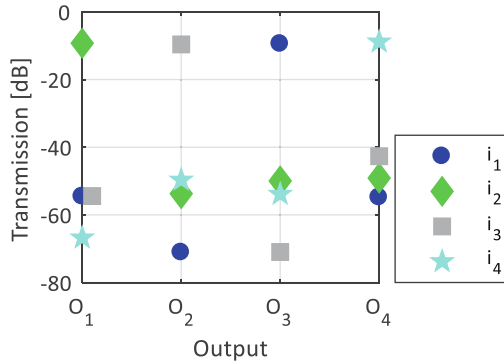


Fig. 4. Transmission of the 4×4 switch with idle switches randomly in bar or cross state. (Color figure online)

5 Conclusions

We have proposed the design of a scalable broadband switching element based on a SWGs-MZI. The proposed device consists of a MZI switch made of two SWG-couplers and a thermo-optical phase shifter. The SWG-based switching cell shows an operating bandwidth of about 150 nm with a worst-case IL ≈ 2 dB and XT = -12 dB. To demonstrate its scalability, the performance of a 4×4 dilated Banyan network with a strictly non-blocking behavior has been calculated. For the whole matrix, a worst-case IL ≈ 9 dB, XT ≈ -34 dB, and ER ≈ 42 dB has been derived with a total power consumption of 46 mW, time response of 9.2 μ s, and footprint ≈ 0.1 mm².

To meet the request of a larger port number in the next generation payloads, the 4×4 switching matrix could be replicated in a layered structure, which could result very promising for several Space applications.

Acknowledgments. The work has been supported by Ministry of Research and University in the framework of New Satellites Generation Components (NSG) project – ARS01_01215.

References

1. Ciminelli, C., Dell’Olio, F., Armenise, M.N.: *Photonics in Space: Advanced Photonic Devices and Systems*. World Scientific (2016). <https://doi.org/10.1142/9817>
2. Ravel, K., et al.: Optical switch matrix development for new concepts of photonic based flexible telecom payloads. In: 2018 ICSO (2019) 111803H
3. Hauschildt, H., et al.: HyDRON: high throughput optical network. In: 2019 IEEE ICSOS, pp. 1–6 (2019)
4. Poudereux, D., Barbero, J., Tijero, J.M.G., Esquivias, I., Mc Kenzie, I.: Evaluation of optical switches for space applications. In: 2016 ICSO, 11180 (2019) 111807H
5. Ciminelli, C., Armenise, M.N.: *Photonic Switches*, Enc. of Optical and Photonic Engineering, 2nd edn. CRC Press (2016)
6. Stabile, R., Albores-Mejia, A., Rohit, A., Williams, K.A.: Integrated optical switch matrices for packet data networks. *Microsys. Nanoeng.* **2**, 1–10 (2016)

7. Soref, R.: Tutorial: integrated-photonic switching structures. *APL Photon.* **3**, 021101 (2018)
8. El-Bawab, T.S. (ed.): *Optical Switching*. Springer US, Boston, MA (2006). <https://doi.org/10.1007/0-387-29159-8>
9. Li, B., Chua, S.J. (eds.): *Optical Switches*. Woodhead Publications (2010)
10. Hinton, H.: *An Introduction to Photonic Switching Fabrics*. Springer US, Boston, MA (1993). <https://doi.org/10.1007/978-1-4757-9171-6>
11. Emelett, S.J., Soref, R.A.: Synthesis of dual-microring-resonator cross-connect filters. *Opt. Express* **13**(12), 4439 (2005)
12. Ramaswamy, V., Divino, M., Standley, R.D.: Modified balanced-bridge switch with two straight waveguides. *Appl. Phys. Lett.* **32**, 644–646 (1978)
13. Lee, G.B., Dupuis, N.: Silicon photonic switch fabrics: technology and architecture. *J. Lightw. Technol.* **37**(1), 6–20 (2019)
14. Yang, H., Kuan, Y., Xiang, T., Zhu, Y., Cai, X., Liu, L.: Broadband polarization-insensitive optical switch on silicon-on-insulator platform. *Opt. Express* **26**, 14340–14345 (2018)
15. Van Campenhout, J., Green, W.M., Assefa, S., Vlasov, Y.: A low-power, 2×2 silicon electro-optic switch with 110-nm bandwidth for broadband reconfigurable optical networks. *Opt. Express* **17**, 24020–24029 (2009)
16. Chen, S., Shi, Y., He, S., Dai, D.: Low-loss and broadband 2×2 silicon thermo-optic Mach-Zehnder switch with bent directional couplers. *Opt. Lett.* **41**, 836–839 (2016)
17. Lalanne, P., Lemerrier-Lalanne, D.: On the effective medium theory of subwavelength periodic structures. *J. Mod. Opt.* **43**, 2063–2085 (1996)
18. Alferness, R.C., Cross, P.: Filter characteristics of codirectionally coupled waveguides with weighted coupling. *IEEE J. Quant. Elec.* **14**, 843–847 (1978)
19. Horikawa, T., Shimura, D., Mogami, T.: Low-loss silicon wire waveguides for optical integrated circuits. *MRS Commun.* **6**(1), 9–15 (2015). <https://doi.org/10.1557/mrc.2015.84>
20. Wang, Y., et al.: Compact broadband directional couplers using subwavelength gratings. *IEEE Photon. J.* **8**(3), 1–8 (2016)
21. Brunetti G., Dell’Olio F., Conteduca D., Armenise M.N., Ciminelli C.: Comprehensive mathematical modelling of ultra-high Q grating-assisted ring resonators. *J. Opt.* **22**, 035802 (2020)
22. Brunetti, G., Marocco, G., Di Benedetto, A., Giorgio, A., Armenise, M.N., Ciminelli, C.: Design of a large bandwidth 2×2 interferometric switching cell based on a sub-wavelength grating. *J. Opt.* **23**, 085801 (2021)
23. Brunetti, G., Sasanelli, N., Armenise, M.N., Ciminelli, C.: High performance and tunable optical pump-rejection filter for quantum photonic systems. *Opt. Las. Tech.* **139**, 106978 (2021)
24. Čtyroký, J., Richter, I., Šiňor, M.: Dual resonance in a waveguide-coupled ring microresonator. *Opt. Quan. Ele.* **38**, 781–797 (2006)
25. Hunter, D.K., Smith, D.G.: New architectures for optical TDM switching. *J. Light. Tech.* **11**, 495–511 (1993)
26. Ma, Y., et al.: Ultralow loss single layer submicron silicon waveguide crossing for SOI optical interconnect. *Opt. Exp.* **21**, 29374–29382 (2013)



A Smart Portable Potentiostat for Point-of-Care Testing

Marco Bassoli¹, Valentina Bianchi¹, Andrea Boni¹, Simone Fortunati²,
Marco Giannetto², Maria Careri², and Ilenia De Munari¹(✉)

¹ Department of Engineering and Architecture, University of Parma, Parma, Italy
{marco.bassoli, valentina.bianchi, andrea.boni,
ilaria.demunari}@unipr.it

² Department of Chemistry, Life Sciences and Environmental Sustainability,
University of Parma, Parma, Italy
{simone.fortunati, marco.giannetto, maria.careri}@unipr.it

Abstract. In this paper, a portable potentiostat, based on the ad-hoc designed Analog Front End (AFE) and equipped with a microcontroller including a Wi-Fi network processor is presented. Unlike the most common portable sensors that lean on a smartphone or personal computers, the device exploits the cloud service for data storage. This feature enhances the device portability and makes it suitable for various semi-quantitative and quantitative analyses. The effectiveness of the proposed sensor was validated with measurements on known concentrations of potassium ferri/ferrocyanide redox probe in aqueous solution. The resulting calibration curve was compared with the results reported in the literature, showing an R² value comparable or even higher. This result is very promising for both semi-quantitative and quantitative analyses.

1 Introduction

The Internet of Things (IoT) can be seen as an extension of the Internet concept to different devices and sensors, providing a high degree of computing and analytical capabilities even to simple objects. Therefore, business use of the IoT is increasing, involving different fields of application ranging from consumer electronics to industrial products. In this context, healthcare devices represent one of the fastest-growing sectors of the IoT market. The most common application is remote monitoring, where the IoT devices can be used to collect health metrics from patients at their homes or as close as possible to them. Therefore, the need to reach a hospital environment is overcome with the implementation of the so-called Point of Care Testing (POCT) paradigm.

Electrochemical biosensors for the identification of clinically relevant biomarkers for disease diagnosis and follow-up and healthcare management are not detached from this scenario. Flexible and portable sensors equipped with IoT features and connected to a cloud service to share the results with physicians/caregivers/users can be designed for *in situ* detection and long-term monitoring of target analytes (e.g. as a follow-up of a particular pathology). In the current COVID-19 pandemic context, for example, the

possibility of carrying out analyses in a rapid and delocalized way would have helped to counter the spread of the virus.

For these purposes, Differential Pulse Voltammetry (DPV), requiring a three-electrode potentiostat, is one of the most used techniques to quantify target analytes. To maximize the device flexibility, both semi-quantitative and quantitative analysis should be performed. As for semi-quantitative analysis, the device should discriminate between two types of samples, positive and negative, through a suitable threshold value. Conversely, in a quantitative approach, the aim is the assessment of the concentration of the tested biomarker. This analysis requires the implementation of a proper calibration function for the processing of the measurement data [1] and the reconfigurability of the device based on the evaluated parameters.

In particular, enabling remote access to the device would allow an easy reconfiguration, which is mandatory in several applications. For example, the calibration procedure or the thresholds calculations could be done once, and the obtained parameters could be stored in a cloud service. Then, the portable device will download the necessary configuration and calibration parameters using relevant data and settings according to the analysis to be carried out in the POCT context.

A flexible device should also perform the tests in places where no internet connection is available. For this reason, the data collected should be processed and stored onboard to be later sent to the cloud when the network link is restored. This solution also provides relevant benefits in terms of power consumption since both the amount of data transferred to the cloud and the front-end radio's overall working time are minimized.

In the literature, portable solutions have been described [2, 3], but designed to be used with a computer connected through a USB port or with a smartphone via Bluetooth (BT) link for data processing.

In this paper, a device designed for POCT scenarios is presented. It is a portable, cloud-based reconfigurable potentiostat for both semi-quantitative and quantitative analysis, exploiting a web-based interface to simplify the configuration process. This interface is available from any PCs, tablets or smartphones, without the need for any custom and application-specific software. The parameters for semi-quantitative or quantitative analyses are stored in a cloud service for subsequent download and sharing from different devices. Furthermore, to increase its flexibility, the elaboration of the acquired data is performed onboard without the need for any Internet connection.

The paper is organized as follows: in Sect. 2, the system architecture is depicted, with both the hardware and web components; in Sect. 3, the tests are described, and the results are presented and discussed. Finally, in Sect. 4, conclusions are drawn.

2 System Architecture

The proposed device is devised to be interfaced to an immunosensor, achieving accuracy comparable to that of common laboratory equipment but with a smaller size for maximum portability. The device is based on a system-on-chip (SOC) integrating a microcontroller and a Wi-Fi network processor. In our previous work [4] a prototype version of the device has been developed. In this paper, some hardware improvements are presented with a first version of the web interface.

The designed hardware consists of a custom-designed AFE (Analog Front-End) [4] and a microcontroller (MCU) with an integrated Wi-Fi transceiver (CC3200 by Texas Instruments [5]). The hardware is combined with a biosensor belonging to the amperometric sensor category. The AFE circuit performs the conditioning of the cell and the measurement of the cell current due to the chemical reaction, whereas the MCU performs signal conversion in the digital domain, data processing, and data transmission to a cloud service (Fig. 1).

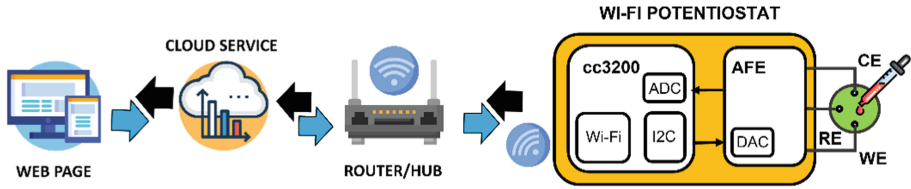


Fig. 1. System architecture.

An amperometric sensor is based on a three-electrode electrochemical cell consisting of a working electrode (WE), a counter electrode (CE), and a reference electrode (RE). The current output (flowing from/into the WE pin) is related to the concentration of the target analyte and cell bias voltage V_{bias} , i.e., the voltage difference between the WE and the RE pin voltage. The chemical reaction is induced by a suitable V_{bias} input. In the proposed design, the DPV technique is used [4]. The details of the DPV implementation are reported in the next section.

The simplified schematic of the implemented AFE is shown in Fig. 2 with the electrical model of the amperometric cell. The opamp in the upper part of the circuit (ACTRL) drives the CE pin to set the RE voltage equal to the V_{Sig} input [6]. This amplifier must exhibit a low output resistance to maintain a good stability margin with a cell capacitance C_{CH} as large as a few microfarads. Furthermore, an input bias current in the nanoampere range is required not to perturb the chemical reaction with current flowing from/into the RE pin. In the proposed design, an AD8608 opamp was used. The cell current is

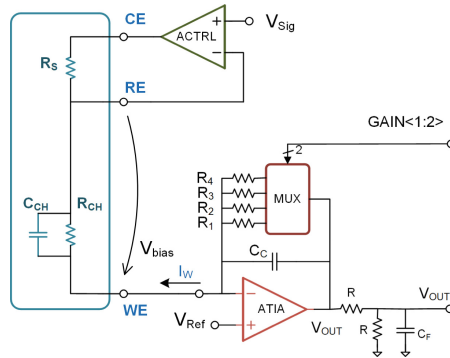


Fig. 2. AFE (simplified schematic) and electrical model of the electrochemical cell.

converted into a voltage signal by a TransImpedance Amplifier (TIA), setting also the reference potential at the WE pin [2, 4, 7]. Therefore, the cell bias voltage V_{bias} corresponds to the difference between the input signal V_{Sig} and the TIA reference voltage V_{Ref} .

This transimpedance amplifier is implemented with an opamp (ATIA), feedback resistors setting the TIA gain, and a feedback capacitance introducing a left-half plane zero in the loop-gain transfer function for adequate phase margin.

The cell current range depends on the chemical reaction and concentration of the analyte to be identified and quantified. Such chemical-dependent parameters and the cell substrate define the value of the RE-to-WE resistance R_{CH} and capacitance C_{CH} . The programmability of the TIA gain is a mandatory feature to make to potentiostat suitable for different applications, cell substrates, and analytes. In the AFE of Fig. 2 the TIA gain is changed with an ADG704 four-channel CMOS multiplexer controlled by the MCU with two control bits ($\text{GAIN} < 1:2 >$). Thus, the TIA gain is programmed by selecting the feedback resistor among four (i.e. R_1 , R_2 , R_3 , and R_4). At the power-on, the maximum gain is chosen to minimize the effect of the quantization error due to the A/D converter (ADC) that digitizes the TIA output voltage. During the measurement, the MCU monitors the ADC output and reduces the TIA gain if its output reaches the saturation limits. In the current implementation, the TIA gain can be varied from 1 M Ω to 33 k Ω .

The proposed solution relies on the cloud platform MATLAB ThingSpeak [9], for data storage and processing, and on static webpages hosted on a remote server for the front-end and user interaction. The developed control app is accessible from any mobile device (e.g. smartphone or tablet) or PC. Depending on the operation to be performed on the system, one out of three types of profiles can be selected: Technician, Caregiver, or User.

The Technician profile is designed for the administrator managing the device and setting the types of analysis to be performed, along with the electrical parameters and the expected response. A dedicated page can be filled, where the values of the voltages applied to the cell (i.e. V_{WE} and V_{RE}) are specified. Different calibration modes can be selected according to the need to carry out a semi-quantitative or quantitative analysis (e.g. a threshold or linear regression).

The Caregiver profile is created for the typical tasks that a healthcare professional should carry out. In a typical scenario, the operator should set up the device at the beginning of a set of measurements by choosing one of the available analyses already defined by the administrator with the Technician profile. Depending on the selected analysis, different calibration processes must be considered before starting a campaign of measurements. For example, for a given semi-quantitative analysis, starting from a few acquisitions using samples with known concentration values, the suitable threshold values are defined. As for quantitative analysis, the parameters of the appropriate regression function must be extrapolated. In both cases, the computation algorithms are running on ThingSpeak platform. The Caregiver must perform the calibration procedure carrying out a predefined number of measurements on calibration standards with known concentration. The concentration values are then reported in the dedicated section of the web page for the evaluation of the calibration function parameters. The procedure can

be completed on the cloud service after the measurement data are uploaded. As soon as the calibration data processing is completed, the calibration parameters are available for download to the device.

The User profile is exploited to perform a new analysis using the calibration terms obtained in the previous step. With this profile, the analysis results are immediately shown for rapid feedback. If no Wi-Fi network is available or the battery charge must be preserved, the analysis and subsequent data processing can be carried out directly onboard, reducing the amount of data to be uploaded. The obtained result can be subsequently reviewed by the user at the dedicated webpage.

Furthermore, on the Historical Records page, all the results of a single patient, for private usage, or the results of different people, are available for the physician who wants to get a complete overview of the patients he/she is treating.

3 Results and Discussion

The proposed device was characterized by applying DPV to known concentrations of the target analyte. When a known voltage bias is forced between WE and RE, the current sourced from the WE pin is measured. In Fig. 3, an example of the V_{bias} signal for the DPV is shown. The fundamental parameters (the pulse interval time t_{pulse} and the recovery interval time t_{int}) are highlighted in the inset. The current is measured at the beginning and the end of each t_{pulse} and the difference between the two values is stored as the DPV cell current at the correspondent V_{bias} level. The maximum of this differential signal is related to the concentration of the target analyte.

The device was tested exploiting different concentration levels of potassium ferri/ferrocyanide redox probe in aqueous 100 mM KCl as the supporting electrolyte. Four concentrations were considered: 0.05, 0.2, 1 and 5 mM. Measurements were performed on disposable carbon screen-printed electrodes from Dropsens (Metrohm) with silver pseudoreference [10]. A conditioning potential ranging from -0.3 to 0.6 V was applied between the WE and RE pins (Fig. 3). After 30 s, when a constant bias voltage is forced for cell preconditioning, the DPV signal is provided to the cell, with 100 ms t_{pulse} and 200 ms t_{int} . The high and low values are progressively increased with 5 mV step between successive periods. This configuration results in a total of 342 measurement points and an overall measurement time of about 1.5 min.

In Fig. 4 examples of current measurement vs. TIA gain resistor change are shown. Figure 4(a) corresponds to the case of a concentration of 5 mM potassium ferri/ferrocyanide, where at the beginning of the acquisition, the TIA gain is set to the maximum value ($1 \text{ V}/\mu\text{A}$) corresponding to a resistor of $1 \text{ M}\Omega$. As the acquisition proceeds, an increasingly higher current is measured at the sensor output, driving the TIA close to the output saturation. The system detects this situation and automatically changes the MUX input in Fig. 2 to progressively decrease the TIA feedback resistance. Instead, in Figure 4(b), the measured differential output current with a two-order smaller concentration (i.e. 0.05 mM) is shown. In this case, given the low peak current observed, the TIA gain is maintained at the maximum value throughout the acquisition. Indeed, setting the initial value to maximum gain ensures that the current is measured with the highest accuracy.

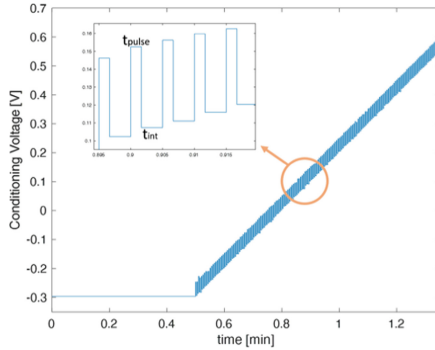


Fig. 3. Conditioning voltage exploited for system validation

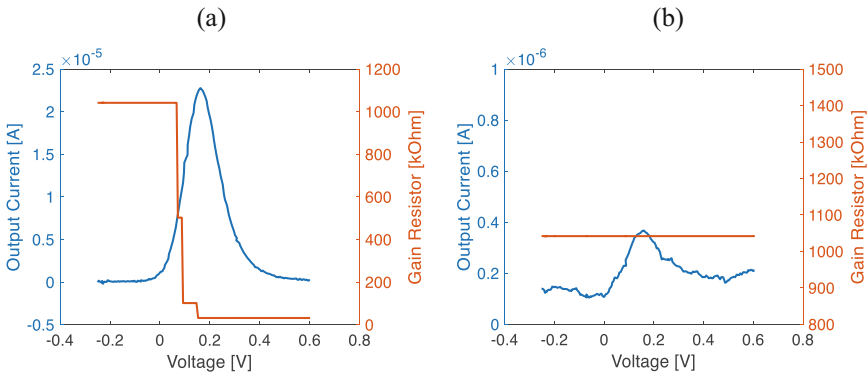


Fig. 4. Current measurement vs Gain Resistor change, a) 5 mM potassium ferri/ferrocyanide b) 0.05 mM potassium ferri/ferrocyanide

Finally, in Fig. 5, the value of the current peaks measured for each sample concentration are reported with the calibration line obtained by measured data fitting. The linear interpolation shows a slope of 4.488 ± 0.211 (with 95% of confidence interval) and an intercept not significantly different from zero. The square of the correlation coefficient R^2 of 0.9998 confirms the good accuracy of the model. This result was compared with others results reported in the literature. In [2] using similar measurement techniques, a value of 0.9976 was found. In [3] an R^2 of 0.995 is reported. Moreover, the authors of [3] compare their results with those obtained with a commercial device finding a value of 0.999. Therefore, the results obtained with the proposed work are very promising and could also allow both semi-quantitative and quantitative analyses with good accuracy.

An evaluation of the device power consumption is reported in [4] where a previous version of the potentiostat is presented. In this paper, the flexibility of the device has been enhanced introducing the automatic configuration of the TIA gain depending on the measured current, through a set of resistances controlled by a MUX having a typical power consumption lower than $1 \mu\text{W}$. In [4] the current gain resistance was manually chosen. Hence, the additional hardware introduced in this new device version has a

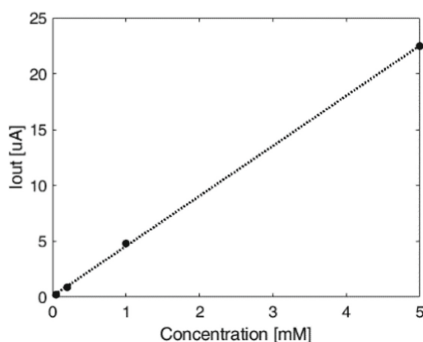


Fig. 5. Linear regression applied to the measurements acquired at different concentrations.

negligible impact on power consumption. Considering two 1.5-V, 2700-mAh standard AA batteries, a battery life of about 3.8 years, can be confirmed [4].

4 Conclusions

In this paper, a smart portable potentiostat exploiting a Wi-Fi connection and based on a cloud service for data storage is presented. Compared to other devices reported in the literature, the proposed solution is characterized by high flexibility and portability, since it does not lean on an additional device (such as a laptop or a smartphone) for data processing, storage, and transmission. Since the Wi-Fi connection is used only to transfer the analysis results, the device is fully operational also in the absence of network coverage. The device characterization shows an accuracy comparable to other works presented in the literature and also to commercial benchtop instruments, proving its feasibility for both semi-quantitative and quantitative analyses.

Funding. This work was supported by the project “Biosensoristica innovativa per i test sierologici e molecolari e nuovi dispositivi PoCT per la diagnosi di infezione da SARS-CoV-2” funded in 2020 by “Bando Straordinario di Ateneo per Progetti di Ricerca Biomedica in Ambito SARS-COV-2 e COVID-19” – University of Parma.

References

1. Bianchi, V., Mattarozzi, M., Giannetto, M., Boni, A., De Munari, I., Careri, M.: A self-calibrating IoT portable electrochemical immunosensor for serum human epididymis protein 4 as a tumor biomarker for ovarian cancer. *Sensors (Switzerland)*, **20**(7) (2020)
2. Adams, S.D., Doeven, E.H., Quayle, K., Kouzani, A.Z.: MiniStat: development and evaluation of a mini-potentiostat for electrochemical measurements. *IEEE Access* **7**, 31903–31912 (2019)
3. Cordova-Huaman, A.V., Jauja-Ccana, V.R., La Rosa-Toro, A.: Low-cost smartphone-controlled potentiostat based on Arduino for teaching electrochemistry fundamentals and applications. *Heliyon* **7**(2), e06259 (2021)

4. Bianchi, V., Boni, A., Fortunati, S., Giannetto, M., Careri, M., De Munari, I.: A Wi-Fi cloud-based portable potentiostat for electrochemical biosensors. *IEEE Trans. Instrum. Meas.* **69**(6), 3232–3240 (2020)
5. CC3200 SimpleLink Wi-Fi Wireless MCU Technical Reference Manual. <https://www.ti.com/product/CC3200>. Accessed 23 Aug 2021
6. Reay, R.J., Kounaves, S.P., Kovacs, G.T.A.: An integrated CMOS potentiostat for miniaturized electroanalytical instrumentation. In: *Proceedings of IEEE International Solid-State Circuits Conference - ISSCC 1994*, pp. 162–163 (1994)
7. Kraver, K.L., et al.: A mixed-signal sensor interface microinstrument. *Sens. Actuators A Phys.* **91**(3), 266–277 (2001)
8. ADG704 - Low Voltage 4 Ω, 4-Channel Multiplexer. Analog Devices, Inc., Norwood (1999)
9. Thingspeak Webpage. <https://www.mathworks.com/products/thingspeak.html>. Accessed 25 May 2021
10. DROPSens screen-printed electrodes web page. https://www.dropsens.com/en/screen_printed_electrodes_pag.html. Accessed 25 May 2021



Experimental Results of Vectorized Posit-Based DNNs on a Real ARM SVE High Performance Computing Machine

Marco Cococcioni¹ , Federico Rossi¹  , Emanuele Ruffaldi² ,
and Sergio Saponara¹ 

¹ Department of Information Engineering, University of Pisa, Pisa, Italy
{marco.cococcioni,sergio.saponara}@unipi.it, federico.rossi@ing.unipi.it

² MMI s.p.a., Calci, Pisa, Italy
emanuele.ruffaldi@mmimicro.com

Abstract. With the pervasiveness of deep neural networks in scenarios that bring real-time requirements, there is the increasing need for optimized arithmetic on high performance architectures. In this paper we adopt two key visions: i) extensive use of vectorization to accelerate computation of deep neural network kernels; ii) adoption of the posit compressed arithmetic in order to reduce the memory transfers between the vector registers and the rest of the memory architecture. Finally, we present our first results on a real hardware implementation of the ARM Scalable Vector Extension.

Keywords: ARM SVE · Vectorization · Alternative representation of reals · Posit arithmetic · HPC

1 Introduction

Nowadays, Deep Neural Networks (DNNs) face new problems and challenges: on one hand, there is a need to reduce network design and computation complexity in order to better accomplish real-time tasks in resource-constrained devices. On the other hand, the trend is to address specific platform accelerators (for example, NVIDIA cuDNN for NVIDIA Graphics Processing Units (GPUs)) to significantly accelerate neural network processing in both the training and inference phases.

DNNs extensively use matrix multiplications, dot products and convolutions, highlighting the need for vectorization routines capable of increasing throughput for these operations. Although the use of GPUs in this field is important, high implementation costs and low-power requirements may prevent such components from being used. Several implementations of vector CPUs are available in most of the common processor architectures: i) Intel/AMD AVX2/SSE for $\times 86$ processors, ii) RISC-V “V” vector extension for the RISC-V architecture, iii) ARM SVE for the ARMv8 architecture. In [1–3] we were able to produce binaries that

employ, respectively, ARM SVE and RISC-V vectorization. In particular, for the ARM SVE platform we were able to enable vectorization in two different tiers: one tier was the auto-vectorization approach that relies on automatic optimization from the compiler (for example, loop unrolling). The second tier involved was the use of intrinsic functions that allowed us to explicitly use ARM SVE instructions in the C++ code.

The idea behind vector extensions is to fit as much data as possible in the vector registers, that acts as very low latency memory for the vector processor. In order to increase the data we can fit in the registers and reduce the memory transfer, it is crucial to minimize the number of bits used to represent the weights of the DNNs. Several alternatives to the standard IEEE 754 32-bit floating point have been proposed: Google (Brain Float—BFloat16—[4]), Intel (Flexpoint—FP16—[5,6]) have already suggested several concepts. BFloat8 [7] is also very interesting, adding the support for stochastic rounding.

The positTM format [8–10] is one of the most promising representations that deviates from the IEEE 754 standard. In machine learning, this kind has been shown to be a great drop-in replacement for 32-bit IEEE 754 floats, using only 16 bits [11–16]. Furthermore, it has been successfully used in low-precision inference down to 8-bit posit representation with minimal network inference accuracy degradation. Moreover, as explained in [12], this number system can be used to create quick, approximated, and efficient activation functions for neural networks such as the sigmoid function by simply using the already existing CPU integer arithmetic operations.

While in [1] we were not able to profile our code on a real hardware (we used the ARM Instruction Emulator for SVE (ARMIE)), in this paper we will instead evaluate the performance of the vectorized extension of the cppPosit C++ posit arithmetic library targeting a real hardware implementation of the ARM SVE architecture.

2 Posit Number and cppPosit Library

Posit numbers are represented by a fixed length format. The overall length (*nbits*) and exponent length (*esbits*) can be modified. The posit format can have a maximum of 4 fields as in Fig. 1. Hereafter we describe the format fields:

- Sign field: 1 bit;
- Regime field: variable length, composed by a string of bits equal to 1 or 0 ended, respectively by a 0 or 1 bit;
- Exponent field: at most *esbits* bits (it can even be absent);
- Fraction field: variable length mantissa (it can even be absent too).

Given a posit(*nbits*, *esbits*), represented in 2’s complement signed integer X and let e and f be exponent and fraction values, the real number r represented by X encoding is:

$$r = \begin{cases} 0, & \text{if } X = 0 \\ \text{NaN}, & \text{if } X = -2^{(nbits-1)} \\ \text{sign}(X) \cdot \text{used}^k \cdot 2^e \cdot (1 + f), & \text{otherwise} \end{cases}$$

Where $useed = 2^{2^{esbits}}$ and k is strictly related to the regime length l and bitstring (b is the bit that composes the string of identical bits, e.g. in 00001 $b = 0$). If $b = 0$ the k is negative, otherwise the k is positive:

$$k = \begin{cases} -l, & \text{if } b = 0 \\ l - 1, & \text{otherwise} \end{cases}$$

In [17] we proved some interesting properties for the configuration ($esbits = 0$). Under this configuration, we could implement fast and approximated versions of common operations. We could evaluate these operations only using the arithmetic-logic unit (ALU) making them faster than the original ones computed using the FPU. These operations are the double and half operators ($2x$ and $x/2$), the inverse operator ($1/x$) and the one’s complement ($1 - x$). In [1] we combined this idea with vectorization, obtaining several posit operations such as ELU and Tanh, exploiting the already existent vector integer operations in the ARM SVE vector environment.

We provide the software support for posit numbers through the `cppPosit` library, developed in Pisa and maintained by the authors of this work. We exploit templatization to configure the posit format. We classify posit operations into four different operational levels, identified with ($\mathcal{L}1$ – $\mathcal{L}4$). Each level has increasing computational complexity (see [17]).

When in levels $\mathcal{L}3$ – $\mathcal{L}4$ we need to use three different back-ends to accelerate posit operations that cannot be directly evaluated via ALU (waiting for proper posit hardware support):

- FPU back-end;
- Fixed back-end, exploiting big-integer support (64 or 128 bits) for operations;
- Tabulated back-end, generating lookup tables for most of the operations (suitable for $\langle [8, 12], * \rangle$ due to table sizes).

3 The Advantages of Vectorized CPUs

The newly introduced ARM SVE is a modern Single Instruction, Multiple Data (SIMD) for the 64-bit ARMv8 instruction set. It is intended as an evolution of the older ARM Neon vector instruction engine. The power of SVE lies in the Vector Length Agnostic (VLA) nature of the engine; indeed, there is no need to specify, at compilation time, the size of the vector registers. This dimension can be retrieved at run-time using a single assembly instruction. This design highly enhance portability of code across different ARM SVE platforms and revisions.

The VLA design is very similar to the one adopted RISC-V vector extension. The main differences between the RISC-V “V” extension and ARM SVE that we believe worth mentioning are:

- Maximum register size: while ARM SVE can only reach 2048-bits, RISC-V “V” can reach up to 16384-bits

- Register grouping: when dealing with different element sizes in the same vector loops, RISC-V can handle the wider element grouping registers so that it can be indexed as it was smaller (e.g. if we want to convert a vector of 16-bit posits to a vector of 32-bit floats).

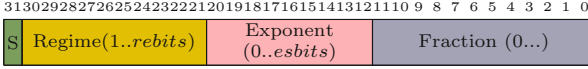


Fig. 1. Illustration of a posit(32, 9) data type. Both the exponent and the fraction field can be absent, for specific configurations having a regime field particularly lengthy.

3.1 Vectorized CPUs and Deep Neural Networks

The most recurrent computations in deep neural networks are [1, 2, 17]:

- GEMM (general matrix-matrix multiplication) (training phase)
- matrix-vector multiplication (inference phase)
- matrix-matrix convolution product (both training and inference phases)
- vector-vector dot product (both training and inference phases, for computing the activations)
- non-linear activation function (both training and inference phases), computed over a vector of activations.

In this work we have used posit, since they allow to save memory for storing the weights. Moreover we and other authors have proved that posit16 is as accurate as 32-bit IEEE Floats for machine learning applications. In machine learning application even a posit8 can be accurate enough compared to 32-bit floats, thus saving $4\times$ storage space (both on disk and, more importantly, on RAM and caches) with minimal accuracy loss.

4 Test-Bench, Methodologies and Benchmarks

In [1] we tested ARM SVE capabilities using the ARM Instruction Emulator. We ran the emulator on a HiSilicon Hi1616 CPU with 32@2.4 GHz ARMv8 Cortex-A72 cores. This emulator was able to trap all the illegal instruction interruptions coming from the execution of binaries compiled using the SVE instruction set extension. These instructions were then executed via software by the ARM Instruction Emulator. During *emulation* we were able to modify the vector register length from 128-bit to 2048-bit.

Instead, in this work we were able to use an actual hardware implementation of the ARM SVE architecture using the HPE Apollo80 machine available at University of Pisa. The Apollo80 is based on the ARMv8 A64FX core [18], the first commercial implementation of the ARM SVE architecture. In particular, the ARMv8 A64FX is the first processor to support the full feature set of the ARM SVE architecture without emulating any instruction.

This platform is particularly interesting since it will be employed in the European Processor Initiative framework [19], and it will be used as a base computing platform for the EUPEX and TEXTAROSSA EuroHPC projects.

In detail, this platform consists of 4 different blades equipped with 48 ARMv8 A64FX Cores with 512-bit vector registers, running, respectively, at 1.8 and 2.0 GHz. Each blade has access to 32 GB of High Bandwidth Memory.

In order to evaluate SVE-related performance on this machine, we used following benchmarks: i) vectorized activation functions only using posits and integer vector instructions, ii) vectorized matrix-matrix multiplication and convolution.

Moreover, we employed posit numbers to compress and decompress data across the kernels, in order to reduce memory transfers by a factor 4 (with $\text{posit}(8, 0)$) or 2 (with $\text{posit}(16, 0)$). Also compression and decompression phases were implemented using vectorization, exploiting vector integer arithmetic.

Benchmarks were compiled using the `armclang++ 20.3` compiler, based on LLVM 9.0.1 and executed on the aforementioned Apollo80 machine, running CentOS Linux release 7.9.2009.

5 Experimental Results and Discussion

Figure 2 shows the performance of the activation function benchmarks on the Apollo80 machine. The benchmarks consisted in the computation of sigmoid and extended linear unit (ELU) on 4096-wide vectors (even if the hardware only supports 512-bit). Each computation was repeated 100 times and the average computation time was reported.

As reported, the computation of the two activation functions using $\text{posit}(8, 0)$ benefits from the reduction in size of the format. This is because most of the steps of the activation function computation is performed using `int8_t` for $\text{posit}(8, 0)$ and `int16_t` for $\text{posit}(16, 0)$.

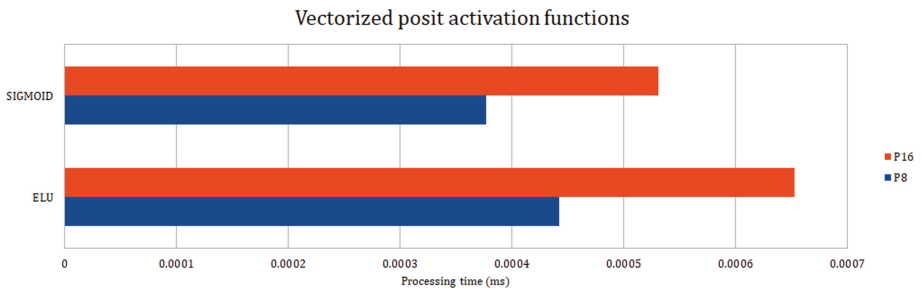


Fig. 2. Processing time of vectorized activation functions on a 4096-element vector with a 512-bit vector register length.

Figure 3 shows the performance of the kernel benchmarks, using $\text{posit}(8, 0)$ and $\text{posit}(16, 0)$. These benchmarks consisted in the computation of: i) dot product between vector of 4096 elements, ii) convolution on a 128×128 image with a

3×3 filter, iii) matrix-matrix multiplication between square matrices of 128×128 elements.

As reported, the benefit in reducing the information size is not as effective as in the previous case. The issue is that in this case, we could not use posits in every step of computation (of course ARM SVE lacks dedicated posit instructions). This means that we used posits just for data compression and decompression at the start and at the end of the kernels. For example, in the convolution kernel, we decompress the posit input to float using our vectorized routine, then we compute the convolution using native vectorized float support from ARM SVE and finally we compress the result back to posits.

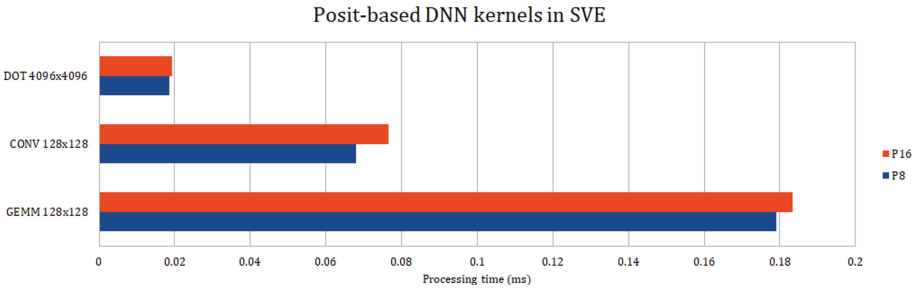


Fig. 3. Processing time of vectorized DNN kernels with a 512-bit vector register length (DOT: vector-vector dot product, CONV: matrix-matrix convolution, GEMM: General matrix matrix multiplication).

Figures 4 and 5 show the measured speedup from the emulated machine to the real hardware implementation. The speedup is computed as $t_{\text{HPe80}}/t_{\text{ARMIE}}$. Since we already proved that we can get better timing performance using $\text{posit}\langle 8, 0 \rangle$ instead of $\text{posit}\langle 16, 0 \rangle$, we reported the speedup relative to $\text{posit}\langle 8, 0 \rangle$ computations. As reported, the speedup spans from at least ~ 11 , in the case of the GEMM function, up to ~ 1500 in the case of the ELU function.

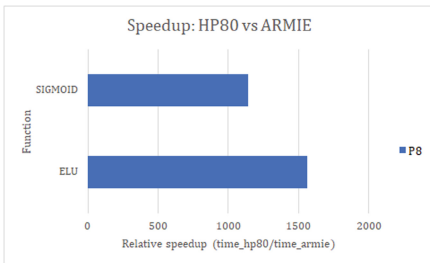


Fig. 4. Relative speedup of activation functions with 512-bit vector register length.

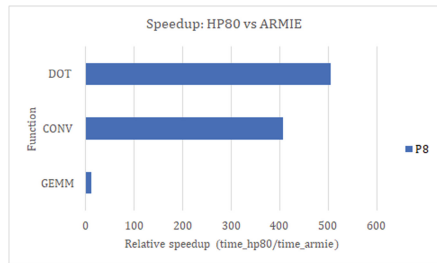


Fig. 5. Relative speedup of DNN kernels with a 512-bit vector register length.

6 Conclusions

In a previous work, we designed posit-based and vectorized operations on an ARM 64bit SVE emulator. The operations considered are the most time consuming ones in deep neural networks. In the present work we compared the impact of vectorization on a real machine. By using such machine, we were able to assess the true speedup due to vectorization, which turned out to be remarkable (with a speedup factor from $11\times$ to $1500\times$, depending on the executed task). Future works will involve the combination of presented algorithms with MPI, to enable multi-processor or multi-node computation of deep neural networks (e.g. exploiting all the blades and cores of the HP80).

Acknowledgments. Work partially supported by H2020 projects (EPI grant no. 826647, <https://www.european-processor-initiative.eu> and TEXTAROSSA grant no. 956831, <https://textarossa.eu>) and partially by the Italian Ministry of Education and Research (MIUR) in the framework of the CrossLab project (Departments of Excellence). We thank the personnel of the Green DataCenter of the University of Pisa (<https://start.unipi.it/en/computingunipi>). In particular, we thank Prof. P. Ferragina, Dr. M. Davini and Dr S. Suin, for having provided us with the computational resources that have been used in the experimental section.

References

1. Cococcioni, M., Rossi, F., Ruffaldi, E., Saponara, S.: Fast deep neural networks for image processing using posits and ARM scalable vector extension. *J. Real-Time Image Process.* **17**(3), 759–771 (2020). <https://doi.org/10.1007/s11554-020-00984-x>
2. Cococcioni, M., Rossi, F., Ruffaldi, E., Saponara, S.: Vectorizing posit operations on RISC-V for faster deep neural networks: experiments and comparison with ARM SVE. *J. Neural Comput. Appl.* **33**, 575–585 (2021). <https://doi.org/10.1007/s00521-021-05814-0>
3. Cococcioni, M., Rossi, F., Ruffaldi, E., Saponara, S.: Faster deep neural network image processing by using vectorized posit operations on a RISC-V processor, In: *Real-Time Image Processing and Deep Learning 2021*, Kehtarnavaz, N., Carlsohn, M.F. (Eds.), International Society for Optics and Photonics. SPIE, vol. 11736, pp. 19–25 (2021). <https://doi.org/10.1117/12.2586565>
4. Burgess, N., Milanovic, J., Stephens, N., Monachopoulos, K., Mansell, D.: Bfloat16 processing for neural networks. In: *2019 IEEE 26th Symposium on Computer Arithmetic (ARITH)*, pp. 88–91 (2019)
5. Koster, U., et al.: Flexpoint: an adaptive numerical format for efficient training of deep neural networks. In: *Proceedings of the 31st Conference on Neural Information Processing Systems (NIPS 2017)* (2017)
6. Popescu, V., Nassar, M., Wang, X., Tumer, E., Webb, T.: Flexpoint: predictive numerics for deep learning. In: *Proceedings of the 25th IEEE Symposium on Computer Arithmetic (ARITH 2018)*, pp. 1–4 (2018)
7. Mellempudi, N., Srinivasan, S., Das, D., Kaul, B.: Mixed precision training with 8-bit floating point (2019)

8. Gustafson, J.L.: *The End of Error: Unum Computing*. Chapman and Hall/CRC (2015)
9. Gustafson, J.L.: A radical approach to computation with real numbers. *Supercomput. Front. Innov.* **3**(2), 38–53 (2016)
10. Gustafson, J.L., Yonemoto, I.T.: Beating floating point at its own game: posit arithmetic. *Supercomput. Front. Innov.* **4**(2), 71–86 (2017)
11. Cococcioni, M., Rossi, F., Ruffaldi, E., Saponara, S.: Novel arithmetics to accelerate machine learning classifiers in autonomous driving applications. In: *Proceedings of the 26th IEEE International Conference on Electronics Circuits and Systems (ICECS 2019)*
12. Cococcioni, M., Rossi, F., Ruffaldi, E., Saponara, S.: A fast approximation of the hyperbolic tangent when using posit numbers and its application to deep neural networks. In: Saponara, S., De Gloria, A. (eds.) *ApplePies 2019*. LNEE, vol. 627, pp. 213–221. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-37277-4_25
13. Cococcioni, M., Ruffaldi, E., Saponara, S.: Exploiting posit arithmetic for deep neural networks in autonomous driving applications. In: *2018 International Conference of Electrical and Electronic Technologies for Automotive*, pp. 1–6. IEEE (2018)
14. Carmichael, Z., Langroudi, H.F., Khazanov, C., Lillie, J., Gustafson, J.L., Kudithipudi, D.: Conference exhibition (DATE), pp. 1421–1426. IEEE (2019)
15. Langroudi, H.F., Carmichael, Z., Gustafson, J.L., Kudithipudi, D.: Positnn framework: tapered precision deep learning inference for the edge. In: *2019 IEEE Space Computing Conference (SCC)*, pp. 53–59. IEEE (2019)
16. Cococcioni, M., Rossi, F., Ruffaldi, E., Saponara, S., de Dinechin, B.D.: Novel arithmetics in deep neural networks signal processing for autonomous driving: challenges and opportunities. *IEEE Signal Processing Magazine.* **24**, 38(1), 97–110 (2020)
17. Cococcioni, M., Rossi, F., Ruffaldi, E., Saponara, S.: Fast approximations of activation functions in deep neural networks when using posit arithmetic, *Sensors*, **20**(5) (2020). www.mdpi.com/1424-8220/20/5/1515
18. Fujitsu Processor A64FX. www.fujitsu.com/global/products/computing/servers/supercomputer/a64fx/ Accessed 4 June (2021)
19. European Processor Initiative, an H2020 project. www.european-processor-initiative.eu/ (2019)



An Open-Source Hardware/Software Architecture for Remote Control of SoC-FPGA Based Systems

Werner Florian^{1,2(✉)}, Bruno Valinoti^{1,2,3}, Luis G. García^{1,2}, Marcos Cervetto³, Edgardo Marchi³, Maria Liz Crespo¹, Sergio Carrato², and Andres Cicuttin¹

¹ The Abdus Salam International Centre for Theoretical Physics—MLAB,
Strada Costiera, 11, 34151 Trieste, TS, Italy

{wflorian,bvalinot,lgarcia1,mcrespo,cicuttin}@ictp.it

² Università degli Studi di Trieste—Dipartimento di Ingegneria e Architettura,
v. Valerio 6/1, 34127 Trieste, TS, Italy
carrato@units.it

³ Instituto Nacional de Tecnología Industrial, Av. Gral. Paz 5445, 1650 San Martín,
BA, Argentina

Abstract. We present an open hardware/software architecture for remote control of Field Programmable Gate Array (FPGA) based Systems on Chip (SoC). These systems, which integrate embedded processors, FPGA fabric, memory blocks and other resources, usually need to be controlled from a computer. The proposed architecture comprises a set of commands, instructions for data movement, and standardized data packets. A minimal set of specifications and design guidelines will effectively separate hardware and software developments granting compatibility to the different subsystems. A simple architectural approach ensures compatibility of computer resident software, embedded processor software, and FPGA designs. The implicit structured design methodology associated with the proposed architecture facilitates remote control as well as maintenance, debugging, and portability among SoC-FPGA vendors. We describe a concrete implementation in order to show how data and instructions can be moved across the whole system.

Keywords: Hardware-software codesign · System-on-chip · Embedded software · FPGA design · Real-time systems · Reconfigurable virtual instrumentation · Data acquisition systems

1 Introduction

Modern complex electronic devices are characterized by the growing integration of different functional units such as multicore microprocessors (μP), random access memory blocks (BRAM), digital signal processors, and FPGA fabrics. The integration of these units in the same chip includes a high degree of internal

interconnection. Systems based on SoC-FPGA provide a great number of computational services, low latency responses and high throughput online data processing, making them very attractive, if not mandatory, for advanced instrumentation and specialized supercomputing infrastructures [1,2]. The huge amount of reconfigurable logic and computational resources allow the implementation of very complex systems, which often require remote control from a computer [3-5] through a standard communication link such as USB or Ethernet.

Despite the complexity of these heterogeneous systems, they can be described in a simple unified way by means of an abstract model [6]. It is then possible to implement a suitable hardware/software architecture along with a set of services and remote control activities.

Although there are some commercial solutions for remote control of systems based on FPGA [7,8], they have severe limitations such as closed sources, limited functionality, and no portability. Furthermore, most of the time SoC-FPGA based system developers implement their custom procedures for remote control. Both commercial and custom solutions are typically incompatible among them due to lack of standards. In order to cope with these problems we propose an open source hardware/software architecture for remote control of systems based on SoC-FPGA.

2 Hardware/Software Architecture

The proposed architecture follows a simple modular approach that foresees the encapsulation of certain aspects that are specific of the target devices, e.g., adopted standard communication links, operating systems, and external hardware.

This architecture requires the whole system to be abstractly represented by a set of functional blocks interconnected via ports associated with a unique address in a global memory map, divided into non-overlapping domains. The activity of such a system can be described by a set of concurrent data exchanges among its functional blocks. The movement of data from one place to another of the global system is described by a special instruction called *Universal Direct Memory Access (UDMA)*.

A generic UDMA instruction can then be expressed as follows:

$$\text{UDMA } \langle \text{src_addr} \rangle \langle \text{dst_addr} \rangle \langle \text{src_inc} \rangle \langle \text{dst_inc} \rangle \langle N \rangle$$

where the associated parameters are respectively: the addresses of the source and destination, the address step increments at the source and destination, and the number of 32-bit words to be transferred.

Even though a global memory map abstracts some implementation details showing addresses contiguously, in general the corresponding functional blocks may be physically distant, belonging to different hardware components.

Special entities are in charge of executing the data movements. Each of these entities, called Local Resource Agent (LRA), has direct and exclusive access to one memory domain of the global map and can modify its content. All LRAs

are interconnected, have a unique ID, and exchange data among them by means of packets. Inside every LRA, a UDMA processor is in charge of executing the UDMA instructions.

In Fig. 1 the common packet structure is shown.

The header contains a common keyword that announces the start of packet for asynchronous communication, the protocol number, the packet type, the priority, and the source and destination IDs of the LRAs.

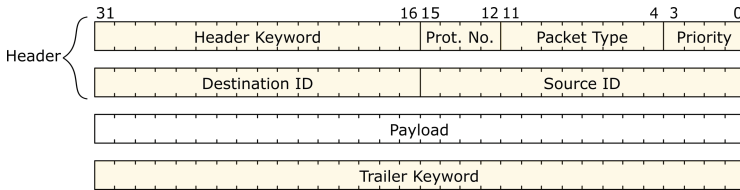


Fig. 1. Common packet structure.

Three essential packet types were identified for the basic operation of the system, as described below:

- *Command Packet*: It consists of a single word containing a code associated with a predefined activity (START, STOP, RESET, etc.) or error messages. It has a reduced size allowing a faster transmission and lower latency.
- *Raw data Packet*: This is the packet used for moving data among LRAs. It contains the data to be written and the destination-related part of a UDMA instruction. Given that the data may require more than one packet, indexing may be used to keep an order throughout the multiple transactions required to complete a data exchange. A data integrity check such a CRC and checksum is implemented for these packets.
- *UDMA Packet*: This packet contains a UDMA instruction which is passed to the UDMA processor inside the LRA. Depending on the source and destination, the instruction might trigger a cross-LRA exchange that will require a single or multiple raw data packets.

3 Single SoC-FPGA Based System

We consider a typical heterogeneous system [9] based on a single SoC-FPGA device connected on one side to a standard PC for remote control and user interface, and on the other side to external hardware, which in general will be specific to the application [2]. The FPGA, the μ P and the PC offer different but complementary computational resources. Maximum performance can be achieved if the whole computational activity is distributed among these three subsystems taking into account their specific characteristics.

Figure 2 schematically shows a typical SoC-FPGA system with its control PC.

Following a modular approach, the SoC-FPGA can be seen as the combination of a μP and a FPGA, interconnected by vendor specific SoC bus.

The FPGA is subdivided into three main modules: (i) an external hardware controller, (ii) a Communication Block (ComBlock) [10], and (iii) the core FPGA design. These modules should have standardized interfaces to facilitate their interconnection and the communication among them. For the internal ports of the modules, the Wishbone (WB) standard bus interface [11] is proposed; the ComBlock is used to interact with the μP , abstracting the SoC bus complexities.

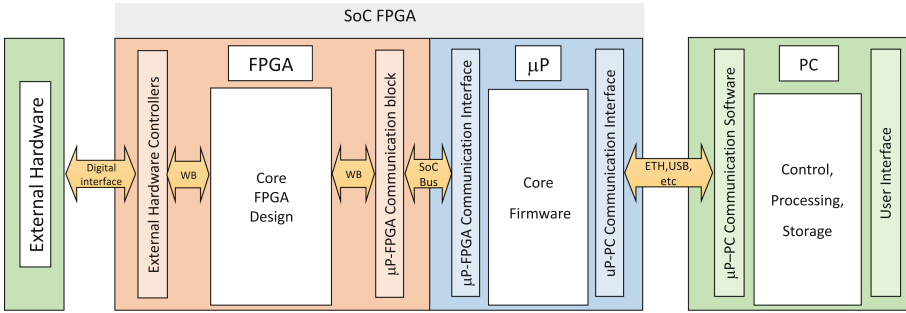


Fig. 2. Block diagram of a typical system based on SoC-FPGA.

Similarly, the implementation of the μP software separates the μP core program from the communication services with the PC and the FPGA. The μP relies on the UDMA firmware [12] for the communication with the PC and the FPGA. The PC resident software consists of a Python based Command Line Interpreter (CLI) to manage the communication and the interaction between the PC and the μP . The control software benefits from Python by relying on its scriptability and wide cross-platform support.

3.1 Implementation Example

To illustrate the proposed architecture we implemented a demonstrative system to move data among different components according to instructions generated in the PC. Figure 3 shows a simplified block diagram of this system composed by two LRAs: the PC and the SoC-FPGA.

Three type of memories were implemented in the FPGA: one BRAM, two FIFOs, and the True Dual Port Ram (TDPRAM) of the ComBlock. These memories along with the Python UDMA CLI assigned memory defined the global memory map. The PC and the μP communicated via packets over TCP-IP. Inside the FPGA, a *WB Interconnect* was used to interface the UDMA processor with the FPGA memory resources. The ComBlock registers were used by the

UDMA firmware to control the state of the UDMA processor. When a packet arrived from the CLI, the μP interpreted the header and extracted the payload. According to the type of the packet and its content, the μP performed different operations. It executed a command or a UDMA instruction, or it passed a UDMA instruction to the FPGA through the ComBlock should the affected memory domain lay on the FPGA subsystem. In this last case, the UDMA processor executed the instruction to move the data as prescribed and, once finished, it used the reserved registers to communicate to the UDMA firmware that the operation was successfully completed.

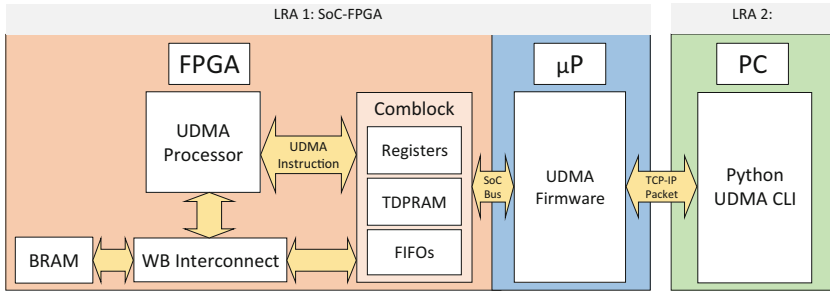


Fig. 3. Block diagram of the implemented system showing LRAs and inner blocks.

A test application was developed on the implemented system. The test consisted of writing data in the BRAM from the PC, and then verifying the written data. This was done by first sending a data packet from the UDMA CLI to the μP . The μP wrote the data in the ComBlock's FIFO and sent a UDMA instruction. Next, the UDMA processor interpreted the UDMA instruction and moved the data from the ComBlock's FIFO to the BRAM. To verify the success of the operation, a UDMA instruction was sent from the UDMA CLI to retrieve the data. The UDMA processor moved the data from the BRAM to the ComBlock's FIFO. Finally, the data was sent to the PC in a data packet by the UDMA firmware. With these mechanisms it was possible to arbitrarily move data between the instantiated memory resources.

The system has been successfully tested in two different FPGA based SoC development boards: the ZedBoard [13] and the CIAA-ACC [14]. The FPGA resources utilization of the *WB Interconnect*, the UDMA processor, and the ComBlock are shown in Table 1 for the CIAA.

The system has been developed to be multi-platform and communication protocol independent following the proposed architecture. Due to the UDMA processor encapsulation, the user can easily add more resources to the global memory map by just connecting them to the *WB Interconnect*.

Table 1. Resource utilization on a CIAA^a (less than 1% of the total)

	LUT	LUTRAM	Flip-flops	Slices	BRAM tiles
WB interconnect	91	0	4	0	0
UDMA processor	193	0	506	87	1
Comblock	221	48	510	126	1

^a The utilization on the Zedboard was practically identical.

4 Conclusions

The proposed hardware/software architecture has shown to be an effective solution for the remote control and debugging of systems based on SoC-FPGA devices. A reduced number of commands and instructions allows moving data across the entire system involving memory elements, microprocessors, re-configurable functional blocks, and a standard computer for remote control.

The architecture modular structure also facilitates the porting of complex designs among different SoC-FPGA vendors and device families. The open source approach enables FPGA designers and embedded software programmers to benefit from the freely available IP blocks and software routines to implement their designs, saving valuable time in dealing with and debugging complex communication mechanisms such as those involving Ethernet connections and SoC-Buses.

The proposed approach seems to be suitable not only for advanced instrumentation but also for high performance computing based on multiple interconnected SoC-FPGA based platforms. The presented architecture is appropriate to efficiently exploit scalable platforms such as clusters of SoC-FPGAs.

References

1. Cicuttin, A., Crespo, M.L., Mannatunga, K.S., Samarawickrama, J.G., Abdallah, N., Sabet, P.B.: HyperFPGA: a possible general purpose reconfigurable hardware for custom supercomputing. In: 2016 International Conference on Advances in Electrical, Electronic and Systems Engineering (2016)
2. Gazzano, J., Crespo, M., Cicuttin, A., Calle, F.: Field-Programmable Gate Array (FPGA) Technologies for High Performance Instrumentation. Advances in Computer and Electrical Engineering. IGI Global (2016). ISBN:9781522502999
3. Cicuttin, A., Crespo, M.L., Mannatunga, K.S., et al.: A programmable system-on-chip based digital pulse processing for high resolution x-ray spectroscopy. In: 2016 International Conference on Advances in Electrical, Electronic and Systems Engineering, pp. 520–525 (2016)
4. Mannatunga, K.S., Ali, S.H.M., Crespo, M.L., Cicuttin, A., Samarawickrama, J.: High performance 128-channel acquisition system for electrophysiological signals. IEEE Access **8**, 366–383 (2020)
5. Velmurugan, S., Rajasekaran, C.: A reconfigurable on-chip multichannel data acquisition and processing (DAQP) system for multichannel signal processing. In: 2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering, pp. 109–114 (2013)

6. Mannatunga, K.S., et al.: Design for portability of reconfigurable virtual instrumentation. In: 2019 X Southern Conference on Programmable Logic (SPL), pp. 45–52 (2019)
7. National Instruments. CompactRIO Systems. <https://www.ni.com/it-it/shop/compactrio.html> (2020)
8. Opal Kelly. FrontPanel. <https://opalkelly.com/products/frontpanel/>
9. Crespo, M.L., Cicuttin, A., Gazzano, J., Calle, F.: Reconfigurable virtual instrumentation based on FPGA for science and high-education. In: Fagerberg, J., Mowery, D.C., Nelson, R.R. (eds.) Field-Programmable Gate Array (FPGA) Technologies for High Performance Instrumentation, chap. 5, pp. 99–123. IGI Global (2016)
10. ICTP MLAB and INTI CMNT. The Core Comblock. <https://gitlab.com/rodrigomelo9/core-comblock> (2021)
11. Free and Open Source Silicon Foundation. WISHBONE System-on-Chip (SoC) Interconnection Architecture for Portable IP Cores (2010)
12. ICTP MLAB. Universal Direct Memory Access - UDMA. <https://gitlab.com/brunovali/udma> (2021)
13. AVNET. ZedBoard. <https://www.avnet.com/wps/portal/us/products/avnet-boards/avnet-board-families/zedboard/>
14. INTI. CIAA-AAC A Open Hardware Card for HPC and Industrial Applications. https://github.com/ciaa/CIAA_ACC_Support



A Self Referencing Technique for the RC-pLMS Adaptive Beamformer and Its Hardware Implementation

Ghattas Akkad¹(✉), Ali Mansour¹, Bachar ElHassan², Elie Inaty³,
and Rafic Ayoubi³

¹ ENSTA Bretagne, Lab-STICC, UMR, 6285 Brest, France
{Ghattas.akkad,Ali.Mansour}@ensta-bretagne.fr

² Faculty of Engineering, Lebanese University, Tripoli, Lebanon

³ Department of Computer Engineering, University of Balamand, Koura, Lebanon
{Elie.Inaty,Rafic.Ayoubi}@balamand.edu.lb

Abstract. In this paper, we propose a self referencing scheme for the reduced complexity parallel least mean square (RC-pLMS) adaptive beamforming algorithm as means of robustness against possible interruptions in the reference signal and its hardware implementation. The RC-pLMS is a single stage, non-blind, least mean square (LMS) algorithm with modified input vectors formed as a linear combination of the current and the previous input sample. In this context, its convergence and its stability are critically dependent on the availability of its reference signal and are known to severally degrade when discontinued. Thus, for robustness against the pre-mentioned and with respect to the RC-pLMS accelerated convergence and low residual error profile, we propose the use of its filtered output, as an alternative learning sequence, whenever the original reference signal is discontinued, i.e. self-referencing. The proposed self referencing approach is evaluated in infinite and finite precision modes on software and on hardware, i.e. Field Programmable Gate Array (FPGA), respectively. Hardware and software simulation validates the RC-pLMS robustness against different reference signal obstruction scenarios, through the use of the proposed self-referencing approach, while maintaining an accelerated convergence behavior, a low complexity architecture and a high precision beam pointing accuracy.

Keywords: LMS · Parallel LMS · Robust adaptive beamforming · Self referencing · RC-pLMS · Digital communication · FPGA

1 Introduction

The unprecedented increase in wireless connected devices, has tightened the constraints on popular adaptive algorithms, such as the Least Mean Square (LMS), the Recursive Least Square (RLS) and their variants [1–9], when targeting beamforming applications [10–12]. Such constraints are reflected by the requirement

of a robust system with an accelerated convergence rate, a high precision beam pointing accuracy and a reduced complexity structure suitable for a hardware implementation [13–16]. Recently, two multi stage adaptive beamforming algorithms has been proposed to eliminate the tradeoff between the LMS convergence speed and its steady state error [13, 17]. These algorithms are the parallel LMS (pLMS) [13] and the reduced complexity parallel LMS (RC-pLMS) [18] and are formed of an LMS/LMS connected by an error feedback and a LMS with modified input vectors, respectively.

However simple and effective, the RC-pLMS is still a non blind adaptive algorithm and its convergence and stability is highly dependent on the presence of an uninterrupted reference signal.

Therefore, our contribution in this paper is summarized as follows: we propose a self referencing technique for the RC-pLMS with the use of its filtered output signal, with respect to its accelerated convergence profile, whenever the original learning sequence is interrupted. Furthermore, we present its hardware architecture and implement it on FPGA to evaluate its computational complexity and performance in finite precision arithmetics.

2 Mathematical Review

In this section, we present a basic review on the RC-pLMS beamformer for a uniform linear antenna (ULA) array of N equally spaced elements and narrow-band complex signals [18]. Let the input vector $\mathbf{x}(k) = [x_1(k), x_2(k), \dots, x_N(k)]^T$ impinging the array from the far field [19], at the discrete time instant k defined by

$$\mathbf{x}(k) = \mathbf{a}_d s_d(k) + \sum_{l=0}^{N-1} \mathbf{a}_{i,l} \mathbf{i}_l(k) + \mathbf{n}(k) \quad (1)$$

where $[\cdot]^T$ represents the matrix transpose, $s_d(k)$ and $\mathbf{i}_l(k)$ are the desired and interfering signals, with $l < N$, \mathbf{a}_d and $\mathbf{a}_{i,l}$ are the $N \times 1$ complex array steering vector for the desired and l^{th} interference signal, respectively, and $\mathbf{n}(k)$ is the complex additive white Gaussian noise (CAWGN) vector [13, 20]. A general form of \mathbf{a} is given by

$$\mathbf{a} = [1, e^{-j2\pi \frac{B \sin(\theta)}{\lambda}}, \dots, e^{-j(N-1)2\pi \frac{B \sin(\theta)}{\lambda}}]^T \quad (2)$$

where θ is the angle of arrival, B is the distance between two consecutive antenna elements, and λ is the signals wavelength [18]. The output of the beamformer subject to a linear combiner [1, 18] is given by

$$y(k) = \mathbf{w}^{\mathbf{H}}(k) \mathbf{x}(k) \quad (3)$$

where $[\cdot]^{\mathbf{H}}$ represents the matrix Hermitian transpose, i.e. transpose and complex conjugation, and $\mathbf{w}(k)$ is the array weight vector.

2.1 RC-pLMS Algorithm

RC-pLMS is formed by one LMS stage whose inputs are obtained as a linear combination of the present and previous sample [18], as shown in Fig. 1. Where the block jz^{-1} represents a multiplication by the imaginary number $j = \sqrt{-1}$ and a one sample delay [18], $\mathbf{u}(k)$ and $S(k)$ are the new modified inputs [18] and e_{RCpLMS} is the overall error, such as

$$\mathbf{u}(k) = \mathbf{x}(k) - j\mathbf{x}(k-1) \quad (4)$$

$$S(k) = d(k) - jd(k-1) \quad (5)$$

$$e_{RCpLMS}(k) = S(k) - \mathbf{w}^H(k)\mathbf{u}(k) \quad (6)$$

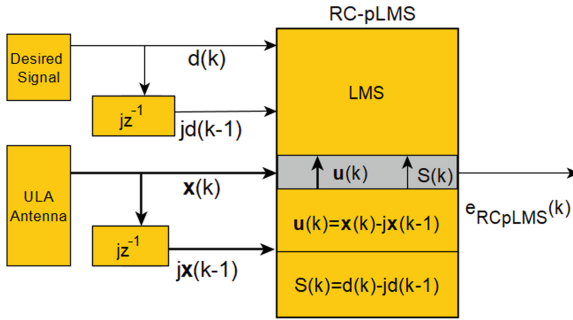


Fig. 1. RC-pLMS block diagram

where $d(k)$ is the original reference signal and $y_r(k)$ is the filtered output signal. The filtered output signal and the weight update equation are given as

$$y_r(k) = \mathbf{w}^H(k)\mathbf{u}(k) \quad (7)$$

$$\mathbf{w}(k+1) = \mathbf{w}(k) + \mu e_{RCpLMS}^*(k)\mathbf{x}(k) \quad (8)$$

where μ is the RC-pLMS step size. The RC-pLMS mean square error (MSE) cost function [18] becomes

$$\begin{aligned} \xi_{RCpLMS} &= E[|e_{RCpLMS}(k)|^2] \\ &= E[|S(k)|^2] - \mathbf{z}^H\mathbf{w}(k) - \mathbf{w}^H(k)\mathbf{z} \\ &\quad + \mathbf{w}^H(k)\mathbf{Q}\mathbf{w}(k) \end{aligned} \quad (9)$$

where $|\cdot|$ denotes the complex modulus, $E[\cdot]$ is the expectation operator, $\mathbf{Q} = \mathbf{Q}(0)$ is the input signal auto-correlation matrix and $\mathbf{z} = \mathbf{z}(0)$ is the cross correlation vector formed by the input $\mathbf{u}(k)$ and desired signal $S(k)$, assuming the process is wide sense stationary (WSS) and all the signals are zero mean. $\mathbf{Q}(0)$ and $\mathbf{z}(0)$ are defined at lag $\tau = 0$ as

$$\mathbf{Q}(\tau) = E[\mathbf{u}(k-\tau)\mathbf{u}^H(k)] \quad (10)$$

$$\mathbf{z}(\tau) = E[S^*(k-\tau)\mathbf{u}(k)] \quad (11)$$

where $*$ denotes complex conjugation, the lag $\tau = k_1 - k_2$ and k_1, k_2 are different time instances from which an observation of the random process is taken.

The RC-pLMS optimal weight vector, at convergence, \mathbf{w}_{op} , [18] is given as

$$\mathbf{w}_{op} = \mathbf{Q}^{-1} \mathbf{z}^H \quad (12)$$

assuming the auto-correlation matrix \mathbf{Q} is invertible.

3 RC-pLMS with Self-referencing

Despite the RC-pLMS superior performance [18], compared to other LMS and RLS variants, it still lacks robustness against interruptions in the reference, learning, signal. Thus, for robustness against the pre-mentioned and with respect to the RC-pLMS accelerated convergence and low steady state error profile [18], we propose the use of its filtered output, $y_r(k)$, as an alternative learning sequence, whenever the original reference signal, $d(k)$, is discontinued, i.e. self-referencing.

Proposition: Assuming convergence, as $k \rightarrow \infty$, we can approximate $y_r(k) \approx d(k)$ and thus, from (5), the approximate reference signal, $S_y(k)$, becomes

$$S_y(k) = y_r(k) - jy_r(k-1) \quad (13)$$

Proof: Assuming convergence, as $k \rightarrow \infty$, and assuming that the environment is WSS and the step size μ is small enough [1, 18], we can approximate $\mathbf{w}(k) \approx \mathbf{w}(k-1) \approx \mathbf{w}_{op}$, hence the filtered output signal, $y_r(k)$, tend to approach $d(k)$ with both interference and noise signals being suppressed [1, 18]. As such, with ϵ as the approximation error and using (13), we can write

$$\hat{S}(k) \approx S_y(k) + \epsilon \quad (14)$$

where $\hat{S}(k)$ is the estimate of $S(k)$. In this context, with respect to (12), the RC-pLMS optimal weight vector can be rewritten as

$$\begin{aligned} \mathbf{w}_{op} &\approx \mathbf{Q}^{-1} \mathbf{E}[\mathbf{u}^H(k)(S_y(k) + \epsilon)] \\ &\approx \mathbf{Q}^{-1} \mathbf{E}[\mathbf{u}^H(k)S_y(k)] + \mathbf{Q}^{-1} \mathbf{E}[\mathbf{u}^H(k)\epsilon] \end{aligned} \quad (15)$$

From (15), it is clear that the RC-pLMS converges to an optimal weight in function of $S_y(k)$ and the approximation error ϵ . As $k \rightarrow \infty$, $\epsilon \rightarrow 0$ and the right hand term of (15) can be omitted. Therefore, we can approximate $y_r(k) \approx d(k)$ at the cost of a larger residual error.

The validity of the proposed technique will be demonstrated experimentally in Sect. 5.

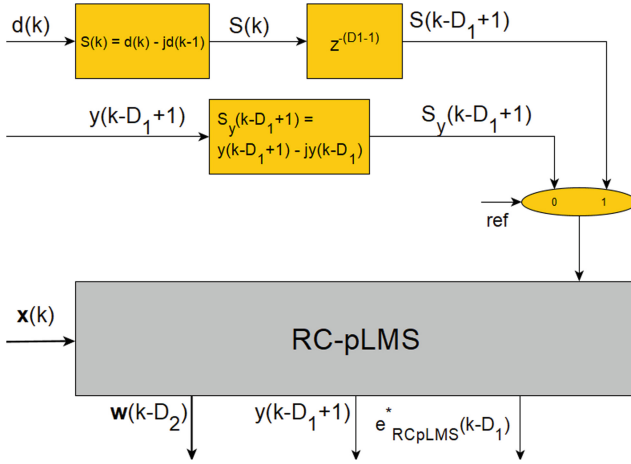


Fig. 2. RC-pLMS-SR top level architecture

4 Hardware Implementation

In this section, we propose a pipeline parallel hardware architecture for the RC-pLMS with self referencing (RC-pLMS-SR). The architecture is then implemented on FPGA in order to evaluate the performance of the proposed self referencing scheme in finite precision arithmetic and highlight its computational simplicity.

While the RC-pLMS is formed from a classical LMS its pipelining remains difficult due to the error feedback in the weight update equation, as shown in (8) [18,21]. As such, assuming the process is WSS and the step size μ is small enough, we make use of the delay and sum relaxed look ahead technique [18,21]. The RC-pLMS Eqs. (6) and (8) are now modified with a delay relaxation of D_1 samples in the error path, D_2 samples in the weight update path and a relaxation of D_3 terms in the look-ahead averaging sum to obtain

$$\mathbf{w}(k+1) = \mathbf{w}(k - D_2) + \mu \sum_{i=0}^{D_3-1} [e_{RCpLMS}^*(k - D_1 - i) \times \mathbf{x}(k - D_1 - i)] \tag{16}$$

$$e_{RCpLMS}(k) = S(k) - \mathbf{w}^H(k - D_2)\mathbf{u}(k) \tag{17}$$

to reduce the multiply accumulate hardware overhead of (16), D_3 is chosen such as $1 \leq D_3 \leq D_2$. With respect to (13), (16) and (17) the top level design becomes as shown in Fig. 2.

Where the multiplexer control signal, ref, is used to alternate between the original reference and the approximate reference signal when needed. The RC-pLMS adaptive beamformer is formed of a linear combiner block and a weight update block whose architecture is detailed in [18].

The presented architecture is implemented using complex arithmetic for $N = 8$ elements and a 18bits signed fixed point in $Q2.15$ format i.e. 1 signed bit, 2 integer bits and 15 precision bits with $D_1 = 4$, $D_2 = 2$ and $D_3 = 1$, i.e. six pipeline stages [18].

Table 1. 8-Input RC-pLMS beamformer synthesis results

Design	LUTs	DSP blocks	Registers
RC-pLMS-SR	929	32	2141
RC-pLMS	905	32	2065

Synthesis results of the RC-pLMS-SR and RC-pLMS are obtained for the Intel Stratix V 5SGXMABN3F45I4 model and shown in Table 1. Thus, it is clear that the RC-pLMS-SR is obtained at the cost of a minor increase in resource utilization, i.e. 1 bit multiplexer and an adder, compared to the original RC-pLMS architecture for a maximum operating frequency of 208.33 MHz [18].

5 Simulation Results and Discussion

A Monte Carlo type simulation is conducted with 500 realizations of 500 samples each for a ULA array of $N = 8$ elements. The input signals are generated as independent random complex Gaussian sequences m of the form $m = \mathcal{N}(0, \sigma_p^2) + j\mathcal{N}(0, \sigma_q^2)$ where $\mathcal{N}(0, \sigma^2)$ denotes normal (Gaussian) distribution with mean 0 and variance σ^2 . σ_p^2 and σ_q^2 are the inphase and out of phase signal variances, respectively. the input signal, \mathbf{x} , is formed of one message signal and two interferences with an angle of arrival (AOA) of 50° , 15° and 75° , respectively and are corrupted by CAWGN with a signal to noise ratio (SNR) of 5 dB. The parameters and initial conditions at $k = 0$ are given as $\mu = 2^{-6} = 0.0156$, $d(-1) = d(0)$, $\mathbf{x}(-1) = \mathbf{x}(0)$, $\sigma_p^2 = 0.05$ and $\sigma_q^2 = 0.01$. The proposed system is evaluated for two different obstruction scenarios, i.e. multiple and permanent.

As shown in Fig. 3, the ref control signal is used to flag reference signal availability where a value of 0 denotes signal discontinuity. The RC-pLMS-SR behavior is studied with respect to the MSE plot and the beam radiation pattern in infinite (software) and finite (hardware) precision modes.

As shown in Fig. 4, only the RC-pLMS-SR maintained a stable MSE convergence behavior for both obstruction cases. Moreover, for the second test case, where the reference signal was permanently discontinued, the RC-pLMS was unable to resume its operation. Hence, the RC-pLMS-SR maintained its convergence profile with respect to the proposed self referencing scheme, validating its robustness.

In order to better highlight the accuracy of the RC-pLMS-SR, the beam radiation pattern is simulated for the infinite and finite precision modes and is shown in Fig. 5.

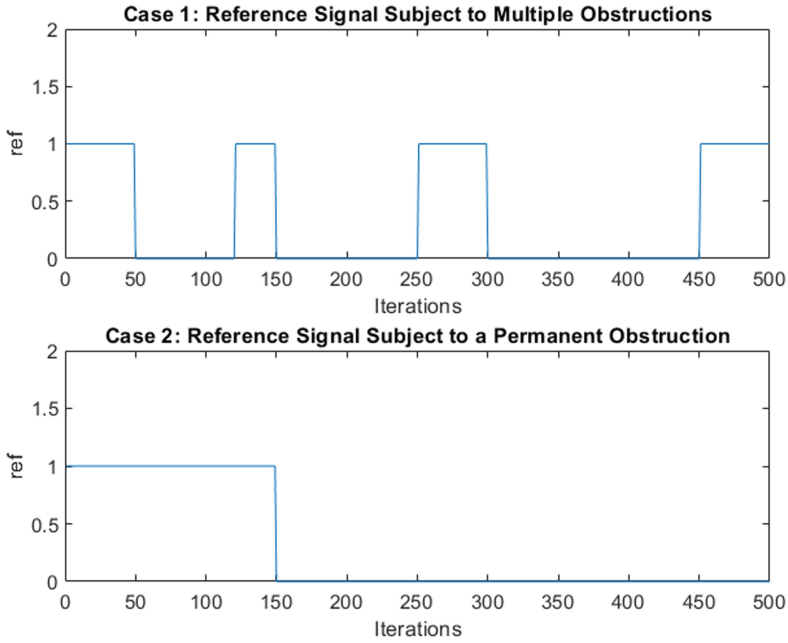


Fig. 3. Different reference signal obstruction scenarios

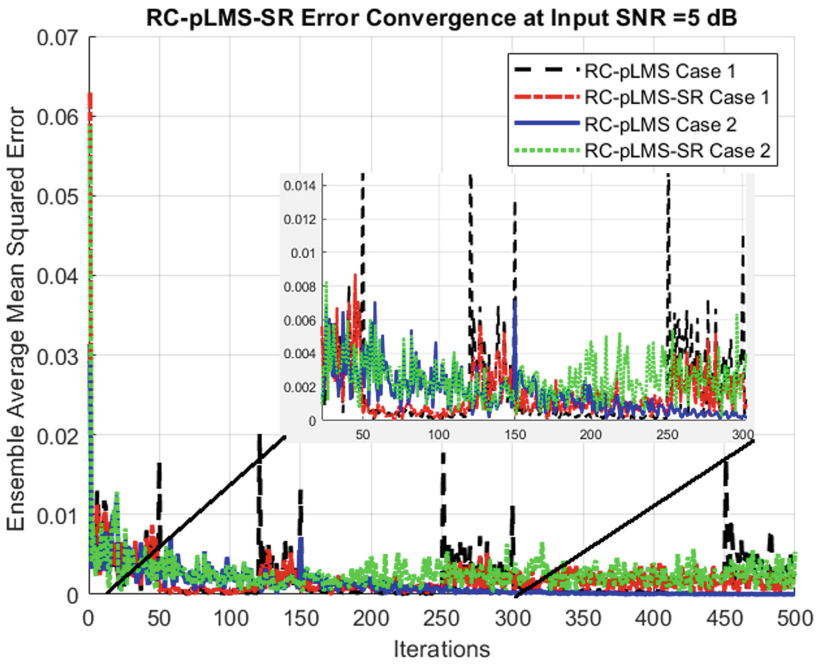


Fig. 4. RC-pLMS-SR MSE convergence at input SNR = 5 dB

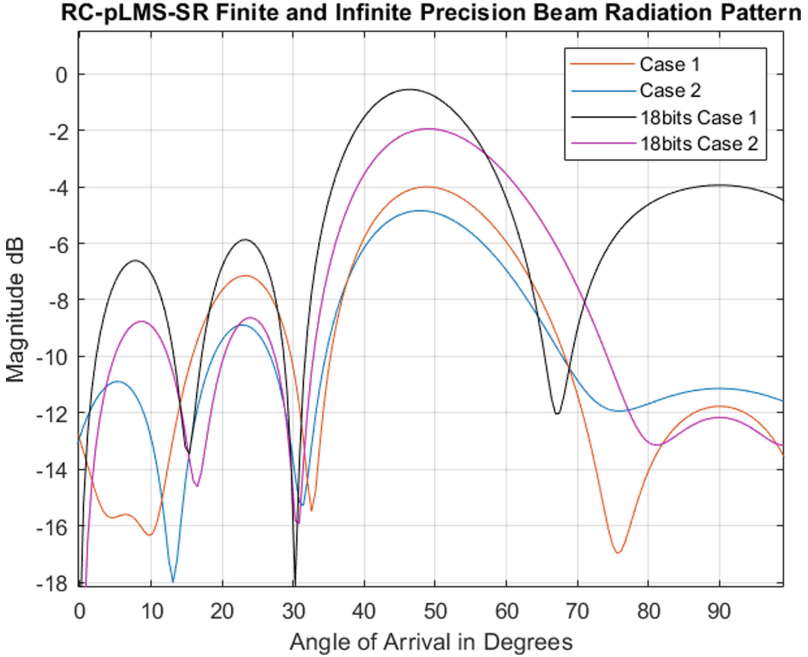


Fig. 5. RC-pLMS-SR finite and infinite precision beam radiation pattern at input SNR = 5 dB

From Fig. 5, it is clear that the infinite precision beam pattern, denoted by Case 1 and Case 2, maintained a satisfactory beam pointing accuracy, with minor misadjustment, by steering the main beam towards the direction of the desired signal, i.e. 50° , and its nulls towards both interferences, i.e. 15° and 75° . The beam misadjustment is a consequence of the approximation adopted in (13) and (15).

Moreover, the RC-pLMS-SR finite precision beam radiation patterns, denoted by 18bits Case 1 and 18bits Case 2, maintained an acceptable main beam pointing accuracy, however at the cost of a much larger error, i.e. $5^\circ - 8^\circ$. The resulting error increase, in finite precision mode, is due to quantization errors, the use of a fixed point $Q2.15$ format and the sum relaxation, of D_3 terms, employed in (16). Therefore, by increasing D_3 , the number of averaging terms is increased and the accuracy is increased. However, at the cost of a larger hardware overhead, i.e. complex adders and multipliers, and a more complex architecture [18].

6 Conclusion and Future Work

In this paper, we proposed a self referencing scheme for the reduced complexity parallel least mean square (RC-pLMS) adaptive beamforming algorithm for

robustness against interruptions in the reference signal and its hardware implementation. The RC-pLMS is a non-blind algorithm and its convergence and stability is initially correlated with the presence of an uninterrupted reference signal. As such, with respect to its accelerated convergence and high accuracy profile, we proposed the use of its filtered output signal as an alternative learning sequence whenever the reference signal is discontinued, i.e. self referencing. Additionally, we presented its hardware implementation on Field Programmable Gate Array (FPGA) to evaluate its performance in finite precision arithmetic. Experimental results, on software and hardware, demonstrated the validity of the proposed self referencing technique for robustness against different reference signal obstruction scenarios. Moreover, synthesis and implementation results highlighted the simplicity of the proposed technique, where it can be obtained at the cost of a negligible increase in resource utilization, i.e. one multiplexer and one adder, compared to the original RC-pLMS architecture.

Acknowledgment. The authors are grateful to AID - DGA (l' Agence de l' Innovation de Defense a la Direction Generale de l' Armement – Minitere des Armees) & ANR (Agence Nationale de le Recherche en France) for supporting our ANR-ASTRID – Project (ANR-19-ASTR-0005-03).

References

1. Srar, J.A., Chung, K., Mansour, A.: Adaptive array beamforming using a combined LMS-LMS algorithm. *IEEE Trans. Anten. Propagat.* **58**(11), 3545–3557 (2010)
2. Aboulnasr, T., Mayyas, K.: A robust variable step-size LMS-type algorithm: analysis and simulations. *IEEE Trans. Signal Process.* **45**(3), 631–639 (1997)
3. Sayed, A.H., Kailath, T.: A state-space approach to adaptive RLS filtering. *IEEE Signal Process. Magaz.* **11**, 18–60 (1994)
4. Xiubing, Z., et al.: A new modified robust variable step size LMS algorithm. In: *Proceedings of the 4th IEEE Conference on Industrial Electronics and Applications (ICIEA 2009)*, pp. 2699–2703. IEEE (2009)
5. Shengkui, Z., Zhihong, M., Suiyang, K.: A fast variable step-size LMS algorithm with system identification. In: *Proceedings of the 2nd IEEE Conference on Industrial Electronics and Applications (ICIEA 2007)*, pp. 2340–2345. IEEE (2007)
6. Kwong, R.H., Johnston, E.W.: A variable step size LMS algorithm. *IEEE Trans. Signal Process.* **40**(7), 1633–1642 (1992)
7. Lee, S., Lim, J.-S., Sung, K.: A low-complexity AFF-RLS algorithm using a normalization technique. *IEICE Electron. Exp.* **6**, 1774–1780 (2009)
8. Paleologu, C., Benesty, J., Ciochina, S.: A robust variable forgetting factor recursive least-squares algorithm for system identification. *IEEE Signal Process. Lett.* **15**, 597–600 (2008)
9. Albu, F., Kadlec, J., Coleman, N., Fagan, A.: The Gauss-Seidel fast affine projection algorithm. In: *IEEE Workshop on Signal Processing Systems, San Diego, CA*, pp. 109–114. IEEE (2002)
10. Mansour, A., Mesleh, R., Abaza, M.: New challenges in wireless and free space optical communications. *Opt. Lasers Eng.* **89**, 95–108 (2017)
11. Hong, W., et al.: Multibeam antenna technologies for 5G wireless communications. *IEEE Trans. Anten. Propagat.* **65**, 6231–6249 (2017)

12. Kim, D., Park, S., Kim, T., Minz, L., Park, S.: Fully digital beamforming receiver with a real-time calibration for 5G mobile communication. *IEEE Trans. Anten. Propag.* **67**, 3809–3819 (2019)
13. Akkad, G., Mansour, A., ElHassan, B., Inaty, E.: A multi-stage parallel LMS structure and its stability analysis using transfer function approximation. In: *Proceedings of the 28th European Signal Processing Conference (EUSIPCO)*, Amsterdam, August 2020, pp. 1–5 (2020)
14. Sarma, R.K., Khan, M.T., Shaik, R.A., Hazarika, J.: A novel time-shared and LUT-less pipelined architecture for LMS adaptive filter. *IEEE Trans. Very Large Scale Integr. Syst.* **28**(1), 1–10 (2019)
15. Zhao, W., Lin, J.Q., Chan, S.C., So, H.K.: A division-free and variable-regularized LMS-based generalized sidelobe canceller for adaptive beamforming and its efficient hardware realization. *IEEE Access* **6**, 64470–64485 (2018)
16. Albu, F., Kadlec, J., Coleman, N., Fagan, A.: Pipelined implementations of the a priori error-feedback LSL algorithm using logarithmic arithmetic. In: *Proceedings of the 2002 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Orlando, FL, May 2002, pp. III-2681–III-2684 (2002)
17. Akkad, G., Mansour, A., ElHassan, B., Srar, J., Najem, M., LeRoy, F.: Low complexity robust adaptive beamformer based on parallel RLMS and Kalman RLMS. In: *Proceedings of the 27th European Signal Processing Conference (EUSIPCO)*, A Coruna, September 2019, pp. 1–5 (2019)
18. Akkad, G., Mansour, A., ElHassan, B.A., Inaty, E., Ayoubi, R., Srar, J.A.: A pipelined reduced complexity two-stages parallel LMS structure for adaptive beamforming. In: *IEEE Transactions on Circuits and Systems I: Regular Papers*, pp. 1–13 (2020)
19. Yedavalli, P.S., Riihonen, T., Wang, X., Rabaey, J.M.: Far-field RF wireless power transfer with blind adaptive beamforming for internet of things devices. *IEEE Access* **5**, 1743–1752 (2017)
20. Akkad, G., Mansour, A., ElHassan, B., Inaty, E., Ayoubi, R.: Two stages parallel LMS structure a pipelined hardware architecture. In: *Proceedings of the 28th European Signal Processing Conference (EUSIPCO)*, Amsterdam, August 2020, pp. 1–5 (2020)
21. Shanbhag, N.R., Parhi, K.K.: Relaxed look-ahead pipelined LMS adaptive filters and their application to ADPCM coder. *IEEE Trans. Circuits Syst. II Analog Digital Signal Process.* **40**(12), 753–766 (1993)



A Data-Driven Method for Reliability Estimation of Auxiliary Power Consumption Prediction in Commercial Electric Vehicles

Tommaso Apicella^(✉), Edoardo Ragusa, Alessio Canepa, and Paolo Gastaldo

Department of Electrical, Electronic and Telecommunications Engineering,
and Naval Architecture, DITEN, University of Genoa, Genoa, Italy
{tommaso.apicella, edoardo.ragusa, alessio.canepa}@edu.unige.it,
paolo.gastaldo@unige.it

Abstract. One of the main issues concerning Battery Electric Vehicles (BEVs) is represented by range anxiety. This problem becomes crucial considering commercial vehicles equipped with electric Power Take Off (ePTO), which acts as power supplier for auxiliary loads. The paper presents a technique to estimate the reliability of power consumption prediction performed on ePTO consumption trends. Results show that the proposed algorithm balances the effects of power consumption prediction error in a more effective way with respect to a baseline solution.

Keywords: Data-driven · BEV · ePTO · Reliability score

1 Introduction

Battery electric vehicles (BEVs) are emerging as the most stable alternative to internal combustion engine (ICE) vehicles. Since one of the most pressuring features of BEVs is residual battery energy (a.k.a. range anxiety problem) [1], it is vital to manage power consumption in the most accurate and reliable way possible. This is not a trivial task because a lot of factors influence BEV power consumption e.g. driver's behavior [2] and vehicle's electronic components [3].

BEV literature recently started adopting machine learning techniques to explore driving range prediction aiming at reducing range anxiety. In general, several features are exploited to predict the driving range such as current, voltage and temperature of battery cells as well as road conditions and speed. Models range from Multiple Linear Regression (MLP) to Decision trees [4, 5]. One of the most exploited information about BEV status is the remaining battery charge i.e. State of Charge (SoC), which has a whole literature for its estimation featuring techniques such as Kalman filter [6], Deep Neural Networks (DNN) [7] and Support Vector Machine (SVM) [8]. Driving range prediction, as well as energy consumption estimation, are often associated with routing solutions to plan the

possible paths [9–11]. Mentioned works focus mainly on the accuracy of adopted machine learning techniques.

These topics are particularly critical for those commercial vehicles equipped with electric Power Take OFF (ePTO) [12]. The ePTO is the element that manages and powers auxiliary devices such as electro-mechanical arms and refrigerators, which can significantly impact BEV power consumption and, therefore, its residual driving range, influencing range anxiety.

This paper presents a data-driven strategy to improve the reliability of ePTO mean consumption prediction relative to the sequent, unseen, driving cycle, through the computation of an associated score. This reliability score is retrieved exploiting solely the statistical properties of the last k power consumption trends and leveraging the similarities among them. Normally, vehicles equipped with ePTO have routine missions or paths, thus the underlying assumption is that the consuming behavior of the unknown load varies slowly between consecutive driving cycles. The proposed method constitutes a lightweight solution, suitable for a wide range of applications since it can adapt to different ranges of power consumption. Moreover, a cost function is created to select the best parameters for the algorithm.

2 Proposal

The target system elaborates ePTO consumption information relative to the last k driving cycles and outputs a prediction of the mean power consumption for the next driving cycle along with its reliability score. Equations 1 and 2 highlight the outputs of the proposed system.

$$\hat{p}_t = f(\mathbf{pow}_{t-1}, \dots, \mathbf{pow}_{t-k}) . \quad (1)$$

$$\hat{r}_t = g(\mathbf{pow}_{t-1}, \dots, \mathbf{pow}_{t-k}) . \quad (2)$$

\hat{p}_t is the mean power consumption estimate for driving cycle t , \hat{r}_t is a value in $[0, 1]$ range representing the reliability score of \hat{p}_t , and \mathbf{pow}_n indicates the statistical features of the power consumption trend associated to n -th driving cycle: mean, standard deviation, kurtosis and skewness. From a qualitative point of view the system is supposed to provide a low reliability score for the next driving cycle (\hat{r}_t) when power consumption prediction error (\hat{e}_t) is high and vice-versa. Prediction error \hat{e}_t represents the discrepancy between mean power consumption prediction \hat{p}_t and the actual mean power consumption value, \hat{l}_t :

$$\hat{e}_t = |\hat{p}_t - \hat{l}_t| . \quad (3)$$

The generation of \hat{p}_t is performed using standard techniques discussed in experimental section (3.2).

Algorithm. The reliability score \hat{r}_t leverages the similarities among the last k driving cycle and it is generated in three steps. First, the euclidean distance matrix (similarity matrix) M of size $k \times k$ is computed as:

$$M(i, j) = \|\mathbf{pow}_i - \mathbf{pow}_j\| \quad \text{where } i, j = 1, \dots, k . \quad (4)$$

Second, the distance among the driving cycles is computed as average of M , \hat{m} . Finally, a filtering function provides \hat{r}_t clipping \hat{m} in $[0, 1]$ range. Two are the considered parametrized functions: the exponential and a piece-wise linear function which is probably the easiest to implement in an embedded system.

Cost Function. Equation 5 formalizes the concept that reliability score and error should have opposite behavior, setting the maximum value for the sum of \hat{e}_t and \hat{r}_t to 1.

$$\hat{r}_t + \hat{e}_t = 1. \quad (5)$$

Parameters selection of \hat{r}_t filtering functions is performed through the minimization in Eq. 6 applying the mentioned concept (Eq. 5) and exploiting validation set formed by N_{val} driving cycles.

$$\min_{param} \frac{1}{N_{val}} \sum_{i=1}^{N_{val}} [1 - (\hat{e}_i + \hat{r}_i(param))]. \quad (6)$$

param is replaced by exponential or linear function parameters σ or B respectively.

3 Experiments

The section is divided in three parts. First, the dataset is described. Second, three methods for power consumption prediction, \hat{p} , are compared. Finally, the proposed algorithm for \hat{r} retrieval is compared against a baseline solution. Comparisons with mentioned state-of-the-art approaches are not straightforward, as the main literature works focus on the accuracy of employed techniques. Moreover, mentioned papers exploit several different information e.g. road condition and parameters related to the vehicle such as motor features and average speed, while from ePTO system's perspective, consumed power represents the only available datum.

3.1 Dataset

Dataset is created using a simulator developed in collaboration with Altra S.p.A.¹ and it is formed by a set of **pow** vectors, each one describing a different driving cycle. To reflect the fact that commercial vehicles usually follow routine missions, a slowly changing discrete distribution D , randomly initialized, is used to sample one out of different power consumption templates T_i . Each T_i is obtained by modeling ePTO as a markov chain featuring two states, ON and OFF, characterised by transition probabilities $\{P_OFF_ON, P_ON_OFF\}$ such that $P_OFF + P_OFF_ON = 1$ and $P_ON + P_ON_OFF = 1$. Once it is sampled, T_i is mixed with additive Gaussian noise and the statistical features of the

¹ Altra S.p.A. is a Legal Entity of CNH Industrial operating in electric commercial vehicles market.

pow vector are retrieved, resulting in the actual driving cycle. After each sampling, the probabilities of the discrete distribution D slightly change applying a random increment/decrement to their value based on percentage boundaries B_i .

The simulator receives the total number N of consecutive driving cycles to create, the ePTO nominal absorbed power, the number of power consumption templates T_i , the percentage boundaries B_i and it outputs the dataset matrix $N \times 4$, where 4 are the statistical features of each **pow** vector.

For each experiment the number N of **pow** vectors is 3000 and 70%, 20%, 10% are the percentages of separation into the training, validation and test sets, respectively, the number of power consumption template is 6, the percentage bands are from 10% to 20% ($B1$), from 20% to 30% ($B2$) and from 30% to 40% ($B3$). The transition probabilities of each template are reported in Table 1.

Table 1. Transition probabilities of the markov chain (ePTO)

	T1	T2	T3	T4	T5	T6
P_OFF_ON	0.2	0.02	0.3	0.02	0.01	0.7
P_ON_OFF	0.02	0.2	0.3	0.01	0.02	0.7

3.2 Power Consumption Prediction

The first power consumption experiment assesses the impact of window's width k on power consumption prediction. This parameter indicates the minimum number of driving cycles that a new vehicle needs to perform to start obtaining the reliability score regarding the mean power consumption prediction.

Since **pow** vectors forming the dataset are retrieved exploiting the slowly changing distribution D , they are consecutive in time and they can be considered as temporal series. Two methods, widely employed in time series analysis, are used to predict power consumption: Sample Moving Average (SMA) and AutoRegressive Integrated Moving Average (ARIMA). In addition, Support Vector Machine for Regression (SVR) is applied as a comparison.

The polynomial degrees of univariate ARIMA are set to 1 after a grid search assessing the performance on the validation set over 20 trials. SVR is trained using standard deviation, kurtosis and skewness as features and targeting mean feature. 5-fold cross-validation is applied to select the model showing the lowest loss. Power forecast, \hat{p}_t , is the average of SVR prediction over the last k cycles. SMA is applied considering the mean of the next driving cycle equal to the mean of the k previous cycles, which is consistent with the employment of D .

The best value for k is chosen through a grid search, with values varying from 2 to 9 with offset 1 and from 10 to 200 with offset 10, and provides the lowest mean validation error. Validation error $\hat{\epsilon}_t^{val}$ is computed as described in Eq. 3.

Table 2 reports the average of best k values, (\hat{k}), obtained over 100 simulations. The table is divided into three sections based on distribution variation

bands (from $B1$ to $B3$). For each section SMA, ARIMA (ARI) and SVR are visible. In general, by increasing the percentage bands, the distribution becomes more erratic and this implies the rise of \hat{k} value, since the algorithms need more samples to reduce error on validation set.

Table 2. Results of k parameter selection.

	B1			B2			B3		
	SMA	ARI	SVR	SMA	ARI	SVR	SMA	ARI	SVR
\hat{k}	4	7	7	14	29	22	30	48	50

The second power consumption experiment evaluates the test set error of each consumption predictor exploiting \hat{k} . Table 3 shows mean and standard deviation of test set power consumption error varying the percentage bands (from $B1$ to $B3$). The error is computed as in Eq. 3 and it is expressed in terms of percentage with respect to ePTO nominal power. As bands value increase, the percentage error increases in mean, because the distribution becomes more and more erratic. In the first band the error is limited, especially for SMA and ARIMA (around 5%). SVR is the method which provides always the highest mean error.

Table 3. Percentage error of consumption predictors on test set

Error %	B1			B2			B3		
	SMA	ARI	SVR	SMA	ARI	SVR	SMA	ARI	SVR
μ	5.3	5.4	7.3	12.8	12.9	13.7	17.1	16.9	17.7
σ	2.9	3.0	3.6	5.3	5.2	4.9	5.1	5.0	4.9

3.3 Reliability Score

Reliability score experiment counts 100 simulations. In each simulation the dataset is created following mentioned procedure (subsect. 3.1).

The baseline method to retrieve reliability score consists on applying algorithm's filtering functions to the standard deviation of the last \hat{k} driving cycles' mean. Validation error is computed exploiting power consumption predictors (Eq. 3) and it is used to retrieve the best values for filtering function parameters, employing both algorithm and standard deviation method, leveraging Eq. 6. By means of Eq. 5 performances are assessed on test set retrieving the discrepancy with respect to 1, which is the target.

Table 4 shows the mean discrepancy from the target (1) in percentage employing algorithm (A) and standard deviation (S) with filtering functions (EXP and LIN). Mean discrepancy with respect to 1 is mostly lower using the algorithm than exploiting standard deviation method, meaning that the reliability acts effectively. Discrepancy's standard deviation is not reported in the table but its maximum value is 5.08%.

Table 4. Mean discrepancy from target

Method	B1			B2			B3		
	SMA	ARI	SVR	SMA	ARI	SVR	SMA	ARI	SVR
A_EXP	6.25	5.33	7.63	10.92	11.95	12.62	16.08	16.34	17.71
S_EXP	5.58	5.41	7.64	11.55	12.76	12.91	16.78	16.92	18.12
A_LIN	4.94	5.28	7.62	10.83	11.92	12.49	15.65	15.95	17.06
S_LIN	4.92	5.33	7.61	11.34	12.55	12.93	16.31	16.56	17.39

Table 5 presents the percentage of algorithm performing better than standard deviation method using exponential (EXP) or linear function (LIN), varying percentage bands (from $B1$ to $B3$). Increasing the band, the improvement is remarkable. Results of SMA in the first band are due to the small \hat{k} parameter.

Table 5. The comparison between algorithm and standard deviation method

Comp.	B1			B2			B3		
	SMA	ARI	SVR	SMA	ARI	SVR	SMA	ARI	SVR
EXP	38	65	53	73	87	62	70	75	66
LIN	40	61	50	76	79	63	70	72	66

4 Conclusions

This work presents a proposal in the context of range anxiety problem, providing an algorithm to compute reliability score associated to power consumption prediction, in case of ePTO equipped BEV. The case of a generic load, powered by the ePTO, is analysed with the assumption of being stable in time and having a measurable consumption. The proposed algorithm exploits statistical information about power consumption trends of the last k driving cycles and provides the reliability score relative to the next cycle. Experiments show that the proposal balances the effects of prediction error in a more effective way with respect to standard deviation method. The proposed solution provides an interpretable score and it features an hardware-friendly implementation. As a future development, different sources of noise could be added to assess the robustness of the technique.

References

1. Pevec, D., Babic, J., Carvalho, A., Ghiassi-Farrokhfal, Y., Ketter, W., Podobnik, V.: Electric vehicle range anxiety: An obstacle for the personal transportation (r)evolution? In: 2019 4th International Conference on Smart and Sustainable Technologies (SpliTech), pp. 1–8 (2019)

2. Kubaisi, R., Gauterin, F., Giessler, M.: A method to analyze driver influence on the energy consumption and power needs of electric vehicles. In: 2014 IEEE International Electric Vehicle Conference (IEVC), pp. 1–4 (2014)
3. Jinil, N., Reka, S.: Deep learning method to predict electric vehicle power requirements and optimizing power distribution. In: 2019 Fifth International Conference on Electrical Energy Systems (ICEES), pp. 1–5 (2019)
4. Sun, S., Zhang, J., Bi, J., Wang, Y.: A machine learning method for predicting driving range of battery electric vehicles. *J. Adv. Transp.* **2019**, 1–14 (2019)
5. Zhao, L., Yao, W., Wang, Y., Hu, J.: Machine learning-based method for remaining range prediction of electric vehicles. *IEEE Access* **8**, 212423–212441 (2020)
6. Mastali, M., Vazquez-Arenas, J., Fraser, R., Fowler, M., Afshar, S., Stevens, M.: Battery state of the charge estimation using kalman filtering. *J. Power Sources* **239**, 294–307 (2013)
7. Chemali, E., Kollmeyer, P.J., Preindl, M., Emadi, A.: State-of-charge estimation of li-ion batteries using deep neural networks: a machine learning approach. *J. Power Sour.* **400**, 242–255 (2018)
8. Álvarez Antón, J., García Nieto, P., de Cos Juez, F., Sánchez Lasheras, F., González Vega, M., Roqueñí Gutiérrez, M.: Battery state-of-charge estimator using the SVM technique. *Appl. Math. Model.* **37**(9), 6244–6253 (2013)
9. Cauwer, C.D., Verbeke, W., Mierlo, J.V., Coosemans, T.: A model for range estimation and energy-efficient routing of electric vehicles in real-world conditions. *IEEE Trans. Intell. Transp. Syst.* **21**(7), 2787–2800 (2020)
10. De Nunzio, G., Thibault, L.: Energy-optimal driving range prediction for electric vehicles. In: 2017 IEEE Intelligent Vehicles Symposium (IV). (2017) 1608–1613
11. Morlock, F., Rolle, B., Bauer, M., Sawodny, O.: Forecasts of electric vehicle energy consumption based on characteristic speed profiles and real-time traffic data. *IEEE Trans. Veh. Technol.* **69**(2), 1404–1418 (2020)
12. Ban, B., Stipetić, S.: Electric multipurpose vehicle power take-off: Overview, load cycles and actuation via synchronous reluctance machine. In: 2019 International Aegean Conference on Electrical Machines and Power Electronics (ACEMP) & 2019 International Conference on Optimization of Electrical and Electronic Equipment (OPTIM), pp. 596–603 (2019)



Compression of NN-Based Pulse-Shape Discriminators in Front-End Electronics for Particle Detection

Romina Soledad Molina^{1,2,3(✉)}, Luis Guillermo Garcia^{1,2},
Iván René Morales², Maria Liz Crespo², Giovanni Ramponi¹,
Sergio Carrato¹, Andres Cicuttin², and Hector Perez⁴

- ¹ Dipartimento di Ingegneria e Architettura, Università degli Studi di Trieste—IPL,
Piazzale Europa, 1, 34127 Trieste, TS, Italy
{rominasoledad.molina,luisguillermo.garciaordonez}@phd.units.it,
{ramponi,carrato}@units.it
- ² The Abdus Salam International Centre for Theoretical Physics—MLAB,
Via Beirut, 31, 34151 Grignano, TS, Italy
{imorales,mcrespo,cicuttin}@ictp.it
- ³ Departamento de electrónica, Universidad Nacional de San Luis—LEIS,
Av. Ejército de los Andes 950, D5700 San Luis, Argentina
- ⁴ Universidad de San Carlos de Guatemala—ECFM Ciudad Universitaria,
Zona 12, Ciudad de Guatemala, Guatemala

Abstract. Water Cherenkov detectors have been widely adopted as a low-cost technique for cosmic rays (CR) studies. Thus, an existing CR readout system has been chosen as the base DAQ (data acquisition) design, which has been paired to a Neural Network (NN) in order to work as a trace/event discrimination block. We present the compression of two NN architectures for particle classification, targeting a low-end System-on-Chip (SoC). The hls4ml package is used to translate the final NN models into a high-level synthesis project. Both NNs were implemented and tested on Xilinx SoC ZC7Z020 Zynq. A comparison of the accuracy of the detection, resource utilization and latency of the two NNs are presented. The results show the benefits of using compression techniques to deploy a reduced model, which provides a good compromise between efficiency, effectiveness, latency, as well as resource utilization.

Keywords: Distillation · Machine learning · SoC · FPGA · Compression

1 Introduction

Data acquisition systems (DAQ) based on FPGAs and System-on-Chip (SoC) are often used in experimental physics. Many of them are employed in high energy physics (HEP) experiments such as COMPASS at CERN [1], or in large-scale deployments in astrophysics such as the Latin American Giant Observatory (LAGO) [2] for cosmic rays studies. These experiments typically use

FPGAs/SoCs for the digital front-end to perform data acquisition and processing. Usually, the amount of resources needed to implement a digital block is correlated with the complexity of the design. However, implementing a proper compression strategy of a digital design, e.g. a trained neural network model, can reduce the amount of resources in such a way that a hardware block initially designed for a high-end FPGA could fit within a low-end SoC fabric, without significant performance loss.

Water Cherenkov Detectors (WCD) consist of a pure water tank used as a scintillator material coupled to a photomultiplier tube, which is connected to a high-voltage power supply and to an analog front-end [2]. The analog data coming from this type of setups are often digitized, captured, and processed within the detector electronics. However, some of the captured data traces may not be relevant, and will have to be deleted in the subsequent offline data analysis.

A WCD DAQ system has been utilized as a test base for the deployment of a neural network design within an existing project. The main goal was to verify the versatility of advanced processing algorithms as co-existing processing blocks. The original system design has been paired with a NN in order to work as a trace/event discrimination block. Machine learning (ML) techniques have been used for offline pulse shape discrimination (PSD), obtaining high accuracy [3–5]. In a previous work [6] a PSD was implemented into UltraScale+ MPSoC through neural network based on a multilayer perceptron (MLP), and FIR filter. The main goal of the compression is to deploy ML models on resource-constrained devices, obtaining smaller and faster models [7]. Among the most used compression techniques are: pruning (P) to reduce the number of parameters of NN, quantization (Q) which implies a reduction in the number of bits used to represent the weights and activations, and knowledge distillation (KD) [8]. The latter is based on a teacher-student approach, where the teacher transfers the knowledge to the student, obtaining smaller and faster models, while keeping similarities with the teacher.

In this framework, the actual DAQ system is implemented in a low-end FPGA platform, where the original design has already taken up part of the available resources: BRAM with 92% of usage, followed by LUT 3%, then Flip-flop 2% and finally DSP 1%. These parameters set the starting point to determine the maximum space that may be occupied by the upgraded design. To implement the PSD using a neural network, compression of the trained model is a key point to fit the complete system in one single chip. We present the compression for a PSD based on NNs, performing a comparison between two architectures: a CNN (convolutional neural network) and a MLP. We also show the corresponding implementations in SoC-FPGA for both inference processes after compression. The aim is to have a complete front-end in a low-end FPGA, with the DAQ system and the dedicated logic for the processing of the incoming signals, achieving a good compromise between efficiency, effectiveness, latency and resources.

2 NN-Based Pulse-Shape Discriminators

The multi-class classification is based on distinguishing three types of signals from the raw data, labeled as 0 (muon), 1 (electron) and 2 (electric discharge).

Each signal has 30 samples, a reduced length compared to those processed in [6]. Two architectures were designed: one based on a MLP and the other based on a CNN, to perform a comparison in terms of accuracy and size of the final trained networks.

To implement knowledge distillation technique, a teacher and a student for both architectures were defined, as shown in Fig. 1, where the former is a larger model than the latter. In the architecture based on CNN, the teacher is mainly composed by a set of 1D-Convolutional layers and fully connected layers to perform the final classification, whereas the student presents a reduced number of layers. In this approach, the number of parameters are 11,621 and 651 respectively. The teacher for the MLP approach is formed by an ensemble of fully connected layers, while the student has 3 layers. With regard to the number of parameters, the former has 94,203 and the latter 2,093.

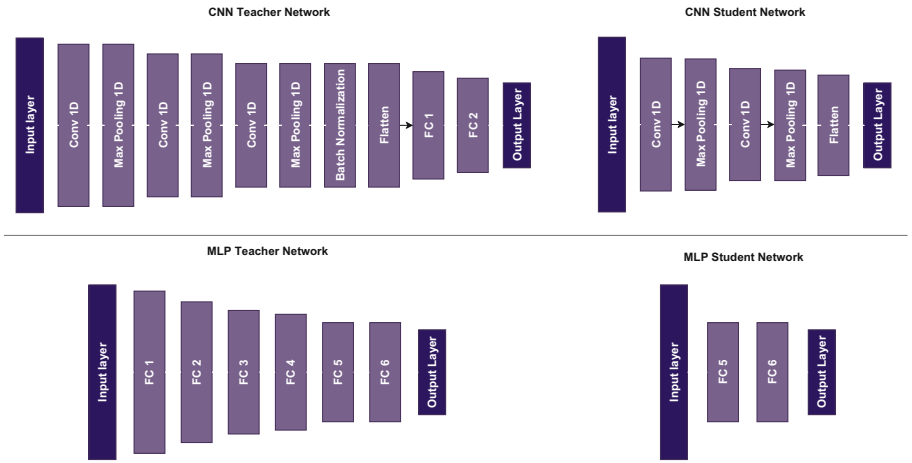


Fig. 1. Teacher and student architectures based on CNN and MLP.

The training was performed using K-Fold Cross-Validation, Adam optimizer with learning rate: 0.0001 and the parameters were configured as follows: (i) for the teacher networks: batch size: 300, epochs: 32, $k = 10$; (ii) for the student networks: batch size: 150, epochs: 16, $k = 3$. For each class, 100,000 signals were used to carry out the training.

The next step in the compression stage involves pruning and quantization in the student networks. We analyzed the distribution of the weights and biases to proceed with the corresponding reduction in the number of bits to decrease the memory footprint in the FPGA. Pruning was performed with target sparsity of 0.5. The final trained and compressed networks were translated into a high-level synthesis (HLS) project using the hls4ml package [9], created to map inference based on learned models in the context of HEP. After the evaluation of the different results obtained with HLS, two designs were selected (one for each NN architecture) and exported as IP cores. Then, the corresponding hardware was generated to configure the FPGA, verifying the inference stages on the board.

In this way, the final hardware utilization and inference times were obtained for each architecture.

3 Experimental Results

We evaluated the impact of compression by implementing both NNs architectures on a CPU Core i7 3.4 GHz 64 GB RAM GeForce GTX 1070, using Python 3.6.7, TensorFlow 1.12 and QKeras [10]. The networks were translated into an HLS project using hls4ml. For the SoC implementation, we used Vitis 2019.2.1 and XC7Z020 family by Xilinx.

For each implementation, Table 1 presents the final accuracy obtained after the different training steps. Both architectures exhibited an overall accuracy higher than 90%, with a small decrease through the compression process.

Table 1. Accuracy for CNN and MLP architectures: Teacher, Student KD (using KD), Student KD-PQ (using KD, P and Q).

CNN			MLP		
Teacher	Student KD	Student KD-PQ	Teacher	Student KD	Student KD-PQ
98.55%	95.37%	94.77%	97.08%	96.36%	95.8%

The final accuracy reached by each trained student after compression is exposed in Fig. 2 through the corresponding confusion matrix. We can notice that, despite having decreased the overall accuracy, the final models have a high rate of correct classification of muons (0) and electrons (1), and are successful in the discrimination of false positives. Furthermore, it is observed that the last column of each teacher’s matrix shows that CNN has a much higher rate of missed detection than MLP. However, for the student’s matrices, this behavior is quite similar in the two architectures.

In regard to the HLS implementation, latency and hardware resources reported by the tool (in a conservative way for FFs and LUTs, whereas for DSPs and BRAMs the utilization is comparable with the final implemented design [9]) are presented in Table 2, with a target clock at 5 ns. For the MLP implementation, IP cores with reuse factor (RF) of 1 (MLP_v1) and 8 (MLP_v2) were synthesized. The former showed an excess of DSP utilization (575%) with reuse factor of 1, which prevents the final implementation on the board. Due to this, the reuse factor was changed to 8, getting a good compromise between latency and resource utilization, where the most stressed resource is DSP with 72% of utilization.

CNN HLS implementations were generated with different combination of optimizations than those generated by default with the hls4ml package, specifying the loops where to use the UNROLL, PIPELINE and ARRAY PARTITION directives. Due to these parallel techniques, the most consumed resources were LUTs and Flip-Flops. In both implementations (CNN_v1 and CNN_v2) the

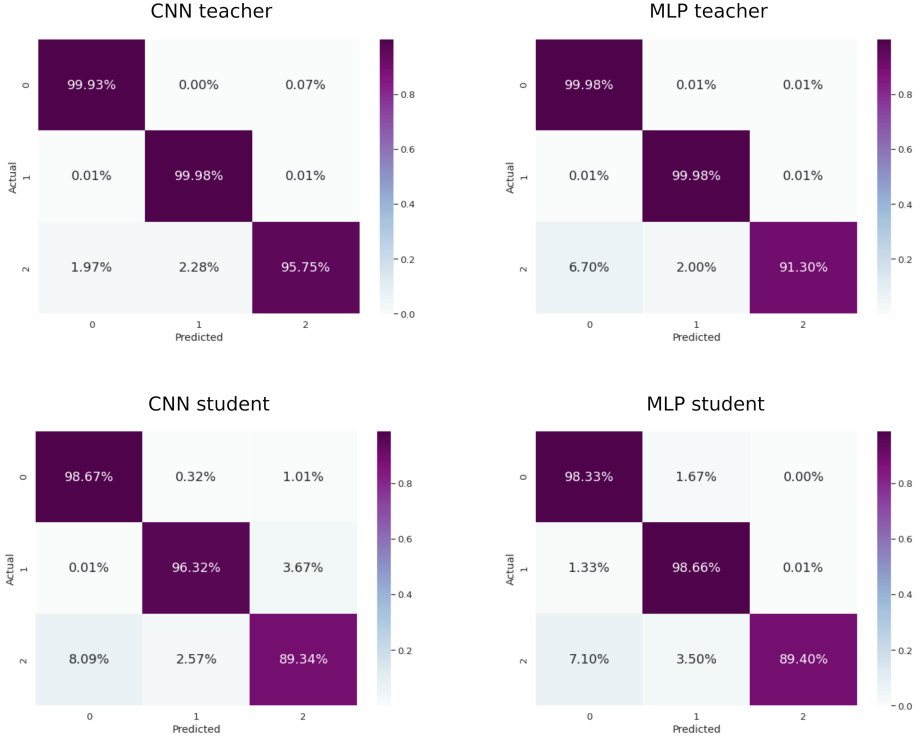


Fig. 2. Confusion matrices for teachers (top). Confusion matrices for students networks after compression (bottom)

latency reported in clock cycles was higher compared to those of the MLP versions. Based on the previous observations, MLP_v2 and CNN_v2 were selected to generate the final IP cores.

Once the IP cores were synthesized and exported as Register Transfer Level (RTL), the final hardware to configure the FPGA was created. An Integrated Logic Analyzer (ILA) core was added to obtain the real number of clock cycles used by PSD IP core for NN-based inference. Despite the fact that, for both the HLS solutions (MLP_v2 and CNN_v2), the resource LUT presented a use larger

Table 2. Hardware utilization and latency (in clock cycles) reported by HLS tools.

	RF	Latency	BRAM	DSP	FF	LUT
MLP_v1	1	33	0%	575%	93%	90%
MLP_v2	8	45	0%	72%	39%	112%
CNN_v1	1	3954	7%	130%	89%	160%
CNN_v2	1	6320	2%	46%	50%	99%

than 100%, Vivado IP Integrator was able to optimize the resource usage for the implemented designs through its synthesis and implementation strategies.

The results obtained after the implementation of the IP cores in SoC-FPGA are shown in Table 3. Lower inference time is observed for the network based on MLP compared to CNN, which is an important point in the actual WCD DAQ. Besides this, the final resource utilization allows to implement both architectures (MLP and CNN) with a DAQ system in a low-end-platform. The most stressed resource for NNs IP cores was DSP (MLP: 73% and CNN: 47%), while for the actual WCD DAQ, it was BRAM with 92% of usage. Indeed, the two designed digital systems are complementary in terms of hardware resources. Regarding True Positive Rate (TPR), we can conclude that the lowest values were reported for the class labeled as 2; this is consistent with the behavior shown in the confusion matrices.

Table 3. Inference time, hardware utilization and TPR for each class after SoC-FPGA implementation

Architecture	Inference time [μs]	Hardware utilization				TPR		
		BRAM	DSP	FF	LUT	Class 0	Class 1	Class 2
CNN	31.75	1%	47%	30%	61%	0.9729	0.9998	0.9387
MLP	0.37	0%	73%	33%	62%	0.9998	0.9671	0.9435

4 Conclusions

In this paper we presented a NN compression of a PSD in an electronic front-end for particle detection. We took advantage of an ensemble of compression techniques to reduce the size of the model, obtaining a good compromise between latency and resources for the final digital design. We showed the effective use of knowledge distillation to compress the models, targeting low-cost SoC devices. We observed that the network based on MLP had better inference time and accuracy than CNN. From the point of hardware usage, CNN employed less resources than MLP. We verified the integration of a PSD based on NNs along with the front-end and DAQ logic for cosmic rays particle analysis on a WCD, in a low-cost SoC device.

References

1. Abbon, P., et al.: The COMPASS experiment at CERN. Nucl. Instrum. Meth. Phys. Res. Sect. A Accel. Spectrom. Detect. Assoc. Equip. **577**, 455–518 (2007)
2. Sidelnik, I., Asorey, H.: LAGO: the Latin American giant observatory. Nucl. Instrum. Meth. Phys. Res. Sect. A Accel. Spectrom. Detect. Assoc. Equip. **876**, 173–175 (2017)

3. Mace, E.K., Ward, J.D., Aalseth, C.E.: Use of neural networks to analyze pulse shape data in low-background detectors. *J. Radioanal. Nucl. Chem.* **318**(1), 117–124 (2018). <https://doi.org/10.1007/s10967-018-5983-1>
4. Holl, P., Hauertmann, L., Majorovits, B., Schulz, O., Schuster, M., Zsigmond, A.J.: Deep learning based pulse shape discrimination for germanium detectors. *Eur. Phys. J. C* **79**(6), 450 (2019). <https://doi.org/10.1140/epjc/s10052-019-6869-2>
5. Droz, D., Tykhonov, A., Wu, X.: Neural networks for electron identification with DAMPE. In: *Proceedings of 36th International Cosmic Ray Conference* (2019)
6. Garcia, L.G., et al.: Muon–electron pulse shape discrimination for water Cherenkov detectors based on FPGA/SoC. *Electronics* **10**(3), 224 (2021)
7. Choudhary, T., Mishra, V., Goswami, A., Sarangapani, J.: A comprehensive survey on model compression and acceleration. *Artif. Intel. Rev.* **53**(7), 5113–5155 (2020)
8. Hinton, G.E., Vinyals, O., Dean, J.: Distilling the knowledge in a neural network. In: *arXiv preprint arXiv:1503.02531* (2015)
9. Duarte, J., et al.: Fast inference of deep neural networks in FPGAs for particle physics. *J. Instrum.* **13**(07), P07027–P07027 (2018)
10. Coelho, C.N., et al.: Automatic heterogeneous quantization of deep neural networks for low-latency inference on the edge for particle detectors. *Nat. Mach. Intell.* **3**, 675–686 (2021)



Assisted Driving for Power Wheelchair: A Segmentation Network for Obstacle Detection on Nvidia Jetson Nano

Gianluca Giuffrida^(✉), Silvia Panicacci, Massimiliano Donati,
and Luca Fanucci

Department of Information Engineering, University of Pisa,
Via G. Caruso 16, 56122 Pisa, Italy

{gianluca.giuffrida,silvia.panicacci}@phd.unipi.it,
{massimiliano.donati,luca.fanucci}@unipi.it

Abstract. Recently, an ever-growing focus and attention is given to accessibility and enabling technologies for people with disabilities. In the mobility field, people with motor skill impairments exploit power wheelchairs to move in indoor and outdoor scenarios. AI-Drive aims to design an assistive power wheelchair for outdoor use, providing obstacle detection and avoidance in urban scenarios, leveraging low-cost digital cameras and artificial intelligence. This paper focuses on the implementation of a convolutional neural network for semantic segmentation of urban scenes, to detect obstacles in an outdoor setting. A U-Net-like architecture was trained on GPU over multiple datasets, representative of the final environment. The selected trained network was then customised to perform inference on the Nvidia Jetson Nano hardware accelerator, to be mounted directly on the wheelchair. The resulting model achieves an accuracy of around 85% and inference time of 35 ms, thus providing a concrete solution towards the target assisted power wheelchair.

Keywords: Convolutional neural network · Segmentation · Assisted driving · Obstacle detection · Embedded hardware accelerator

1 Introduction

In the recent years, society has taken huge steps forward in the field of accessibility and assistive technologies, designing for example power wheelchairs, prosthetic and robotic limbs, screen readers, and voice recognition programs [9, 13]. However, some issues cannot be solved by the actual technologies. For example, people forced to wheelchairs may not have the fine motor skills to safely drive mobility-assisting devices in case of obstacles or other impediments, due to their physical disabilities. In fact, power wheelchairs available on the market could help people with soft disabilities, but they do not foresee multiple impairments, such as restricted field of vision and diminished perception [12].

Multiple researches have tried to empower power wheelchairs with autonomous driving tools, detecting and avoiding obstacles to assure safe experience also to people with severe disabilities [2, 6, 11]. On the other hand, most of these designs use a wide range of sensors to perform the environmental mapping and/or detect dynamic obstacles, e.g., ultrasonic sensors and LIDAR technology [1], with consequently high costs. In addition, they are usually developed for indoor use, making the power wheelchairs less useful outdoor.

AI-Drive wants to address the limitations presented above by exploring a novel approach to the problem of assisted driving: employing digital cameras, one of the most readily-available and cost-effective technologies in the current market, to design and implement a smart power wheelchair, which detects and avoids obstacles bypassing the user's joystick controller through Artificial Intelligence (AI) techniques [5]. Figure 1 shows the high-level system architecture.

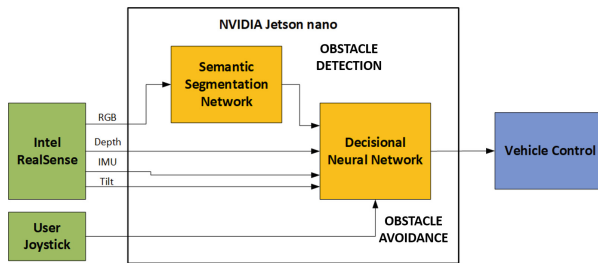


Fig. 1. AI-drive high-level system architecture.

This paper presents the obstacle detection module, implemented through a semantic segmentation Convolutional Neural Network (CNN). Given the pre-processed visual data acquired by the digital camera, the CNN detects the presence and location of any obstacles currently in view. This algorithm runs on a hardware accelerator embedded on the power wheelchair, ensuring that its execution time and latency is small enough to perform the required control action in a timely and safe manner.

After this introduction, Sect. 2 describes CNN implementation (dataset, metrics, network performance on GPU), while the porting on the target hardware accelerator and results are discussed in Sect. 3. Finally, the conclusions are drawn in Sect. 4.

2 Neural Network Architecture Implementation

To efficiently detect obstacles and successively take an action, it is important to segment the RGB image coming from the digital camera, assigning a class to each pixel of the input image. This task, previously executed via computer vision algorithms, is performed by means of supervised AI algorithms, simplifying the computational effort while maintaining high accuracy. In order to maximise the network accuracy for a specific task, choosing a suitable dataset is essential.

Multiple datasets are freely and publicly available for image segmentation tasks. Some of them cover a wide range of classes in several unrelated fields (e.g., COCO [7] and PascalVOC [4]); others are specific for a single task: for urban scene parsing problems, Cityscapes [3], and Mapillary Vistas [8] are two of the most used datasets. CamVid was not considered in this work due to its very small number of images. Cityscapes contains 3475 fine-annotated urban-scene images, taken in different cities in Germany. Each image and corresponding segmented map have a size of $1024 \times 2048 \times 3$, re-scaled to $512 \times 1024 \times 3$. Labels originally belong to 34 classes, but since for this project not all classes were needed, we grouped them into 10 classes. Cityscapes was also entirely pre-processed, changing the image ratio from 2:1 to 4:3, to be compliant with digital camera aspect ratio and double the images within the dataset. Finally, it was further augmented with horizontal flip, generating the so-called augmented Cityscapes dataset. Mapillary Vistas is a large-scale dataset, containing images taken all over the world, at different time of the day and throughout the year, with different perspectives. Images, originally with different dimensions and aspect ratios, were cropped into squares of size $512 \times 512 \times 3$. The final number of training and validation samples was 20471 and 2100, respectively. The original 66 annotated classes were then grouped into 16 classes.

To assess the NN results and its feasibility of executing on an embedded hardware accelerator, we defined four evaluation metrics: i) inference time: the application is time-constrained and, since the maximum wheelchair speed is 10 km/h, to recognise an obstacle ~ 20 cm away ~ 14 fps are required (inference time must be lower than 70 ms); ii) network size: the CNN runs on an embedded hardware accelerator; iii) categorical accuracy: accuracy in detecting correct pixel categories; iv) mean Intersection over Union (mIoU): mean of accuracy per class.

The architecture selected for obstacle detection was derived by U-Net, which is designed to solve binary segmentation problems, exploiting small size datasets [10]. In this work, it was customised to make it suitable for a multi-class problem, increasing the final layer filters from 2 to the number of output classes and changing the final layer activation from a sigmoid to a softmax. The parameters of this U-Net-like NN were about 31 millions. It was trained over the three datasets (Cityscapes, augmented Cityscapes and Mapillary Vistas) on a Tesla P100 GPU. Since the network architecture was the same, the resulting inference time did not change with the training dataset and was about 41 ms. Accuracy and mIoU results are shown on Table 1.

Table 1. Results of U-Net-like architecture over the three datasets.

Dataset	Accuracy (%)	mIoU (%)
Cityscapes	81.28	75.30
Augmented Cityscapes	91.99	87.73
Mapillary Vistas	88.11	80.07

Finally, the three trained CNNs were tested on real images, with the sidewalk in the foreground. Even if the CNN trained on the augmented Cityscapes dataset achieved the best performances on the test set, it did not perform well with images outside the dataset (Fig. 2(a)), while the CNN trained with Mapillary Vistas dataset achieved quite good results also on real images (Fig. 2(b)). This behaviour can be explained by the images contained in the dataset: Cityscapes mainly contains images taken from the road, while Mapillary Vistas contains images taken from different framing. For this reason, we selected the CNN trained with the latter to be ported on the embedded hardware accelerator.

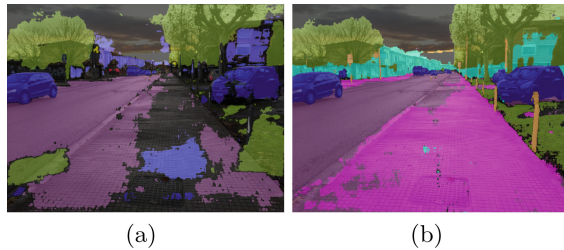


Fig. 2. Results on a real image of U-Net-like NN trained with (a) augmented Cityscapes dataset, and (b) Mapillary Vistas dataset.

3 Neural Network Porting on Hardware Accelerator

The target embedded hardware accelerator selected for this project was the Nvidia Jetson Nano 4 GB developer-board, because it fits the requirements for both the networks of the project. It is built around a 64-bit quad-core ARM Cortex A57 CPU and a Tegra $\times 1$ GPU (Maxwell architecture) with 128 CUDA cores and 4 GB of LPDDR4 memory. The low number of CUDA cores and the fact that the GPU shares its memory with the CPU make the Jetson Nano consume only a few Watts of power. Other accelerators were taken into account, but the limited computational power or memory constraints imposed by them did not allow to implement the entire system (Google Coral) or to reach the desired speed (Raspberry Pi + NCS).

After selecting the embedded hardware accelerator, the CNN architecture and the final dataset for training, some further customisations were required to run the network model on the target and meet the requirements. In particular, in its current state, the CNN was only able to achieve a frame rate of about 25 fps on GPU, which has significantly more computational power available with respect to the Jetson Nano. Moreover, it has much smaller amount of memory available for the model (RAM and storage), making it unable to run the original model with more than 30 million parameters. To solve this issue, we kept the model architecture, while reducing the number of filters per layer. In particular,

the number of initial filters was decreased, thus decreasing also the number of input channels in the following layer. Then, we reduced the number of filters in the first convolutional layer from 64 to 32, 16 and 8, generating three smaller CNNs, proportionally re-scaling the whole filters in the network. Their results are presented in Table 2.

Table 2. Results of the four U-Net-like architectures with different number of initial filters over the Mapillary Vistas dataset.

In. Filters	# Param. (M)	Inference time (ms)	Accuracy (%)	mIoU (%)
64	31.04	41	88.11	80.07
32	7.76	24	87.63	79.85
16	1.94	21	86.54	78.34
8	0.49	19	85.63	76.89

As expected, reducing the number of filters has a positive impact on inference time, but a consequent drop in accuracy and mIoU: the CNN with the smallest number of parameters (about one eighth of the original) experiences a drop of $\sim 3\%$ in both accuracy and mIoU, with a gain of 22ms in inference time. Since the model selected in this phase is only one part of the entire system, we selected the smallest CNN (with 8 initial filters) to be ported on the Jetson Nano.

The porting process was executed via TensorRT native API, which automatically quantises the weights, from floating point 32 to floating point 16. Porting and quantisation did not affect accuracy and mIoU, which remained the same ($\sim 85\%$ and $\sim 76\%$, respectively), as shown in Fig. 3. On the contrary, the inference time increased from 19ms on GPU to 35ms on Jetson Nano, however meeting the requirement of having at least 14 fps.

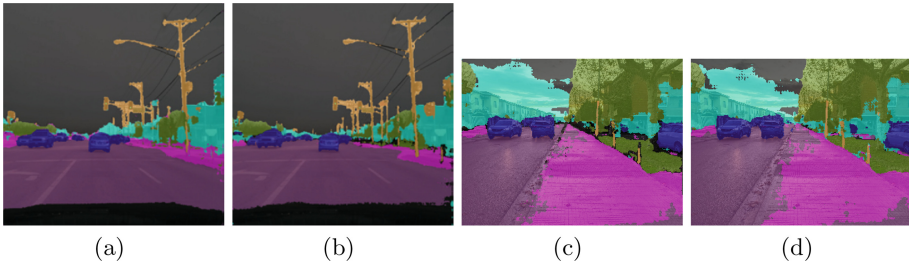


Fig. 3. Inference of U-Net-like NN with 8 initial filters over a Mapillary Vistas image on (a) Tesla P100 and (b) Jetson Nano, and on a real image on (c) Tesla P100 and (d) Jetson Nano.

4 Conclusions

AI-Drive aims to develop a smart power wheelchair for indoor/outdoor use, with obstacle detection and avoidance to help and simplify the motor impaired people's life. The system exploits a low-cost digital camera for the scene representation and runs artificial intelligence algorithms on a Nvidia Jetson Nano, directly mounted on the power wheelchair, to detect obstacles and avoid them, bypassing the user's decisions in case of danger.

This paper presents the implementation of the obstacle detection module, which segments the pre-processed image acquired by the digital camera through a convolutional neural network, identifying obstacles in the view. A U-Net-like architecture was trained on GPU over three different datasets. The network trained over the Mapillary Vistas dataset achieved the best results and was then optimised to run on the target hardware accelerator. The resulting model achieves about 85% of accuracy at more than 28 fps on the Jetson Nano, making it suitable for usage on a space-constrained system like a power wheelchair.

References

1. Behroozpour, B., Sandborn, P.A., Wu, M.C., Boser, B.E.: Lidar system architectures and circuits. *IEEE Commun. Mag.* **55**(10), 135–142 (2017)
2. Cimurs, R., Lee, J.H., Suh, I.H.: Goal-oriented obstacle avoidance with deep reinforcement learning in continuous action space. *Electronics* **9**(3), 411 (2020)
3. Cordts, M., et al.: The cityscapes dataset for semantic urban scene understanding. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3213–3223 (2016)
4. Everingham, M., Eslami, S.A., Van Gool, L., Williams, C.K., Winn, J., Zisserman, A.: The pascal visual object classes challenge: a retrospective. *Int. J. Comput. Vis.* **111**(1), 98–136 (2015)
5. Falzone, G., Giuffrida, G., Panicacci, S., Donati, M., Fanucci, L.: Simulation framework to train intelligent agents towards an assisted driving power wheelchair for people with disability. In: *Proceedings of the 13th International Conference on Agents and Artificial Intelligence - Volume 1. ICAART, SciTePress* (2021)
6. Leaman, J., La, H.M.: A comprehensive review of smart wheelchairs: past, present, and future. *IEEE Trans. Hum. Mach. Syst.* **47**(4), 486–499 (2017)
7. Lin, T., et al.: Microsoft COCO: common objects in context. *arXiv* [arXiv:1405.0312](https://arxiv.org/abs/1405.0312) (2014)
8. Neuhold, G., Ollmann, T., Rota Bulò, S., Kotschieder, P.: The Mapillary Vistas Dataset for semantic understanding of street scenes. In: *Proceedings of the IEEE International Conference on Computer Vision*, pp. 4990–4999 (2017)
9. Panicacci, S., et al.: Empowering deafblind communication capabilities by means of AI-based body parts tracking and remotely controlled robotic arm for sign language speakers. In: Saponara, S., De Gloria, A. (eds.) *ApplePies 2019. LNEE*, vol. 627, pp. 381–387. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-37277-4_44
10. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) *MICCAI 2015. LNCS*, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28

11. Şahin, H.İ., Kavsaoğlu, A.R.: Autonomously controlled intelligent wheelchair system for indoor areas. In: 2021 3rd International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), pp. 1–6. IEEE (2021)
12. Simpson, R., LoPresti, E., Cooper, R.: How many people would benefit from a smart wheelchair? *J. Rehabil. Res. Dev.* **45**(1), 53–71 (2008)
13. US Department of Health and Human Services: What are some types of assistive devices and how are they used? (2018). <https://www.nichd.nih.gov/health/topics/rehabtech/conditioninfo/device>. Accessed June 2021



Analysis of Thermal-Induced Shunt Current Sensor Errors in a Low-Cost Battery Management System

Alessandro Verani^(✉), Roberto Di Rienzo, Federico Baronti,
Roberto Roncella, and Roberto Saletti

Dipartimento Ingegneria dell'Informazione, Università di Pisa,
Via Caruso 16, 56122 Pisa, Italy

alessandro.verani@phd.unipi.it, {roberto.dirienzo,
federico.baronti,roberto.roncella,roberto.saletti}@unipi.it
<http://www.dii.unipi.it>

Abstract. Lithium-Ion batteries are becoming a standard solution for automotive applications. These batteries must be equipped with a Battery Management System that monitors and controls the battery avoiding hazardous situations. The battery current measurement is crucial for both safety and control tasks. Shunt-based current sensors are widely used in automotive applications, thanks to the affordability and linearity of the measurement. This article discusses an unexpected current measurement error occurred during the test of a prototypal Battery Management System for a 12 V cranking battery. The current error only shows up during the cell balancing procedure, pointing to a thermal related effect. A theoretical assessment of the problem suggests the thermoelectric effect as the best candidate to explain the error. The thermal simulation of the circuit validated the hypothesis with perfect match. Finally, the guidelines for a correct shunt choice in a Battery Management System are provided.

Keywords: Lithium-Ion batteries · Battery management system · Shunt resistor · Thermoelectric effects

1 Introduction

Lithium-Ion batteries are becoming a standard solution for powertrain and cranking in electric vehicles. However, such batteries require the presence of a Battery Management System (BMS). The BMS is asked to monitor and control the battery to ensure the operation of the cells inside their Safe Operating Area. This task is usually accomplished by Battery Monitor ICs, which acquire the voltages, temperatures and current of the cells, and act on the power switch to disconnect the battery and prevent harmful conditions. Furthermore, the BMS accomplishes other key tasks: charge equalization, battery state estimation, logging, and communication [1]. Charge equalization is needed to reduce charge

mismatches between the battery cells, ensuring the largest usable capacity of a group of series-connected cells [2]. There are two possible approaches: active and passive balancing [3]. The former involves a cell-to-cell charge exchange from the most charged cells to the lower charged ones. It is very efficient but requires complex and expensive hardware solutions [4]. Instead, passive approaches discharge the most charged cells using bleeding resistors. This leads to power dissipation, but the circuit is much simpler and cheaper than in active approaches. Battery applications in which cost is an important design metric, such as automotive, commonly adopt passive approaches. Generally, the power dissipated by each bleeding resistor is in the range 0.5–1.5 W. It goes without saying that thermal management of the BMS circuit board must be dealt with care, because of the presence of power devices like power switch, fuse, and balancing resistors. The position of these devices on the Printed Circuit Board (PCB) is key for a good layout design and could affect both the safety and the performance of the BMS.

The BMS also performs dedicated algorithms to estimate the State of Charge (SoC) of the battery. They generally apply the Coulomb counting technique consisting of the time integration of the current values. Unfortunately, the estimation error could be very large if an offset affects the current measurement. For this reason, the selection of a proper current sensor is an important task in the development of a low-cost BMS. Two different types of current sensors are generally adopted for battery applications: Hall-effect-based current sensors and shunt resistors. The former rely on the magnetic field generated by the current flowing in a conductor to measure the current itself according to the Hall effect. This method ensures galvanic isolation between the measurement system and the measured conductor and no voltage loss, but its cost could have a significant impact when high current range and accuracy at least of 1% are required [5]. On the other side, the shunt method relies on a resistor series connected to the battery, which measures the current according to the Ohm's Law [6]. Two different categories of shunt resistors can be found on the market: high precision and small package. The high precision shunts often include the analog to digital front-end and directly communicate with a microcontroller using a standard communication protocol. Their cost is comparable to the Hall effect ones. Instead, small package shunts consist of the resistor only. The relevant measurement circuit must be added. Luckily, Battery Monitor ICs are usually equipped with an analog front-end able to acquire the voltage drop on a shunt resistor. The small package shunt is a solution commonly adopted as current sensor for automotive applications, with benefits in terms of dimension and cost. However, there are critical aspects that can be very deleterious in these applications, if neglected. For example, Sect. 2 reports the case of a prototypal BMS for a 12 V cranking battery, in which an unexpected battery current measurement error was observed. The paper shows that the theoretical and simulation analysis of the problem led to its identification, as reported in Sect. 3, in which the results obtained are discussed. Finally, some conclusions are summarized in the final section.

2 Case Study: 12 V Cranking Battery BMS

Our case study is a prototypal BMS for a 12 V cranking battery composed of four series-connected Lithium-Ion cells. The application requires the measurement of a maximum absolute current of 200 A. The BQ76920 IC from Texas Instruments is used as battery monitor to acquire the cell temperatures, voltages and current. This device provides two sense pins that can directly be connected to the shunt resistor terminals. The maximum recommended input voltage for the current sensing is ± 200 mV and the voltage resolution (V_{LSB}) is $8.44 \mu\text{V}$. The shunt resistor was chosen as a constantan foil to keep the shunt cost low. Its resistance R_S is $0.245 \text{ m}\Omega$. The passive balancing approach is adopted in this BMS, and two parallel-connected 22Ω bleeding resistors are used for each cell. The layout of the BMS PCB is schematized in Fig. 1(a), in which the bleeder resistors groups, the power path, the switch, the constantan shunt, and the control section of the board are represented. The balancing resistors groups are labelled with B_1 , B_2 , B_3 , B_4 , and the two crosses on the bottom side of the shunt highlight the shunt terminals and the sampling points of the voltage drop. The PCB has a standard 1.6 mm thickness and the copper layer is $105 \mu\text{m}$ thick.

Given the specifications, the current measurement I_{MEAS} for a given voltage ΔV across the shunt resistor is:

$$I_{MEAS} = \left\lfloor \frac{\Delta V + \frac{V_{LSB}}{2}}{V_{LSB}} \right\rfloor \frac{V_{LSB}}{R_S} \quad (1)$$

The prototype did not show any issue regarding the current measurement apart when balancing is activated, even without battery current. Figure 1(b) shows the measured current and the activation status of the bleeder resistor groups B_1 and B_3 , which were activated during the test. The measured current shows an unacceptable drift that introduces a huge offset error.

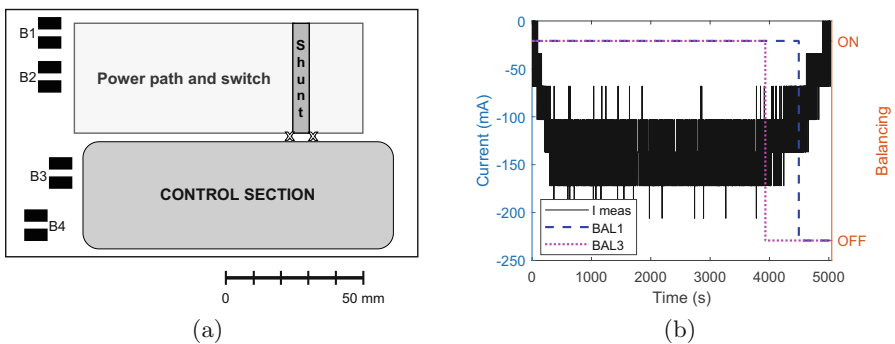


Fig. 1. PCB layout (a) and measured current and balancing status (b)

The error seems related to thermal effects since the transient is slow, with characteristic times of some minutes. Another evidence of a possible thermal

induced effect is the fact that the error fades away in comparable time as soon as the bleeding resistors are deactivated.

3 Analysis of the Shunt Current Sensor Fail

Thermal considerations are crucial for an accurate current measurement when a shunt resistor sensitive to the temperature is used. The resistance value of a resistor depends on its temperature. The variation is proportional to the Temperature Coefficient of Resistance (TCR) that depends on the resistor material. Materials with low TCR, such as constantan or manganin, are adopted as shunt to overcome the problem [7]. A proper shunt design avoids materials with high TCR or applies a temperature compensation.

However, the current measurement problem found in our prototype cannot be due to a TCR effect. The shunt is made of constantan and the error shows up with no current also. Thermoelectric effects, in particular the Seebeck effect, could be the source of the offset error. An electromotive force arises in a loop of two dissimilar conductive materials when the two junctions are at different temperatures. The Seebeck effect is well known but it is often neglected in BMS current sensors, because the electromotive force generated is generally lower than 1 mV. Moreover, the auto heating due to the Joule effect on the shunt resistor does not generate an appreciable temperature mismatch between the two resistor terminals. Instead, the Seebeck effect contribution may be significant when, as in our case study, there are other components that dissipate energy and the shunt resistor value is below 1 m Ω to reduce the voltage loss and the heat generation due to the Joule effect. The thermoelectric voltage generated in such particular conditions could be large enough to be revealed by the analog to digital front-end, and an unexpected additional current measurement error would show up. Algorithms that compensate self-heating and thermoelectric voltages have already been presented in the literature [8–10], but they are hard to implement in a BMS, mainly for cost-related reasons.

3.1 Simulation of the Case-Study

A simple model of the BMS PCB was described and simulated with the COMSOL Multiphysics software. The aim was to estimate the thermal distribution on the PCB and the related thermoelectric voltages caused by the activation of the balancing resistors. An air flux is applied to the external surfaces of the board. The selected air temperature is 298 K, as well as the initial temperature of the PCB, and the convective heat transfer coefficient is 5 Wm⁻²K⁻¹. A time transient analysis with time steps of 10 s was carried out, reproducing the balancing conditions that led to the discovery of the failure. A maximum time step of 50 s was set as additional simulation constraint for the solver. Figure 2(a) shows the thermal distribution on the surface of the board after 1 h simulated time in which B₁ and B₃ dissipate 1 W each one. The temperature gradient along the

board shows that the shunt sampling points are subjected to different temperatures and the Seebeck effect on the shunt is possible. The thermoelectric voltage between the two sampling points was calculated taking into account the Seebeck coefficient of the constantan. Then, it was converted in a current according to the shunt Ohm's Law. Finally, the thermal induced "fake" current was compared with the one acquired by the BMS and reported in Fig. 1(b). The comparison is shown in Fig. 2(b).

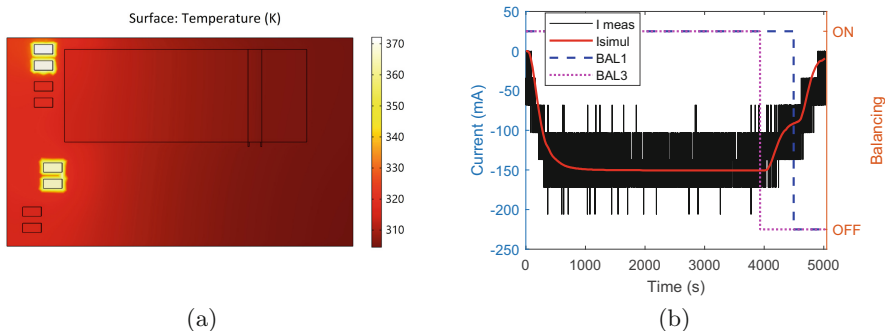


Fig. 2. Simulation results: (a) Thermal distribution on the board; (b) Current values comparison

The simulation result impressively matches the measured one. The analysis proves that the Seebeck effect on the shunt was the source of the current measurement error found when balancing was activated.

The effect of the other combinations of active balancing resistors were also simulated. Table 1 reports the temperature variation ΔT and the thermoelectric voltage ΔV between the shunt resistor sampling points after the initial transient. The thermal induced "fake" current I_{SIMUL} obtained using the Ohm's Law is also reported for all the simulated scenarios.

It should be noted that this thermal induced offset, integrated over time, could lead to a significant error in SoC estimation. This conclusion is even more true in aged battery packs, where the mismatch of the cells leads to more frequent use of balancing. It is worth to say that this offset error is mainly due to the material adopted as shunt resistor in our case study. In fact, the copper-constantan junction shows one of the highest Seebeck coefficients. On the other hand, constantan foils are cheaper than other materials, like manganin. Moreover, the constantan shunt shows very low TCR and stands as a very good choice for BMS shunts, because its resistance variations with the temperature are almost negligible in this application.

Furthermore, the simulation results suggest a correlation between the distance of an active balancing group from the power path and the thermal induced current error, as expected. The less is the distance, the more is the error. As a general rule, the shunt should be placed far away from the balancing resistors

Table 1. Effect of the balancing scenarios

Scenario				$\Delta T [^{\circ}\text{C}]$	$\Delta V [\mu\text{V}]$	$I_{SIMUL} [\text{mA}]$
B4	B3	B2	B1			
OFF	OFF	OFF	ON	-0.56	-20.67	-84.36
OFF	OFF	ON	OFF	-0.56	-20.05	-81.85
OFF	OFF	ON	ON	-1.12	-40.56	-165.56
OFF	ON	OFF	OFF	-0.47	-16.35	-66.73
OFF	ON	OFF	ON	-1.03	-36.88	-150.52
OFF	ON	ON	OFF	-1.03	-36.27	-148.03
OFF	ON	ON	ON	-1.58	-56.63	-231.16
ON	OFF	OFF	OFF	-0.43	-14.52	-59.26
ON	OFF	OFF	ON	-0.99	-35.06	-143.10
ON	OFF	ON	OFF	-0.97	-34.45	-140.61
ON	OFF	ON	ON	-1.54	-54.83	-223.79
ON	ON	OFF	OFF	-0.9	-30.76	-125.55
ON	ON	OFF	ON	-1.45	-51.16	-208.81
ON	ON	ON	OFF	-1.44	-50.55	-206.33
ON	ON	ON	ON	-1.99	-70.80	-288.96

and the other devices that dissipate power, and should be positioned in such a way to avoid a thermal gradient between its terminals.

4 Conclusions

This work has presented the analysis of an unexpected current measurement error noticed in a prototypal BMS for a 12 V cranking battery equipped with a shunt current sensor. The current measurement is always correct unless than during the charge balancing phases. The possible causes of the error that affects the SoC estimation have been analyzed, and the thermoelectric effect seems the most likely cause. The BMS board was modeled and simulated with the COMSOL Multiphysics software. The simulation highlighted that the thermoelectric effect generated by the thermal gradient due to the activation of the passive balancing system is compatible with the error found. Multiple simulations were carried out to estimate the thermal induced measurement error in all the possible balancing combinations.

Moreover, the analysis allows us to draw guidelines for a better choice of the shunt resistor for low-cost BMS that can be summarized in:

- The shunt resistor should be placed as far as possible from components that dissipate power, i.e., balancing resistors, main power switch, and fuse.
- The shunt resistor should be placed to avoid thermal gradients between its terminals in any working condition.

- Given the shunt resistance value, the shunt material should be traded-off between cost, TCR coefficient (to avoid temperature compensation), and Seebeck coefficient (to reduce thermoelectric effects).

As future work the authors will focus on the development of a design tool to estimate the thermal effect of the BMS power devices on a shunt resistor based current measurement system. The aim of this design tool is to simplify the correct placement of the power devices reducing the thermal induced current error.

References

1. Rahimi-Eichi, H., Ojha, U., Baronti, F., Chow, M.Y.: Battery management system: an overview of its application in the smart grid and electric vehicles. *IEEE Ind. Electron. Mag.* **7**(2), 4–16 (2013)
2. Di Rienzo, R., Zeni, M., Baronti, F., Roncella, R., Saletti, R.: Passive balancing algorithm for charge equalization of series connected battery cells. In: *Proceedings of the 2020 2nd IEEE International Conference on Industrial Electronics for Sustainable Energy Systems, IESES 2020*, pp. 73–79 (2020)
3. Raman, S.R., Xue, X.D., Cheng, K.W.: Review of charge equalization schemes for Li-ion battery and super-capacitor energy storage systems. In: *2014 International Conference on Advances in Electronics, Computers and Communications, ICAECC 2014*, pp. 8–13 (2015)
4. Daowd, M., Omar, N., van den Bossche, P., van Mierlo, J.: A review of passive and active battery balancing based on MATLAB/Simulink. *Int. Rev. Electr. Eng.* **6**(7), 2974–2989 (2011)
5. Ziegler, S., Woodward, R.C., Iu, H.H.C., Borle, L.J.: Current sensing techniques: a review. *IEEE Sens. J.* **9**(4), 354–376 (2009)
6. Ripka, P.: Electric current sensors: a review. *Measur. Sci. Technol.* **21**(11), 112001 (2010)
7. Destefan, D.E., Stant, R.S., Ramboz, J.D.: AC and DC shunts - can you believe their specs? In: *Conference Record - IEEE Instrumentation and Measurement Technology Conference*, vol. 2, pp. 1577–1582 (2003)
8. Weßkamp, P., Melbert, J.: High-accuracy current measurement with low-cost shunts by means of dynamic error correction. *J. Sens. Sens. Syst.* **5**(2), 389–400 (2016)
9. Ziegler, S., Woodward, R.C., Iu, H.H.C., Borle, L.J.: Investigation into static and dynamic performance of the copper trace current sense method. *IEEE Sens. J.* **9**(7), 782–792 (2009)
10. Grundkötter, E., Weßkamp, P., Melbert, J.: Transient thermo-voltages on high-power shunt resistors. *IEEE Trans. Instrum. Meas.* **67**(2), 415–424 (2018)



Microaggregation Optimisation Through Random Cluster Shuffling

Armando Maya-López¹, Fran Casino², Agusti Solanas¹,
and Antoni Martínez-Ballesté¹(✉)

¹ Department of Computer Engineering and Mathematics, Universitat Rovira i Virgili, Av. Països Catalans 26, 43007 Tarragona, Catalonia, Spain
{armando.maya, agusti.solanas, antoni.martinez}@urv.cat

² Athena Research Center, Athens, Greece
francasino@unipi.gr

Abstract. The Internet of Things facilitates the collection of large amounts of data: sensors, smartphones, and even home appliances, generate a data deluge about individuals, their context and the events in their daily life. Providers can analyse these data in order to extract patterns and increase knowledge about their services, either on their own or by transferring datasets to third parties. To mitigate the Big Brother effect, *i.e.* to preserve the individuals' right to privacy, techniques in the scope of Statistical Disclosure Control (SDC) must be applied. Microaggregation, is one of the best-known methods in the SDC arena. However, its results are far from optimal. In this paper, we introduce Random Cluster Shuffling, a new post-processing method that aims at improving the results of microaggregation techniques. We describe the proposal and present some results that support the potential of our approach.

Keywords: Microaggregation · Post-processing · Statistical disclosure control

1 Introduction

The Internet of Things (IoT) paves the way for the deployment of large and complex highly-sensorised scenarios that aim at offering a wide range of services: from efficient environment monitoring and control, to smart transportation, including ambient assisted living, smart buildings, cognitive health and other areas that will be a reality in the coming years. At the low-level of an IoT application, the *perception layer*: *i.e.*, a layer of sensors that acquire data at a high rate. At the top-level, the concept of *cloud*, that stores and provides access to the data, while offering a variety of services. 5G communication technologies are the icing on the cake: beyond providing smartphone users with low-latency, high-bandwidth communications, they will enable the rapid interconnection of virtually thousands of devices. In consequence, an unprecedented amount of big data, typically in the form of multivariate microdata, will be generated by the IoT ecosystem.

Notwithstanding, *per se*, datasets are not useful, unless they are analysed using techniques like data mining (*e.g.*, to determine behaviour of attributes, to identify patterns, etc.), process mining (*e.g.*, to discover processes, to check the conformance between existing process models and those reflected in the data, etc.) and, ultimately, feeding machine learning systems. Companies can analyse *their* data on their own, but they can also delegate (*i.e.* release) datasets to third parties. Note that sensitive information can be inferred from records and values in the dataset: consumer habits, location tracking, health issues, etc. In order to mitigate the *Big Brother* effect, data protection regulations aim at protecting data misuse, specially those related to individuals. In this line, individuals must be clearly informed about the use and lifecycle of the data. In addition, techniques like anonymisation or pseudonymisation, *i.e.*, replacing personally identifiable information (*e.g.*, name, car plate number, electrical supply contract numbers, etc.) by artificial identifiers, must be applied before releasing data [1].

1.1 Microaggregation and MDAV

Although identifiers are replaced or removed, there is some potential risk of linking confidential information to identities, either by inferring knowledge using *quasi-identifiers* (*i.e.*, a set of parameters that, when combined, can lead to the identification of an individual) from the “protected” dataset or by using record linkage techniques. A naive approach to solve this issue would be to encrypt the whole dataset but, unfortunately, searching and computing over encrypted data is not straightforward and, hence, data utility is clearly reduced.

Statistical Disclosure Control (SDC) attempts to balance utility and privacy, by preserving the statistical properties of a dataset, while minimising the risk of linking confidential information to individuals. Microaggregation is a family of SDC methods for microdata which consists of building a k -partition (by creating clusters or groups of records, with each one with between k and $2k-1$ elements) and then replacing each original value by the average (*i.e.*, centroid) of the values in the corresponding cluster.

Optimal microaggregation is defined as the one yielding a k -partition maximizing the within-clusters homogeneity. The *sum of squared errors* (SSE) is commonly used for measuring the homogeneity in each cluster, accumulating, for all clusters, the difference between each record and the centroid (*i.e.*, the average) of its corresponding cluster:

$$SSE = \sum_{i=1}^c \sum_{j=1}^{n_i} (x_{i,j} - \hat{x}_i)(x_{i,j} - \hat{x}_i)',$$

where $x_{i,j}$ is the j -th record in cluster i , and \hat{x}_i is the average record of cluster i , in a k -partition of c clusters, each one having n_i records. Hence, microaggregation techniques pursue the finding of a k -partition with minimum SSE but, unfortunately, optimal microaggregation is an NP-hard problem [2]. Along with the SSE, a normalised measure of *information loss* (IL) is also used to evaluate microaggregation: IL is the ratio between SSE and the *total sum of squares*

(SST), *i.e.*, an accumulated sum of errors computed as if only a single cluster exists.

There is a wide variety of heuristics to suboptimally solve the microaggregation problem, being the Maximum Distance to Average Vector (MDAV), proposed by Domingo-Ferrer et al. [3] one of the most well-known methods. It provides a sub-optimal k -partition at $\mathcal{O}(n^2)$ computational cost by iteratively creating clusters of k members considering the furthest records from the dataset centroid.

1.2 Contribution and Plan of the Paper

This paper presents *Random Cluster Shuffling* (RCS), a new post-processing technique aiming at improving MDAV's results. Hence, in a first step, the dataset will be microaggregated using MDAV; in a second step, RCS will be applied to improve microaggregation. Section 2 describes the post-processing method. Preliminary results are shown in Sect. 3. Finally, Sect. 4 concludes the paper.

2 Random Cluster Shuffling

Cluster optimisation is a topic widely explored in the literature; specifically, its application to the microaggregation problem (where cluster size is constrained), has been explored in [4], in which the Step 1 of our proposal is inspired. The main disadvantage of cluster post-processing techniques is their fast convergence into local minima, a fact related to their greedy nature. RCS modifies the basic algorithm to overcome such disadvantage by creating an event that, with a certain probability of occurrence, modifies the clusters created during the initial post-processing step to minimize the SSE. RCS, described in Algorithm 1, consists of the following steps:

1. For each element of the k -partition, the algorithm evaluates if extracting it from its current cluster C and assigning it to the nearest one improves the SSE. According to C 's cardinality, two situations can occur: if $\text{card}(C) = k$, we dissect the whole cluster and perform the previous evaluation for each of its records; if $\text{card}(C) > k$, we will only evaluate the current record¹.
2. After the evaluation of each record, and given a shuffling probability and a maximum number of events, a shuffling event may occur. If such event occurs, one of the clusters created is randomly selected (C_i).
3. The cluster whose centroid is nearest to C_i is selected. Both clusters will be merged into a single cluster, in which the records will be sorted with respect to the centroid of the new cluster.

¹ Our method differs from [4] in the order in which clusters are evaluated. In the former, clusters are selected sequentially, while in the latter, the next cluster is the one with the highest SSE.

4. The new cluster can contain between $2k$ and $4k-2$ records and, thus, the cluster will be divided into new clusters, so as to satisfy the cardinality constraints of optimal microaggregation. The resulting formed clusters replace the merged ones.
5. The stopping condition of the algorithm evaluates the SSE improvement at each iteration and, if it is below 0.0001, the algorithm finishes.

Algorithm 1. Random Cluster Shuffling

```

D : Dataset with n p-dimensional data points
k : Minimum cardinality constraint
C : Matrix with clusters from microaggregated dataset
S : Shuffling event probability
N : Shuffling event maximum number
repeat
  for record in D do
     $C_i(\text{record}) \leftarrow \text{SearchRecordCluster}(C, \text{record})$ 
    if  $\text{card}(C_i(\text{record})) < k$  then
      DeleteRecord(C, record)
       $\text{cent} \leftarrow \text{ComputeCentroid}(C)$ 
       $C \leftarrow \text{AddRecordToNearestCentroid}(C, \text{cent}, \text{record})$ 
    else
       $c_i \leftarrow \text{BreakCluster}(C_i(\text{record}))$ 
       $\text{cent} \leftarrow \text{ComputeCentroids}(C)$ 
       $C' \leftarrow \text{AssignElementsToNearestCentroids}(C, \text{cent}, c_i)$ 
      if  $\text{SSE}(C) \geq \text{SSE}(C')$  then
         $C \leftarrow C'$ 
      end if
    end if
  if RandomEvent(S) and EventCount  $\leq N$  then
     $C_i \leftarrow \text{SelectRandomCluster}(C)$ 
     $C_j \leftarrow \text{SelectNearestCluster}(C, C_i)$ 
     $C_m \leftarrow \text{MergeClusters}(C_i, C_j)$ 
     $C_m \leftarrow \text{SortElementsFromCentroid}(C_m)$ 
     $C_s \leftarrow \text{SplitIntoClustersOfSizeCloseToK}(C_m, k)$ 
     $C \leftarrow \text{ReplaceByNewCluster}([C_i, C_j], C_s)$ 
    EventCount ++
  end if
end for
until NoSignificantImprovement
 $D' \leftarrow \text{ReplaceClustersByCentroid}(D, C)$ 
return (D')

```

3 Experiments

We used three benchmark datasets in our experiments. The Census dataset was obtained using the public *Data Extraction System of the U.S. Bureau of the Census* and contains 1,080 records with 13 numerical attributes. In addition, we used two datasets composed of OpenStreetMap GPS traces collected from two different cities, Barcelona and Madrid. The Barcelona dataset contains 969 records with 30 GPS points. In the case of the Madrid dataset, it contains 959 records with 30 GPS points. In both trajectory-based datasets, each record

Table 1. Outcomes of the different settings of RCS applied after the MDAV clustering. The “average” column denotes the percentage of IL (the best results for each k and dataset are highlighted in green). Gray rows correspond to the outcomes of the original MDAV method.

Dataset	Method	Shuffling probability	Max events	k = 3		k = 5		k = 10	
				Average	σ	Average	σ	Average	σ
Census	MDAV	NA		5.692	NA	9.088	NA	14.156	NA
	MDAV - RCS	0	0	5.483	NA	8.450	NA	12.774	NA
		1/1000	10	5.538	0.051	8.299	0.032	12.446	0.028
		1/1000	20	5.617	0.066	8.401	0.068	12.918	0.064
		10/1000	10	5.559	0.050	8.513	0.052	12.775	0.050
		10/1000	20	5.485	0.057	8.328	0.055	12.708	0.056
		100/1000	10	5.530	0.056	8.512	0.020	12.750	0.064
		100/1000	20	5.691	0.043	8.555	0.057	12.885	0.052
Barcelona	MDAV	NA		2.567	NA	4.285	NA	7.699	NA
	MDAV - RCS	0	0	1.682	NA	2.723	NA	4.849	NA
		1/1000	10	1.697	0.025	2.725	0.067	4.779	0.034
		1/1000	20	1.690	0.053	2.718	0.020	4.745	0.031
		10/1000	10	1.701	0.051	2.714	0.021	4.824	0.053
		10/1000	20	1.705	0.042	2.734	0.042	4.957	0.054
		100/1000	10	1.723	0.025	2.749	0.055	4.755	0.072
		100/1000	20	1.752	0.049	2.753	0.028	4.969	0.066
Madrid	MDAV	NA		3.188	NA	5.288	NA	8.611	NA
	MDAV - RCS	0	0	2.634	NA	4.218	NA	7.591	NA
		1/1000	10	2.700	0.077	4.154	0.036	6.690	0.065
		1/1000	20	2.663	0.039	4.251	0.060	7.323	0.046
		10/1000	10	2.642	0.011	4.294	0.060	7.233	0.034
		10/1000	20	2.672	0.023	4.215	0.027	7.260	0.033
		100/1000	10	2.668	0.055	4.248	0.045	7.135	0.024
		100/1000	20	2.585	0.028	4.274	0.060	7.006	0.065

consists of 30 coordinate points representing latitude and longitude (*i.e.*, a total of 60 values per record). More details on the datasets can be obtained in [5] and [6].

We used the R package `sdcMicro` [5] for the MDAV implementation. Since our optimisation algorithm randomly selects the clusters which will shuffle elements, the experiments have been repeated 5 times. The outcomes of our post-processing approach are reported in Table 1.

From the results in Table 1 we can affirm that post-processing always improves the solution offered by the traditional MDAV method, since it allows the creation of clusters of variable size that are able to better capture the structure of the data. For $k = 3$, we observe that random events can negatively affect the configuration of the clusters, influencing the local minimum where the solution fall. For $k = 5$ and $k = 10$, the cluster’s configuration allows the impact of these modifications to be diminished. Apparently, the algorithm improves the results of the base post-processing method in configurations with higher cardinality, although this preliminary results should be extended by using different microaggregation methods to analyze if this effect is also related to the dataset distribution of records (*e.g.* if data are clustered or scattered), as seen in [4].

In general, we can observe that, while the SSE of Census and Madrid was improved substantially, the effect of the RCS post-processing method in the Barcelona dataset was outstanding, with improvements of approximately 40% over the original MDAV. Overall, we can conclude that a low probability in the occurrence of the shuffling events allows the algorithm to continue improving the SSE values without falling into a local minimum. If these events occur with a high probability, the number of cluster modifications in the early stages of post-processing is drastically increased, resulting in a potentially unrecoverable SSE increase.

4 Conclusion

In this article, we have proposed a novel post-processing method that applies a heuristic to reduce the SSE of a clustered dataset. Given a shuffling probability and a maximum number of events, our approach exchanges elements between a randomly selected cluster and its closest one to find an alternative k -partition with lower SSE. As observed in our preliminary experiments, the outcomes support the potential of our approach. Future work will focus on studying the impact of a more extensive set of parameter configurations and the use of other benchmark datasets. Moreover, we will apply our approach to different microaggregation methods to observe the relation between such methods and the potential improvement leveraged by our post-processing approach.

Acknowledgements. This work was supported by the European Commission under the Horizon 2020 Programme (H2020), as part of the project *LOCARD* (<https://locard.eu>) (Grant Agreement n. 832735), by the Spanish Ministry of Science & Technology with project IoTrain RTI2018-095499-B-C32, by the Government of Catalonia with projects 2017-SGR-896 and ACTUA 2020PANDE00103, and by Universitat Rovira i Virgili with project 2017PFR-URV-B2-41.

References

1. Zigomitros, A., Casino, F., Solanas, A., Patsakis, C.: A survey on privacy properties for data publishing of relational data. *IEEE Access* **8**, 51071–51099 (2020)
2. Domingo-Ferrer, J., Sebé, F., Solanas, A.: A polynomial-time approximation to optimal multivariate microaggregation. *Comput. Math. Appl.* **55**(4), 714–732 (2008)
3. Domingo-Ferrer, J., Martínez-Ballesté, A., Mateo-Sanz, J., Sebé, F.: Efficient multivariate data-oriented microaggregation. *VLDB J.* **15**, 355–369 (2006)
4. Solanas, A., Pujol, G., Martínez-Ballesté, A., Mateo-Sanz, J.M.: A post-processing method to lessen k -anonymity dissimilarities. In: 2008 Third International Conference on Availability, Reliability and Security, pp. 1060–1066 (2008)
5. Templ, M.: Statistical disclosure control for microdata using the R-package *sdcmicro*. *Trans. Data Priv.* **1**(2), 67–85 (2008)
6. Maya-López, A., Casino, F., Solanas, A.: Improving multivariate microaggregation through hamiltonian paths and optimal univariate microaggregation. *Symmetry* **13**(6) (2021)



Preliminary Design of a Flexible Test Station for Second-Life Battery Development

Andrea Carloni¹(✉), Stefano Consta², Manlio Pasquali², Federico Baronti¹, Roberto Di Rienzo¹, Roberto Roncella¹, and Roberto Saletti¹

¹ University of Pisa, Dipartimento di Ingegneria dell'Informazione,
Via Caruso 16, 56122 Pisa, Italy

{andrea.carloni,federico.baronti,roberto.dirienzo,
roberto.roncella,roberto.saletti}@unipi.it

² ENEA, Dipartimento Tecnologie Energetiche e Fonti Rinnovabili, Via Anguillarese,
301, 00123 Roma, Italy

{stefano.consta,manlio.pasquali}@enea.it

Abstract. Second-life batteries are storage systems composed of many second-hand lithium cells picked up from batteries that reached the end-of-life in high power and energy applications, such as electric vehicles. Indeed, at least 80% of the available capacity is still available and can be reused in other applications that require lower energy and power densities, such as in stationary energy storage systems. Although this concept can produce environmental and economic improvements, it has not penetrated the market yet because many issues must be solved before the massive introduction of second life batteries. This work presents the preliminary design of a flexible test station for second-life battery modules. This station aims at providing a tool to accelerate the investigation on the second-life battery issues. The station is highly flexible because is allocated in a 19-inch rack architecture that allows seamless board swap and replacement. Indeed, the battery cells and the electronic control boards can easily and quickly be replaced by extracting the specific rack drawers and inserting new ones with different features.

Keywords: Second-life battery · Test station · Battery management system

1 Introduction

Thanks to the high power and energy density of Lithium batteries, the Electric Vehicles (EV) market is rapidly growing, and 85 million electric cars are expected worldwide in 2030 [1]. Generally, an EV battery pack is considered exhausted and is replaced with a new one when it loses 20% of the rated capacity, according to the US Advanced Battery Consortium (USABC) [2]. Since the

retired battery still contains at least 80% of the original capacity [3,4], it can be reused in other applications as Second Life Battery (SLB). SLBs fit to applications that require a lower energy and power density than automotive, such as the Stationary Energy Storage Systems (SESS) [1,2]. Thus, refurbishing and reusing of SLBs may postpone the recycling process leading to many environmental and economical improvements. Indeed, recycling of lithium battery is still not mature as it is for lead-acid ones, in terms of both facilities and international rules [2]. Most of the exhausted battery packs risk of being inefficiently wasted in the landfills after the first End-Of-Life (EOL) point is reached. The scientific literature on this field is quickly increasing but many research topics are still open, and the focus on SLBs will become central for the scientific and industrial communities in the next future. For these reasons, this work presents the preliminary idea of a flexible second-life test station, which can simplify the SLB exploitation process. The station allows the investigation of different SLB configurations by simply exchanging or replacing second-life cells in a seamless way. The rest of the paper is organized as follows. Section 2 lists the main issues related to SLBs. Section 3 shows the main features and the architecture of the proposed design. Finally, Sect. 4 gives the conclusions.

2 Second-Life Battery Typical Issues

The second-life process usually starts when a battery loses 20% of its rated capacity. Then, it is removed from the first application, and the battery could be sent to facilities specialized on the refurbishing and repurposing processes. As SESSs usually contain batteries with high voltage and capacity, the second-hand cells should be grouped and series-connected in modules. In their turn, modules are series-connected to form a string, and several strings are parallel-connected to form a battery cabinet controlled by a single power system [2]. This architecture is not the best for the usable capacity, but it guarantees a high system reliability because a faulty string can be excluded without interruptions of the battery service [5]. Therefore, an SLB is generally a complex system, and its Battery Management System (BMS) is usually implemented with a hierarchical architecture. Module Management Units (MMU), String Management Units (SMU) and the Battery Management Unit (BMU) manage and control the SLB at different hierarchy levels splitting the BMS complexity [2]. However, many engineering challenges emerge during the SLB assembling process. The lack of cells with standard features makes mixing and matching them into the second-life module hard to achieve. Furthermore, the cells of a first-life module are often welded making the disassembling process difficult. Additionally, cells with characteristics as much as possible similar should be grouped into an SLB pack [6]. The available capacity, the series-resistance and the State-of-Health (SoH) are the most important characteristics used during the refurbishing process. In any case, the weakest cell will limit the overall pack performance in terms of power, charge capability and reliability [6]. Even if the grouping process is well executed, the different histories of the cells may cause dissimilar aging trends while

the SLB is working [7], and the cell parameters may consistently diverge one to the others affecting the SLB performance and reliability [8]. For these reasons, a second-life BMS should provide additional features than a normal BMS [2]. One of the most important features is the dynamic active balancing of the cells, and the execution of sophisticated estimation algorithms to accurately determine the SLB State of Charge (SoC) and SoH [1,2]. The active balancing optimizes the battery performance when the residual capacity and the SoC of the SLB cells consistently differ from one to the other. Active balancing means moving charges from the cell with the highest SoC (and residual capacity) to the one with the lowest one. Instead, estimation algorithms must use battery models that contain the electrical parameters of every cell of the pack to take into account the mismatch between them. Faulty and dangerous events can be avoided if the BMS accurately knows the SoC and SoH of each cell. Although many architectures and algorithms are presented in the literature, these topics are still open to further research [1,2]. For these reasons, the second-life BMS complexity may be higher than that of the first-life BMS. The overall system cost is consequently affected. Doubts concerning the economic viability of these repurposing processes are still ongoing. Haram *et al.* estimate the SLB final price from 44 \$/kWh to 300 \$/kWh after the review of many works [1]. Since a brand-new lithium battery cost is around 150 \$/kWh [1], the cost factor could be an important issue. Indeed, the SLB final price should be lower than 60 % of a similar brand-new battery to be appealing on the market [9]. Given the brief review reported above, our paper proposes the preliminary design of a flexible second-life test station that gives the possibility to investigate in fast and easy way many SLB frameworks, BMS functionalities and cell types. It will be shown that the term flexibility concerns both the structure of a module and the general framework of the test station. Thus, the test station may accelerate the investigation on the SLB open issues cited above.

3 Preliminary Design of the Flexible Test Station

The battery module is the basic block of the developed test framework, and it is shown in Fig. 1(a). The modules are composed of up to 12 series-connected second-life cells and the MMU. The latter is divided into three different boards: the control unit, the power electronics and the active balancing system. The first board manages the physical quantities measured on each cell, communicates the data to the upper hierarchical level of the BMS and controls the other two boards. The power electronic board contains power relays that connect and disconnect the module from the string, and it is equipped with a current sensor to measure the module current. The active balancing board dynamically moves the charge from one cell to another. Two communication buses connect the cells to the MMU control unit. The idea is that the functions routinely carried out by the MMU refer to the primary bus as communication mean. For example, the voltages, the current and the temperature readings pass through this bus. On the other hand, the secondary bus brings to the control unit advanced measurements that

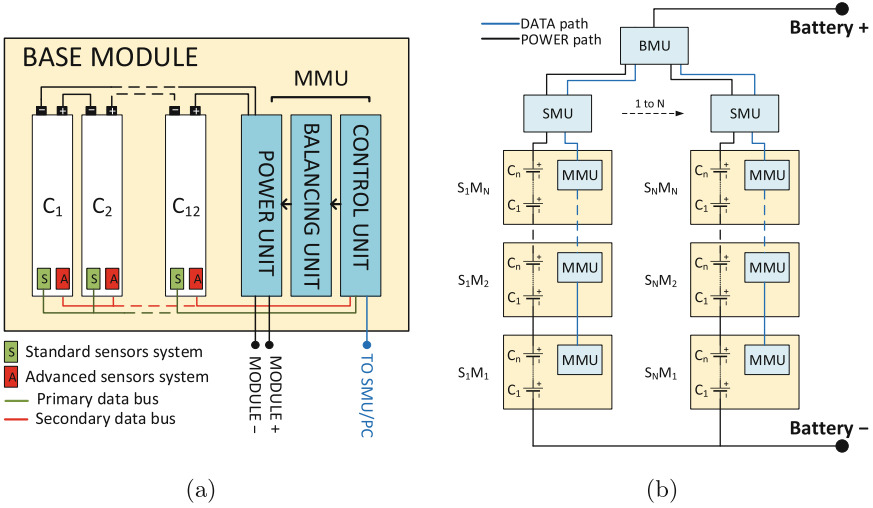


Fig. 1. A simplified scheme of the basic block of the proposed test station in (a); The flexible framework in (b).

the SLB pack could require to operate preventive strategies and avoid unsafe conditions. For example, additional temperature measurement points or swelling sensors may be inserted to identify possible dangerous degradation of the cells. Furthermore, the module can be used stand-alone by directly connecting it to a personal computer, or it can be series and parallel-connected to form the typical SLB structure depicted in Fig. 1(b). Figure 2 shows the 3D design of the developed module. It fits in a standard 19-inch rack. Each lithium cell and the three MMU boards have dedicated drawers that moves on two guideways contacting the back panel board. The back panel is provided with connectors for each drawer. The cells are series-connected with busbars on the back panel that also carries the primary and secondary communication buses described above. The module design is flexible because the operator can easily replace the single cells and the MMU boards by extracting a drawer and substituting it with another.

Therefore, a second hand 20 Ah battery previously used for peak-shaving in a cable car is chosen as case-study to size the first version of the test station. In particular, the cell holder is composed of two PCBs that form a sandwich-like structure, which embraces the pouch cell and fixes it to the drawer. The PCB under the cell provides the electrical contacts that connect the cell terminals to the back panel. The PCB over the cell fixes the cell to the other board in the sandwich by spring screws. Furthermore, it contacts the cell surface with a heat-conductive layer that helps to bring out the cell heat. The control unit implements an architecture based on a stack monitor and a microcontroller [5]. The former measures the voltage, the temperature and current of each cell and sends all of the data to the latter. The microcontroller communicates with the Personal Computer (PC) that acts as user interface, manages the stack moni-

tor, implements the battery estimation algorithms, and controls the power and balancing units. The power unit adopts a fuse and power relay to disconnect the module to the string. Furthermore, a precharge circuit limits the current inrush when the module is connected to a capacitive load. The active balancing board moves charge from the most charged cell to the least charged one with the same architecture shown in [10, 11]. The board contains a matrix of diverters that select a specific cell of the module. Finally, it is worth highlighting that the flexibility of the proposed framework allows the investigation of many configurations by simply replacing cells or boards with different functional units that share the same electrical interfaces. One or more cells can be bypassed to test the module with a different voltage. Furthermore, the battery structure in Fig. 1(b) is flexible, and sundry frameworks can be explored. For instance, several modules can also be parallel-connected and then series-connected to form a string.

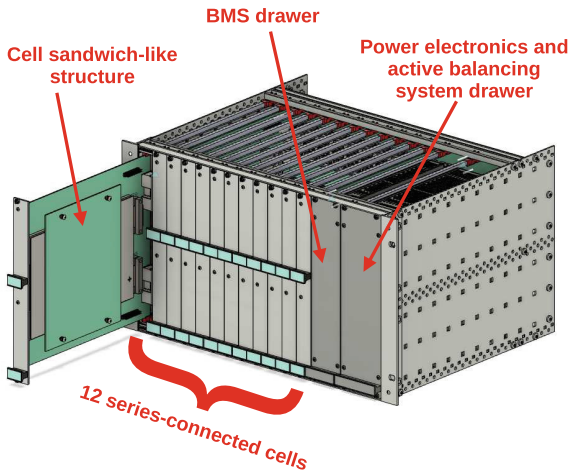


Fig. 2. 3D design of the flexible second-life battery module

4 Conclusion

This work presents the preliminary design of a second-life battery test station. The module is the main block of the proposed framework, and it is composed of up to 12 second-hand cells slots and a module management unit. Each cell and the module management unit boards are allocated into a 19-inch rack with dedicated drawers. Indeed, the operator can replace the cells or the MMU functionality by substituting the corresponding drawer. In that way, it is possible to test different second life battery configurations, study the battery system performance and find the optimal configuration with the same testing framework. The module can be series and parallel connected with the other ones to form a string reaching the requirements of a stationary energy storage system. At the same time, the module is designed to be stand-alone and can directly be connected to a personal computer that manages it. The availability of a test station

flexible and easily reconfigurable might accelerate the development of properly configured second-life battery packs. Several issues about second-life batteries have not been solved yet, and more investigations are required before the commercialization of these systems. Work is in progress for the first characterization of a module composed of second-life Li-poly pouch cells.

Acknowledgments. The project was partially funded by MiSE and ENEA in the frame of the Project “1.2 Sistemi di accumulo e relative interfacce con le reti”, and supported by CrossLab project, University of Pisa, funded by MIUR Department of Excellence program.

References

1. Haram, M.H.S.M., Lee, J.W., Ramasamy, G., Ngu, E.E., Thiagarajah, S.P., Lee, Y.H.: Feasibility of utilising second life EV batteries: Applications, lifespan, economics, environmental impact, assessment, and challenges. *Alexandria Eng. J.* **60**(5), 4517–4536 (2021)
2. Hossain, E., Murtaugh, D., Mody, J., Faruque, H.M.R., Sunny, M.S.H., Mohammad, N.: A comprehensive review on second-life batteries: Current state, manufacturing considerations, applications, impacts, barriers & potential solutions, business strategies, and policies. *IEEE Access* **7**, 73215–73252 (2019)
3. Gary Hunt, I.N.E.: *Usabc electric vehicle battery test procedures manual revision 2* (1996)
4. Wood, E., Alexander, M., Bradley, T.H.: Investigation of battery end-of-life conditions for plug-in hybrid electric vehicles. *J. Power Sour.* **196**(11), 5147–5154 (2011)
5. Rienzo, R., et al.: Low cost and flexible battery framework for micro-grid applications. In: Saponara, S., De Gloria, A. (eds.) *ApplePies 2020*. LNEE, vol. 738, pp. 246–251. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-66729-0_29
6. Warner, J.: *The Handbook of Lithium-Ion Battery Pack Design* 1st Edition. Elsevier Science (2015)
7. Martinez-Laserna, E., et al.: Technical viability of battery second life: a study from the ageing perspective. *IEEE Trans. Ind. Appl.* **54**(3), 2703–2713 (2018)
8. Heymans, C., Walker, S.B., Young, S.B., Fowler, M.: Economic analysis of second use electric vehicle batteries for residential energy storage and load-levelling. *Energy Policy* **71**, 22–30 (2014)
9. Mathews, I., Xu, B., He, W., Barreto, V., Buonassisi, T., Peters, I.M.: Technoeconomic model of second-life batteries for utility-scale solar considering calendar and cycle aging. *Appl. Energy* **269**, 115127 (2020)
10. Carloni, A., Baronti, F., Di Rienzo, R., Roncella, R., Saletti, R.: An open-hardware and low-cost maintenance tool for light-electric-vehicle batteries. *Energies* **14**(16), 4962 (2021)
11. Carloni, A., Baronti, F., Di Rienzo, R., Roncella, R., Saletti, R.: Preliminary study of a novel lithium-ion low-cost battery maintenance system. In: Saponara, S., De Gloria, A., (eds.) *ApplePies 2020*. LNEE, vol. 738, pp. 241–245. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-66729-0_28



Novel Setup to Extend the Temperature Characterization Range of a Sodium-Metal Halide Battery

Gianluca Simonte, Roberto Di Rienzo, Ian Biagioni, Federico Baronti, Roberto Roncella, and Roberto Saletti^(✉)

Dipartimento di Ingegneria dell'Informazione, University of Pisa,
Via Caruso 16, 56122 Pisa, Italy
roberto.saletti@unipi.it

Abstract. Sodium-Metal Halide Batteries might be a good candidate to compete with Lithium-ion batteries for stationary applications. However, their characteristics cannot be fully exploited by the Battery Management System if an accurate battery model to be used in the control algorithms is lacking. Electrical battery models trade off simplicity and accuracy, but they are not yet able to reproduce the Sodium-Metal Halide battery behavior in every conditions. This work presents a novel experimental setup to characterize a commercial Sodium-Metal Halide Battery in extended temperature range. The battery was deprived of its embedded Battery Management System that was replaced with a custom one to freely control the temperature. Finally, the preliminary results of characterization tests on extended temperature range are described and analyzed.

Keywords: Temperature experimental setup · Sodium metal halide battery · Zebra battery · Salt battery

1 Introduction

The continuous growth of the Lithium-Ion battery market suggests the opportunity to explore other battery technologies that may reduce the use of Lithium and support the coverage of all the future battery applications [1]. One of the most promising technology is the Sodium-Metal Halide Battery (SMHB) that has a comparable energy density with the Lithium-Ion one and is composed of cheaper and inherently safe elements [2]. The drawback is the need to work with temperatures between 250 °C and 350 °C. It is thus equipped with a heater and considerable thermal insulation solutions. Although this kind of battery is not optimal for all the applications, it could gain a significant market share, particularly in stationary applications, i.e. as energy storage for telecom and renewable energy applications [2,3], if its characteristics would be fully exploited.

The main problem of this technology is represented by the complexity of the inner chemical reactions, which require accurate and complex models to predict its behavior. Some accurate chemical [4] and mathematical [5] models have

been presented in the literature. However, their complexity does not allow affordable implementations in the Battery Management System (BMS). Therefore, the control and estimation algorithms executed in the BMS cannot rely on accurate models and all the potential features of the battery cannot be fully exploited [6]. For this reason, electrical models [7, 8], usually derived from the Lithium-Ion ones, have been studied in the last years. These models are easy to implement in the BMS [9], but they are not able to accurately model the SMHB behavior in every possible condition. Several characterization test campaigns were carried-out by different research groups in the last years. The aim was to validate and then improve the electrical model reported in Fig. 1 and presented in [8].

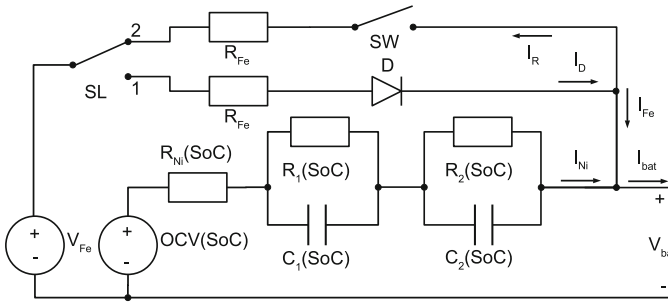


Fig. 1. Electrical model of a Sodium-Metal Halide Battery proposed in [8].

The same model was used in the experimental investigation reported in [10, 11], in which the authors developed a procedure to identify the model parameters starting from Pulsed Current Tests (PCTs) applied to the commercial battery FZSonick 48TL200. The PCT is a battery characterization test very popular and powerful. It consists of current charge/discharge pulses followed by a rest time [12].

The research works cited above neglect the dependence of the model parameters with the temperature [13]. This assumption can introduce a significant error in the model output, because the SMHB temperature shows large changes in the normal battery use [10]. However, the studies cited above use a commercial battery with its embedded BMS. The BMS independently controls and manages the battery and does not leave the possibility of carrying out tests in which the battery internal temperature is set according to desired conditions.

To overcome this problem, a new experimental setup able to completely control the commercial battery FZSonick 48TL200, by substituting its embedded BMS with a custom one was developed. We thus obtain the full control of the battery temperature, and the characterization tests can be carried out in an extended temperature range not possible with the original BMS behavior.

The experimental setup is described in Sect. 2, while Sect. 3 reports the preliminary results of the characterization test campaign carried out in extended

temperature range with the custom BMS. Finally, the paper conclusions are summarized in the last section.

2 Novel Experimental Setup

As mentioned before, the experimental setup was developed to characterize the commercial battery FZSonick 48TL200 without its BMS. This battery is composed of 5 identical strings which consist of 20 series-connected cells each. The string has a nominal capacity of 40 Ah, a nominal voltage of 48 V, and can be discharged with a maximum current of 30 A. Moreover, the battery is equipped with 3 K thermocouples to measure the internal temperature and a heater consisting of 3 resistors of $22\ \Omega$, 150 W to maintain the battery in the required operating temperature range.

The custom BMS replacing the original one must in any case guarantee that the battery works in the Safe Operating Area (SOA) of voltage, current and temperature that is equal to $40\div 53.35\ \text{V}$, $-8\div 30\ \text{A}$ (positive in discharge phase), and $265\div 350\ ^\circ\text{C}$, respectively.

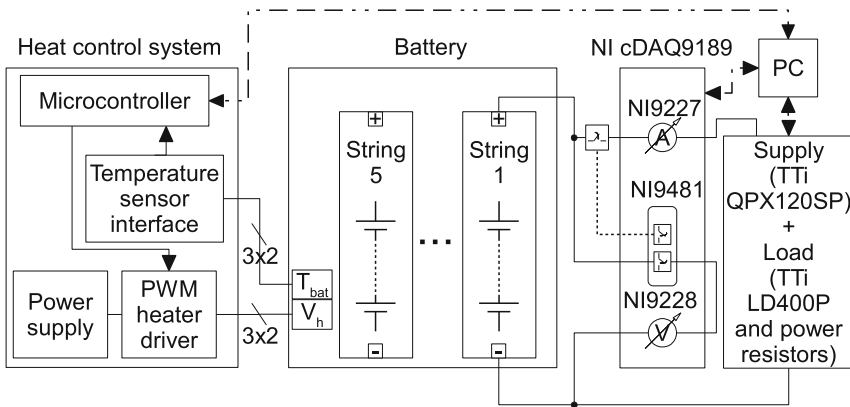


Fig. 2. Block diagram of the developed experimental setup.

The block diagram of the experimental setup is reported in Fig. 2. It was designed to control one string at a time to reduce its complexity. This is not a problem because the strings are electrically independent from each other. Instead, the heater is common to all the strings that are all maintained at the same temperature. Therefore, a custom Heat Control System was developed to control the heater. It is based on a microcontroller that acquires the 3 thermocouple signals to obtain the internal battery temperature and controls the 3 heater resistors with three Pulse-Width Modulation (PWM) signals. The duty-cycle of these signals is defined by a PID controller. The board is powered by an

external power supply that avoids draining the heater current from the battery. Moreover, the microcontroller communicates with a PC LabVIEW interface that represents the custom BMS and implements all the main safety and control features required for the safe battery operation. It disconnects the string with the NI9481 module, should the battery come out of the SOA. The LabVIEW interface also controls a power supply and an electrical load to charge and discharge the string under test, and a NI cDAQ system to measure the string voltage and current values. Thus, the obtained experimental setup introduces the valuable advantage of allowing an independent control of the battery temperature and adds an additional degree of freedom in the battery characterization experiments.

3 Characterization Tests as a Function of Temperature

As mentioned before, the PCT is one of the most used tests to characterize the battery [12]. This test consists of a series of constant current pulses, of value I_{pulse} and duration t_{pulse} , followed by a rest period of t_{rest} . The pulses are applied at constant battery temperature T_{bat} . Usually, the evolution of the battery voltage in the rest times is used to identify the parameters of the model that account for the relaxation phenomena. The constant current pulses are used to determine the battery series resistance and to gradually change the Depth of Discharge (DoD) of the battery. DoD is defined as $1 - Q/Q_n$, where Q is the charge stored in the battery and Q_n is the battery capacity. Starting from the DoD definition, I_{pulse} and t_{pulse} are chosen to discharge a specific quantity of charge that corresponds to a ΔDoD which is equal to $(I_{pulse} \cdot t_{pulse})/Q_n$. Instead, t_{rest} depends of the battery under test and is chosen large enough to acquire the battery voltage relaxation voltage after the application of the current pulses. The value of t_{rest} trades off the transient completion with the test duration.

3.1 Preliminary Tests Analysis

The results of some PCTs carried-out with different battery temperatures are reported in this section to highlight the different battery responses. These experiments may also help in the identification of a more accurate model of the battery, given the fact that the most recent models presented in the literature still lacks accuracy under certain conditions and do not take into account the possible temperature variations [11]. This preliminary PCTs were carried out taking the temperature as independent variable. Instead, I_{pulse} , ΔDoD and t_{rest} were kept constant and equal to 8 A, 10 %, and 2 h, respectively. Fig. 3 shows the string voltage and currents waveforms of three PCTs carried out at 270 °C, 300 °C and 330 °C. The tests were executed on string #1, starting from a fully charged condition. 9 pulses are applied to explore the DoD range from 0 to 90%. It is worth noting that the battery voltage response differs from test to test showing that the temperature determines non-negligible effects on the battery performance. We observe the voltage dips after the current pulse application show a dependence of the series resistance with the temperature. The relaxation phenomena also

show different responses, particularly with the last current pulses that explore DoD values above 70%. Therefore, the preliminary conclusion we can draw is that the battery models currently adopted need to be upgraded with some more or less sophisticated temperature-dependent parameters.

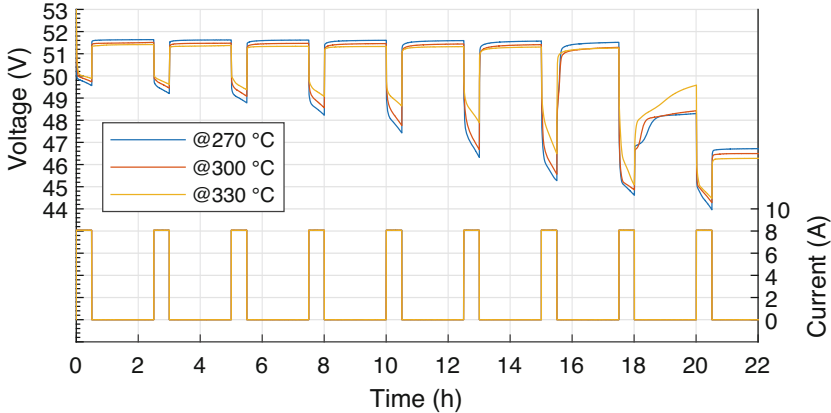


Fig. 3. 3 PCTs showing the different battery behavior as a function of temperature

It is not straightforward to justify the behavior shown by the experiments. Surely, deeper and extensive tests are needed. The test parameters I_{pulse} , ΔDoD , t_{rest} and the battery temperature, as finally allowed by the novel setup, should be varied to explore the battery response in a wider set of operating conditions, to allow the improvement of the presented battery models or even propose new and more accurate ones.

4 Conclusion

This paper shows the experimental setup used to characterize a commercial Sodium-Metal Halide Battery in which the embedded Battery Management System has been removed and replaced with a custom one, based on a microcontroller board and a LabVIEW application running on a PC. The replacement allows setting up and keeping the battery temperature at an established value and to explore the battery response in all the temperature working range. The custom BMS guarantees battery operation in the Safe Operating Area of voltage, current and temperature, without risks. The modified battery system is inserted in an experimental setup that allows the execution of charge/discharge profiles with an accurate monitoring of the battery electrical quantities. Finally, the novel setup is used to carry out characterization tests at different battery temperatures. The preliminary experimental results show that the battery response depends on the battery temperature in both the series resistance and the time-dependent part of the circuit. Thus, the presently adopted battery models need

to be upgraded to take temperature into account. The explanation of the temperature dependent phenomena showed in the tests needs a larger collection of data at various temperatures, which the novel experimental setup is able to provide.

References

1. Narins, T.P.: The battery business: Lithium availability and the growth of the global electric car industry. *Extract. Ind. Soc.* **4**(2), 321–328 (2017)
2. Benato, R., et al.: Sodium-nickel chloride (Na-NiCl₂) battery safety tests for stationary electrochemical energy storage. In: 2016 AEIT Int. Ann. Conf. (AEIT), 1–5 (2016). <https://doi.org/10.23919/AEIT.2016.7892756>
3. Gaillac, L., Skaggs, D., Pinsky, N.: Sodium nickel chloride battery performance in a stationary application. In: INTELEC, International Telecommunications Energy Conference (Proceedings) (2006)
4. Eroglu, D., West, A.C.: Modeling of reaction kinetics and transport in the positive porous electrode in a sodium-iron chloride battery. *J. Power Sour.* **203**, 211–221 (2012)
5. Bracco, S., Delfino, F., Trucco, A., Zin, S.: Electrical storage systems based on Sodium/Nickel chloride batteries: a mathematical model for the cell electrical parameter evaluation validated on a real smart microgrid application. *J. Power Sour.* **399**(January), 372–382 (2018)
6. Boi, M., Battaglia, D., Salimbeni, A., Damiano, A.: A novel electrical model for iron doped-sodium metal halide batteries. *IEEE Trans. Ind. Appl.* **55**(6), 6247–6255 (2019)
7. O’Sullivan, T.M., Bingham, C.M., Clark, R.E.: Zebra battery technologies for the all electric smart car. In: International Symposium on Power Electronics, Electrical Drives, Automation and Motion, 2006. SPEEDAM 2006, vol. 2006, no. November, pp. 244–248 (2006)
8. Boi, M., Battaglia, D., Salimbeni, A., Damiano, A.: Non-linear electrical model for iron doped sodium metal halides batteries. In: 2018 IEEE Energy Conversion Congress and Exposition. ECCE **2018**, pp. 2039–2046 (2018)
9. Sun, K., Shu, Q.: Overview of the types of battery models. In: Proceedings of the 30th Chinese Control Conference, CCC 2011, pp. 3644–3648 (2011)
10. Baronti, F., Di Rienzo, R., Roncella, R., Simonte, G., Saletti, R.: Experimental characterization of a commercial sodium-nickel chloride battery for telecom applications. In: Lecture Notes in Electrical Engineering **627**, pp. 285–291 (2019)
11. Di Rienzo, R., Simonte, G., Biagioni, I., Baronti, F., Roncella, R., Saletti, R.: Experimental Investigation of an Electrical Model for Sodium-Nickel Chloride Batteries. *Energies* **13**(10), 2652 (2020)
12. Huria, T., Ceraolo, M., Gazzarri, J., Jackey, R.: High fidelity electrical model with thermal dependence for characterization and simulation of high power lithium battery cells. In: *2012 IEEE International Electric Vehicle Conference, IEVC 2012* (2012)
13. Benato, R., Dambone Sessa, S., Necci, A., Palone, F.: A general electric model of sodium-nickel chloride battery. In: AEIT 2016 - International Annual Conference: Sustainable Development in the Mediterranean Area, Energy and ICT Networks of the Future (2016)



An Effective Approach to the Cross-Border Exchange of Digital Evidence Using Blockchain

Pablo López-Aguilar^{1,2}  and Agusti Solanas^{1,2} 

¹ Department of Computer Engineering and Mathematics,
Universitat Rovira i Virgili, Av. Països Catalans 26,
43007 Tarragona, Catalonia, Spain

² Anti-Phishing Working Group - Europe (APWG.EU), Av. Diagonal 621-629,
08028 Barcelona, Catalonia, Spain
pablo.lopezaguilar@apwg.eu

Abstract. Cybercrime has become a prevalent threat over the last decade. Despite the numerous efforts devoted by national and supranational institutions, disparities on regulations along with the decentralisation of information networks have increased the complexity for judiciary forces, and all actors involved in the investigation process, of fighting against cybercriminals. With the aim to increase the security level of the actors involved in judicial processes and assure the proper collection and integrity of digital evidence, in this article we discuss how blockchain technology could help to ensure the chain of custody throughout the flow of forensic analyses. Likewise, besides the successful instruments promoted by the EU aimed at facilitating cross-border exchange of data, we identify further actions to improve the exchange of digital evidence among all entities involved in the investigation. In this context, we highlight the approach presented by the EU project LOCARD that provides a collaborative and distributed platform to automate the management of digital evidence with blockchain technology guaranteeing, therefore, the integrity and transparency of the cross-jurisdictional chain of custody.

Keywords: Digital evidence · Chain of custody · Forensic flow · Blockchain

1 Introduction

The presentation of digital evidence in a court of justice has always been challenging. The non-localised nature of cybercrime along with the very different regulations to handle digital information represent an arduous challenge for the European Union. The European Investigation Order (EIO) introduced by the Directive 2014/41 [1] or the guideline provided by the Convention on Cybercrime [2] are some examples of successful instruments aimed at helping countries to achieve a common framework fostering cross-border cooperation. In this

The authors are supported by the EU with project LOCARD (G.A no. 832735).

line, the European NIS Directive [3] provides legal measures to comprehensively enhance the level of cybersecurity in the EU. This regulatory instrument seeks to foster the strategic cooperation and exchange of information among Member States. However, despite these efforts, the disparities on how forensic reports are presented to different stakeholders along with the lack of standardisation on cross-border forensic procedures result in a considerable room for improvement.

This article addresses current issues and challenges in the management of digital evidence across the forensic workflow. The rest of the article is organised as follows: Sect. 2 presents some of the current problems with sharing digital evidence among jurisdictions and comments on some of the most relevant instruments aiming to facilitate cross-border data exchange. Section 3 discusses the use of blockchain for sharing digital evidence and highlights the European project LOCARD [4] as a solution to face some of the current challenges in this field. Finally, Sect. 4 finishes with some recommendations to guarantee evidence integrity for presentation in courts of justice along with suggestions to perform further research in the context of digital forensics.

2 The Problem of Sharing Digital Evidences

The proper handling of digital forensic procedures for crimes investigations is extremely relevant in our digitized world. Crimes such as financial frauds, intellectual property theft and even terrorism have become transnational and require efficient regulations to improve the cooperation and successful prosecution among forensic investigators, Law Enforcement Agencies (LEAs) and judicial actors from different countries and jurisdictions. However, the lack of unified cross-border forensics procedures along with the legal ambiguities between jurisdictions do not facilitate the handling and admissibility of digital evidence in the courts of justice.

With the aim to address the several challenges and facilitate the process of exchanging cross-border evidence, the internal bodies of the EU promoted the EIO to facilitate cross-border access to electronic evidence and overcome the complexity and fragmentation represented by the Mutual Legal Assistance [5] system. This single instrument replaces the different rules and systems to obtain all kinds of evidence. On the other hand, the Council of Europe is fostering the implementation of the Budapest Convention [2], which is a powerful framework to promote international cooperation between the signers of this treaty. Moreover, at a global scale, the ISO standards [6] (i.e., ISO27037, ISO 27042, and ISO 27043) provide guidance for investigators on how to perform a digital forensics analysis of evidence.

Despite these efforts and the available instruments, there are still several disparities on the management of digital evidence and presentation of forensic reports among different jurisdictions. Moreover, there is a demand for education and training of all actors involved in forensics, from legal and technical perspectives. Thus, the harmonisation of digital forensic procedures along with the development of new technologies such as blockchain to support chain of custody

management can contribute to enhance digital forensics management procedures and face some of the legal and technical challenges related to the integrity and cross-border validity of the data collection and evidence control.

3 Blockchain for Digital Evidence Exchange

The relationship between blockchain and cryptocurrencies has generated some shadows on this technology, by associating it with speculative contexts related to the volatility of some cryptocurrencies, high energy consumption in mining processes, fraud in ICO processes, or use in deceptive activities on the “Dark web”. However, as showed by recent studies [11–14], these shadows are not part of its essence, but of its misuse. The use of blockchain for chain of custody assurance offers also differentiating values. It provides traceability, security, time and cost efficiency, but it may also foster ethical values such as truthfulness, transparency, democratization or trust.

In this line and with the aim to facilitate criminal investigations and guarantee evidence integrity for presentation in a court of justice, the LOCARD project [4] is developing an innovative platform to manage digital evidence using blockchain technology. Thus, the use of an immutable chain of custody would increase trust and provide transparency across the forensic workflow.

3.1 The LOCARD Project

The European project LOCARD is inspired by the French criminologist Edmond Locard [7] whose most important contribution to forensic sciences was the Principle of Exchange generally understood as: *any human action, let alone a violent action that is a crime, cannot take place without leaving some trace.*

Following this approach, LOCARD, that started on May 2019 and will deliver its final outcomes on April 2022, provides a collaborative and distributed platform to automate the collection and documentation of every digital evidence using blockchain technology to handle potential digital evidence to be able to present them in a court of justice. In particular, the project brings an immutable chain of custody of the digital evidence for every crime under investigation, therefore, allowing the international usage of the digital evidence and accommodating independent users, jurisdictions, territories and global regulations.

The high-level architecture of the project is depicted in Fig. 1, where entities from different jurisdictions may join the platform. Thus, each entity can implement none, a single or multiple blockchain nodes based on their profile and demands. For example, although LEAs might request access to the blockchain and require one or multiple nodes (depending on the performance needed), prosecutors may only request access to the platform to analyse the information and not to upload data on the blockchain.

As highlighted in Fig. 1, an Area Coordinator (Lead LEA) is assigned to each jurisdiction and he/she is responsible for determining the access rules and permissions for other entities willing to join the platform. In addition, the Lead LEA

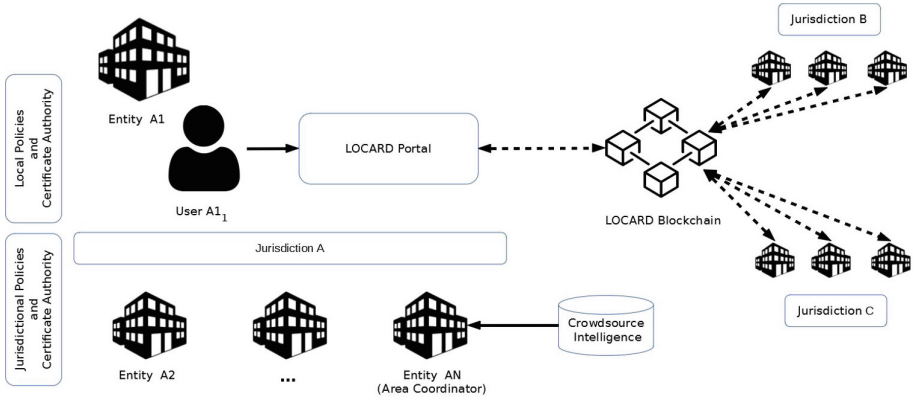


Fig. 1. LOCARD high-level architecture - figure defined in the Deliverable 3.5 “Reference Architecture” [8].

is responsible for collecting the information of the victim using the Crowdsource Intelligence Module [8] and distribute the input case through the corresponding entities. Likewise, all transactions or data recorded on the blockchain are executed by smart contracts [15,16] and aim to facilitate, verify, and negotiate the contract agreements across jurisdictions.

The implemented blockchain technology in the platform is Hyperledger [9] which is publicly available by the Linux Foundation and enables the creation of public and private channels [17]. On the one hand, the use of public channels would be recommended to foster collaboration among users investigating similar matters but without the need for disclosing any confidential information or details about the investigation. On the other hand, the use of private channels would be highly recommended for sharing data among users who have a similar level of clearance and are collaborating under the same investigation. With the aim to ensure the admissibility of the digital evidence in court, and its compliance with the legislation, the forensic flow followed in the platform corresponds to the procedures recommended by Interpol [10]. Thus, the flow of LOCARD is highlighted in Fig. 2, where dashed lines represent the blockchain interactions with the system. As depicted in the figure, each interaction with the blockchain is performed after each of the following steps:

1. Case opening: register case after request assessment,
2. Evidence registration via crawler or standalone forensic tools,
3. Evidence analysis with forensic tools,
4. Data cross checking with intelligence (optional),
5. Communication and collaboration (optional), and,
6. Report creation.

Following the previously mentioned sequence and as represented in Fig. 2, the Crowdsource Intelligence (1) seeks to collect suspicious information from the

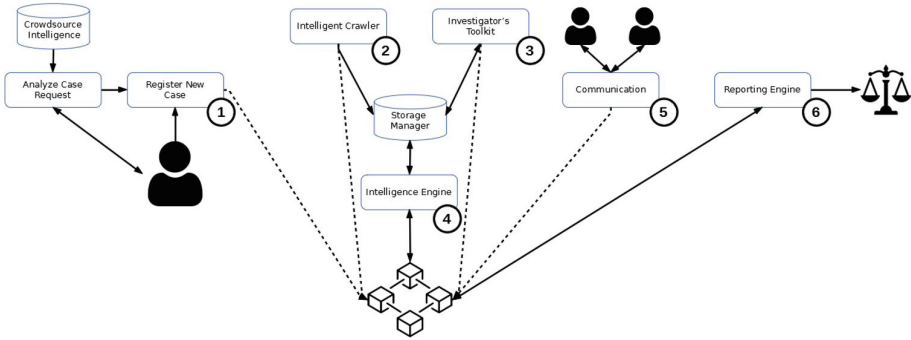


Fig. 2. LOCARD forensic flow - figure defined in the Deliverable 3.5 “Reference Architecture” [8].

victims (i.e., abuse in social networks or illegal streaming content) and forward it to the corresponding Lead LEA. This information is used by the Intelligent Crawler (2) to search for more specific information that is then processed by the offline forensic tools (3) (Investigator’s Toolkit) allowing, thus, investigators to perform their usual data analysis activities. With the aim to facilitate the investigation, the Intelligence Engine (4) provides intelligence and extracts correlations of the existing data. The Communication Engine (5) allows communication among LOCARD users to request specific details of the investigations.

Finally, the Reporting Engine (6) collects all transactions from the blockchain from a specific case and produces a report reflecting its status. The report might include not only the hashes of the transactions with the corresponding forensic outcomes, but also people assigned to the case, collected evidences, and all iterations performed during the investigation.

4 Conclusions and Further Work

The non-localised nature of data along with the several disparities on regulations open a new window for cybercriminals, who see endless opportunities to commit their malicious actions. As reflected in this article, the instruments brought by national and international institutions seek to adapt to this new paradigm and aim to enhance and homogenise the management and handling of digital evidence procedures. From a technical perspective, the adoption of new digital forensic technologies undoubtedly opens new benefits for the prosecution of cybercrime. Likewise, with the increasing applications of blockchain technology, the use of smart contracts to enable secure data exchange, and remove the need for a third party has become an interesting alternative to existing systems.

Although the adoption of these technologies will enable the management of these investigations in a more efficient and flexible manner, the high sensitivity of information obtained in digital investigations might also raise ethical dilemmas.

Hence, the development and use of ethical-compliant technologies will require a close collaboration between legal and ethical experts.

In this line, the development of a global and trusted platform (ethically compliant) aimed at guaranteeing the integrity of a cross-jurisdiction chain of custody, will greatly facilitate cross-border investigations supporting, thus, the Budapest Convention on Cybercrime and the NIS Directive. Likewise, the LOCARD standard will facilitate the cross-border cooperation and exchange of digital evidence among all entities (in particular, LEAs) involved in the investigation and, therefore, contribute to elevate the level of security across the globe. In the mid term, the implementation of a “*golden standard*”, promoted by a global body, aimed at unifying forensic procedures could be also seen as a promising solution to solve most of the current disparities in this context.

References

1. Directive 2014/41/EU of the European Parliament and of the Council of 3 April 2014 regarding the European Investigation Order in criminal matters. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32014L0041>
2. Budapest Convention and related standards. <https://www.coe.int/en/web/conventions/full-list/-/conventions/treaty/185>
3. Directive (EU) 2016/1148 of the European Parliament and of the Council of 6 July 2016. <https://digital-strategy.ec.europa.eu/en/policies/nis-directive>
4. LOCARD: Lawful evidence cOLlecting and Continuity pLAtfoRm Development. <https://locard.eu>
5. Mutual Legal Assistance. <https://ec.europa.eu/info/law/cross-border-cases/judicial-cooperation/types-judicial-cooperation/mutual-legal-assistance-and-extradition>
6. ISO - Standards. <https://www.iso.org/standards.html>
7. Locard, E.: <https://www.encyclopedia.com/science/encyclopedias-almanacs-transcripts-and-maps/locard-edmond>
8. Deliverable 3.5 - Reference Architecture. <https://locard.eu/outcomes/public-deliverables/104-work-package-3>
9. Hyperledger blockchain technology. <https://www.linuxfoundation.org/projects/case-studies/hyperledger/>
10. Global Guidelines for Digital Forensics Laboratories (INTERPOL), p. 27. https://www.interpol.int/content/download/13501/file/-INTERPOL_DFL_GlobalGuidelinesDigitalForensicsLaboratory.pdf
11. Tasatanattakool, P., Techapanupreeda, C.: Blockchain: challenges and applications. In: 2018 International Conference on Information Networking (ICOIN), pp. 473–475. IEEE, January 2018
12. Xu, X., Weber, I., Staples, M.: Architecture for Blockchain Applications. Springer, Cham (2019). <https://doi.org/10.1007/978-3-030-03035-3>
13. Alladi, T., Chamola, V., Parizi, R.M., Choo, K.K.R.: Blockchain applications for industry 4.0 and industrial IoT: a review. IEEE Access **7**, 176935–176951 (2019)
14. Pilkington, M.: Blockchain technology: principles and applications. In: Research Handbook on Digital Transformations. Edward Elgar Publishing (2016)
15. Deliverable 4.3 - State of the Art report on blockchain technologies. <https://locard.eu/outcomes/public-deliverables/105-work-package-4>

16. Wang, S., Ouyang, L., Yuan, Y., Ni, X., Han, X., Wang, F.Y.: Blockchain-enabled smart contracts: architecture, applications, and future trends. *IEEE Trans. Syst. Man Cybern. Syst.* **49**(11), 2266–2277 (2019)
17. Hyperledger Fabric channels. <https://hyperledger-fabric.readthedocs.io/en/release-2.2/channels.html>



TinyML Platforms Benchmarking

Anas Osman, Usman Abid, Luca Gemma^(✉), Matteo Perotto,
and Davide Brunelli

Department of Industrial Engineering, University of Trento, 38123 Povo, Italy
{[anas.osman](mailto:anas.osman@unitn.it),[usman.abid](mailto:usman.abid@unitn.it),[luca.gemma](mailto:luca.gemma@unitn.it),[matteo.perotto](mailto:matteo.perotto@unitn.it),
[davide.brunelli](mailto:davide.brunelli@unitn.it)}@unitn.it

Abstract. Recent advances in state-of-the-art ultra-low power embedded devices for machine learning (ML) have permitted a new class of products whose key features enable ML capabilities on microcontrollers with less than 1 mW power consumption (TinyML). TinyML provides a unique solution by aggregating and analyzing data at the edge on low-power embedded devices. However, we have only recently been able to run ML on microcontrollers, and the field is still in its infancy, which means that hardware, software, and research are changing extremely rapidly. Consequently, many TinyML frameworks have been developed for different platforms to facilitate the deployment of ML models and standardize the process. Therefore, in this paper, we focus on benchmarking two popular frameworks: Tensorflow Lite Micro (TFLM) on the Arduino Nano BLE and CUBE AI on the STM32-NucleoF401RE to provide a standardized framework selection criterion for specific applications.

Keywords: TinyML · Microcontrollers · Tensorflow Lite Micro · CUBE AI · IoT

1 Introduction

Machine Learning (ML) is at the forefront of innovation in technical and scientific applications, creating new insights on new and existing applications. Typically, running and analyzing large amounts of data on a complex ML algorithm requires a significant amount of resources and capabilities, which have been existing as barriers to the mainstreaming of ML in industry. However, with the improvement of powerful and energy-efficient embedded devices, ML inference is possible at the edge, and enables data analysis on the device as an alternative to the data exchange between servers and devices for decision making [1]. Initially, the field of edge ML focused on mobile inference, which ultimately led to several improvements for machine learning models such as quantization, sparsity, and pruning [2]. Recently, as IoT systems became mainstream, industry interest in extending Edge ML to microcontrollers grew to create a whole new potential for Edge ML as TinyML. The main goal of TinyML is to deploy ML models on ultra-low power devices to perform inference and achieve robust performance

while breaking the power consumption barrier that has previously hindered such systems. TinyML eliminates the need for cloud server connectivity and improves responsiveness and privacy measures while running using a coin size battery.

Furthermore, the field is emerging and still in its infancy, with potential for innovative and state of the art applications to unlock its full potential [3]. Nonetheless, TinyML is already being implemented in many applications to provide smarter sensor technology that enables advanced monitoring to improve productivity and safety in many sectors. For example, predictive maintenance and monitoring of wind turbines is normally a cumbersome task as in most cases these turbines are located in remote areas and failures result in a long downtime. However, when predictive maintenance is implemented, downtime is significantly reduced, resulting in noticeable cost savings and an overall increase in quality and reliability. In [4], an Australian start-up company has developed a novel IoT device that can autonomously and simultaneously monitor the turbine during operation. The device is able to detect and report potential problems before they occur in the turbine’s system. TinyML is also widely used in smart agriculture, as it is done for example by the Plant Village team, which developed an app that helps farms detect and treat potential diseases that affect crops [5]. In the health field, Solar Scare Mosquito focused on developing an IoT robotic platform that uses low-power, low-speed communication protocols to detect and warn of a potential mass breeding of mosquitoes [6].

The contribution of this paper is driven by the need to provide a standard framework and platform for TinyML use cases to build a foundation that drives the development of ML on edge devices. In particular, a comparison is made between two popular frameworks: Tensorflow Lite Micro (tested on an Arduino Nano BLE) and CUBE AI (tested on an STM32 NucleoF401RE) based on two TinyML applications.

This paper is structured as follows: Section 2 presents a summary overview of TinyML frameworks. In Sect. 3, we provide a complete breakdown of benchmarking setting and tools implemented. Finally, the benchmarking is applied by comparing the two frameworks in Sect. 4 and conclusions are drawn in Sect. 5.

2 TinyML Frameworks

Due to the unlimited potential and great interest in TinyML to revolutionize various industries, many libraries and tools are constantly being developed and deployed to facilitate the implementation of ML algorithms on constrained platforms. TinyML frameworks can be divided into three different categories.

The most trivial approach refers to converting existing trained models to overcome MCU limitations. Thus, these tools typically use inference tools derived from well-known ML libraries such as TensorFlow [11], Scikit-Learn [12] or PyTorch [13], and port their code to run on devices with scarce resources.

The second category is based on the implementation of ML libraries specifically designed for MCUs to provide them with offline training and inference capabilities. It allows models to generate from data retrieved on the go from

their device, which can immensely improve the model’s accuracy and enables the implementation of unsupervised learning algorithms.

Finally, the last technique relies on the possibility of integrating a fully dedicated co-processor to support the main computing unit in ML-specific tasks. This strategy allows for an increase in computing power, although it is the least common approach because it significantly increases the price and complexity of processing platforms. Big tech companies are helping expand the TinyML ecosystem by contributing to open source development libraries [14]. Google has launched TensorFlow Lite for microcontrollers [15], which includes a set of tools to optimize TensorFlow models to make them run on mobile and embedded devices. The key to reducing the size and complexity of the framework is that it keeps only important features on the platform and eliminates less important ones. For instance, it fails to perform a full training of a model, but capable of making inferences on models that have already been trained on a cloud computing platform. This framework is based on two elements: the model converter, which converts TensorFlow models into optimized binary code that can be used on low-power MCUs, and the model interpreter, which executes the code generated by the converter.

The optimized models, which support a range of algorithms from the NN class, can run on several platforms, including smartphones, embedded Linux systems, and MCUs. In the case of MCUs, the optimized code is written in C++ and requires 32-bit processors. It has been successfully deployed on devices, such as the Arduino Nano, and other architectures, such as the ESP32 with ARM Cortex-M series processors. Given the prominence of Arduino, a special library for this platform is available through their IDE. STMicroelectronics is among the well-known electronics manufacturer that has developed specific libraries for its devices. Specifically, the STM32Cube.AI Toolkit [16] allows the integration of pre-trained NNs into STM32 ARM CortexM-based microcontrollers. It generates from the NN models provided by Tensorflow and Keras [17] STM32-compatible C code or from models in the standard ONNX format. As an interesting feature, STM32Cube.AI allows the execution of large NNs by storing weights and activation buffers in external flash memory or RAM. In addition, Microsoft has also contributed to the TinyML scene with the release of its open-source library Embedded Learning (ELL) [18].

This framework enables the design and deployment of pre-trained ML models on constrained platforms, such as ARM Cortex-A and Cortex-M based architectures like Arduino, Raspberry Pi and micro:bit. ELL acts as an optimizing cross-compiler that runs on a regular desktop computer and outputs C++ code that can be executed on the targeted single-board computer. The API of ELL can be used for both C++ and Python and uses pre-trained NN models provided by the Microsoft Cognition Toolkit (CNTK). The toolkit ARM-NN was introduced by ARM for integrating ML into their devices [19]. In addition to open source offerings, some institutions and companies have also launched privately licensed products. The Fraunhofer Institute for Microelectronic Circuits and Systems (IMS) has developed Artificial Intelligence Library for Embedded Systems

(AifES) running on even the smallest microcontrollers [20]. However, despite the variety of frameworks presented, they focus on only one type of ML algorithm, namely NN. Researchers and industry leaders have recently considered other ML techniques, such as decision trees, Naive Bayes classifier, k-Nearest Neighbors (k-NN), and others. For example, MicroML [21] is a novel technique that allows porting Support Vector Machine (SVM) and Relevance Vector Machine (RVM) algorithms to C code that can be used on a variety of MCUs, e.g. Arduino, ESP8266, ESP32 and others with C support. It supports the widely used scikit-learn toolkit and converts models generated by this library for use on 8-bit microcontrollers with 2 KB of RAM. A similar tool is m2cgen [22], which can transform the data from models formed with scikit-learn into native code, e.g. Python, C, Java. In this case, both the number of compatible algorithms and target programming languages is even larger than in m2cgen. Table 1 summarizes the main features of the frameworks considered in this section.

Table 1. Framework comparison

Framework	Algorithms	Compatible platforms	Output languages	External libraries	Availability
TFLM	Neural networks	ARM Cortex-M	C++ 11	Tensor Flow	Open Source
STM Cube AI	Neural networks	STM32	C	Keras TensorFlow Lite Caffe ConvNetJs Lasagne	STM32 Devices only
ELL	Neural networks	ARM Cortex-M ARM Cortex-A	C/C++	CNTK ARMDarknet ONNX	Open Source
ARM-NN AI	Neural networks	ARM Cortex-A ARM Mali ARM Ethos	C	TensorFlow Caffe ONNX	Open Source
AifES	Neural networks	Raspberry Pi Windows (DLL) ARM Cortex-M4	C	TensorFlow Keras	Private License
MicroMLGen	SVM RVM	Arduino ESP32 Arduino ESP8266	C	Scikit-learn	Private License
m2cgen	Linear regression Logistic regression Neural networks SVM Decision tree	Multiple constrained nonconstrained platforms	Python C C # Java	Scikit-learn	Private License

3 Benchmarking Setting

Machine learning benchmarks fall somewhere on the continuous sequence between low-level and application-level evaluation. Low-level benchmarks attempt target kernels at the core of many ML performance analysis, such as matrix multiplication, but hide critical elements such as memory bandwidth or model-level optimizations. Conversely, application-level benchmarks can hide the benchmark’s goal behind other stages of the application pipeline. Our TinyML benchmark targets model inference and memory occupation. This section outlines the benchmarking setting for our two use cases. Each benchmark targets a specific use case with a different dataset, modelled on two separate targets. To perform this comparison with due diligence, we ensured that all parameters,

device specifications, data used to train the model and model architecture were identical for both platforms.

3.1 Gesture Recognition Use Case

Motivation. Gestures are expressive, meaningful body movements that involve physical movements of the fingers, hands, arms, or body to communicate and interact with the environment. New technology trends are driving the need to integrate such applications as part of smart systems to establish gesture recognition applications on tiny embedded devices.

Dataset. As for the dataset, there are a number of open-source datasets relevant to TinyML use cases. However, we build a distinct dataset similar to the technique used by authors in [23], using the inertial measurement unit (IMU) via the LSM9DS1 sensor on the Arduino Nano 33 BLE. Figure 1 shows the spectral features extracted from the dataset of characters O, H, G, and C for acceleration data in the X, Y, and Z planes. From Fig. 1 we can also identify the significant difference between the acquired data, which confirms the robustness of the dataset [23]. For each character of the 26 alphabet letters, 100 samples were acquired with a sampling frequency of 100 Hz, and each sample had an acquisition duration of 4000 ms, as shown in Fig. 2.

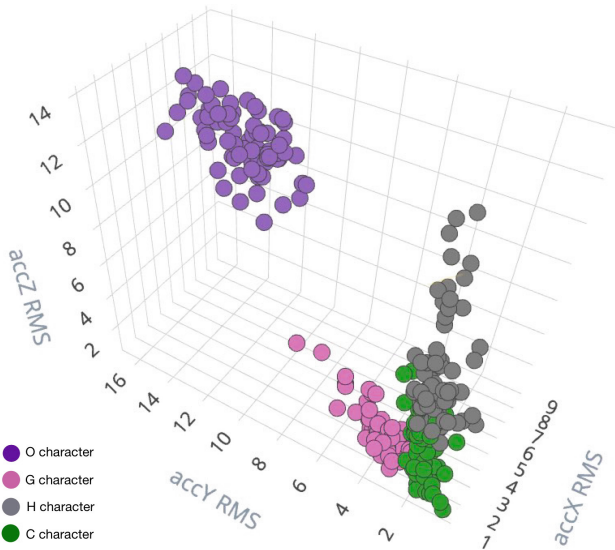


Fig. 1. Spectral features of the dataset.

Network Architecture. For our application, we opted for a Convolutional Neural Network (CNN) to train our model compatible with TinyML deployment, as illustrated in Fig. 3. The model is trained using Keras and Tensorflow Lite (TFL) libraries, which are compatible with both devices. In the case of the Arduino Nano, the model is converted to TFL using the Python API of the TFL converter. Then, our Keras model is written to disk in the form of a FlatBuffer, a special file format designed to save space [24, 25].

Model Optimization. There are two methods to choose from when optimizing a model: quantization and pruning. For the purpose of our application, we chose quantization. Quantization is still an active research topic, and there are many different options [26, 27]. With dynamic quantization from float32 to int8, we were able to achieve promising results, as the model size was significantly reduced compared to the original version. In doing so, we managed to maintain reasonable accuracy when testing from 346 KB to 275 KB for TFLM and to 192 KB for CUBE AI, while maintaining 85% accuracy.

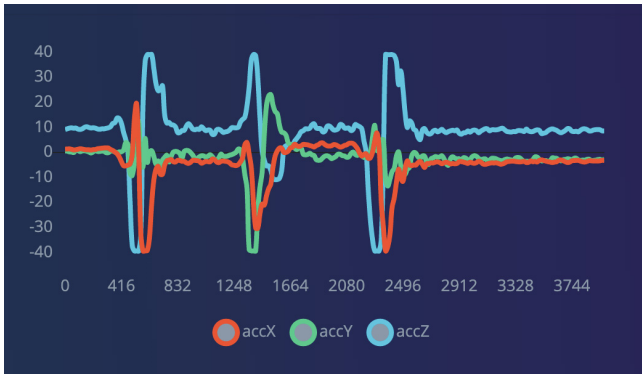


Fig. 2. N character data representation.

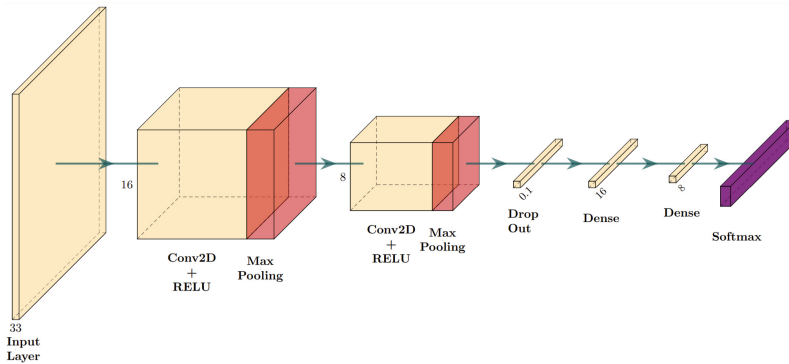


Fig. 3. Network architecture

3.2 Wake Word Spotting Use Case

Motivation. Wake Word Spotting (WWS), also known as Key Word Spotting, is a highlighted application and early use of TinyML, because voice command is an important aspect of human-machine interaction. The WWS application aims to run fully trained ML models on low-power devices to continuously monitor the environment for the wake-up word that triggers a particular functionality or service.

Dataset. Dataset acquired from an open-source voice commands dataset featuring 500 speech samples with a duration of 1000 ms for 10 different wake-up words (UP, DOWN, YES, NO, GO, STOP, LEFT, RIGHT, ON, OFF) [28].

Network Architecture. Similar to the Gesture Recognition application, we select a Convolutional Neural Network (CNN) to train our model compatible for TinyML deployment. The model is trained using Keras and Tensorflow Lite (TFL) libraries compatible with both devices.

Model Optimization. Using the dynamic quantization feature from float32 to int8, we were able to achieve promising results as the model size was significantly decreased in comparison to the unoptimized size. The post training model was reduced from 650 KB to 288 KB for TLM and 247 KB for X-Cube-AI.

3.3 Inference

After the model is converted, it is used on the two selected microcontrollers. For the Arduino platform, the C++ library for microcontrollers compatible with TFL is used to load the model and make predictions. The model is integrated as part of our applications shown in Fig. 4 for the Gesture Recognition application to infer input data and display the output through the serial port. For the STM32 platform, the X-Cube AI supports both TFL and Keras model formats, allowing great flexibility in deploying the model on the microcontroller. We also generate C code using the platform-tools to represent and allocate all the resources of the model. Then, inference is performed based on given test data representing the characters and predictions are made based on the recognised character with the highest probability, which then indicates the accuracy of the model.

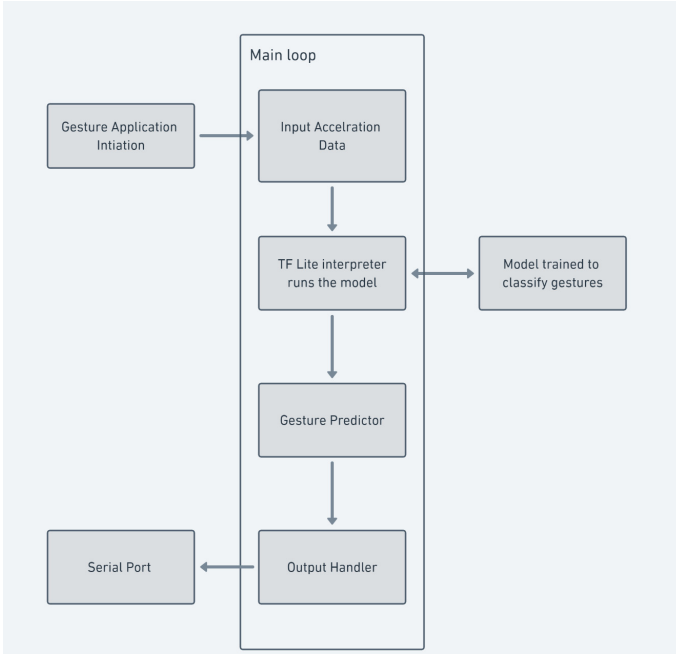


Fig. 4. Application structure

4 TFLite-Micro vs X-CUBE-AI

For the comparison between the two frameworks, we chose two different micro-controllers that support the frameworks. The Arduino Nano 33 BLE and the STM NUCLEO F401RE were selected, as indicated in Table 2. After successfully running and releasing our models on both platforms using the respective supported frameworks, as can be seen in Table 3, the flash memory required by the application deployed on the X-CUBE-AI framework is significantly smaller than the same application deployed on the Arduino platform using the TFLM framework. Moreover, the inference time on the X-CUBE-AI framework is also significantly less in the first application but almost the same as TFLM for the second application.

Table 2. Comparison between the two devices

Device	MCU	CPU	Clock	Memory	Framework
Arduino Nano BLE 33	nRF52840	32-bit ARM Cortex M4	64 MHz	1 MB	TFLM
STM32 NUCLEO-F401RE	LQFP64	32-bit ARM Cortex M4	84 MHz	512 KB	X-CUBE-AI

Table 3. Comparison between the two frameworks for the two use cases

Gesture recognition application			
Framework	Memory	DSP	Inference time
TFLM	275 KB	28 ms	30 ms
X-CUBE-AI	192 KB	5 ms	9 ms
Wake word spotting application			
Framework	Memory	DSP	Inference time
TFLM	288 KB	187 ms	193 ms
X-CUBE-AI	247 KB	162 ms	211 ms

5 Conclusion

Overall, the CUBE AI has a fairly straightforward system with a powerful interface that provides many tools for optimizing and handling the model and even generating code. On the other hand, the TFLM is more complex and requires many compromises to use the model successfully. Nevertheless, the TFLM outperforms the CUBE AI platform in terms of wide availability, because it is open-source and supports many devices. After running our two trained models on the two devices, the results show that CUBE AI performs better than the Tensorflow Lite Micro models, from the size differences to the more robust performance. However, it has the disadvantage of being supported only for STM devices and being software-oriented.

Finally, through the two discussed and compared applications in this paper, we conclude that the CUBE AI framework is better suited for memory-limited and performance-intensive TinyML applications. Future work would include further implementation of the two frameworks on more platforms through different performance demanding applications.

References

1. Merenda, M., Porcaro, C., Iero, D.: Edge machine learning for AI-enabled IoT devices: a review. *Sensors* **20**(9), 2533 (2020)
2. Han, S., Mao, H., Dally, W.J.: Deep compression: compressing deep neural networks with pruning, trained quantization and Huffman coding. arXiv preprint [arXiv:1510.00149](https://arxiv.org/abs/1510.00149) (2015)
3. Banbury, C.R., et al.: Benchmarking TinyML systems: challenges and direction. arXiv preprint [arXiv:2003.04821](https://arxiv.org/abs/2003.04821) (2020)
4. Iot device detects wind turbine faults in the field by Tomlombardo. Engineering.Com (2021). <https://www.engineering.com/story/iot-device-detects-wind-turbine-faults-in-the-field>
5. <https://grow.google/intl/europe/story/transforming-farmers%E2%80%99-lives-with-just-a-mobile-phone>
6. Solar Scare Mosquito 2.0. Hackaday.Io (2021). <https://hackaday.io/project/174575-solar-scare-mosquito-20>

7. Mitra, S., Acharya, T.: Gesture recognition: a survey. *IEEE Trans. Syst. Man Cybern. Part C (Appl. Rev.)* **37**(3), 311–324 (2007)
8. Rishikanth, C., et al.: Low-cost intelligent gesture recognition engine for audio-vocally impaired individuals. *IEEE Global Humanitarian Technology Conference (GHTC 2014)*. IEEE (2014)
9. Scherer, M., et al.: TinyRadarNN: combining spatial and temporal convolutional neural networks for embedded gesture recognition with short range radars. *IEEE Internet Things J.* **8**(13), 10336–10346 (2021)
10. Coffen, B., Mahmud, M.S.: TinyDL: edge computing and deep learning based real-time hand gesture recognition using wearable sensor. In: *2020 IEEE International Conference on E-health Networking, Application & Services (HEALTHCOM)*. IEEE (2021)
11. Abadi, M., et al.: TensorFlow: a system for large-scale machine learning. In: *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 2016)* (2016)
12. Pedregosa, F., et al.: Scikit-learn: machine learning in Python. *J. Mach. Learn. Res.* **12**, 2825–2830 (2011)
13. Paszke, A., et al.: PyTorch: an imperative style, high-performance deep learning library. *arXiv preprint arXiv:1912.01703* (2019)
14. Sanchez-Iborra, R., Skarmeta, A.F.: TinyML-enabled frugal smart objects: challenges and opportunities. *IEEE Circ. Syst. Mag.* **20**(3), 4–18 (2020)
15. David, R., et al.: TensorFlow lite micro: embedded machine learning on TinyML systems. *arXiv preprint arXiv:2010.08678* (2020)
16. Middelkamp, A.: Online. *Praktische Huisartsgeneeskunde* **3**(4), 3–3 (2017). <https://doi.org/10.1007/s41045-017-0040-y>
17. Gulli, A., Pal, S.: *Deep Learning with Keras*. Packt Publishing Ltd., Birmingham (2017)
18. *Embedded Learning Library: The Embedded Learning Library - Embedded Learning Library (ELL)*. Microsoft.Github.io (2021). <https://microsoft.github.io/ELL/>
19. *ARM-NN: ARM-Software/Armn*. Github (2021). <https://github.com/ARM-software/armnn>
20. *AIFES: Artificial intelligence for embedded systems - Aifes - Fraunhofer IMS*. Fraunhofer-Institut Für Mikroelektronische Schaltungen Und Systeme IMS (2021). <https://www.ims.fraunhofer.de/de/Geschaeftsfelder/Electronic-Assistance-Systeme/Technologien/Artificial-Intelligence-for-Embedded-Systems-AIFES.html>
21. *MicroML: Eloquentarduino/Micromlgen*. Github (2021). <https://github.com/eloquentarduino/micromlgen>
22. *m2cgen: Bayeswitnesses/M2cgen*. Github (2021) <https://github.com/BayesWitnesses/m2cgen>
23. Perotto, M., Gemma, L., Brunelli, D.: Non-invasive air-writing using deep neural network. In: *2021 IEEE International Workshop on Metrology for Industry 4.0 & IoT (MetroInd 4.0 & IoT)*. IEEE (2021)
24. Murshed, M.G., et al.: Machine learning at the network edge: a survey. *arXiv preprint arXiv:1908.00080* (2019)
25. Stanislava, S.: TinyML for ubiquitous edge AI. *arXiv preprint arXiv:2102.01255* (2021)
26. Heim, L., et al.: Measuring what really matters: optimizing neural networks for TinyML. *arXiv preprint arXiv:2104.10645* (2021)
27. Krishnamoorthi, R.: Quantizing deep convolutional networks for efficient inference: a whitepaper. *arXiv preprint arXiv:1806.08342* (2018)
28. Warden, P.: Speech commands: a dataset for limited-vocabulary speech recognition. *arXiv preprint arXiv:1804.03209* (2018)



Automatic Design Space Exploration of Redundant Architectures

Antonio Tierno^{1(✉)}, Giuliano Turri¹, Alessandro Cimatti²,
and Roberto Passerone¹

¹ University of Trento, 38123 Trento, Italy
antonio.tierno@unitn.it

² Fondazione Bruno Kessler, 38123 Trento, Italy

Abstract. In this paper, we propose an approach to automatic Design Space Exploration of redundant Embedded System architectures that is heavily based on their reliability assessment. Given a high-level description of the system, we model the deviation from its nominal behavior, and compute the set of all fault configurations, also referred to as cut-sets, in order to extract a reliability function for the architecture under analysis. We use the reliability function, together with the evaluation of other design objectives, to compare different redundant configurations, thus supporting the exploration of the design space.

Keywords: Embedded systems · Fault-tolerance · Redundancy allocation · Reliability analysis · Design space exploration · Synthesis

1 Introduction

An approach that is usually applied to increase system reliability consists in integrating the system with additional redundant components, which can take over in case of failure of the primary ones. In this paper, we extend the work proposed by Bozzano et al. [1] in order to propose a fully automated approach to the reliability assessment of complex redundant architectures and to support the Design Space Exploration (DSE) w.r.t. multiple (conflicting) design objectives. Existing techniques to analyze redundant architectures are based on Monte Carlo simulations [4], but simulation-based approaches do not provide exhaustive evaluation of the system. Other techniques widely used in the industry for reliability analysis are based on Markov Decision Process and Probabilistic Petri Nets [3], but they do not provide a completely automated process. Another technique uses MILP solvers, but may incur large runtimes [5]. We encode the problem with a symbolic technique that allows us to compare different redundant architectural configurations independently of the specific values of failure probability. We also consider other non-functional requirements such as cost, power dissipation, and size area, addressing therefore a Multi-Objective Optimization Problem (MOOP). The problem under study can be formulated as a Combinatorial Optimization Problem (COP). There is a considerable amount of research available

in the literature, but often only a portion of the problem is actually explored, either by applying simple and, for real-life problems, unrealistic constraints and objectives, referring to a custom architecture, using a static scheduling algorithm for exploration, assuming observable errors. Furthermore, crucial activities such as extraction of models, analysis of alternatives, and evaluation of properties are not yet totally automated. By using a *Model-driven* approach, complexities of ES are managed at the highest level of abstraction (system level) by using models as key artefacts throughout the development process. The proposed method is general and can be applied to similar constraint based problems that can exist in different fields.

2 Proposed Method

Given a basic (non-redundant) system defining the architecture and the behavior of a component-based system, and a finite set of available redundant patterns, we propose a *multi-objective DSE* process that *automatically* selects the appropriate set of fault-tolerant techniques to be applied to the basic system in order obtain a redundant scheme, optimizing simultaneously a collection of objective functions (see Fig. 1). Although non-functional requirements (like reliability or performance) are highly dependent on the implementation, the introduction of this abstract information at the early stages of a design can be of great benefit, by giving an indication about the impact of a redundant solution, and discarding unsuitable design points. Our method is organized in the following steps.

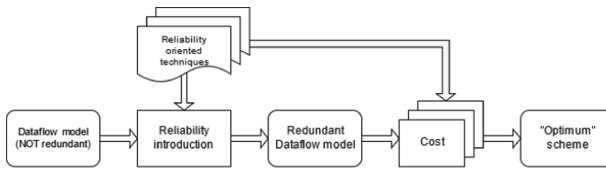


Fig. 1. Redundant system-level synthesis flow

STEP 1: Modeling the system architecture. We describe the system as a directed graph where nodes represent the components of the architecture and edges describe how components are connected.

STEP 2: Construction of a library of redundant design patterns. Widely used and proven solutions in the field of ES have to be collected from the literature, as shown in Fig. 2. The library can be considered as extensible: we have the ability to add new patterns, as well as new design objectives.

STEP 3: Modeling the redundant architectures. Each component C_i of the basic system architecture has a set of assignable patterns $Lib_{C_i} = \{P_1, P_2, \dots, P_n\}$. The tuple (C_i, P_j) denotes the assignment of pattern P_j to the component C_i that we can represent with a *configuration variable* using bit vectors with binary

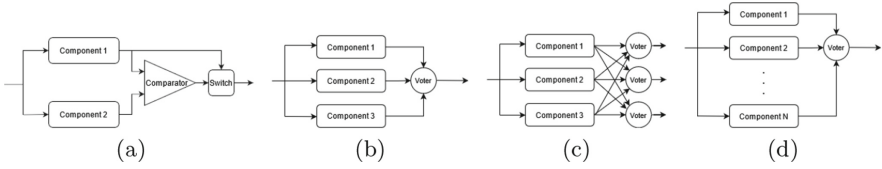


Fig. 2. Redundant design patterns: (a) comparator, (b) Triple Modular Redundancy (TMR) with one voter, (c) TMR with three voters, (d) M-oo-N.

encoding. Each configuration variable translates into a building block of the redundant architecture. Each block is characterized by one or more inputs, one or more outputs, some faulty variables, and a behavior that with our formalism will be represented as a set of Satisfiability Modulo Theory (SMT) constraints describing the input-output relationship. SMT is the problem of deciding the satisfiability of Boolean combinations of propositional atoms and theory atoms. Thus, SMT is a generalization of the Boolean satisfiability problem (SAT). Compared to SAT, SMT allows a richer representation of formulas, as it can reason about equality, linear arithmetic, bit-vectors, and other first-order theories. In order to represent the architecture, we use the theory of equality logic with uninterpreted functions (EUF) that extends Boolean logic and adds the equality predicate ($=$). Using a formal notation, a block of a redundant architecture is a tuple $\langle \vec{I}, \vec{O}, \vec{F}, \pi \rangle$ where:

- \vec{I} is the vector of inputs
- \vec{O} is the vector of outputs
- \vec{F} is the set of faults events
- $\pi(\vec{I}, \vec{O}, \vec{F})$ is an SMT formula

The single components are modeled as combinatorial elements equipped with sets of Boolean fault variables, which determine the behavior when one or more faults occur. Each module within the component has two separate behaviors: nominal (M_N) and faulty (M_F), both represented using EUF theory. Each module receives the input values of the computation (of type real), a Boolean parameter *can fail* that enables the component to have internal failures, and the *nominal behaviour* of the computation. In addition, it has a local variable *is faulty* that keeps track of the current behavior (nominal or faulty). In short, we extend each

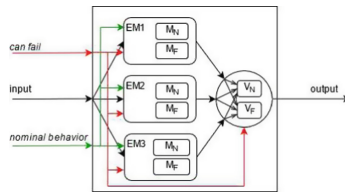


Fig. 3. Example of building block of redundant architecture

module with a fault model, obtaining an extended module (EM). Figure 3 shows the modeling technique described above applied to a TMR. By combining all valid (C, P) allocations, we obtain the set of the redundant architectures. For instance, Figure 4 illustrates the case of a system of three components, namely $\{C_1, C_2, C_3\}$, connected in series, using a library of seven patterns, namely $\{P_1, \dots, P_7\}$, assuming that $\{P_1, P_2, P_3\}$ are suitable patterns for C_1 , $\{P_4, P_5\}$ are suitable for C_2 , and $\{P_6, P_7\}$ are suitable for C_3 . Combinations of building blocks that generate the redundant architectures can be modeled in a single formula specifying that the output of one block is the input of the next according to the configuration, i.e., for each configuration and for each $block_{ij}$ connected to $block_{i'j'}$ holds that:

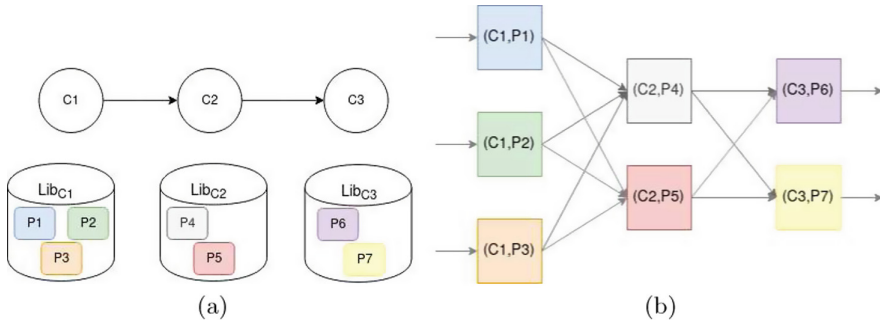


Fig. 4. (a) Components and patterns, (b) possible combinations of (C, P) allocations

$$\bigwedge_{i=1}^N ((C_i, P_j) \implies (block_{ij}.out = block_{i'j'}.in)), \forall j : j \in Lib_{C_i}, \quad (1)$$

It must be noted that the formulae involved are composed by *configuration variables* that determine selected patterns, and *fault variables* that depend on that configuration (see Fig. 5). Modeling the modules with nominal and faulty behaviors gives us the possibility to describe both reference and faulty systems. The reference architecture is instantiated by providing FALSE as *can fail* parameter to all components, while the faulty description is obtained by setting it to TRUE. By providing the same inputs to the two architectures, we can compare them by evaluating the difference in the outputs. A deviation from nominal behavior is also referred to as Top-Level Event (TLE). If the model has a single output, then TLE is the deviation between the two copies, if there are multiple outputs then the TLE is the disjunction of all the output deviations.

STEP 4: Computation of cut-sets. The system composition described above is a selective switch that allows us to generate the set of conditions that may

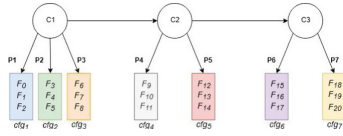


Fig. 5. Configuration and fault variables

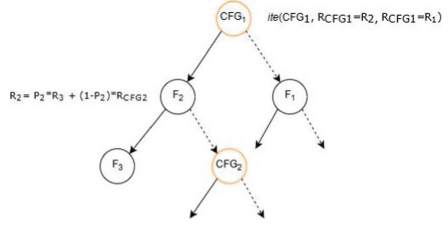


Fig. 6. Extraction of reliability formula

cause the nominal and faulty system to provide different outputs under the same inputs, i.e., the set of fault configurations and fault variables, also called *cut-sets*, that lead to wrong result. Every cut-set can be represented, via a propositional formula, as a conjunction of components’ faults, and the set of configurations as a disjunction of cut-sets. The problem of extracting the cut-sets can be therefore encoded as an AllSMT for the theory of EUF, i.e., computing all models of (1) with respect to the set of decision variables.

STEP 5: Assessment of Reliability and other non-functional parameters. We extract a symbolic reliability function, mapping the probability of fault of the basic components to the probability that the overall architecture deviates from the expected behavior. To do that, we traverse the Binary Decision Diagrams (BDD) that represents the cut-sets. Thanks to the recursive semantics of BDDs, we can calculate the fault probability of the entire system by means of the following recursive equations, where n is a BDD node, n_1, n_2 the sub-nodes, and F_q is the probability of occurrence of the fault variable (see Fig. 6).

$$Pr\{0\} = 0, Pr\{1\} = 1, Pr\{F_q\} = F_q \cdot BddProb(n_1) + (1 - F_q) \cdot BddProb(n_2)$$

For the other cost functions, the assessment is quite simple as we can consider the cost, power dissipation, and size of the redundant system equal to the sum of the cost, power dissipation, and size of the single composing redundant patterns.

STEP 6: Perform Optimization. To find the allocation (C, P) that optimizes the objective functions of the redundant system we take advantage of the existing highly-optimized SMT solvers without having to dive into their intricate implementations.

3 Experimental Evaluation

The steps described in the previous section have been implemented in Python, exploiting the pySMT [2] library for SMT formulae manipulation and solving. Figure 7(a) illustrates an example of a basic ES architecture composed of six components, each of which has two suitable redundant patterns, namely a TMR

with one voter and a TMR with three voters. The optimization problem involves a reliability function to be maximized and a cost function to be minimized. Assigning arbitrary values of fault probability and cost, our method produced eight solutions that define the best trade-off between the two competing objectives. The objective function values of these solutions are shown in Fig. 7(b) and suggested that the most reliable solution is also the most expensive, and dually the cheapest is the one with highest value of fault probability. Figure 7(c) shows the Pareto set. Figures 7(d) and 7(e) show two of the alternative solutions.

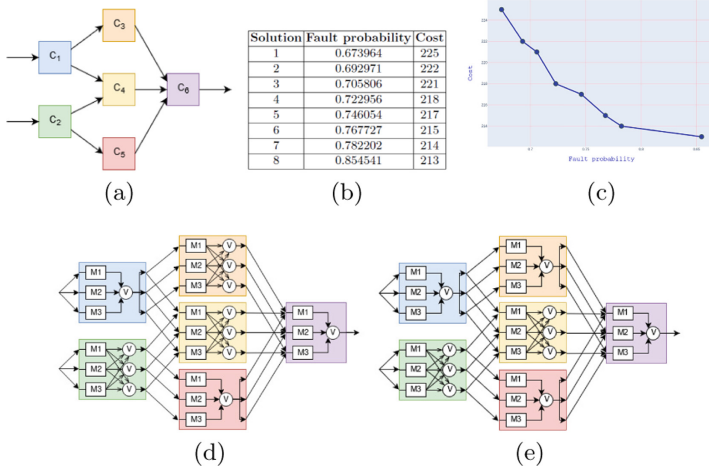


Fig. 7. Example result: (a) basic system, (b) set of solutions, (c) Pareto front (d), (e) redundant schemes (corresponding to solutions 4 and 5).

References

- Bozzano, M., Cimatti, A., Mattarei, C.: Formal reliability analysis of redundancy architectures. *Formal Aspects Comput.* **31**(1), 59–94 (2019). <https://doi.org/10.1007/s00165-018-0475-1>
- Gario, M., Micheli, A.: PySMT: a solver-agnostic library for fast prototyping of SMT-based algorithms. In: *SMT Workshop*, vol. 2015 (2015)
- Kwiatkowska, M., Norman, G., Parker, D.: PRISM: probabilistic model checking for performance and reliability analysis. *ACM SIGMETRICS Perform. Eval. Rev.* **36**(4), 40–45 (2009)
- Lee, S., Jung, J.I., Lee, I.: Voting structures for cascaded triple modular redundant modules. *IEICE Electron. Express* **4**(21), 657–664 (2007)
- Nuzzo, P., Bajaj, N., Masin, M., Kirov, D., Passerone, R., Sangiovanni-Vincentelli, A.L.: Optimized selection of reliable and cost-effective safety-critical system architectures. *IEEE Trans. Comput. Aided Des. Integr. Circ. Syst.* **39**(10), 2109–2123 (2020)



Visible Light Communication for Intermittent Computing Battery-Less IoT Devices

Alessandro Torrisi^(✉), Federico Baggio, and Davide Brunelli

Department of Industrial Engineering, University of Trento, 38123 Povo, Italy
{alessandro.torrisi,davide.brunelli}@unitn.it,
federico.baggio@studenti.unitn.it

Abstract. Eliminating the dependency on batteries as primary energy sources boosts the Internet of Things (IoT) to scale up to billion devices. New low power communication technologies combined with new energy harvesting techniques can significantly improve the energy efficiency of IoT battery-less devices. We present a communication system based on Visible Light Communication (VLC) embedded in a lighting system that enables data transmission and energy harvesting. The transmitter is a powerful LED light source. We avoid light variation and flickering of the source while transmitting data. Moreover, we embed the receiver in a cheap and compact solution by using only a single MCU and very few external components such as small-size photovoltaic cells. Experimental results show that our approach is a viable solution for powering the IoT battery-less devices of the future.

1 Introduction

The omnipresence of lighting systems in indoor environments makes Visible Light Communication (VLC) a noticeable solution for Internet of Things (IoT) applications. LED-based systems can be adopted to improve the lighting system efficiency [1] and enable communication over battery-less devices [2]. Avoiding battery dependencies as the primary energy source means overcoming the hard limit to the growth of the billion devices trend [3]. However, to consolidate the spread of battery-less devices, several aspects must be considered. First, the alternative energy source should exploit various energy harvesting technologies relying on the surrounding environmental energy sources (such as light and solar [4], radio-frequency (RF) [5], bacteria species [6]). As the energy source is not always available and can be considered as sporadic, the battery-less device must deal with frequent power failures and so intermittent operations [7, 8]. The operation is cycled through waiting for energy availability and the very action moment, when the energy is sufficient [9, 10]. Exploiting light energy to power IoT devices is a viable solution [11], even with indoor lighting systems where visible light is used as a communication channel [12, 13]. Communication over visible light is achieved by modulating the light. The challenge for lighting systems is to maintain a certain comfort level, so the modulation process must be “invisible” to human eyes. Thus, meaning the absence of light flickering and variations during data transmission.

1.1 The State-of-the-Art

IoT devices claim to communicate information towards the Internet. In battery-less devices, communication is crucial due to the ultra-low power constraints. Intermittent computing battery-less devices drastically reduce the power consumption thanks to optimized power management [4, 6, 8]. Reliable communication under ultra low power constraints is still an open challenge. One of the challenges to tackle for battery-less devices is the synchronization over the sensing network [16, 17]. In this scenario, we believe light communication would be a promising solution. Thus, we focus our preliminary investigation on VLC. LED lighting systems are suitable to achieve data transmission [18] even with high data rates [14, 20]. The second challenge to overcome is related to battery-less device communication between nodes [7]. Nowadays, light communication is enabled only between a specific illuminator and the end node even with complex modulation [13]. Furthermore, bidirectional communication is achieved thanks to visible light backscatter [15, 19]. Finally, hybrid systems combining different communication technologies are gaining momentum. Different sensors exploit different communication channels such as RF backscatter [11]. Others combine the technologies to allow devices to communicate between nodes and achieve tag-to-tag communication [12].

1.2 The Problem Statement and Contributions

Eliminating batteries brings IoT devices to new applications fields. Communicating information with light is an appealing solution to reduce the IoT device's power consumption, thanks to low power receiver circuit, and harvesting energy directly from the light communication channel. We present a system design composed of a LED transmitter (i.e., the illuminator) emulating a light bulb, and an all-in-one receiver solution based on a single microcontroller unit (MCU) and few external components such as a photodiode, or alternatively a small size photovoltaic cell, and a RC network. To accomplish this result the first task to do is to run the MCU in low power mode and reduce the overall system power consumption to the lowest 28.4 mW (i.e., 8.6 mA @ 3.3 V). Thus, allowing the system to be powered entirely with super small sized photovoltaic cells (46 mm × 15 mm). However, in dark conditions the system can still experience power failures as the only power source comes from light. The introduction of a small energy buffer (i.e. a supercapacitor) and the use of intermittent computing architecture can solve this further issue, allowing the system to be operated with memory consistencies even in the case of an intermittent power failure. Our approach covers the preliminary measurement to support communication and energy harvesting from the light communication channel. Moreover, we focus on the light modulation and coding technique to ensure a quality of service in the lighting system, avoiding light flickering and variation due to data transmission.

2 Systems Design

First, we focus on the light produced by the LED, which should not flicker during the data transmission. We also focus on simple modulation schemes to keep the receiver side as simple as possible.

2.1 VLC Transmitter

The transmitter embeds light sourcing and data into the same visible light channel. The Manchester coding of the signal prevents the light intensity from varying during data transmission, thanks to the constant average value of the symbols. The symbols are generated by modulating the LED current between a high level and a low level. These two levels must be well separated from each other, as this directly impacts the communication capability and light flickering. Moreover, the LED should always be polarized to speedup switching transients, thus we decide on two current levels different from zero, also reducing the flycker intensity.

The schematic in Fig. 1a reports the LED current modulation circuit (i.e. the VLC transmitter). The transmitter is driven by an MCU STM32L476RG, acting on the SW1 and SW2 inputs. The chosen LED is a Cree XHP50A-0S-01-0D0BJ440E¹ with secta maximum power of up to 18 W, a forward current of up to 1.5 A and a forward voltage below 11.5 V. By using a 12 V supply in combination with the LED forward voltage the overall power dissipation on the current limiting resistors is kept roughly at 5% to the power reaching the LED.

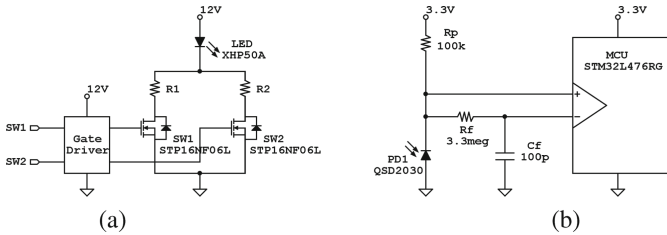


Fig. 1. Schematic of the VLC transmitter (a) and VLC receiver (b).

2.2 VLC Receiver

We embed the VLC receiver in a cheap and compact solution using only a single MCU, the STM32L476RG, and very few components (as shown in the schematic Fig. 1b). As a sensing element to receive the light signals, we select a photodiode QSD2030². The MCU internal comparator is used to discretize the analog signal generated by the photodiode. It is important to remark that the inverting input operates as a moving threshold for the comparator. Thus, we can distinguish the modulated signals even in slowly changing light intensity (e.g. due to shadows or surrounding lights turned on or off). It is important to tune the filter cut-off frequency in order to attenuate the modulation frequency. When the illuminator is transmitting data packets, the comparator output produces a sequence of high to low and low to high transitions. Finally, the MCU, through a timer and an interrupt, captures these transitions and reconstructs the message.

¹ https://www.mouser.it/datasheet/2/810/NewEnergy_XHP50_Modules_DataSheet-2326278.pdf. Accessed June 2021.

² <https://www.onsemi.com/pdf/datasheet/qsd2030-d.pdf>. Accessed June 2021

To enable energy harvesting, we also test a small sized photovoltaic cell composed of four single elements in parallel³ of 46 mm × 15 mm in total dimensions. The photovoltaic cell claims to produce up to 44 mW with optimal illumination conditions. We have to consider the possibility of operation with light sporadicity and intermittent operation. With this approach, we have to introduce an energy buffer (e.g., a supercapacitor) that charges when the node is sleeping and provides the proper energy level for the limited time activation.

3 Result

As aforementioned, the Manchester coding of the signal prevents the light intensity from varying during data transmission. Moreover, the higher the light the light modulation, the higher the received signal. However, light intensity can change between idle period and the transmission one. A trade-off between light variations is required so we performed several tests.

First, we have to consider the basic operation, thus having one switch always closed and the other providing the light modulation when data transmission is needed. At this point, it is clear a light variation appears as the average light intensity changes when transmission is activated by reaching a new higher value. The first possible way to solve this problem is to reduce the difference between the two current levels to reduce the light intensity variation between data communication and idle.

Table 1 collects the data for different tests involving different current limiting resistors. Note that during the test SW1 is always closed and SW2 is modulated when data is transmitted. Since we are interested in visual light comfort, we pay attention to light variation recorded by human eyes. Even with the smallest 4.9% current difference, the light intensity variation can be noticed to our eyes. Moreover, this configuration is particularly unfavourable as the two current levels are really close to each other, thus reducing the voltage swing at the receiver side.

A different solution is mandatory to accomplish visual comfort and a sufficient light modulation variation. We propose to send a series of alternated symbols 0 and 1 when the system is idle. Figure 2a reports the waveforms at the receiver side: V_{out} is the comparator output; V_{pd} is the photodiode cathode voltage; V_{filter} is the voltage after the RC filter. Data transmission with the preamble starts with the first longer pulses. Due to oscilloscope probe connection and loading effect ($R_f = 3.3 \text{ M}\Omega$, $R_{probe} = 10 \text{ M}\Omega$), it can be seen that the voltage after the filter is not centered with the photodiode one. Figure 2b reports evidence of a minor problem solved by the new solution. Due to the time response of the RC filter used for the moving comparator threshold, the comparator loses the first pulses until the filter reaches the steady state.

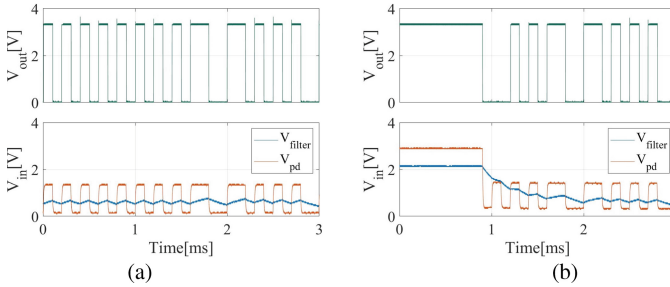
3.1 Characterization: Photodiode

We performed some measurements using the setup with two current levels (high 400 mA and low 212 mA) and a modulation frequency of 5 kHz at different distances in a range

³ <https://www.np.micro-semiconductor.hk/datasheet/b4-KXOB22-01X8F.pdf>. Accessed June 2021.

Table 1. LED currents with different transmitter configurations (supply voltage of 12 V).

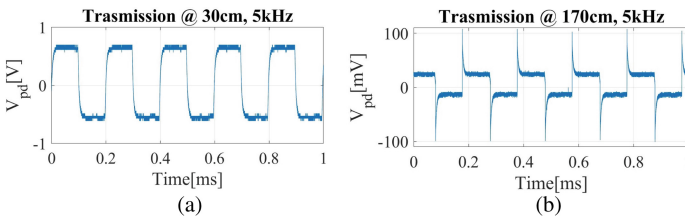
$R_1[\Omega]$	$R_2[\Omega]$	$I_{R1}[\text{mA}]$	$I_{R1+R2}[\text{mA}]$	$\Delta I[\%]$
4.7	2.2	212	400	+88.7
4.7	4.7	212	319	+50.5
3.3	4.7	267	367	+37.5
3.3	47	267	280	+4.9

**Fig. 2.** Data transmission with (a) and without (b) alternated 1 and 0 during idle.

between 30 cm and 170 cm (as an upper limit). Figure 3 reports the voltage waveforms at the receiver side and at the photodiode cathode (AC coupled). The voltage swing decreases as the distance increases. Moreover, when the distance is larger and the voltage swing is smaller, transition spikes appear due to the changes in the input bias current of the comparator.

Figure 4a reports further experiment parameters such as the received average illuminance and the average photodiode cathode voltage V_{dc} . In particular, when the distance is shorter, the photodiode current is larger, thus the average voltage is lower.

We performed a second test by fixing the distance to the largest 170 cm and using different frequencies. The results are shown in Fig. 5. There are two main problems related to the comparator that limit the maximum distance and frequency: generation of voltage spikes due to changes in the input bias current; large response time (propagation delay) of the comparator due to low overdrive voltage.

**Fig. 3.** Voltage waveform at the photodiode cathode at 30 cm (a) and 170 cm (b) with a modulation frequency of 5 kHz (AC coupled).

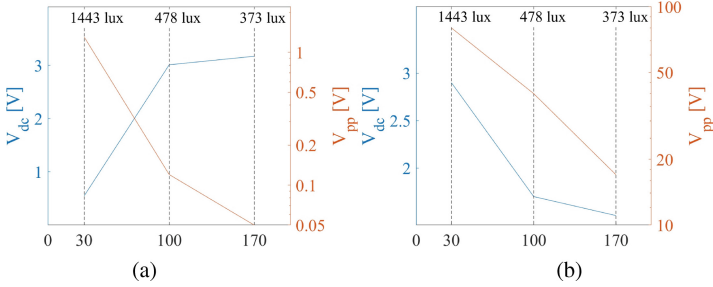


Fig. 4. Photodiode (a) and photovoltaic cell (b) average voltages (V_{dc}) and peak to peak AC values (V_{pp}) at different distances with a background of 313 lux.

In Fig. 5b, we can see that when in V_{comp} there is a high to low transition, and vice versa, a spike appears in V_{pd} . Due to the high input impedance provided by the photodiode polarization network, the comparator input bias current generates these spikes. The second problem is related to the propagation delay of the comparator due to the low input overdrive voltage. The comparator tends to become slower when the difference between the two input signals is lower. These two problems combined are shown in Fig. 5b. At high frequency, a large distance and low overdrive voltages, the comparator is too slow to properly translate the signal. For this setup, at 170 cm the frequency limit is 20 kHz. Finally, an important remark is that the peak to peak steady state photodiode voltage remains constant over the frequencies as the distance is fixed and the light intensity is not changing.

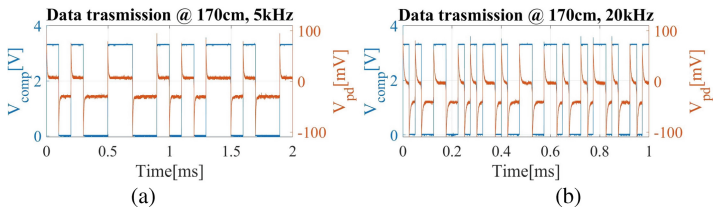


Fig. 5. Voltage waveform at the photodiode cathode and comparator output at 170 cm with different modulation frequencies: 5 kHz (a), 20 kHz (b).

3.2 Characterization: Photovoltaic Cell

Another set of tests are performed to measure the limits of the photovoltaic cell used to receive the signal and harvest energy. In Fig. 4b results at different distances are collected. The behavior of V_{dc} is the opposite with respect to the one seen for the photodiode. Comparing the distance and peak-to-peak voltage, the photovoltaic cell sensitivity is smaller than the photodiode one. For distances larger than 100 cm V_{pp} is too small to make the signal useful for the comparator.

In Fig. 6, we compare the more comfortable condition for the photovoltaic cell and the limit one. These two measurements are affected by noise but still the signal can be

detected properly. The noise contribution at roughly 30 kHz is still present even if the cell is in dark conditions; thus, it is related to the setup environment.

In conclusion, the photovoltaic cell behaves differently with respect to the photodiode, providing an energy source and increasing the output voltage. In opposition, the photodiode reveals a way faster time response, thus allowing for higher frequency modulation. However, the photodiode does not provide any energy. On the contrary, it draws current through the polarization resistor.

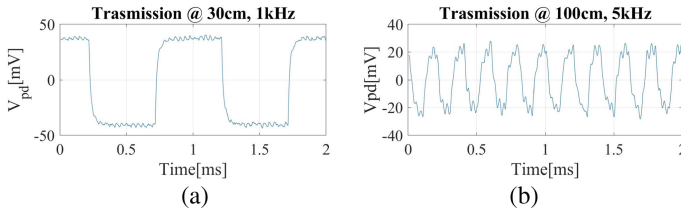


Fig. 6. Voltage waveform at photovoltaic cell: (a) best condition at 30 cm, 1 kHz, (b) worst condition 100 cm, 5 kHz

4 Conclusions

We developed a test setup for preliminary measurements on VLC in battery-less scenarios with a valid solution to accomplish the quality of service and visual light comfort. Finally, we develop a compact VLC receiver based only on a few components, comparing the results with a photodiode and a photovoltaic cell.

The future perspective is to combine VLC with other harvesting and communication technology. An example would be an autonomous wake-up system (similar to wake-up radio) to further improve system efficiency and reliability in intermittent operation. These will be the tasks for future work.

References

1. Rehman, S.U., Ullah, S., Chong, P.H.J., Yongchareon, S., Komosny, D.: Visible light communication: a system perspective—overview and challenges. *Sensors*. **19**, 1153 (2019). <https://doi.org/10.3390/s19051153>
2. Palacín, M.R., de Guibert, A.: Why do batteries fail? *Science* **351**, 1253292 (2016)
3. Nordrum, A.: The internet of fewer things. *IEEE Spectr.* **53**(10), 12–13, October 2016. <https://doi.org/10.1109/MSPEC.2016.7572524>
4. Nardello, M., Desai, H., Brunelli, D., Lucia, B.: Camaroptera: a Batteryless long-range remote visual sensing system. In: *Proceedings of the 7th International Workshop on Energy Harvesting & Energy-Neutral Sensing Systems*, pp. 8–14. Association for Computing Machinery, New York (2019)
5. Smith, J.R.: *Wirelessly Powered Sensor Networks and Computational RFID*. Springer Science & Business Media, New York (2013). <https://doi.org/10.1007/978-1-4419-6166-2>

6. Sartori, D., Brunelli, D.: A smart sensor for precision agriculture powered by microbial fuel cells. In: 2016 IEEE Sensors Applications Symposium (SAS), pp. 1–6 (2016)
7. Hester, J., Sorber, J.: the future of sensing is batteryless, intermittent, and awesome (2017). <https://doi.org/10.1145/3131672.3131699>
8. Hester, J., Sorber, J.: Flicker: rapid prototyping for the batteryless Internet-of-Things. In: Proceedings of the 15th ACM Conference on Embedded Network Sensor Systems, pp. 1–13. Association for Computing Machinery, New York (2017)
9. Balsamo, D., Weddell, A.S., Merrett, G.V., Al-Hashimi, B.M., Brunelli, D., Benini, L.: Hibernus: sustaining computation during intermittent supply for energy-harvesting systems. *IEEE Embedded Sys. Lett.* **7**, 15–18 (2015)
10. Yıldırım, K.S., Majid, A.Y., Patoukas, D., Schaper, K., Pawelczak, P., Hester, J.: InK: reactive kernel for tiny batteryless sensors. In: Proceedings of the 16th ACM Conference on Embedded Networked Sensor Systems, pp. 41–53. Association for Computing Machinery, New York (2018)
11. Varshney, A., Soleiman, A., Mottola, L., Voigt, T.: Battery-free visible light sensing. In: Proceedings of the 4th ACM Workshop on Visible Light Communication Systems (VLCS 2017), pp. 3–8. Association for Computing Machinery, New York (2017). <https://doi.org/10.1145/3129881.3129890>
12. Galisteo, A., Varshney, A., Giustiniano, D.: Two to tango: hybrid light and backscatter networks for next billion devices. In: Proceedings of the 18th International Conference on Mobile Systems, Applications, and Services (MobiSys 2020). Association for Computing Machinery, pp. 80–93. Association for Computing Machinery, New York (2020). <https://doi.org/10.1145/3386901.3388918>
13. Shao, S., Khreishah, A., Elgala, H.: Pixelated VLC-backscattering for self-charging indoor IoT Devices. *IEEE Photon. Technol. Lett.* **29**(2), 177–180 (2017). <https://doi.org/10.1109/LPT.2016.2631946>
14. Lee, Y., Lai, J., Yu, C.: The LED driver IC of visible light communication with high data rate and high efficiency. In: 2016 International Symposium on VLSI Design, Automation and Test (VLSI-DAT), pp. 1–4 (2016). <https://doi.org/10.1109/VLSI-DAT.2016.7482534>
15. Li, J., Liu, A., Shen, G., Li, L., Sun, C., Zhao, F.: Retro-VLC: enabling battery-free duplex visible light communication for mobile and IoT applications. In: Proceedings of the 16th International Workshop on Mobile Computing Systems and Applications (HotMobile 2015), pp. 21–26. Association for Computing Machinery, New York (2015). <https://doi.org/10.1145/2699343.2699354>
16. Geissdoerfer, K., Zimmerling, M.: Bootstrapping battery-free wireless networks: efficient neighbor discovery and synchronization in the face of intermittency. In: Proceedings of the 18th Symposium on Networked Systems Design and Implementation (NSDI 2021). USENIX Association. <https://www.usenix.org/conference/nsdi21/presentation/geissdoerfer>
17. Torrisi, A., Brunelli, D., ildirim, K.S.: 2020. Zero power energy-aware communication for transiently-powered sensing systems. In: Proceedings of the 8th International Workshop on Energy Harvesting and Energy-Neutral Sensing Systems (ENSsys 2020), pp. 43–49. Association for Computing Machinery, New York (2020) <https://doi.org/10.1145/3417308.3430269>
18. Teixeira, L., Loose, F., Barriquello, C.H., Reguera, V.A., Costa, M.A.D., Alonso, J.M.: Review of LED drivers for visible light communication. In: IECON 2019 - 45th Annual Conference of the IEEE Industrial Electronics Society, 2019, pp. 4274–4279 (2019). <https://doi.org/10.1109/IECON.2019.8927340>

19. Xu, X., et al.: PassiveVLC: enabling practical visible light backscatter communication for battery-free IoT applications. In: Proceedings of the 23rd Annual International Conference on Mobile Computing and Networking (MobiCom 2017). Association for Computing Machinery, New York, NY, USA, pp. 180–192. DOI:<https://doi.org/10.1145/3117811.3117843>
20. Rajagopal, N., Lazik, P., Rowe, A.: Hybrid visible light communication for cameras and low-power embedded devices. In: Proceedings of the 1st ACM MobiCom Workshop on Visible Light Communication Systems (VLCS 2014). Association for Computing Machinery, New York, NY, USA, pp. 33–38 (2014). <https://doi.org/10.1145/2643164.2643173>



Resource Optimization in MEC-Based B5G Networks for Indoor Robotics Environment

Tadeus Prastowo^(✉), Ayub Shah, Luigi Palopoli, and Roberto Passerone

Department of Information Engineering and Computer Science,
University of Trento, Via Sommarive 9, 38123 Trento, Italy
{tadeus.prastowo, ayub.shah, luigi.palopoli, roberto.passerone}@unitn.it
<https://www.disi.unitn.it>

Abstract. The deployment of fifth-generation (5G) and beyond 5G (B5G) networks is the ambitious objective of modern research on future mobile networks that are evolving to support computation-intensive and communication-sensitive applications. Such applications (e.g., autonomous vehicles, industrial automation, and remote surgery) impose diverse quality-of-service (QoS) requirements on the network in terms of processing, latency, reliability, and bandwidth, and will require ultra-reliable low-latency communication (URLLC), paving the way for multi-access edge computing (MEC). Our work considers a dynamic indoor B5G network in a robotic scenario where agents continuously need MEC services and migrate from one cell to another to perform their tasks in an ultra-dense cell environment. Assuming that every MEC service is a virtual machine (VM) to execute in one of the cells with the possibility of migrating the VM to another cell by paying some cost, we formalize the joint problem of (1) placing/migrating the VMs to respect their end-to-end communication latency requirements and (2) allocating their computation and communication bandwidth as a mixed-integer linear program (MILP). An MILP solver is then used to find the optimal VM placements/migrations and bandwidth allocations over a time horizon.

Keywords: 5G · MEC · MILP · URLLC · Real-time

1 Introduction

The 5G (fifth-generation) and B5G (beyond 5G) network infrastructures will play a vital role in modern technological evolution, including industry 4.0 and autonomous vehicles. In particular, B5G services will require diverse QoS (quality-of-service) in terms of processing, latency, reliability, and bandwidth [1] and place heavy communication and computation loads on the network as exemplified by URLLC (ultra-reliable low-latency communication) services whose domain is the convergence of different areas, including control, robotics, signal processing, artificial intelligence, and data analysis. As different areas converge

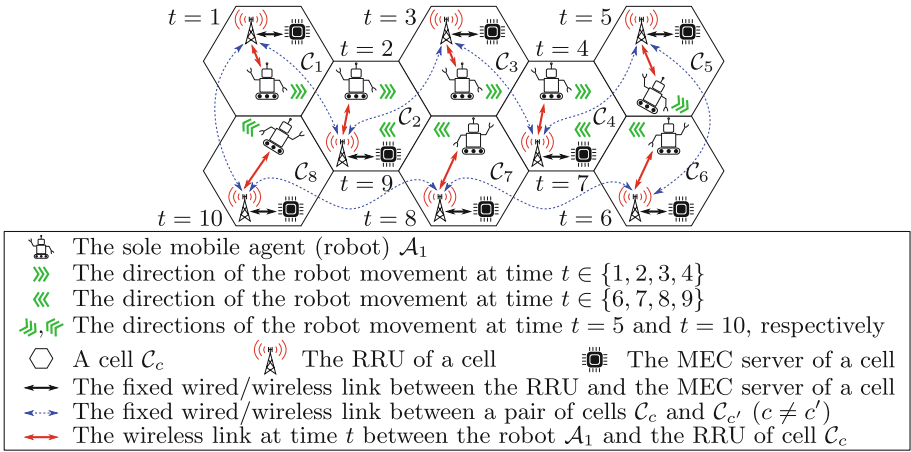


Fig. 1. A robot serviced by MEC-equipped cells finishes a circuit every 10 time units.

on URLLC services, the services are inherently heterogeneous, increasing the complexity of managing the computing and communication resources available in MEC (multi-access edge computing) servers at the network edge. Therefore, the success of deploying heterogeneous services on B5G networks crucially depends on the optimal management of MEC servers.

Many researchers have investigated MEC for offloading computation and optimizing resource allocations in static and dynamic scenarios. Markov decision process has been proposed to solve the problem of placing/migrating services and allocating their resources in an uncertain environment [2, 3]. However, the proposals do not jointly optimize the computation and communication resources at the network edge, which is crucial as the tasks offloaded by emerging applications become more considerable and heterogeneous in terms of their QoS requirements. Hence, [4–6] propose solutions for a static user environment, and closer to our research objective, [7–10] propose optimization algorithms and iterative procedures that take into account the optimal resource allocations for offloaded tasks subject to latency and energy consumption constraints when the users are mobile. However, none holistically takes into account all characteristics that we consider on an ultra-dense cell environment populated by agents with very predictable mobility that continuously offload to MEC servers (see Fig. 1) heterogeneous tasks with diverse QoS requirements.

Therefore, our contributions are twofold. First, Sect. 2 presents a novel methodology for flexibly and optimally managing task offloading in the context of RT (real-time) applications that require heterogeneous computing and communication services. Second, Sect. 3 shows an application of the methodology on the network shown in Fig. 1. Finally, Sect. 4 presents our conclusions.

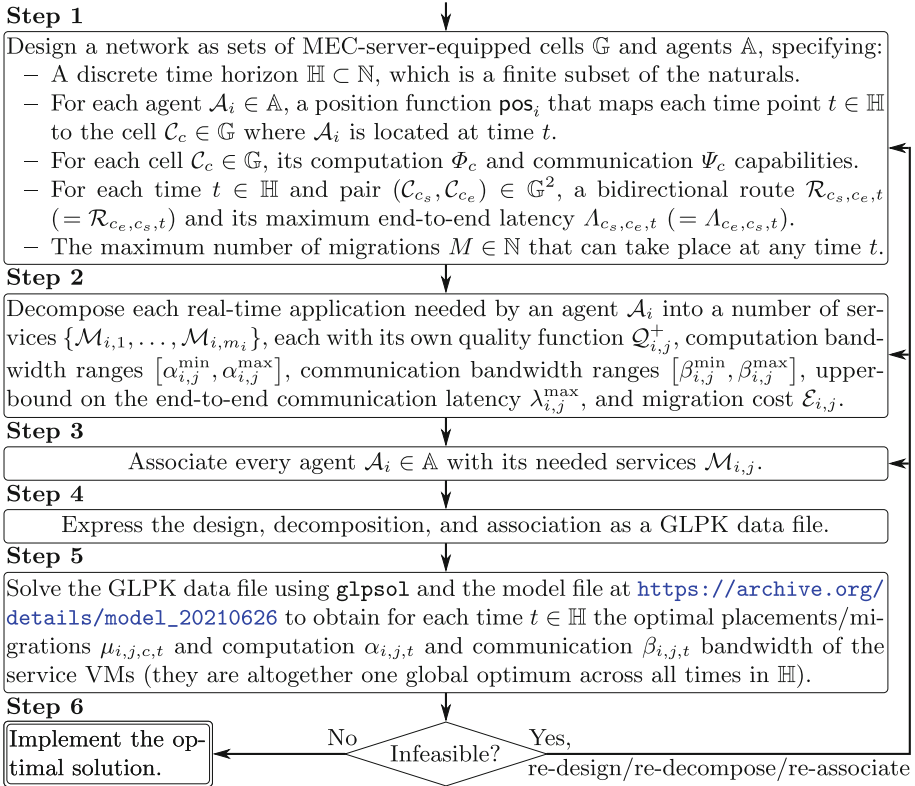


Fig. 2. Our methodology to flexibly and optimally manage network edge resources.

2 Methodology

Figure 2 shows the methodology that we propose to flexibly and optimally manage task offloading in the context of RT applications that require heterogeneous computing and communication services. The flexibility is achieved by first decomposing an RT application into a number of services, each implemented by a single VM (virtual machine), that impose different upper-bounds on their end-to-end communication latency and different ranges of computation and communication bandwidth. Then, each service quantifies its QoS as a function of the experienced end-to-end latency and allocated computation and communication bandwidth. By the quality functions quantifying the QoS of the different services, the optimality of managing task offloading can then be measured against the greatest possible total QoS achievable in the lifetime of the RT application. To illustrate the steps of the proposed methodology, the scenario depicted in Fig. 1 is used as a running example in the rest of this paper.

Table 1. One possible set of end-to-end routes for the network in Fig. 1 for all $t \in \mathbb{H}$.

$\mathcal{R}_{c_s, c_e, t}$	$c_e = 1$	$c_e = 2$	$c_e = 3$	$c_e = 4$	$c_e = 5$	$c_e = 6$	$c_e = 7$	$c_e = 8$
$c_s = 1$	$\{C_1\}$	$\{C_1, C_2\}$	$\{C_1, C_2, C_3\}$	$\{C_1, C_2, C_3, C_4\}$	$\{C_1, C_2, C_3, C_4, C_5\}$	$\{C_1, C_8, C_7, C_6\}$	$\{C_1, C_8, C_7\}$	$\{C_1, C_8\}$
$c_s = 2$	$\mathcal{R}_{1,2,t}$	$\{C_2\}$	$\{C_2, C_3\}$	$\{C_2, C_3, C_4\}$	$\{C_2, C_3, C_4, C_5\}$	$\{C_2, C_3, C_4, C_5, C_6\}$	$\{C_2, C_1, C_8, C_7\}$	$\{C_2, C_1, C_8\}$
$c_s = 3$	$\mathcal{R}_{1,3,t}$	$\mathcal{R}_{2,3,t}$	$\{C_3\}$	$\{C_3, C_4\}$	$\{C_3, C_4, C_5\}$	$\{C_3, C_4, C_5, C_6\}$	$\{C_3, C_4, C_5, C_6, C_7\}$	$\{C_3, C_2, C_1, C_8\}$
$c_s = 4$	$\mathcal{R}_{1,4,t}$	$\mathcal{R}_{2,4,t}$	$\mathcal{R}_{3,4,t}$	$\{C_4\}$	$\{C_4, C_5\}$	$\{C_4, C_5, C_6\}$	$\{C_4, C_5, C_6, C_7\}$	$\{C_4, C_5, C_6, C_7, C_8\}$
$c_s = 5$	$\mathcal{R}_{1,5,t}$	$\mathcal{R}_{2,5,t}$	$\mathcal{R}_{3,5,t}$	$\mathcal{R}_{4,5,t}$	$\{C_5\}$	$\{C_5, C_6\}$	$\{C_5, C_6, C_7\}$	$\{C_5, C_6, C_7, C_8\}$
$c_s = 6$	$\mathcal{R}_{1,6,t}$	$\mathcal{R}_{2,6,t}$	$\mathcal{R}_{3,6,t}$	$\mathcal{R}_{4,6,t}$	$\mathcal{R}_{5,6,t}$	$\{C_6\}$	$\{C_6, C_7\}$	$\{C_6, C_7, C_8\}$
$c_s = 7$	$\mathcal{R}_{1,7,t}$	$\mathcal{R}_{2,7,t}$	$\mathcal{R}_{3,7,t}$	$\mathcal{R}_{4,7,t}$	$\mathcal{R}_{5,7,t}$	$\mathcal{R}_{6,7,t}$	$\{C_7\}$	$\{C_7, C_8\}$
$c_s = 8$	$\mathcal{R}_{1,8,t}$	$\mathcal{R}_{2,8,t}$	$\mathcal{R}_{3,8,t}$	$\mathcal{R}_{4,8,t}$	$\mathcal{R}_{5,8,t}$	$\mathcal{R}_{6,8,t}$	$\mathcal{R}_{7,8,t}$	$\{C_8\}$

The design in Step 1 assumes a set of mobile agents \mathbb{A} that roam a cellular network \mathbb{G} where each cell $C_c \in \mathbb{G}$ is equipped with a MEC server. Hence, if multiple RRUs (remote radio units/heads) are served by one MEC server, then the RRUs altogether are taken as one cell. For example, the network shown in Fig. 1 has $\mathbb{G} = \{C_1, \dots, C_8\}$. The design then considers a finite discrete time horizon \mathbb{H} . For example, for the case shown in Fig. 1, using $\mathbb{H} = \{1, \dots, 11\}$ is sufficient to obtain a complete task-offloading plan as the sole agent repeats its complete tour every 10 time units. The design then specifies a position function $\text{pos}_i : \mathbb{H} \rightarrow \mathbb{G}$ for each agent $\mathcal{A}_i \in \mathbb{A}$. For example, the sole agent \mathcal{A}_1 in Fig. 1 has its position function pos_1 defined as $\text{pos}_1(t) = C_t$ if $1 \leq t \leq 6$, $\text{pos}_1(t) = C_{11-t}$ if $t \in \{7, 9\}$, and $\text{pos}_1(t) = C_{7+\frac{t-8}{2}}$ if $t \in \{8, 10\}$. Denoting the latest time in \mathbb{H} as t^* (i.e., $t^* = \max_{t \in \mathbb{H}} t$), the design reserves enough computation and communication bandwidth to allow at most M services to migrate at any time $t'' \in \mathbb{H}$ with $t'' < t^*$. In reserving the computation and communication bandwidth, the design makes it so that M migrations taking place at any time t'' complete within one time unit Δ (i.e., $t'' + \Delta = t'' + 1 \in \mathbb{H}$), which in reality can be some seconds or minutes depending on the cell sizes, agent speeds, and migration technique. Hence, when the design specifies for each cell C_c its computation capability Φ_c and its communication capability Ψ_c , the specified capabilities exclude the bandwidth reserved for migrations. For example, if the network shown in Fig. 1 reserves Δ_1 computation bandwidth and Δ_2 communication bandwidth to allow at most one migration at any time t'' (i.e., $M = 1$), and every cell C_c in the network has the same capabilities with $\Phi_c = 100$ GIPS (gigainstructions/second) and $\Psi_c = 1000$ Mbps (megabits/second), then every

cell is indeed capable of computing at $(100 + \Delta_1)$ GIPS and communicating at $(1000 + \Delta_2)$ Mbps. Afterwards, the design designates every end-to-end route $\mathcal{R}_{c_s, c_e, t}$ as a nonempty subset of \mathbb{G} . Specifically, $\mathcal{R}_{c_s, c_e, t} = \{\mathcal{C}_c\}$ refers to both the wireless link between an agent \mathcal{A}_i and the RRU of cell \mathcal{C}_c and the wired/wireless link between the RRU and the MEC server of \mathcal{C}_c . On the other hand, $\mathcal{R}_{c_s, c_e, t} \supseteq \{\mathcal{C}_{c_s}, \mathcal{C}_{c_e}\}$ refers to the wired/wireless links between either \mathcal{A}_i in \mathcal{C}_{c_s} and the MEC server at \mathcal{C}_{c_e} or \mathcal{A}_i in \mathcal{C}_{c_e} and the MEC server at \mathcal{C}_{c_s} . For example, one possible design of the routes for the mesh network in Fig. 1 is shown in Table 1. Finally, the design specifies for each route $\mathcal{R}_{c_s, c_e, t}$ its end-to-end latency $\Lambda_{c_s, c_e, t}$, which always includes the latency in the agent-RRU wireless link and, additionally if $c_s \neq c_e$, the total latency of every wired/wireless links between \mathcal{C}_{c_s} and \mathcal{C}_{c_e} . For example, the end-to-end latency of each route in Table 1 can be stated in terms of hop count so that $\Lambda_{c_s, c_e, t} = |\mathcal{R}_{c_s, c_e, t}|$ (e.g., $\Lambda_{1,1,t} = 1$ hop).

Table 2. An example of three heterogeneous services used by the agent in Fig. 1.

$\mathcal{M}_{1,j}$		$j = 1$	$j = 2$	$j = 3$
Computation	Min ($\alpha_{1,j}^{\min}$)	25 GIPS		
bandwidth	Max ($\alpha_{1,j}^{\max}$)	75 GIPS		
Communication	Min ($\beta_{1,j}^{\min}$)	250 Mbps		
bandwidth	Max ($\beta_{1,j}^{\max}$)	750 Mbps		
Latency upper-bound ($\lambda_{1,j}^{\max}$)		2 hops	4 hops	6 hops
Normalized migration cost		20%	50%	80%
Quality function	$(\mathcal{Q}_{1,j}^+(\alpha_{1,j,t}, \beta_{1,j,t}, \lambda_{1,j,t}))$	$U_{1,j}^{(1)}(m_{1,j}^{(1),1}(\alpha_{1,j,t} - L_{1,j}^{(1),0})) + C_{1,j}^{(1)} +$ $U_{1,j}^{(2)}(m_{1,j}^{(2),1}(\beta_{1,j,t} - L_{1,j}^{(2),0})) + C_{1,j}^{(2)} +$ $U_{1,j}^{(3)}(m_{1,j}^{(3),1}(\lambda_{1,j,t} - L_{1,j}^{(3),0})) + C_{1,j}^{(3)}$		
Weight and offset	$\alpha_{1,j,t}$, i.e., $k = 1$ ($U_{1,j}^{(k)}, C_{1,j}^{(k)}$)	(1, 0)		
	$\beta_{1,j,t}$, i.e., $k = 2$ ($U_{1,j}^{(k)}, C_{1,j}^{(k)}$)	(1, 0)		
	$\lambda_{1,j,t}$, i.e., $k = 3$ ($U_{1,j}^{(k)}, C_{1,j}^{(k)}$)	(1, 1)		
Interval and gradient with $l = 1$	$\alpha_{1,j,t}$, $([L_{1,j}^{(k),l-1}, L_{1,j}^{(k),l}], m_{1,j}^{(k),l})$	([25, 75], $\frac{1}{50}$)		
	$\beta_{1,j,t}$, $([L_{1,j}^{(k),l-1}, L_{1,j}^{(k),l}], m_{1,j}^{(k),l})$	([250, 750], $\frac{1}{500}$)		
	$\lambda_{1,j,t}$, $([L_{1,j}^{(k),l-1}, L_{1,j}^{(k),l}], m_{1,j}^{(k),l})$	([1, 2], -1)	([1, 4], $-\frac{1}{3}$)	([1, 6], $-\frac{1}{5}$)

The decomposition in step 2 assumes that each service communicates with no other service but an agent. Furthermore, since each service $\mathcal{M}_{i,j}$ at any time $t \in \mathbb{H}$ needs some computation bandwidth $\alpha_{i,j,t}$ and some communication bandwidth $\beta_{i,j,t}$ and expects to experience some end-to-end latency $\lambda_{i,j,t}$ in the communication between $\mathcal{M}_{i,j}$ and \mathcal{A}_i , each service specifies their bounds $\alpha_{i,j}^{\min}$, $\alpha_{i,j}^{\max}$, $\beta_{i,j}^{\min}$, $\beta_{i,j}^{\max}$, and $\lambda_{i,j}^{\max}$ such that the correct functioning of the service necessitates $\alpha_{i,j}^{\min} \leq \alpha_{i,j,t} \leq \alpha_{i,j}^{\max}$, $\beta_{i,j}^{\min} \leq \beta_{i,j,t} \leq \beta_{i,j}^{\max}$, and $\lambda_{i,j,t} \leq \lambda_{i,j}^{\max}$. Each service $\mathcal{M}_{i,j}$ then quantifies its QoS at any time $t \in \mathbb{H}$ based on $\alpha_{i,j,t}$, $\beta_{i,j,t}$, and $\lambda_{i,j,t}$ using a quality function $\mathcal{Q}_{i,j}^+$. Every quality function should map to the

same range (e.g., $[0, 1]$) so that they are comparable and every migration cost $\mathcal{E}_{i,j}$ is easily expressible in the range. For example, suppose the agent \mathcal{A}_1 shown in Fig. 1 is served by an RT application that is decomposed into three services $\mathcal{M}_{1,1}$, $\mathcal{M}_{1,2}$, and $\mathcal{M}_{1,3}$ with different QoS characteristics shown in Table 2. Since $\mathcal{Q}_{1,j}^+$ shares the same range $[0, 3]$ due to $25 \text{ GIPS} \leq \alpha_{1,j,t} \leq 75 \text{ GIPS}$, $250 \text{ Mbps} \leq \beta_{1,j,t} \leq 750 \text{ Mbps}$, and $1 \leq \lambda_{1,j,t} \leq \lambda_{1,j}^{\max}$, the migration costs are easily derived to be $\mathcal{E}_{1,1} = 0.6$, $\mathcal{E}_{1,2} = 1.5$, and $\mathcal{E}_{1,3} = 2.4$, which are 20%, 50%, and 80% of 3, respectively. On the other hand, the association in step 3 assumes that each service is used exclusively by one agent, and hence, if multiple agents need the same service, then the VM implementing the service is duplicated for each agent (e.g., if \mathcal{A}_1 and \mathcal{A}_2 need the same service, then the VM of the service is duplicated as $\mathcal{M}_{1,j}$ for \mathcal{A}_1 and as $\mathcal{M}_{2,j'}$ for \mathcal{A}_2 for some j and j').

```

1  set Hbb 1 2 3 4 5 6 7 8 9 10 11; set Gbb 1 2 3 4 5 6 7 8;
2  param Phi default 100; param Psi default 1000; param M 1;
3  # The model file states that unspecified Rcal[s,e,1] defaults to {s,e} while
4  # Rcal[s,e,t'] for every t'>1 defaults to Rcal[s,e,t'-1].
5  set Rcal[1,3,1] 1 2 3;   set Rcal[2,4,1] 2 3 4;   set Rcal[3,5,1] 3 4 5;
6  set Rcal[1,4,1] 1 2 3 4; set Rcal[2,5,1] 2 3 4 5; set Rcal[3,6,1] 3 4 5 6;
7  set Rcal[1,5,1] 1 2 3 4 5; set Rcal[2,6,1] 2 3 4 5 6;
8  set Rcal[1,6,1] 1 8 7 6;   set Rcal[2,7,1] 2 1 8 7;   set Rcal[3,7,1] 3 4 5 6 7;
9  set Rcal[1,7,1] 1 8 7;   set Rcal[2,8,1] 2 1 8;   set Rcal[3,8,1] 3 2 1 8;
10 set Rcal[4,6,1] 4 5 6;   set Rcal[5,7,1] 5 6 7;
11 set Rcal[4,7,1] 4 5 6 7; set Rcal[5,8,1] 5 6 7 8;
12 set Rcal[4,8,1] 4 5 6 7 8; set Rcal[6,8,1] 6 7 8;
13 # Unspecified RcAllambda[s,*,t'] for t'>1 defaults to RcAllambda[s,*,t'-1].
14 param RcAllambda [1,*,1] 1 1, 2 2, 3 3, 4 4, 5 5, 6 4, 7 3, 8 2
15                    [2,*,1] 2 1, 3 2, 4 3, 5 4, 6 5, 7 4, 8 3
16                    [3,*,1] 3 1, 4 2, 5 3, 6 4, 7 5, 8 4
17                    [4,*,1] 4 1, 5 2, 6 3, 7 4, 8 5
18                    [5,*,1] 5 1, 6 2, 7 3, 8 4
19                    [6,*,1] 6 1, 7 2, 8 3
20                    [7,*,1] 7 1, 8 2
21                    [8,*,1] 8 1;
22 param pos [1,*,1] 1 1, 2 2, 3 3, 4 4, 5 5, 6 6, 7 4, 8 7, 9 2, 10 8, 11 1;
23 set Mbb (1,*) 1 2 3; param Ecal [1,*,1] 1 0.6, 2 1.5, 3 2.4;
24 param alpha_min default 25; param alpha_max default 75;
25 param beta_min default 250; param beta_max default 750;
26 param lambda_max [1,*,1] 2, 2 4, 3 6;
27 param U [1,1,*] 1 1, 2 1, 3 1 [2,1,*] 1 1, 2 1, 3 1 [3,1,*] 1 1, 2 1, 3 1;
28 param n [1,1,*] 1 1, 2 1, 3 1 [2,1,*] 1 1, 2 1, 3 1 [3,1,*] 1 1, 2 1, 3 1;
29 param L [1,0,1,*] 1 25, 2 25, 3 25 [1,1,1,*] 1 75, 2 75, 3 75
30   [2,0,1,*] 1 250, 2 250, 3 250 [2,1,1,*] 1 750, 2 750, 3 750
31   [3,0,1,*] 1 1, 2 1, 3 1 [3,1,1,*] 1 2, 2 4, 3 6;
32 param C [1,1,*] 1 0, 2 0, 3 0 [2,1,*] 1 0, 2 0, 3 0 [3,1,*] 1 1, 2 1, 3 1;
33 param m [1,1,1,*] 1 0.02, 2 0.02, 3 0.02 [2,1,1,*] 1 0.002, 2 0.002, 3 0.002
34   [3,1,1,*] 1 -1, 2 -0.333333, 3 -2; end;
```

Fig. 3. The data file `data.glp` in Step 4 of our methodology application on Fig. 1.

The expression of the design, decomposition, and association in step 4 can be manual as in Fig. 3 or automated by some sophisticated network design software. We use GLPK (<https://gnu.org/software/glpk>) due to its flexibility and extensibility. To save time, CPLEX (<https://ibm.com/analytics/cplex-optimizer>) can be substituted for GLPK in step 5. The model file in step 5 expresses our formulation of an MILP (mixed-integer linear program) whose objective is to maximize

$\sum_{\mathcal{M}_{i,j} \in \mathbb{M}, t \in \mathbb{H}} (\mathcal{Q}_{i,j}^+(\alpha_{i,j,t}, \beta_{i,j,t}, \lambda_{i,j,t}) - \mathcal{E}_{i,j} h_{i,j,t})$ where \mathbb{M} is the set of all services, $h_{i,j,t_0} = 0$ for $t_0 = \min_{t \in \mathbb{H}} t$ and, for $t' > t_0$, $h_{i,j,t'} = 1$ if $\mathcal{M}_{i,j}$ was on the MEC server of another cell at time $t' - 1$ (i.e., $\mathcal{M}_{i,j}$ migrated at time $t' - 1$ and completed by time t') but $h_{i,j,t'} = 0$ otherwise. The decision variables $\alpha_{i,j,t}$, $\beta_{i,j,t}$, $\lambda_{i,j,t}$, and $h_{i,j,t}$ depend on the decision variable $\mu_{i,j,c,t}$ (not shown) that specifies which cell \mathcal{C}_c has $\mathcal{M}_{i,j}$ running on its MEC server at time t because $\mu_{i,j,c,t}$ determines $\alpha_{i,j,t}$ and $\beta_{i,j,t}$ due to the limited computation Φ_c and communication Ψ_c capabilities, $\lambda_{i,j,t}$ due to the mobility of \mathcal{A}_i and latency $\Lambda_{c_s, c_e, t}$, and $h_{i,j,t}$ due to $h_{i,j,t'} = 1$ whenever $\mu_{i,j,c,t'} \neq \mu_{i,j,c,t}$.

3 Application

To demonstrate the proposed methodology, we apply the methodology in Fig. 2 on the network in Fig. 1 with the resulting data file in step 4 given in Fig. 3, and henceforth, every line number refers to Fig. 3. The 11 time points and the 8 cells $\mathcal{C}_{c_1}, \dots, \mathcal{C}_{|\mathbb{G}|}$ are specified in line 1 (`Hbb t0 ... t*` and `Gbb c1 ... c|G|`) with their equal capabilities specified in line 2 (`Phi default Φc`, `Psi default Ψc`, and `M M`). The routes shown in Table 1 are specified in line 3–12 (`Rcal[cs, ce, 1]` `Rcs, ce, 1`) with their latency in lines 13–21 (`RcalLambda:end` (`Rcallambda [ch, *, 1]` `ch` `Λch, ch, 1`, $(c_h + 1)$ `Λch, ch+1, 1`, \dots , `|G| Λch, |G|, 1`). The mobility of the agent \mathcal{A}_1 is then specified in line 22 (`pos [i, *]` `t0 posi(t0)`, \dots , `t* posi(t*)`), while the services in Table 2 are specified in line 23 (`Mbb (i, *) ji,1 ... ji,mi`) with their migration costs in line 23 (`Ecal [i, *]` `ji,1 Ei,ji,1`, \dots , `ji,mi Ei,ji,mi`), computation bounds in line 24 (`alpha_min default αi,jmin` and `alpha_max default αi,jmax`), communication bounds in line 25 (`beta_min default βi,jmin` and `beta_max default βi,jmax`), latency bounds in line 26 (`lambda_max [i, *]` `ji,1 λi,ji,1max`, \dots , `ji,mi λi,ji,mimax`), and quality functions in lines 27–34 (`U [k, i, *]` `ji,1 Ui,ji,1(k)`, \dots , `ji,mi Ui,ji,mi(k)`, `n [k, i, *]` `ji,1 l`, \dots , `ji,mi l`, `L [k, l', i, *]` `ji,1 Li,ji,1(k),l'`, \dots , `ji,mi Li,ji,mi(k),l'`, `C [k, i, *]` `ji,1 Ci,ji,1(k)`, \dots , `ji,mi Ci,ji,mi(k)`, and `m [k, l', i, *]` `ji,1 mi,ji,1(k),l'`, \dots , `ji,mi mi,ji,mi(k),l'`). Finally, to get a complete task-offloading plan that is repeatable every 10 time units, the following constraint is added to the model file in step 5 so that each service is on the same server both at time 1 and 11: `rep{(i, j) in Mbb, c in Gbb}: mu[i, j, c, 1] = mu[i, j, c, 11];`.

In step 5, the optimal solution can be obtained by running GLPK as `glpsol -m model.glp -d data.glp -o out.sol`. To substitute CPLEX for GLPK, the input file `in.lp` can be obtained by running GLPK as `glpsol --check --wlp in.lp -m model.glp -d data.glp`. Using CPLEX CCSATML (i.e., 20.1.0 for GNU/Linux) by running `cplex -c "read in.lp" mipopt "write out.sol"` on Ubuntu 16.04.7 on a Lenovo E40-80 laptop with a 4-core Intel Core i3-5010U (2×64-bit 2.1-GHz cores, 2 threads/core), the optimal solution (`out.sol`) was obtained in 78.96 s (13677.56 ticks) without memory swap to disk. The result of post-processing `out.sol` to start step 6 is as follows where $\alpha_{1,j,t} = 75$ GIPS, the average $\lambda_{1,1,t}$, $\lambda_{1,2,t}$, and $\lambda_{1,3,t}$ are 1.4, 2.3, and 2.8 hops, respectively, the

numbers in parentheses are $\beta_{1,j,t}$ in tens of Mbps, and the colors are alternated across a row whenever the value changes to double check that $M = 1$ is respected:

t	1	2	3	4	5	6	7	8	9	10
\mathcal{A}_1	\mathcal{C}_1	\mathcal{C}_2	\mathcal{C}_3	\mathcal{C}_4	\mathcal{C}_5	\mathcal{C}_6	\mathcal{C}_4	\mathcal{C}_7	\mathcal{C}_2	\mathcal{C}_8
$\mathcal{M}_{1,1}$	$\mathcal{C}_1(50)$	$\mathcal{C}_3(25)$	$\mathcal{C}_3(25)$	$\mathcal{C}_4(25)$	$\mathcal{C}_6(25)$	$\mathcal{C}_6(25)$	$\mathcal{C}_4(25)$	$\mathcal{C}_7(50)$	$\mathcal{C}_1(50)$	$\mathcal{C}_1(50)$
$\mathcal{M}_{1,2}$	$\mathcal{C}_8(25)$	$\mathcal{C}_8(25)$	$\mathcal{C}_5(25)$	$\mathcal{C}_5(25)$	$\mathcal{C}_5(25)$	$\mathcal{C}_5(50)$	$\mathcal{C}_5(50)$	$\mathcal{C}_5(25)$	$\mathcal{C}_5(25)$	$\mathcal{C}_8(25)$
$\mathcal{M}_{1,3}$	$\mathcal{C}_2(25)$	$\mathcal{C}_2(50)$	$\mathcal{C}_2(50)$	$\mathcal{C}_2(50)$	$\mathcal{C}_2(50)$	$\mathcal{C}_2(25)$	$\mathcal{C}_2(25)$	$\mathcal{C}_2(25)$	$\mathcal{C}_2(25)$	$\mathcal{C}_2(25)$

4 Conclusions

For a dynamic indoor B5G network roamed by agents that continuously offload their heterogeneous RT tasks in an ultra-dense cell environment, we have proposed a methodology for flexibly and optimally managing the computing and communication resources at the network edge, demonstrating its application on a mesh network roamed by one agent that needs three services with distinct end-to-end latency upper-bounds. The demonstration scale is small for clarity, but in our ongoing investigation on the scalability of the methodology, a 10-by-10 grid (100 cells) with a mesh topology roamed by 40 agents that each needs 3 services can be solved within 5% of the optimum in few minutes by CPLEX 20.1.0.

Acknowledgments. This work has received funding from the Italian Ministry of Education, University and Research (MIUR) through the PRIN project no. 2017NS9FEY entitled “Realtime Control of 5G Wireless Networks: Taming the Complexity of Future Transmission and Computation Challenges”. The views and opinions expressed in this work are those of the authors and do not necessarily reflect those of the funding institution.

References

1. Akyildiz, I.F., Kak, A., Nie, S.: 6G and beyond: the future of wireless communications systems. *IEEE Access* **8**, 133995–134030 (2020)
2. Wang, S., Uргаonkar, R., Zafer, M., He, T., Chan, K., Leung, K.: Dynamic service migration in mobile edge computing based on Markov decision process. *IEEE/ACM Trans. Netw.* **27**(3), 1272–1288 (2019)
3. Zhang, Y., Niyato, D., Wang, P.: Offloading in mobile cloudlet systems with intermittent connectivity. *IEEE Trans. Mob. Comp.* **14**(12), 2516–2529 (2015)
4. Subramanya, T., Harutyunyan, D., Riggio, R.: Machine learning-driven service function chain placement and scaling in MEC-enabled 5G networks. *Comput. Netw.* **166**, 106980 (2020)
5. Berno, M., Alcaraz, J.J., Rossi, M.: On the allocation of computing tasks under QoS constraints in hierarchical MEC architectures. In: 4th International Conference on Fog and Mobile Edge Computing (FMEC), pp. 37–44 (2019)
6. Kirov, D., Nuzzo, P., Passerone, R., Sangiovanni-Vincentelli, A.L.: Optimized selection of wireless network topologies and components via efficient pruning of feasible paths. In: Proceedings of the 55th Design Automation Conference, 24–28 June 2018

7. Zhan, W., Luo, C., Min, G., Wang, C., Zhu, Q., Duan, H.: Mobility-aware multi-user offloading optimization for mobile edge computing. *IEEE Trans. Veh. Technol.* **69**(3), 3341–3356 (2020)
8. Thananjeyan, S., Chan, C.A., Wong, E., Nirmalathas, A.: Mobility-aware energy optimization in hosts selection for computation offloading in multi-access edge computing. *IEEE Open J. Commun. Soc.* **1**, 1056–1065 (2020)
9. Wu, C.L., Chiu, T.C., Wang, C.Y., Pang, A.C.: Mobility-aware deep reinforcement learning with glimpse mobility prediction in edge computing. In: *IEEE International Conference on Communications (ICC)*, pp. 1–7 (2020)
10. Chien, W.C., Huang, S.Y., Lai, C.F., Chao, H.C., Hossain, M.S., Muhammad, G.: Multiple contents offloading mechanism in AI-enabled opportunistic networks. *Comput. Commun.* **155**, 93–103 (2020)



Signal Alignment Problems on Multi-element X-Ray Fluorescence Detectors

Francesco Guzzi^{1,2}(✉), George Kourousias², Fulvio Billé², Gioia Di Credico³,
Alessandra Gianoncelli², and Sergio Carrato¹

¹ Image Processing Laboratory, University of Trieste, Trieste, Italy
francesco.guzzi@elettra.eu

² Elettra Sincrotrone Trieste, Basovizza, Italy

³ Department of Statistics, University of Trieste, Trieste, Italy

Abstract. X-ray fluorescence (XRF) is a spectroscopic technique with applications in several fields, such as biology, food science and forensics. Often setups have multi-element detectors in order to improve the signal-to-noise ratio. The relative set of produced spectra have to be aligned to a reference spectrum, in a procedure that is referred to as energy calibration, which is necessary for the fitting. Automated methods fail and a manual procedure is typically employed instead. In this paper, we discuss the signal alignment problem of such systems and we illustrate the preliminary results of a new automated method for linear XRF spectra alignment, which potentially can be used also for other time-series like data.

Keywords: LEXRF · Signal alignment · Energy calibration · SDDs

1 Introduction

X-ray fluorescence (XRF) is a spectroscopic technique that is currently applied in a large variety of fields [1], spanning from biology, cultural heritage, food science and space exploration [2]. XRF spectra are commonly exploited to gather information on the composition of a specimen, as the X-ray fluorescent emissions appear at specific photon energies, making elemental identification a complex but definitely solvable inverse process. Due to its nature, the Low-Energy variant of the technique (LEXRF) poses many more experimental difficulties [3] but is crucial for the identification of light elements. Silicon Drift Detectors (SDD) [4] are currently the most employed to measure the photon energy for each detected photon as it arrives; indeed, the amount of photo-produced electrons is a function of the pair creation energy and the incident photon energy, which can then be calculated by measuring the total accumulated charge of the signal pulse [4]. If a micro-focused X-ray beam (Fig. 1 a) is raster-scanned onto a sample surface (Fig. 1 b) [5], the elemental information (provided by the acquired spectra) is also spatially dependent, producing an “elemental map” of the specimen

itself (Scanning Fluorescence X-ray Microscopy). The spectroscopic information is sometimes correlated with the outcome of another imaging technique, such as Scanning Transmission X-ray Microscopy (STXM) that instead provides the sample absorption information (Fig. 1 e) [2]. Setups such as the ones employed at the TwinMic spectro-microscopy beamline [5] of the Elettra Sincrotrone Trieste synchrotron facility are specially designed to simultaneously perform both the experiments, by employing a transverse XRF detector geometry (Fig. 1). Being the produced XRF radiation isotropic, in order to increase the counting efficiency (but also to acquire topographic information [6]) it is mandatory to increase the observation solid angle, preferably not by enlarging the detector area, but by increasing the number of channels [7] (Fig. 1 c1), arranging many small detectors around the specimen (Fig. 1, panel c2). Thus, to obtain meaningful information, especially for low counts spectral lines, each spectrum needs to be summed channelwise. High pixel count planar ring configurations [8] have been superseded by a slanted design with many SDDs suspended around the sample. As a result, both the detector capacity (and thus the Equivalent Noise Charge) and the risk of pile up are reduced. TwinMic currently employs 8 SDDs which cover a solid angle of 0.255 sr, meaning that 4% of the isotropically emitted photons are collected [5]. New detectors and instrumentation design are under active development and will provide a total surface area that covers 22.4% of the entire hemisphere; the proposed solution will employ a total of 32 detectors [7]. Even more complex design with 64 channels are under development [9].

A set of quantised spectra (Fig. 1 d) is the starting point for the data analysis (element fitting). Unfortunately, due to the heterogeneity in the signal processing chain, the effective value of each energy bin for each channel is slightly different, meaning that no channel-wise operation (e.g. summation/average) can be directly carried out on the raw spectra. The spectra are not only energy-shifted but a scale transform is also required for the x-axis to correctly align the signals. Many automatic techniques can be fooled [10]: being collected from different viewpoints, each spectrum is only similar to the others and possesses only few common features. Manual alignment is thus currently the preferred method, which however becomes more tedious and error-prone as the number of channels increases.

In this paper, we describe the problems of multi-channel energy calibration and present the preliminary results obtained by employing an automatic multi-stage technique that is promising for our class of signals.

2 Alignment Problem and State of the Art Solutions

Mono-dimensional signal alignment, or in this context *energy calibration*, is a problem that is faced in many fields, from speech processing to time series analysis. As for image registration, the degrees of freedom define the underlying correction model. In our context, we are interested in a transformation of the form [11]:

$$f'_k = a \cdot f_k + b \quad (1)$$

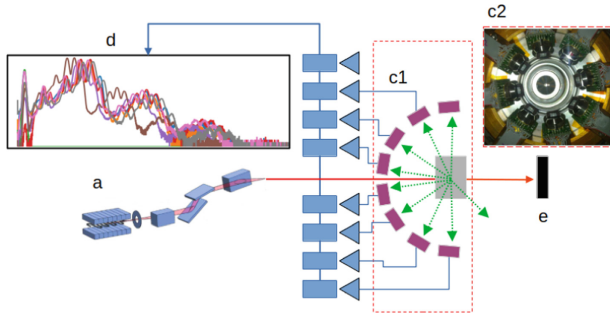


Fig. 1. Transverse geometry spectro-microscopy beamline: an X-ray beam is produced, focused (a) and projected onto a sample (b). X-ray fluorescence is partially detected by the detector arrangement (c1, c2). The processed signals are raw spectra which require an energy calibration procedure (d) prior to fitting. The absorption image is recorded on a different detector (e).

where f_k is the k th energy bin in the original axis, f'_k is the bin in the transformed axis, a is a scaling parameter and b controls the shift. The amount of recent work in the literature related to spectra alignment is the measure of how the problem is not completely solved yet [10]. Indeed, even in the case of commercial multi-channel systems, the acquisition software provides a GUI for semi-automatic methods, which rely on the human-made selection of features in the signals (peaks). Similar reasoning is applied also for open-source solutions, such as PyMCA [12], a widely used software for XRF processing. From this perspective it is apparent that a common approach to signal alignment is based on three phases: 1) peak detection; 2) peak matching among different signals and 3) transform parameters estimation from the peaks list. Manual intervention is obviously required only in the first two phases where the user is asked to select corresponding features among many spectra. However, automatic methods following the same reasoning are also reported, e.g. in [13]; a second common solution to signal alignment is referred to as “Dynamic Time Warping” [14], a non-linear warp procedure that is typically employed in speech processing. The method is designed to align two signals by creating a warping path in the 2D synthetic space of the x coordinates, which describes the joint coordinates transformation that makes the two signals appear more similar. Unfortunately, this approach is not very suitable for our purpose as it is designed for a pairwise alignment and no fixed reference is maintained. A second major drawback is the severe deformation of the signals (staircase artefact, Fig. 2 panel e). A third common way to align the signals is to formalise the process in an optimisation framework, where a loss function such as MSE or a form of correlation is iteratively minimised (maximised). In [11] the authors propose the Pearson correlation as a loss function and employ a windowed non-linear refinement near the peaks. A different type of iterative process is instead used in [15] which is a hybrid technique, where an artificial signal is generated by finding the locations

of the peaks on the reference signal and parameters of (1) are then iteratively searched to maximise the correlation between any other spectra with this artificial reference signal.

3 Proposed Method

In this paper we propose the use of a two-step procedure which employs two different iterative processes. Similarly to an image registration procedure [16], the main alignment step is carried out within an optimisation framework which uses a multi-resolution approach, from the coarsest to the finest resolution. This pyramidal approach is essential to capture the global envelope of each spectrum and avoid local minima in the loss function, notwithstanding the differences among the curves. Additionally we propose the use of the Normalised Mutual Information (NMI) as the optimisation function [17], which measures how a given function value in s_1 can be predicted by the corresponding value in s_2 . It can be calculated by (2):

$$NMI(s_1, s_2) = \frac{H(s_1) + H(s_2)}{H(s_1, s_2)} \quad (2)$$

where $H(s_1)$ and $H(s_2)$ are respectively the entropies of the two signals s_1 and s_2 , while $H(s_1, s_2)$ is the joint entropy. The minimisation can be carried out by negating (2). For each channel (detector), all the acquired spectra at any pixel are summed up to produce the cumulative spectra to align. The algorithm starts by choosing one signal of the set as the common reference. This choice can be made upon the number of features in a spectrum, or by choosing the spectrum with highest energy. To prepare for the second step, a peak finding routine is used on the per-channel mean, helping to reject false positives for the peaks. The actual second alignment uses the peaks-made synthetic signal as described in [15], exploiting a grid search approach. Each pairwise alignment is carried out independently from each other through a multi-process parallel software architecture. This is crucial for a large number of spectra such as in [7].

4 Results

The proposed algorithm was tested on a set of two cumulated XRF scans [18] which we released online [19]. The spectrum obtained at each pixel is summed to create the cumulated signal for each channel-detector. Figure 2 shows how the raw signals in panel a are aligned cascading the two algorithms as described in the previous section (panel b). Algorithm [15] alone is easily fooled by signal variations, leading to an incorrect alignment for one of the pairs (panel c). Conversely, the iterative algorithm we formulated is more capable of finding a coarser transform which can be sufficient in some cases (panel d). It has to be noted that the signal of channel 2 (Fig. 2) is composed only by one peak, the first. Figure 2 panel e shows instead how a DTW-based algorithm fails in correctly aligning the two signals, as it fits the single peaks on the signal of detector 2 on the entire envelope of the two central peaks centered around bin 1000.

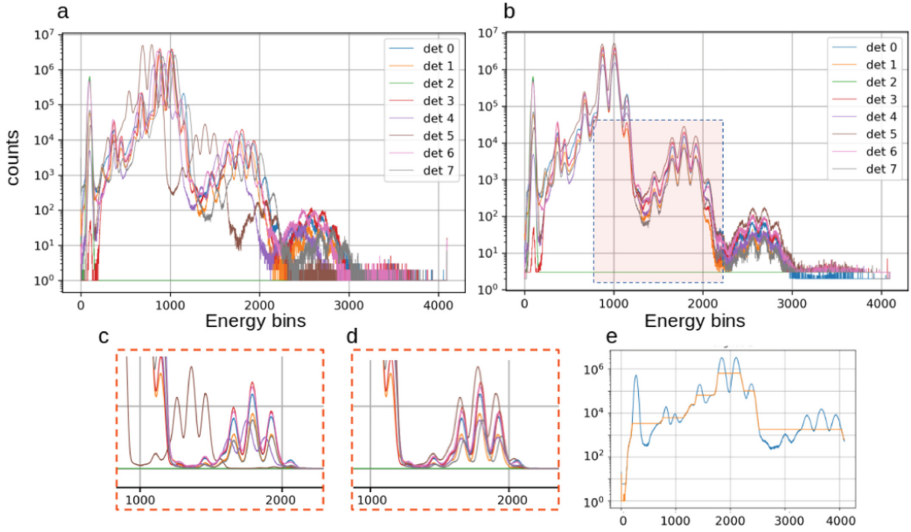


Fig. 2. Alignment of dataset 1. (a): raw spectra; (b) output of the automatic alignment; (c) and (d): output of the two algorithms if used singularly; (e): failing DTW procedure for detector 2.

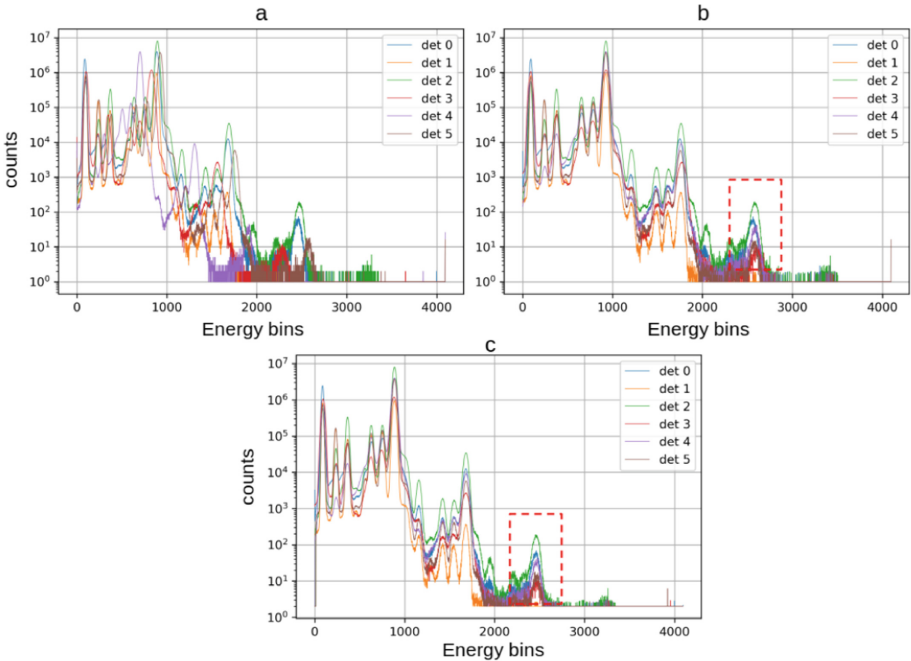


Fig. 3. Alignment of dataset 2: (a) raw spectra; (b): manual alignment; (c): proposed procedure, which better fits the data than (b).

Figure 3 panel a shows the output of the alignment procedure for the dataset 2; as can be seen in panel c the alignment is even better than the one done manually (panel b), which has been carried out by selecting only the most visible peaks. Even if the procedure is iterative, the alignment of an entire dataset of 8 spectra takes less than 5 s on a PC equipped with 8 (threaded) cores (Intel Xeon e3-1245), running Ubuntu 18.04, Python 3.8, Numpy 3.20.2 and SciPy 1.6.2.

5 Conclusions

In this paper we presented the problem of energy calibration, a crucial pre-processing step in LEXRF analysis, especially in the case of a multi-detector system. We described the elements that typically fools many automated procedures and finally propose a method, whose preliminary results are really promising. As we encourage the signal processing community to work on this problem, we release two dataset [19] acquired during a real LEXRF experiment performed with synchrotron radiation light.

References

1. Grieken, R.V. (ed.): Handbook of X-Ray Spectrometry. CRC Press, Boca Raton (2001)
2. Gianoncelli, A., et al.: Recent developments at the TwinMic beamline at ELETTRA: an 8 SDD detector setup for low energy X-ray fluorescence. *J. Phys. Conf. Ser.* **425**(Part 18), 8–11 (2013)
3. Gianoncelli, A., et al.: Simultaneous soft X-ray transmission and emission microscopy. *Nucl. Instrum. Methods Phys. Res., Sect. A* **608**(1), 195–198 (2009)
4. Strüder, L., et al.: Development of the silicon drift detector for electron microscopy applications. *Microsc. Today* **28**(5), 46–53 (2020)
5. Gianoncelli, A., et al.: Current status of the TwinMic beamline at Elettra: a soft X-ray transmission and emission microscopy station. *J. Synchrotron Radiat.* **23**(6), 1526–1537 (2016)
6. Kourousias, G., et al.: XRF topography information: simulations and data from a novel silicon drift detector system. *Nucl. Instrum. Methods Phys. Res. Sect. A* **936**(2018), 80–81 (2019)
7. Bufon, J., et al.: Towards a multi-element silicon drift detector system for fluorescence spectroscopy in the soft X-ray regime. *X-Ray Spectrom.* **46**(5), 313–318 (2017)
8. Longoni, A., et al.: A new XRF spectrometer based on a ring-shaped multi-element silicon drift detector and on X-ray capillary optics. *IEEE Trans. Nucl. Sci.* **49**(3), 1001–1005 (2002)
9. Rachevski, A., et al.: The XAFS fluorescence detector system based on 64 silicon drift detectors for the SESAME synchrotron light source. *Nucl. Instrum. Methods Phys. Res. Sect. A* **936**, 719–721 (2019)
10. Vu, T.N., et al.: Getting your peaks in line: a review of alignment methods for NMR spectral data. *Metabolites* **3**(2), 259–276 (2013)
11. Kourousias, G., et al.: Automated nonlinear alignment of XRF spectra. *X-Ray Spectrom.* **46**(1), 44–48 (2017)

12. Solé, V., et al.: A multiplatform code for the analysis of energy-dispersive x-ray fluorescence spectra. *Spectrochim. Acta, Part B* **62**(1), 63–68 (2007)
13. Hsiao, C.H., et al.: A dynamic data correction method for enhancing resolving power of integrated spectra in spectroscopic analysis. *Anal. Chem.* **92**(19), 12763–12768 (2020)
14. Tomasi, G., et al.: Correlation optimized warping and dynamic time warping as pre-processing methods for chromatographic data. *J. Chemom.* **18**(5), 231–241 (2004)
15. Monchamp, P., Cetto, L., Zhang, J.Y., Henson, R.: Signal processing methods for mass spectrometry. *Syst. Bioinform.: Eng. Case-Based Approach*, 101–124 (2007)
16. Evangelidis, G.D., et al.: Parametric image alignment using enhanced correlation coefficient maximization. *IEEE Trans. Pattern Anal. Mach. Intell.* **30**(10), 1858–1865 (2008)
17. Studholme, C., et al.: An overlap invariant entropy measure of 3D medical image alignment. *Pattern Recogn.* **32**(1), 71–86 (1999)
18. Kourousias, G., et al.: Compressive sensing for dynamic XRF scanning. *Sci. Rep.* **10**, 9990 (2020)
19. Online dataset. <https://doi.org/10.5281/zenodo.5036780>. Accessed 28 June 2021



Low-Level Advanced Design of True Random Number Generators Based on Truly Chaotic Digital Nonlinear Oscillators in FPGAs

Tommaso Addabbo, Ada Fort, Riccardo Moretti^(✉), Marco Mugnaini, and Valerio Vignoli

Department of Information Engineering and Mathematics, University of Siena,
Via Roma 56, 53100 Siena, (SI), Italy
{addabbo,ada,moretti,mugnaini,vignoli}@diism.unisi.it
<http://www.diism.unisi.it/>

Abstract. Recently, a new class of circuits named Digital Nonlinear Oscillators (DNOs) has been proposed for the design of fully digital True Random Number Generators (TRNGs). In this work we discuss the low-level advanced design of TRNGs based on chaotic DNOs, specialized for FPGAs. In detail, starting from a specific DNO topology, we discuss technical solutions to implement these systems exploiting FPGA device primitives. The proposed solutions have been characterized by means of exhaustive measurement campaigns to assess and investigate the impact on the entropy of both chip-to-chip and intra-device variability.

1 Introduction

In information security, True Random Number Generators (TRNGs) are circuits designed to generate truly random binary sequences to be used in cryptographic protocols. These circuits implement entropy sources, being able to generate unpredictable random bits *deemed to be sufficiently secure* for the considered application [1, 2].

A relevant effort has been made by researchers to investigate solutions suitable for being implemented in digital hardware. Traditionally, such sources of entropy, designed in silicon integrated circuits, are based on the exploitation of electronic noise and meta-stable cells, combined in different ways [3–17]. TRNGs must be distinguished from Pseudo Random Number Generators (PRNGs), that are deterministic digital circuits aiming to simulate a truly random source, i.e., not capable to generate information by themselves [1, 2, 18].

Recently, a new class of circuits named Digital Nonlinear Oscillators (DNOs) has been proposed for the design of fully digital TRNGs [19–22]. As defined in [19–22], DNOs are networks of electronic digital circuits, each one originally designed to behave as an asynchronous logic gate, implementing autonomous nonlinear dynamical systems, exhibiting oscillations in the time-continuous

domain. In [19] it has been shown that DNOs can define dynamical systems *supporting structurally stable chaotic dynamics*.

In this work we discuss the low-level advanced design of TRNGs based on chaotic DNOs, specialized for FPGAs.

2 DNOs as Dynamical Systems

Conceptually, the topology of a DNO can be represented as the interconnection of subcircuits called Elementary Logic Blocks (ELBs).

In Fig. 1a a DNO made of six ELBs is presented. The asynchronous domain, in which the DNO operates as the entropy source, has been separated by the surrounding synchronous domain, in which a digital state machine can be used to implement a TRNG. The two domains are joined by low-complexity synchronous circuitry that, in the simplest case, performs the uniform 1-bit sampling of the DNO *analog* dynamics, e.g., by means of a single D flip flop. The interface between the synchronous and the asynchronous domains represents another possible source of randomness, since the flip-flop can be affected by metastability. This condition represents an advantage for the circuit purposes, because it increases the entropy provided by the source.

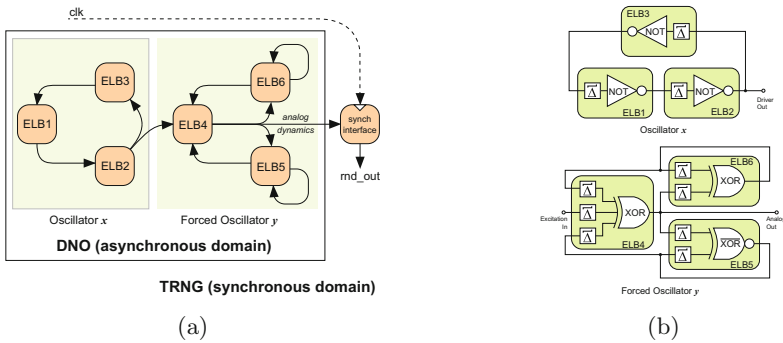


Fig. 1. (a) A Digital Nonlinear Oscillator (DNO) made of six Elementary Logic Blocks (ELBs). The synchronous and asynchronous domains are joined by low-complexity interfacing circuitry (a single D flip flop in the simplest case). (b) DNO based on the topology shown in Fig. 1a suitable for being implemented in FPGAs. Each ELB incorporates a logic functionality and the active digital routing of signals, reported at their inputs.

The specific internal structure of the DNO reported in the figure can be interpreted as a coupled oscillators. A special case of coupled oscillators is obtained when an autonomous dynamical system \mathbf{x} is used to generate a driving signal exciting a second dynamical system \mathbf{y} :

$$\begin{cases} \dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}), \\ \dot{\mathbf{y}} = \mathbf{g}(\mathbf{x}, \mathbf{y}), \end{cases} \quad (1)$$

being $\mathbf{x} : \mathbb{R} \rightarrow \mathbb{R}^N, \mathbf{y} : \mathbb{R} \rightarrow \mathbb{R}^M$ real-valued functions of time t , and \mathbf{f}, \mathbf{g} nonlinear smooth real-valued functions of \mathbf{x} and \mathbf{y} , respectively. If $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ and $\dot{\mathbf{y}} = \mathbf{g}(\mathbf{0}, \mathbf{y})$ define two periodic dynamical systems, we may call \mathbf{y} in (1) the *forced oscillator*, being \mathbf{x} the *forcing periodic driver*.

As it can be seen in Fig. 1b, the driving oscillator is composed of three ELBs, and is topologically equivalent to a Ring Oscillator made of three inverting gates. On the other hand, the forced oscillator \mathbf{y} is composed by three ELBs implementing *xor* or *nxor* boolean gates. As represented in the figure, each ELB incorporates both a logic functionality and the active digital routing of its input signals, represented in the scheme as active delays. In the next Section, the low-level FPGA implementation of the DNO is discussed in detail.

3 FPGA Hardware Implementation

The low-level design of the circuit shown in Fig. 1b in a FPGA, discussed hereafter, is a constrained implementation, in which we took control of the hardware resources utilization, using device primitives. In this section we discuss how this constrained implementation should be performed. For lack of space, the HDL description of the investigated system implemented referring to a Xilinx Artix 7 FPGA is reported in [23].

The ELBs low-level design have been carried out using device primitives. The access to these resources is guaranteed invoking the UNISIM library by Xilinx, in any entity directly involved in the design [23]. The logic functionality of any ELB can be implemented by means of one LUT: this means that the DNO in Fig. 1b requires a total of six LUTs (less than the resources available in two Artix 7 slices). For clarity of presentation, each ELB in [23] has been associated to a VHDL entity. In the code, special directives have to be used to force the resource utilization at specific chip locations. Since the DNO is an asynchronous circuit, registers have not to be used. The active routing, necessary to interconnect the ELBs, completes the DNO architecture.

In DNO design, *combinatorial loops are mandatory*. Normally, combinatorial loops should be avoided in digital designs. They occur when a combinatorial logic feed back to itself without registers, potentially creating logic race conditions or spoiling the timing analysis during the design phases. For these reasons, most design tools generate Design Rules Check (DRC) errors during the synthesis. To allow combinatorial loops, specific directives to enable the synthesis of intentional loops in asynchronous digital structures have been provided [23].

To its minimal terms, the synchronization interface in Fig. 1a can be reduced to a single D flip-flop performing, at the same time, the 1-bit A/D conversion and the uniform sampling of the chaotic signal (provided by the ELB4 in Fig. 1a). To optimize hardware utilization, the 1-bit register FF primitive has been located in the same slice of the 6LUT primitive implementing the logic functionality of the ELB4. The Synchronization Interface takes part in the static timing analysis of the full design, and its constrained location may have an impact on the successful meet of timing constraints. In complex projects, this issue has to be carefully addressed by the designer.

Finally, in a FPGA the configurable routing is organized by means of programmable switches and connection boxes, according to a hierarchical architecture offering local and regional connectivity. Once that the ELBs have been placed in specific chip locations, the final routing is left to the compiler. To minimize the impact of the DNO on the general project routing, it is recommended to have the ELBs concentrated in few slices, placed close to each others. As discussed in the next Section, this also mitigates the impact of hardware variability on the entropy levels of the TRNG.

4 Impacts of Variability: Experiments

The entropy that can be extracted from a chaotic system is always sensitive to the perturbation of its dynamical parameters. The magnitude of the perturbation is linked to its effects on the entropy in a nonlinear way and, with the exception of few cases, the issue has to be investigated by means of numerical simulations or experiments.

In this Section, we present an exhaustive measurement campaign based on experiments devised to assess the impact of both the chip-to-chip variability and the intra-device variability on the entropy. To this aim, we designed 16 instances of the DNO, in different chip areas of the FPGA, repeating the measurements for six different chips (using the same slice locations), reaching a total of 96 DNO instances. The measurements have been performed implementing the system in Xilinx Artix 7 xc7a35 FPGAs, using an internal sampling clock frequency of 400 MHz.

For each case we estimated the Average Shannon Redundancy (ASR) evaluated on binary words of 10 bits, defined as $ASR_{10} = 1 + \frac{1}{10} \sum_{i=1}^{2^{10}} P(w_i) \log_2 P(w_i)$, [bit/sym], where $P(w_i)$ is the generation probability for $w_i \in \{0, 1\}^{10}$, being the summation extended to the 1024 possible binary words of 10 bits. The estimations were obtained acquiring streams of 1 million bits per experiment at room temperature.

4.1 Condensed Layout

Logic resources in a Xilinx Artix 7 FPGA are organized as a matrix of Configurable Logic Blocks (CLBs), each one containing two slices, and each slice being composed of four 6-input Look Up Tables (LUTs) and eight storage elements.

In the upper subplots of Fig. 2 we reported the experimental results highlighting the effects on the ASR of both the chip-to-chip variability and the intra-device variability, for a condensed layout in which the entire DNO in Fig. 1a was concentrated in two slices (same CLB), including the sampling flip flop D. The percentile levels L_x , expressed in bit/sym for $x = 10, 50, 80, 90, 95$, were estimated on the base of the entire data set (96 DNO instances). Red square symbols were used to highlight the chip location 1 (A) or the chip number 1 (B).

As it can be appreciated from the figure, 90% of the chaotic DNOs are capable to provide outstanding levels of ASR_{10} below $L_{90} = 0.077$ bit/sym with a

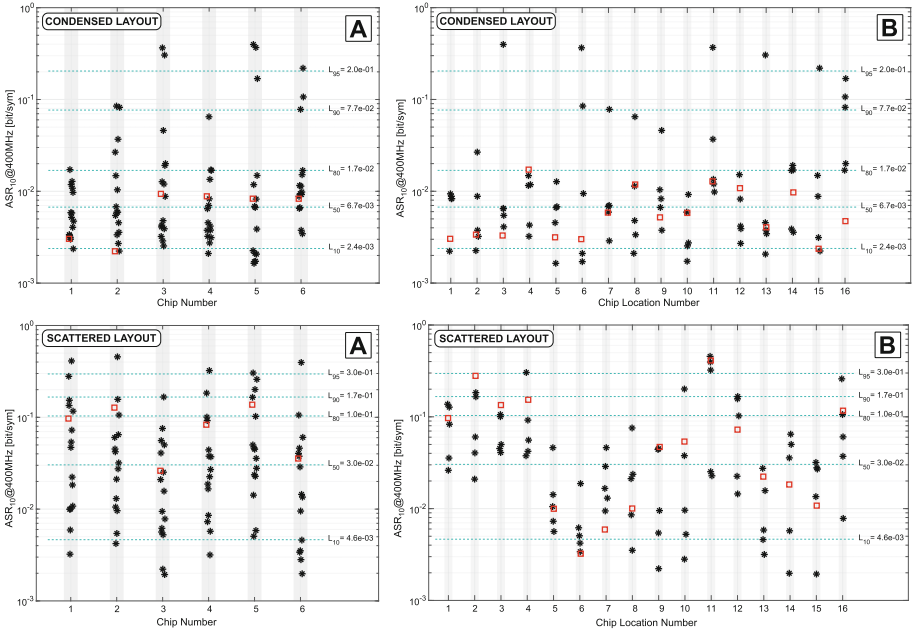


Fig. 2. Effects on the ASR of chip-to-chip and intra-device variabilities, for a condensed and a scattered layout. Red square symbols were used to highlight the chip location 1 (A) or the chip number 1 (B). (Color figure online)

sampling frequency of 400 MHz. This corresponds, in principle, to 369.2 Mbit/s of truly random information generated with the minimal usage of only two FPGA slices.

4.2 Scattered Layout

The effects on the ASR of both the chip-to-chip variability and the intra-device variability have been evaluated as a function of the routing complexity. As confirmed by experimental results, adopting a scattered layout in which the ELBs are separated by longer distances in the CLBs matrix increases the impact of variability. The same experiments presented in the previous Subsection have been repeated adopting a scattered layout in which each LUT of the DNO is positioned in a different FPGA CLB.

When adopting a scattered layout, the ELBs are connected through an higher number of switch boxes. As it can be appreciated from the lower subplots of Fig. 2, this solution worsen the entropy, in average, enhancing the effects of variability. However, it is worth noting that the best cases in both of the layout (10th percentile) share similar levels of ASR.

In this case 90% of the chaotic DNOs are capable to provide levels of ASR₁₀ below $L_{90} = 0.17$ bit/sym with a sampling frequency of 400MHz. This corre-

sponds, in principle, to 332.0 Mbit/s of information. The results are still exceptional, considering the reduced amount of resource utilization.

5 Conclusion

In this work we have discussed the low-level advanced design of TRNGs based on chaotic DNOs, specialized for FPGAs. In detail, starting from a specific DNO topology, we have discussed technical solutions to implement these systems exploiting FPGA device primitives, performing constrained layout implementations. The proposed solutions, capable to provide high entropy levels at a minimal cost of resources utilization, have been characterized by means of exhaustive measurement campaigns to assess and investigate the impact on the entropy of both chip-to-chip and intra-device variability.

According to the observed results, compact layouts allow the system to achieve higher performances in terms of generated entropy. This implies that we can variate the circuit performance simply acting on the routing. An interesting aspect to be investigated would be the effect of the chosen layout on the resulting system power consumption. This kind of investigation will be realized in the future works, as we plan to adapt the presented low-level design approaches for the design of fully digital ASIC based TRNGs.

References

1. Acosta, A., Addabbo, T., Tena-Sanchez, E.: Embedded electronic circuits for cryptography, hardware security and true random number generation: an overview. *Int. J. Circuit Theory Appl.* **45**(2), 145–169 (2017)
2. NIST Special Publication 800-22 Rev. 1a: A statistical test suite for random and pseudorandom number generators for cryptographic applications, April 2010. <https://nvlpubs.nist.gov/nistpubs/legacy/sp/nistspecialpublication800-22r1a.pdf>
3. Kwok, S.H.M., Lam, E.Y.: FPGA-based high-speed true random number generator for cryptographic applications. In: *TENCON 2006– IEEE Region 10 Conference*, pp. 1–4 (2006)
4. Öztürk, H.S., Ergün, S.: A digital random number generator based on chaotic sampling of regular waveform. In: *2020 IEEE 63rd International Midwest Symposium on Circuits and Systems (MWSCAS)*, pp. 178–181 (2020)
5. Demir, K., Ergun, S.: Random number generators based on irregular sampling and Fibonacci-Galois ring oscillators. *IEEE Trans. Circuits Syst. II Express Briefs* **66**(10), 1718–1722 (2019)
6. Anandakumar, N.N., Sanadhya, S.K., Hashmi, M.S.: FPGA-based true random number generation using programmable delays in oscillator-rings. *IEEE Trans. Circuits Syst. II Express Briefs* **67**(3), 570–574 (2019)
7. Carreira, L.B., Danielson, P., Rahimi, A.A., Luppe, M., Gupta, S.: Low-latency reconfigurable entropy digital true random number generator with bias detection and correction. *IEEE Trans. Circuits Syst. I Regul. Pap.* **67**(5), 1562–1575 (2020)
8. Sivaraman, R., Sridevi, A., Rajagopalan, S., Janakiraman, S., Rengarajan, A.: Design and analysis of ring oscillator influenced beat frequency detection for true random number generation on FPGA. In: *2019 International Conference on Computer Communication and Informatics (ICCCI)*, pp. 1–6, January 2019

9. Tao, S., Yu, Y., Dubrova, E.: FPGA based true random number generators using non-linear feedback ring oscillators. In: 2018 16th IEEE International New Circuits and Systems Conference (NEWCAS), pp. 213–216, June 2018
10. Sui, C., Bai, S., Zhu, T., Cheng, C., Beetner, D.: New methods to characterize deterministic jitter and crosstalk-induced jitter from measurements. *IEEE Trans. Electromagn. Compat.* **57**(4), 877–884 (2015)
11. Raitza, M., Vogt, M., Hochberger, C., Pionteck, T.: Raw 2014: random number generators on FPGAs. *ACM Trans. Reconfigurable Technol. Syst.* **9**(2), 15:1–15:21 (2015)
12. Golic, J.D.J.: New methods for digital generation and postprocessing of random data. *IEEE Trans. Comput.* **55**(10), 1217–1229 (2006)
13. Wang, X., et al.: High-throughput portable true random number generator based on Jitter-Latch structure. *IEEE Trans. Circuits Syst. I Regul. Pap.* **68**(2), 741–750 (2021)
14. Tsoi, K.H., Leung, K.H., Leong, P.H.W.: Compact FPGA-based true and pseudo random number generators. In: 11th Annual IEEE Symposium on Field-Programmable Custom Computing Machines, FCCM 2003, April 2003, pp. 51–61 (2003)
15. Hata, H., Ichikawa, S.: FPGA implementation of metastability-based true random number generator. *IEICE Trans. Inf. Syst.* **E95.D**(2), 426–436 (2012)
16. Wiczorek, P.Z.: An FPGA implementation of the resolve time-based true random number generator with quality control. *IEEE Trans. Circuits Syst. I Regul. Pap.* **61**(12), 3450–3459 (2014)
17. Wu, X., Li, S.: A new digital true random number generator based on delay chain feedback loop. In: IEEE International Symposium on Circuits and Systems (ISCAS), vol. 2017, pp. 1–4 (2017)
18. Addabbo, T., Alioto, M., Fort, A., Rocchi, S., Vignoli, V.: Low-hardware complexity PRBGs based on a piecewise-linear chaotic map. *IEEE Trans. Circuits Syst. II Express Briefs* **53**(5), 329–333 (2006)
19. Addabbo, T., Fort, A., Moretti, R., Mugnaini, M., Takaloo, H., Vignoli, V.: A new class of digital circuits for the design of entropy sources in programmable logic. *IEEE Trans. Circuits Syst. I Regul. Pap.* **67**(7), 2419–2430 (2020)
20. Addabbo, T., Fort, A., Moretti, R., Mugnaini, M., Vignoli, V., Garcia-Bosque, M.: Lightweight true random bit generators in PLDs: figures of merit and performance comparison. In: 2019 IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1–5, May 2019
21. Addabbo, T., Fort, A., Mugnaini, M., Vignoli, V., Garcia-Bosque, M.: Digital non-linear oscillators in PLDs: Pitfalls and open perspectives for a novel class of true random number generators. In: 2018 IEEE International Symposium on Circuits and Systems (ISCAS), pp. 1–5, May 2018
22. Addabbo, T., Fort, A., Moretti, R., Mugnaini, M., Vignoli, V.: Analysis of a circuit primitive for the reliable design of digital nonlinear oscillators. In: 2019 15th Conference on Ph.D. Research in Microelectronics and Electronics (PRIME), pp. 189–192, July 2019
23. Addabbo, T., Fort, A., Moretti, R., Mugnaini, M., Vignoli, V.: DNO Xilinx Artix 7 hardware implementation. <http://www3.diism.unisi.it/~addabbo/ApplePies2021/HDL.html>



Design and Implementation of an FPGA-Based CNN Hardware Accelerator Using Partial Reconfigurability: The CloudScout Case Study

Corrado Comino^(✉), Tommaso Pacini, Emilio Rapuano, and Luca Fanucci

Department of Information Engineering, University of Pisa,
G.Caruso 16, 56122 Pisa, Italy
corrado.comino@gmail.com, {tommaso.pacini,emilio.rapuano}@phd.unipi.it,
luca.fanucci@unipi.it
<https://www.dii.unipi.it>

Abstract. This article proposes a method to design and implement an FPGA-based hardware accelerator for Convolutional Neural Networks (CNNs) exploiting Partial Reconfigurability (PR). The design strategy was applied to the CloudScout CNN case study, a network developed in the frame of the Φ -sat-1 ESA mission to perform cloud-detection aboard the satellite. The presented design based on Partial Reconfigurability was implemented, validated and characterized on a Xilinx ZCU106 Evaluation Board. The system was then compared with an alternative FPGA implementation reported in the literature to evaluate the obtained results. The comparison shows that the PR-based approach allows decreasing the resource utilization of the architecture, thus improving the network portability on smaller size FPGAs.

Keywords: CNN · FPGA · Hardware accelerator · On-the-edge · Partial reconfigurability · Resource optimization

1 Introduction

In recent years, Convolutional Neural Networks (CNNs) have acquired a key role in pattern recognition applications such as image classification [6], object detection [7] and video analysis [8]. These algorithms ensure high accuracy and timing performances at the cost of computational power. Such feature constitutes a real issue when the processing of these models has to be moved on-the-edge, i.e. close to the source of information [9]. In this scenario, the devices employed for the acceleration of these networks are usually resource-constrained and with limited power budgets. In addition, when the size of data processed and the CNN structure become more complex, it is necessary to implement CNNs in a faster

way [4], i.e. using hardware-acceleration techniques. Among the feasible accelerators, FPGAs represent a valuable solution thanks to their trade-off between flexibility and performance [10]. Furthermore, these devices offer the opportunity to develop custom hardware accelerator for the targeted network and to integrate it in the same chip with dedicated hardware performing accessory tasks as pre-processing and inter-blocks communication. The major challenges in the development of FPGA-based accelerators for CNN regards the gap between the limited resource availability of the device and the model complexity.

This work proposes a design strategy to implement CNN on a FPGA exploiting Partial Reconfigurability (PR) [1], a technique that allows to dynamically reconfigure a defined partition of FPGA resources while the remaining device operates without interruptions. Exploiting this functionality, this work proposes an accelerator employing custom hardware for each convolutional layer of the CNN, aiming to maximize the efficiency in the resource utilization. This method was used to develop a PR-based implementation for the CloudScout CNN [2], an on-the-edge network developed in the frame of the Φ -sat-1 ESA mission to perform cloud-detection directly on board the satellite. The implemented accelerator was then compared with an alternative FPGA implementation of CloudScout reported in the literature [3] in order to evaluate the benefits of the proposed methods. The implementation of the hardware-accelerator was deployed on the Xilinx ZCU106 [11].

The main contents of this article are:

- Identification of a design strategy to develop a CNN architecture exploiting FPGA Partial Reconfigurability.
- Accurate manual floorplanning to achieve the best trade-off between performances and resource utilization.
- Hardware implementation of the developed design on the Xilinx ZCU106 evaluation board, validation and characterization of the system.
- Benchmark between the proposed design and the state-of-the-art FPGA accelerator for the CloudScout CNN [3].

2 Methods

Partial Reconfigurability (PR) is a modern FPGA functionality, consisting in the ability of dynamically reconfigure a defined partition of FPGA resources while the remaining device continues to operate ordinarily. PR allows to implement different functionalities on the partition in different time intervals, allowing a time-shared use of FPGA resources. In order to *partially reconfigure* the FPGA, a file containing the new configuration of the resources contained in the reconfigurable partition (RP) has to be loaded in a specific FPGA configuration port, this file is called *partial bitstream*.

In a hardware design exploiting partial reconfigurability, there is a collection of *reconfigurable modules* (RM) for each reconfigurable partition (RP) and a *static logic* design. The RMs represent every possible configuration of the RP.

The static logic, instead, describes the remaining logic of the whole hardware design.

This article proposes a method to design and implement a hardware-accelerator for CNN exploiting the PR functionality. In particular, the aim of the work is to reduce the hardware resources required to implement the CloudScout CNN [2], the network taken as case study. Figure 1 presents the architecture of the proposed design. For each convolutional layer of the CNN, a specific accelerator was designed. Several layer-specific optimizations were made to efficiently reduce the resource usage and to maximize the network performances of each accelerator. For instance, layer-based tailor-made scheduling and data storing method were designed and employed, exploiting the degrees of freedom on the number of input data processed in parallel and the data storing order in the on-chip memory. Furthermore, the reconfigurable design allows to save further on-chip memory, storing the convolutional filters of the only layer in execution, avoiding the waste in storing all the network filters. The discussed optimizations allow to reduce the resources required for the implementation, partially mitigating the effect on the network elaboration time. The architecture of each layer exploits almost the same FPGA resources, in order to obtain balanced layer-specific accelerators to insert in the same reconfigurable partition (RP).

The accelerator processes one layer of the network at a time, performing data transfers with an external DDR memory via an Advanced eXtensible Interface (AXI) bus [12]. The DDR memory stores the input and output matrices of the layer under execution and the partial bitstreams of each layer accelerator. The on-chip FPGA resources not included in the RP, the so called *static logic*, are used to implement the logic responsible for the communication interface and the control of the accelerator status. During the inference, the layer-optimized designs are progressively alternated on the RP in order to process the cascade of layers composing the CNN.

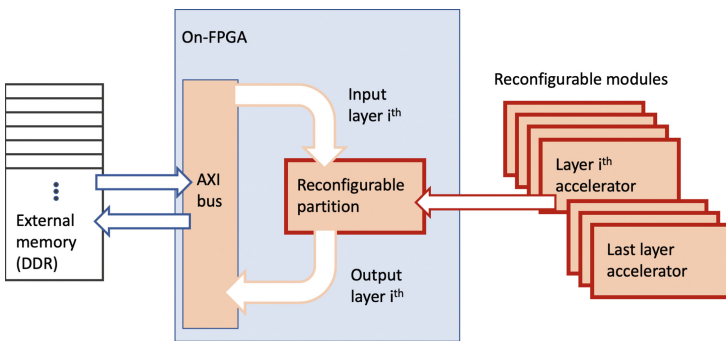


Fig. 1. System architecture

The presented methodology was applied to the CloudScout CNN in order to validate the design flow. Once the static logic and the layer-specific accelerators

are fully described and synthesized, the next step is the implementation of the design on FPGA. In this phase, a pivotal role is covered by the floorplanning task, in which the designer has to manually draw the boundaries of the reconfigurable partition (RP). The choice of this area directly affects the resource usage of the whole system. Therefore, starting from the analysis of the required resources for the implementation of each layer accelerator, a minimum portion of FPGA must be defined as RP. The first step consists in studying the resource footprint of the FPGA and choose the best fit for the RP, depending on the locations of the main resources used by the layers accelerators. The chosen one is presented in Fig. 2, the white boundary collects all the resources allocated to the RP. The coloured blocks inside and outside the RP show respectively the resources used to implement the most resource-greedy layer and the static logic of the CloudScout CNN design.

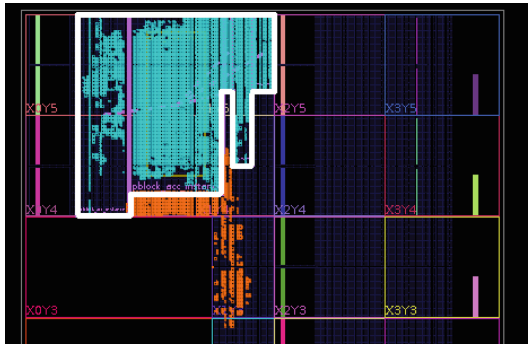


Fig. 2. Manual floorplanning for the reconfigurable partition (RP)

3 Results

The proposed design was implemented on the Xilinx ZCU106, an evaluation board hosting a Xilinx MPSoC (Multiple Processors System on Chip) with a XCZU7EV-2ffvc1156 FPGA. After the validation of the implemented design on FPGA, which concerned the test of the system functionality for a given set of images, the design was characterized in terms of resource usage. In Table 1, the resources employed in the reconfigurable partition and the static logic, together with the whole resource usage are presented.

The on-chip memory and the number of Digital Signal Processors (DSP) represent the critical factor in the implementation of CNNs on FPGA. Therefore, the information about the resource usage presented in Table 1 can be collected in an aggregate result showing the number of *DSP* and the *memory footprint* required for the design implementation. The total memory footprint consists of the sum of the contributions of each RAM block employed, expressed in Mb.

Table 1. Implementation resource usage.

	Reconfigurable partition	Static logic	Total usage
LUT	50080	4077	54157
Flip Flop	100160	3952	104112
LUTRAM	23720	776	24496
BlockRAM	32	4	36
UltraRAM	12	0	12
DSP	262	0	262

This result was employed in the comparison between the proposed accelerator and an existing custom accelerator [3] for the CloudScout CNN on FPGA in terms of resource usage, inference time and power consumption (Table 2).

Table 2. Resource utilization comparison in PR and non-PR based designs.

	PR-based design	State-of-the-art design [3]
Memory footprint	6,14 Mb	17,63 Mb
DSP	262	1163
Inference time	672,4 ms	141,7 ms
Power consumption	3,15 W	3,4 W

The comparison shows a considerable reduction in the resource usage, in particular the PR-based design allows to use 22,5% of the DSP and 34,8% of the memory footprint with respect to the custom accelerator presented in [3]. The portability of the network on smaller-size FPGAs is significantly improved. Indeed, exploiting the partial reconfigurability, the CloudScout model can be implemented on a smaller-size rad-hard FPGA with respect to the original design, providing a valuable system cost reduction without any system performance degradation.

The inference time represents the period required for the classification of a single image by the CNN. In the proposed design, it is the sum of the elaboration time of the layer-specific hardware accelerators and the reconfiguration time of the reconfigurable partition. The inference time of the proposed design is almost five times larger than the one measured for the accelerator presented in [3]. This result does not represent a critical factor for the specific application, due to the larger time between two consecutive classifications.

The last figure of merit is the power consumption, which was measured during a continuous cycle of inference and averaged. In the PR-based design, the power consumption is equal to the 93% of the value for the compared design. The reduction of this factor guarantees an important advantage in power-constrained applications such as space missions.

4 Conclusion

This work presents a methodology to design and implement an FPGA-based hardware accelerator for convolutional neural networks. In particular, the proposed design exploits FPGA partial reconfigurability to minimize the resource usage, improving its portability on smaller FPGAs. This has a particular value for space applications, such as the addressed case study, where rad-hard devices are extremely expensive. The results of the comparison with the accelerator proposed in [3] confirmed the reduction in the resource usage of the PR-based design and in the power consumption at the cost of a longer inference time, that for the specific case study does not represent a limit for the on-board processing of acquired images.

Acknowledgments. This work has been partially funded by the European Space Agency under contract number 4000129792/20/NL and by the European Union's Horizon 2020 innovation action under grant agreement number 761349, TETRAMAX (Technology Transfer via Multinational Application Experiments).

References

1. Vipin, K., Fahmy, S.A.: FPGA dynamic and partial reconfiguration: a survey of architectures, methods, and applications. *ACM Comput. Surv.* **51**(4), 1–39 (2018). <https://doi.org/10.1145/3193827>
2. Giuffrida, G., et al.: CloudScout: a deep neural network for on-board cloud detection on hyperspectral images. *Remote Sens.* **12**(14), 2205 (2020). <https://doi.org/10.3390/rs12142205>
3. Rapuano, E., et al.: An fpga-based hardware accelerator for cnns inference on board satellites: benchmarking with myriad 2-based solution for the cloudscout case study. *Remote Sens.* **13**, 1518 (2021)
4. Zhang, Q., et al.: Recent advances in convolutional neural network acceleration. *Neurocomputing* **323**, 37–51 (2018). <https://doi.org/10.1016/j.neucom.2018.09.038>
5. Albawi, S., et al.: Understanding of a convolutional neural network. In: International Conference on Engineering and Technology (ICET) (2017). <https://doi.org/10.1109/ICEngTechnol.2017.8308186>
6. Lee, H., Kwon, H.: Going deeper with contextual CNN for hyperspectral image classification. *IEEE Trans. Image Process.* **26**(10), 4843–4855 (2017). <https://doi.org/10.1109/TIP.2017.2725580>
7. Nguyen, D.T., Nguyen, T.N., Kim, H., Lee, H.: A high-throughput and power-efficient FPGA implementation of YOLO CNN for object detection. *IEEE Trans. Very Large Scale Integr. (VLSI) Syst.* **27**(8), 1861–1873 (2019). <https://doi.org/10.1109/TVLSI.2019.2905242>
8. Nishani, E., Çiço, B.: Computer vision approaches based on deep learning and neural networks: deep neural networks for video analysis of human pose estimation. In: 2017 6th Mediterranean Conference on Embedded Computing (MECO), pp. 1–4 (2017). <https://doi.org/10.1109/MECO.2017.7977207>
9. Shi, W., Dustdar, S.: The promise of edge computing. *Computer* **49**(5), 78–81 (2016). <https://doi.org/10.1109/MC.2016.145>

10. Zhang, C., et al.: Optimizing FPGA-based accelerator design for deep convolutional neural networks, pp. 161–170. Association for Computing Machinery, New York (2015). <https://doi.org/10.1145/2684746.2689060>
11. Xilinx. ZCU106 Evaluation Board. User guide. UG1244 (v1.4)
12. Xilinx. AXI Reference Guide. UG761(v13.1). Accessed 7 Mar 2011



Exploring GPS L1 C/A Fast Acquisition with COTS FPGA

Andrea Romani^{1,2(✉)}, Franco Bigongiari^{2(✉)}, and Luca Fanucci^{1(✉)}

¹ University of Pisa, Pisa, Italy

luca.fanucci@unipi.it

² SITAEL S.p.A., Via A. Gherardesca 5, 56121 Pisa, Italy

{andrea.romani, franco.bigongiari}@sitael.com

Abstract. In this work we create an environment to validate and implement a GPS L1 C/A fast acquisition core on COTS FPGAs. The frequency domain PCPS algorithm is studied on a Scilab test bench and the results are compared with the outputs of the GNSS-SDR simulator. Then all the PCPS algorithmic blocks are designed and fitted on an FPGA. The hardware core is successfully used to acquire GPS L1 C/A signal records. Timing and area results are analyzed to understand how to extend the design to modern GPS signals.

1 Introduction

L1 C/A is a signal transmitted by the GPS satellite constellation to provide timing and positioning services. Studying and improving its acquisition possibilities is crucial to conscientiously design the next-generation GNSS receivers. In fact, modern signals like L5, L1C and L2C uses correlation codes that are ten times longer than the ones used by L1 C/A. This means that the actual architectures have to be analyzed to understand how they can be extended to the modern GPS system. This work has therefore two main purposes. The first is to build-up a design methodology to study GNSS acquisition algorithms. The second is to implement a GPS acquisition block on COTS FPGA to observe its speed, area and timing results with an eye on modern signals.

The first part of a GNSS receiver is the Acquisition Block. Its role is to detect the presence or the absence of signals coming from a given GNSS satellite. The navigation data transmitted on the L1/CA signal is correlated with a 1023-bit C/A code. This operation allows all the GPS L1 C/A signals to share the same frequency band because of the orthogonality of C/A codes. So the problem to face is to develop a CDMA system that performs a correlation between the received signal and a local replica of the target satellite C/A code. The system must consider the two main problems that affects GPS signals: Doppler shift and code displacement forces the system to perform a large number of correlations before locking a satellite. In fact, the unknown relative satellite-receiver speed causes a shift in the signal frequency that shall be removed before computing the correlation. This may force the system to perform multiple frequency shifts on the signal and to compute a correlation for each one until a match with the

local code is found. Moreover, the signal at the input of the receiver is not time-synchronized with the local code replica so the correlation should virtually be performed for each code bit displacement. We decided to address this problem with the Parallel Code Phase Search (PCPS) algorithm, that can simultaneously test all the code displacements so that it need to perform a sweep on the only values of frequency shift. This method is generally considered the fastest one [1] and it is based on the fact that the correlation between two sequences can be computed in the frequency domain. Equation 1 shows the mathematical steps of the PCPS algorithm. ($z[n]$ is the circular correlation between the received signal $x[n]$ and the local code $c[n]$; $x[n]$ and $c[n]$ are N-samples sequences; the \mathcal{F}^{-1} operator is used for the Inverse Discrete Fourier Transform operation (IDFT); $X[k]$ and $C[k]$ are the Discrete Fourier Transforms (DFT) performed on $x[n]$ and $c[n]$).

$$z[n] = \sum_{m=0}^{N-1} c^*[m]x[n+m] = \mathcal{F}^{-1}\{C[k]^*X[k]\} \quad (1)$$

First of all the input signal is mixed with a frequency of the Doppler sweep to shift the input to the base band. Then the Fourier transforms are computed on the incoming sequences. The code transform is conjugated and multiplied with the other one. Then the IDFT is performed on the resulting sequence. The n_{th} sample of the output sequence is the correlation value for a code- signal misalignment of n bits. So the algorithm gets the maximum value of that sequence for every value of the frequency sweep. If the resulting correlation value is greater than a predefined threshold, it returns the information that the satellite is visible.

2 Functional Validation

Our environment was designed to always have a set of results to be compared with the outputs of our custom blocks. The GNSS-SDR software was chosen as a reference receiver. We used two main input data records to feed the reference and DUT blocks. The first is a 100 s acquisition of raw GNSS signal samples, collected by an RF front-end centered at 1,575.42 MHz, delivering baseband samples at 4 MS/s in an interleaved I/Q 16-bit integer format. The second is a set of samples generated by the GPS-SDR-SIM software with the same format but sampled at 5 MS/s. We implemented the PCPS block in a Scilab simulation and we compared the results with GNSS-SDR. Finally we implemented the design on an FPGA and compared the results with the previously developed tools.

The designed Scilab environment is composed of three functional blocks, as shown in Fig. 1.

1. **input_conversion.sc** acquires 1ms of data from the source.dat file and converts it in a format compatible with the PCPS algorithm. 1ms of data has 4000 in-phase (I) samples and 4000 in-quadrature (Q) samples. So the code acquires 8000 samples in a vector and deinterleaves it to obtain a vector for the I component and a vector for the Q component. The PCPS block needs

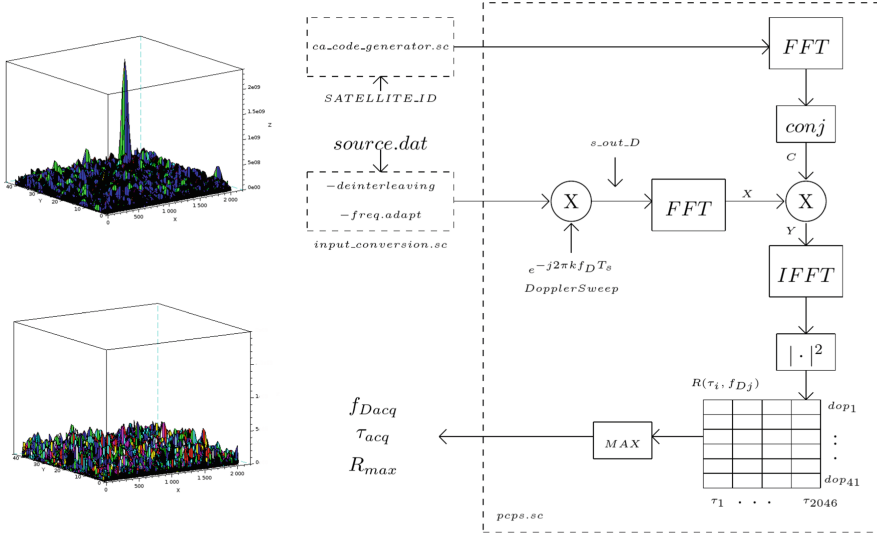


Fig. 1. General block diagram of the simulation and correlation plots.

to compute the product of the Fast Fourier Transform of the PRN and of the incoming signal. So the length of the PRN array and the length of the acquired data must match. The length of the PRN array is 2046 elements to obtain a code mismatch precision of a half chip. So we decimated the acquired data to match that length with the Scilab function $[y] = \text{intdec}(x, \text{freq_out}/\text{freq_in})$, that is designed to change the sampling rate of a signal. Finally we merged the I and Q vectors in a complex data vector formed by elements in the format I+jQ.

2. `ca_code_generator.sc` takes as input the `SATELLITE_ID` value and returns the corresponding PRN C/A correlator. The C/A PRN codes have a period of 1023 chips and they are transmitted at 1.023 Mchip/s. These codes are generated combining two bit streams out of two 10 stage linear feedback shift registers (LFSR). Each LFSR has got its own polynomial generator that states how to compute the reinsertion value of the shift register.
3. `pcps.sc` searches the presence of the given satellite in the input signal by executing the PCPS algorithm. It performs a doppler frequency sweep of 41 frequencies from 0 Hz to 10 kHz with steps of 250 Hz. For each step of the sweep we multiply the incoming sequence `complex_data[k]` for the factor $e^{-j2\pi(dop)kT_s}$ to shift the signal in frequency of the Doppler factor. Then we used the `fft()` and `ifft()` functions to go through the steps of the algorithm. At the end of each step the output vector is saved in a 3D-matrix, which has the frequencies in the y-axis, the code displacements in the x-axis and the correlation values in the z-axis. The next block searches for the maximum correlation values. The found f_D and τ are the output of the system.

In the left side of Fig. 1 it can be seen the output results: the upper tridimensional plot shows a correlation peak when a satellite is transmitting in the input signal. The lower one is the output when the system is looking for a not in-view satellite.

3 Hardware Implementation

Our system is designed to be implemented on a USRP N210 device, which has a Spartan 3A 3400 DSP FPGA. The designed system is divided in three main areas, as shown in Fig. 2.

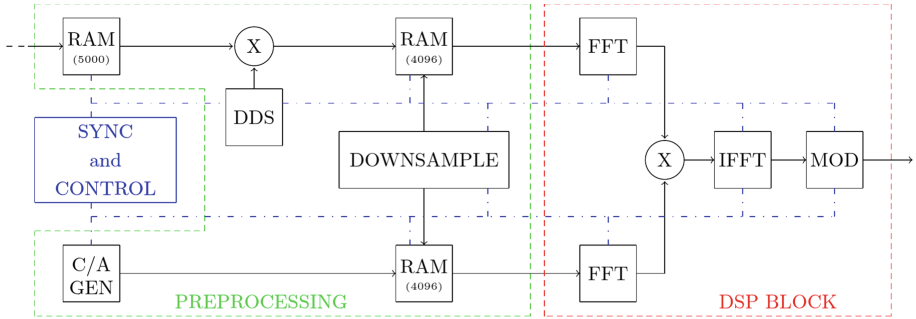


Fig. 2. General architecture blocks.

- Signal buffering and pre-processing block:** One millisecond of 5 MHz IQ samples is acquired, then another down conversion is performed to remove the Doppler effect on the signal. This operation is performed with a Direct Digital Synthesizer. In fact, the Scilab simulation held to validate the PCPS algorithm performs the Doppler shifting by multiplying a complex input signal $I + jQ$ with a complex exponential function $e^{-j2\pi k f_D T_s}$, where f_D is the value of the frequency shift that we want to obtain and T_s is the sampling period of the incoming signal. This exponential function can be expressed by the Euler formula, as shown in Eq. 2.

$$e^{-j2\pi k f_D T_s} = \cos(2\pi f_D t_k) - j \sin(2\pi f_D t_k) \quad (2)$$

This means that the complex input signal has to be multiplied with a sequence of complex numbers whose real part is the cosine of a specific time-varying quantity and whose imaginary part is an inverted version of a sine wave. So the DDS has been set to generate two waveforms ($\cos(x)$ and $-\sin(x)$) with a resolution of 0.4 Hz and an SFDR of 50 dB. The multiplication is performed with a custom complex multiplier.

The local replica of the PRN code are then generated. Finally the signal and the code are decimated from 5000 to 4096 samples to fit the correct size of the fft's in the DSP block. This section is implemented in fixed point arithmetic.

- **DSP block:** This section executes the Parallel Code Phase Search algorithm. Two parallel fft's are performed on the code and on the signal, then a complex multiplication is performed on the two results. Then the inverse fft is performed on the multiplication result. The FFT cores are implemented by Xilinx LogiCORE IP Fast Fourier Transform modules. Finally the modulus of the result is calculated implementing the Robertson's Approximation, where $A = \sqrt{I^2 + Q^2}$ is approximated with $A_R = \text{MAX}(|I| + 0.5|Q|, |Q| + 0.5|I|)$. This section is implemented in single precision, floating point arithmetic to fix to 32 the big growing on the number of bits necessary to represent the results. This was the fastest way to achieve the result. A fixed point approach with truncation and scaling should be investigated in future developments to further improve system performance.
- **Sync and Control block:** It is a finite state machine that generates the signals that enables each block with the correct timing. It ensures the correct communication of each block in the different elaboration stages.

4 Results

The GNSS-SDR software was used to find the in-view satellites in the input record, then the same record was given as input of the Scilab and hardware models to compare the correlation vectors. The output correlation vectors are compared in Fig. 3. It can be seen that both systems returned a correlation peak in correspondence of the same code displacement. The system was able to correctly detect the correlation peak for every satellite in the input record. The time required to have a correlation vector for 4096 different code bins and for a single frequency bin is 1 ms for signal acquisition plus 1.29 ms for elaboration. This result can be further optimized using non floating point architectures. However, this architecture can reach its ideal theoretical performances by pipelining the different stages [2], making the FFT computational time negligible by increasing the throughput of the architecture. The whole design uses the 85% of FPGA Block RAM resources, the 33% of logic slices and the 46% of DSP48 blocks. A single FFT core uses the 24% of BRAM units, the 7% of logic slices and the 7% of DSP48 blocks. It can be deduced that the target device can not contain more than one elaboration branch (no frequency bins can be tested in parallel). However many optimizations have been presented [6] to reduce the length of the FFTs involved in the PCPS algorithm.

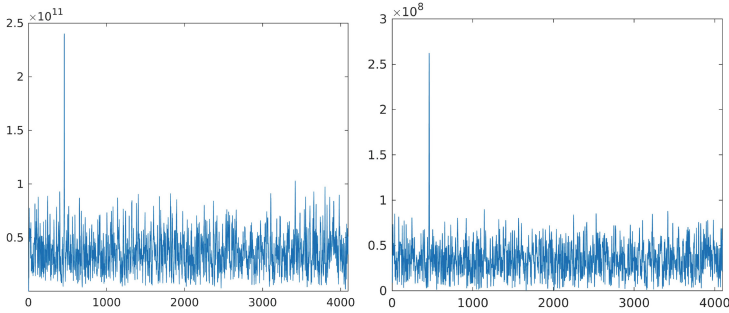


Fig. 3. Hardware and Software correlation vectors: sat ID 9, Dop shift = 2750 Hz.

5 Conclusions

An FPGA-accelerated hardware correlation unit has been developed to acquire the GPS L1 C/A signal. The system has proven to be able of finding the in-view satellites in a record of GPS signals. The adopted kind of architecture can not be directly applied to the next generation GPS signals, since they use ten times longer codes, made longer by the use of secondary codes. This means that the FFT blocks would not fit in a basic COTS FPGA. However, the designed system can be considered as the basic building block of the next generation GPS signals, as their correlators can be reduced to the iteration of smaller FFTs [3–5].

References

1. Rovelli, D., Crosta, P., Iacone, P., Rovini, M., Gentile, G., Fanucci, L.: Acquisition speed-up engine for GNSS signals. In: 5th ESA Workshop on Satellite Navigation Technologies and European Workshop on GNSS Signals and Signal Processing (NAVITEC), Noordwijk 2010, pp. 1–8 (2010). <https://doi.org/10.1109/NAVITEC.2010.5708006>
2. Sajabi, C., Chen, C.H., Lin, D.M., Tsui, J.B.Y.: FPGA frequency domain based GPS coarse acquisition processor using FFT. In: IEEE Instrumentation and Measurement Technology Conference Proceedings. Sorrento 2006, pp. 2353–2358 (2006). <https://doi.org/10.1109/IMTC.2006.328619>
3. Leclère, J., Botteron, C., Farine, P.: High sensitivity acquisition of GNSS signals with secondary code on FPGAs. *IEEE Aerosp. Electron. Syst. Mag.* **32**(8), 46–63 (2017). <https://doi.org/10.1109/MAES.2017.160176>
4. Zeng, Q., Tang, L., Zhang, P., Pei, L.: Fast acquisition of L2C CL codes based on combination of hyper codes and averaging correlation. *J. Syst. Eng. Electron.* **27**(2), 308–318 (2016). <https://doi.org/10.1109/JSEE.2016.00031>
5. Zhou, J., Liu, C.: Joint data-pilot acquisition of GPS L1 civil signal. In: 2014 12th International Conference on Signal Processing (ICSP), Hangzhou, pp. 1628–1631 (2014). <https://doi.org/10.1109/ICOSP.2014.7015271>
6. Leclere, J., et al.: FFT splitting for improved FPGA-based acquisition of GNSS signals. *Int. J. Navig. Obs.* (2015)



Feasibility Study of a Unified Fast Acquisition Core for Modern GPS Signals

Andrea Romani^{1,2(✉)}, Franco Bigongiari^{2(✉)}, and Luca Fanucci^{1(✉)}

¹ University of Pisa, Pisa, Italy

² SITAEL S.p.A., Via A. Gherardesca 5, 56121 Pisa, Italy

{andrea.romani,franco.bigongiari}@sitael.com, luca.fanucci@unipi.it

Abstract. This work shows how the acquisition of all modern GPS signals can be reduced to the iteration of a single FFT-based computational unit. In particular, acquisition algorithms for L5, L1C and L2C are compared to show how they can be solved with a series of 40920-point correlation units, leading to the possible hardware implementation of an efficient multi-band acquisition block.

Keywords: GPS · GNSS · Fast acquisition

1 Introduction

Many acquisition algorithms have been presented to acquire modern GPS signals like L5, L1C and L2C, as they are being transmitted by not yet fully deployed GPS constellations. These signals are characterized by the use of very long codes, making fast acquisition algorithms very resource demanding. So it is important to highlight the structural similarities in fast acquisition algorithms to allow the design of a correlation unit that can be shared by all the signals. In the following sections the parameters of modern GPS signals are summarized and the principles of signal acquisition are presented. Signal-specific acquisition algorithms are then summarized, compared and numerically applied to real signals to show how a single correlation unit can be iterated to efficiently acquire all of them.

2 Modern GPS Signals

Every GPS satellite signs its navigation data by modulo-2 adding a code on it. The code is a binary sequence of a much higher frequency (compared to the data) with strong autocorrelation properties. This technique allows a receiver to extract the navigation data of the target satellite by correlating again the received signal with a local code replica. A modern GPS signal is generally composed by two channels. The *Data Channel* is used to transmit the navigation data correlated by a spreading code. The *Pilot Channel* is used to transmit a data-less code to aid the acquisition process. The correlation properties of each code, the *Primary Code*, can be enhanced by repeating it for a number of times

expressed by the length of a *Secondary Code*. In particular, the code (c) resulting from an N_p bits *Primary Code* (p) and an N_s bits *Secondary Code* (s) is given by Eq. 1, where the operator ‘ \otimes ’ is the Kronecker product.

$$c = p \otimes s = \begin{bmatrix} s_0 p \\ s_1 p \\ \vdots \\ s_{(N_s-1)} p \end{bmatrix} = \begin{bmatrix} s_0 p_0 \\ \vdots \\ s_0 p_{(N_p-1)} \\ s_1 p_0 \\ \vdots \\ s_{(N_s-1)} p_{(N_p-1)} \end{bmatrix} \quad (1)$$

- GPS L2C uses a data channel in which the navigation message is correlated with a 10230 bit code (CM code). The pilot channel uses codes of 767250 bits (CL code) and it is chip-by-chip multiplexed with the data channel before transmission on the same carrier phase.
- GPS L5 uses a data channel in which the navigation message is correlated with a primary code of 10230 bits (I5 code) that is repeated by a 10 bit secondary code (10 bit Neuman-Hofman code). The pilot channel uses codes of 10230 bits (Q5 code) repeated by a 20 bit Neuman-Hofman code. The two channels are transmitted on quadrature carriers.
- GPS L1C uses a data channel in which the navigation message is correlated with a 10230 bit code. The pilot channel uses a 10230 bits primary code and a 1800 bit secondary code (overlay code). Then the two channels are BOC-modulated to reduce the interference with the other signals on the L1 band. Finally the two channels are modulated on the same carrier phase and they are added together.

3 Signal Acquisition

Acquiring a GPS signal means to determine the presence/absence of the transmission of a target satellite in the received signal. This process is performed comparing the received code with a local code replica via the circular cross-correlation operation. It is known as a two-dimensional process as two parameters have to be determined to find a satellite transmission in the incoming signal:

- *Doppler Shift* has to be determined and removed to perform a valid correlation. In fact, incoming signal is affected by frequency shift because of the relative speed between the satellite and the receiver.
- *Code Phase* has to be estimated because the correlation of the signal with the local replica can give a peak only when the two signals are aligned.

It is generally accepted that the fastest way to perform signal acquisition is to compute the correlation in the frequency domain, as the correlation value can be

computed in parallel for all the code phases, in a single iteration of the algorithm, for a given Doppler shift.

$$z[n] = \sum_{m=0}^{N-1} c^*[m]x[n+m] = \mathcal{F}^{-1}\{C[k]^*X[k]\} \quad (2)$$

Equation 2 shows how the circular correlation $z[n]$ between the received signal $x[n]$ and the local code $c[n]$ can be performed both in the time domain and in the frequency domain. $x[n]$ and $c[n]$ are N-samples sequences, the \mathcal{F}^{-1} operator is used for the Inverse Discrete Fourier Transform operation (IDFT), $X[k]$ and $C[k]$ are the Discrete Fourier Transforms (DFT) performed on $x[n]$ and $c[n]$. Equation 2 can be expressed using the Fast Fourier Transform (FFT) algorithm to perform the correlation process:

$$z[n] = IFFT\{FFT\{c[n]\}^*FFT\{x[n]\}\} \quad (3)$$

A block that performs this operation with N-point FFTs can be called *N-point correlation unit*.

4 State of the Art Algorithms

Three state-of-the-art acquisition algorithms are now exposed to highlight the basis of their integration on a common core. The first, here called *Correlation Splitting* algorithm, can be applied to all the signals that have primary and secondary codes. The others are specifically targeted to the acquisition of L2C and L1C.

4.1 Correlation Splitting

The method presented in [1] allows to perform a correlation over the primary code of a GNSS signal, combining the results looking at the bits of the secondary code. In fact, if the signal has got a primary code of length N_p and a secondary code of length N_s , the acquisition would have to be performed with a single $(N_p \cdot N_s \cdot 2)$ -points correlation unit if the whole code is directly correlated with the signal. That approach is generally too resource demanding to be suitable for FPGA implementations. The Correlation Splitting method, instead, allows to compute the correlation value with $N_s \cdot N_s$ iterations of a $(2N_p \cdot 2)$ -points correlation unit.

4.2 Hypercoding

The principle of Hypercodes is to linearly add together small portions of a long code to obtain a new, shorter code (the Hypercode) [2]. In concept, different segments of the CL code are mutually orthogonal. Thus, creating a linear superposition of multiple segments results in a shorter code containing all of the information in the total CL code, but in a shorter form. An Hypercode is generated dividing the original code in N_s segments. Then M couples of segments are

added together to form the hypercode, as shown in Fig. 1. A zero-padded version of signal is correlated with different hypercodes until a peak is found. Then the signal is correlated again with the segments that forms the hypercode to deduce the exact code displacement.

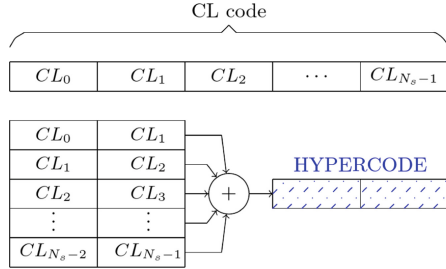


Fig. 1. Hypercode generation process.

4.3 Joint Data-Pilot Acquisition

A set of codes that accounts for both the data code and pilot code together is proposed in [3] to exploit all the power transmitted on the two channels of L1C. In this case the presence of a secondary code on the pilot channel is accounted generating two different code replicas. In fact, the secondary code is constant over a data bit period and it can assume only two values. Then the two codes are frequency-shifted on the subcarriers given by the BOC modulation, resulting in four correlating codes. So the acquisition is performed over a primary code period with four different codes.

5 Numerical Analysis

In this section the presented algorithms are numerically applied to their target signals to highlight their common numerical characteristics.

5.1 Correlation Splitting Applied on L5

A direct correlation with 20 ms of signal is required without this method. It would lead to an enormous $N_s \cdot N_p \cdot 2 = 409200$ points FFT block. The factor 2 comes from the need to double the code rate. In fact, the correlation between a code and an its half-chip misaligned replica is not zero. This means that, after acquiring the signal, a tracking loop can continuously correlate the signal with an half-chip misaligned replica and compare the value obtained with an aligned replica. This operation helps to understand if the receiver is loosing the lock with the signal because of a code displacement. This method allows to reach the final result for a single frequency bin combining $N_s \cdot N_s = 400$ of $2N_p \cdot 2 = 40920$ point FFT calculations. However the power of 2 nearest to 40920 is $65536 (= 2^{16})$, so a

65536-point FFT is needed if a radix-2 algorithm is used to implement the FFT. This block is too big for FPGA implementation. The process can be optimized with the FFT Splitting method presented in [4], that allows to compute 40920-points FFTs with a sequence of smaller FFTs.

5.2 Hypercoding Applied on L2C

Also with this method it is needed the computation of 40920 points FFTs. In fact the 767250-bit CL code can be divided in $N_s = 75$ parts, obtaining a 10230-bit segment. Then, the segment doubles its length when it is zero padded and when it is oversampled to have an half-chip resolution, reaching 40920 bits. The signal segment has to be correlated 15 times with the hypercodes if 5 code segments are added together to form an hypercode. Then the signal has to be correlated again with the 5 code segments that forms the hypercode that gave a peak. So this algorithm can be solved iterating a 40920-point correlation unit for 20 times. The FFT Splitting method can also be used here to efficiently perform that computation using three 16384 FFTs or five 8192 FFTs [4].

5.3 Joint Data-Pilot Acquisition Applied on L1C

Each code version that correlates with the signal has got the same 10230-bit length of the primary codes in both data and pilot channels. The codes have to be represented with 20460 bits to reach an half-chip precision. So correlation units of 20460-points FFTs are needed to perform the acquisition task. Doubling the data length or zero-padding the signal allows L1C acquisition to be fitted on 40920-points correlation units, as for L5 and L2C. The algorithm is performed iterating 4 times the correlation unit, as 4 local code versions are used.

6 Concept of Unified Core Architecture

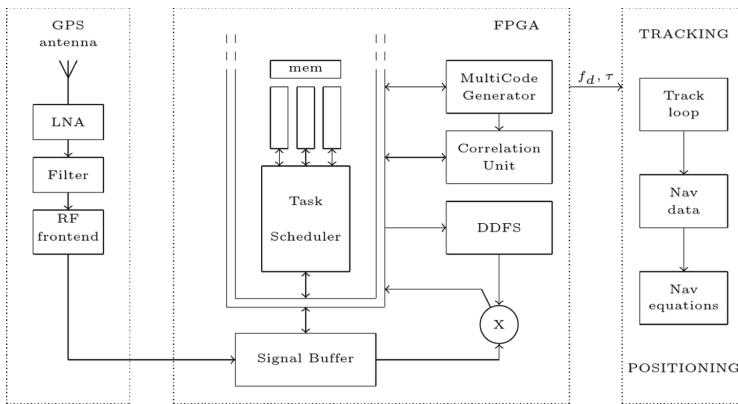


Fig. 2. Unified acquisition architecture.

An acquisition architecture based on a single correlation unit can be investigated as we proved that all the modern GPS signals can be acquired iterating the same computational unit. In fact, a 40920-point FFT-based correlation unit can be designed and implemented on hardware to solve the basic correlation operations (Fig. 2). The correlation unit shall be as fast and optimized as possible because the whole system performances are based on its area and timing results. A 40920-point correlation unit can be efficiently implemented with a 9-FFT architecture [4]. This solution uses 18% of the memory resources of a Stratix V GX FPGA, configuring the Intel FFT IP Core with a 16-bit, streaming architecture. Another key element for the core feasibility is the memory subsystem. In fact, the acquired signal and the final correlation vector have to be stored in a memory. Its length is determined by the L5 acquisition algorithm and it is quantified in $2N_s N_p$ complex samples (16 bits for both real and imaginary parts). This results in the occupation of 800 M20K FPGA blocks for each buffer (60% of the total memory resources). The remaining memory can be used to store partial 40920 bits correlation vectors (3% for each partial result). A code generation block has to be implemented to give a code that is specific for the target satellite and the target signal. This block shall include three similar LFSR-based architectures. The whole unit has to be driven by a task scheduler which handles the partial results as prescribed by the actual executing algorithm. For example it shall weight the results with the secondary code bins and integrate it to perform L5 acquisition. It shall also search for a peak in the hypercode correlation and communicate with the code block to trigger the generation of the right code.

7 Conclusions

In this study we have shown how three acquisition algorithms can be used to acquire the GPS L5, L2C and L1C signals by the iteration of 400, 20 and 4 identical 40920-point correlation units. A general feasible architecture has been proposed for the future implementation of a multi-band GPS receiver capable to acquire all of the modern GPS signals with an area and timing optimized computational core.

References

1. Leclère, J., Botteron, C., Farine, P.: High sensitivity acquisition of GNSS signals with secondary code on FPGAs. *IEEE Aerosp. Electron. Syst. Mag.* **32**(8), 46–63 (2017). <https://doi.org/10.1109/MAES.2017.160176>
2. Zeng, Q., Tang, L., Zhang, P., Pei, L.: Fast acquisition of L2C CL codes based on combination of hyper codes and averaging correlation. *J. Syst. Eng. Electron.* **27**(2), 308–318 (2016). <https://doi.org/10.1109/JSEE.2016.00031>
3. Zhou, J., Liu, C.: Joint data-pilot acquisition of GPS L1 civil signal. In: 2014 12th International Conference on Signal Processing (ICSP), Hangzhou, pp. 1628–1631 (2014). <https://doi.org/10.1109/ICOSP.2014.7015271>
4. Leclere, J., et al.: FFT splitting for improved FPGA-based acquisition of GNSS signals. *Int. J. Navig. Obs.* (2015)



Evaluating Body Movement and Breathing Signals for Identification of Sleep/Wake States

Maksym Gaiduk^{1,2(✉)}, Ralf Seepold^{1,3}, Natividad Martínez Madrid^{3,4},
Thomas Penzel^{5,6}, Lucas Weber¹, Massimo Conti⁷, Simone Orcioni⁷,
and Juan Antonio Ortega²

¹ HTWG Konstanz, Alfred-Wachtel-Street 8, 78462 Konstanz, Germany
{maksym.gaiduk,ralf.seepold}@htwg-konstanz.de

² University of Seville, Avda. Reina Mercedes S/N, Seville, Spain

³ I.M. Sechenov First Moscow State Medical University, Moscow, Russian Federation

⁴ Reutlingen University, Alteburgstr. 150, 72762 Reutlingen, Germany
natividad.martinez@reutlingen-university.de

⁵ Charité Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany
thomas.penzel@charite.de

⁶ Saratov State University, Saratov, Russian Federation

⁷ Università Politecnica Delle Marche, Via Brecce Bianche 12, 60131 Ancona, Italy
{m.conti,s.orcioni}@univpm.it

Abstract. Recognition of sleep and wake states is one of the relevant parts of sleep analysis. Performing this measurement in a contactless way increases comfort for the users. We present an approach evaluating only movement and respiratory signals to achieve recognition, which can be measured non-obtrusively. The algorithm is based on multinomial logistic regression and analyses features extracted out of mentioned above signals. These features were identified and developed after performing fundamental research on characteristics of vital signals during sleep. The achieved accuracy of 87% with the Cohen's kappa of 0.40 demonstrates the appropriateness of a chosen method and encourages continuing research on this topic.

1 Introduction

Sleep is an essential part of our life. It affects our wellbeing, mood, and indeed health [1]. Its influence on the risk of coronary heart disease [2], hypertension [3], and other health issues [4] is documented in numerous scientific studies. Several features can be measured during sleep to provide its characteristic [5]. One of the important ones is sleep quality that can be measured using several methods [6]. Another essential parameter is sleep duration, and there are two main groups of approaches for its measurement: objective and subjective [7]. However, it is known that there is some disagreement between both these methods [8, 9], among other things considering a measurement of wake after sleep onset (WASO) [10].

Objective measurement is based on the usage of devices measuring physiological signals and providing a quantitative output; in sleep/wake states recognition, actigraphy

is commonly used [11]. Another traditional approach is polysomnography (PSG). This method has the advantage of high accuracy (it is a gold standard approach) and is performed according to the guidelines of the American Academy of Sleep Medicine (AASM) [12]. However, it also has several disadvantages, including reduced comfort for the user, high material and personnel costs [13]. Therefore, the development of an alternative method that would overcome these problems is a relevant research topic.

Identification of sleep and wake states is essential not only for the determination of total sleep time (TST) but also of WASO, of number of nighttime awakenings, etc. Therefore, it is essential not only to identify the beginning and the end of the sleep but also to determine all periods of (short) wakefulness over the night.

This work aims to develop an approach for the epoch by epoch identification of sleep/wake states, analyzing only the signals that can be obtained in a contactless way: movement and breathing. This would lead to an increase of the comfort for the users, and in case of using the systems placed under the mattress like described in [14–16] would also substantially reduce the necessary costs and enlarge the convenience of system use, which would subsequently lead to the rise of the technology acceptance [17].

2 Methods

In this section, the used methods are being described. To keep a clear structure, this section is divided into two subsections: “Multinomial Logistic Regression” (MLR), where the selected algorithm is presented, and “Derived Vital Parameters” with the description of the calculated features used as the input for the MLR model.

2.1 Multinomial Logistic Regression

It is known that changes of vital parameters (among others breathing and movement) during sleep correlate with a current sleep state [18, 19]. For the mathematical description of this correlation, regression analysis may be used. After performing literature research, the decision was made to use MLR, a particular case of regression analysis [20].

MLR allows us to calculate the probabilities of dependent variables/classes (Sleep/Wake states in our case), interpreting the values of independent variables that in our case may be represented by vital parameters. However, to obtain the results of higher quality, it is recommended not to apply the MLR model directly to respiration and movement signals but to utilize derived out of these signals parameters [19]. One of the advantages of MLR compared to some other logistic regression models is that it may be used if there are more than two possible discrete outcomes [21].

2.2 Derived Vital Parameters

Thorough literature research was conducted to find an optimal set of vital bio parameters. Among other things, the specifics of the changes in biological values during sleep were extensively studied. Particular focus was made on the parameters that can be recorded non-invasively. After reviewing the possible hardware solutions for such recordings (e.g. [14, 16]), the decision was made to use derived parameters from the movement and

respiration signals. These signals may be recorded non-obtrusively and in a comfortable for a user way [14, 15].

Since the movement takes place in three axes (X, Y, Z), they must all be considered to calculate the respective feature. Therefore, average Body Movement (BM) per epoch (30 s – expected value in sleep studies) is used as input for the MLR model, which is calculated as follows:

$$BM = \frac{1}{n} \sum_{i=0}^{n-1} \sqrt{X_i^2 + Y_i^2 + Z_i^2} \quad (1)$$

where n is the number of body movements during the epoch and X_i , Y_i and Z_i are calculated as a difference of the sensor's position between the current and previous moment.

Two derived out of breathing signal parameters were selected for the usage. Their mathematical representation, according to [22] is the following:

$$V_{in} = \text{median} \left(\sum_{i=1}^k S_{x \in \Omega_i^{in}} \right) \quad (2)$$

$$T_{sdm}(k) = \frac{\text{median}(t_1, t_2, \dots, t_n)}{IQR(t_1, t_2, \dots, t_n)} \quad (3)$$

where k is the number of breathing cycles in one epoch (of 30 s), $S_1, S_2, \dots, S_x, \dots, S_K$ are the volumes of the respiratory effort signal, Ω_i^{in} is the corresponding breathing inhalation, t_1, t_2, \dots, t_n is the sequence of exhalations in one epoch and containing n troughs. IQR is their interquartile range calculated as the difference between the 1st and the 3rd quartile [22].

In addition to these three mentioned above-derived parameters, the decision was made to introduce one novel feature BMV , designed based on the performed study of changes of bio vital parameters during sleep. It includes already presented features BM (Eq. (1)) and V_{in} (Eq. (2)). It emphasizes the fact that body movement is being decreased and breathing effort amplitudes are getting more stable (especially in NREM stages) in the “Sleep” state compared to the “Wake” [22]. It is calculated according to the equation:

$$BMV = \ln \frac{BM(k)}{V_{in}(k) + BM(k)} \quad (4)$$

where k is the actual time period.

3 Results

The evaluation of the proposed algorithm was performed using the PSG dataset from the Center of Sleep Medicine at Charité clinic in Berlin¹. This approach provides exact

¹ Initial study was carried out in the Interdisciplinary Center for Sleep Medicine of Charité-Universitätsmedizin Berlin, Charitéplatz 1, D-10117 Berlin (Germany).

data and is previously analyzed by sleep experts, which was relevant for the evaluation of the system. However, PSG is recorded obtrusively, but also other contactless methods [15, 16] may be used for the obtaining of the necessary signal. In total, 18 193 epochs of 30 s each were used as a test dataset, corresponding to more than 150 h of sleep. Additionally, about 100 h of sleep recordings were used for the training of the algorithm (randomly selected subjects after separation male/female). Every epoch was tagged by sleep medicine doctors with the corresponding sleep stage. For the aims of this work N1, N2, N3, and REM stages were combined into a “Sleep” state.

The study participants had an average age of 38.6 ± 14.5 years old, and their average BMI was 24.4 ± 4.9 kg/m². As far as known, they did not have any significant health disorders, and the ratio of male and female subjects who participated in the conducted study was similar.

All described in this article procedures involving human subjects were approved by the Institutional Review Board of the Charité-Universitätsmedizin Berlin (application number: EA1/320/114).

The confusion matrix with the results of Sleep/Wake states classification is presented in Table 1. The rows represent the classification done by the sleep specialists and the columns – predictions done by the developed algorithm. The numbers are representing the count of epochs identified in the corresponding state.

Table 1. Results of sleep/wake states identification

Stage expert	Stage SW-algorithm		
	Wake	Sleep	Total
Wake	1110	1009	2119
Sleep	1441	14633	16074
Total	2551	15642	18193

Its dominance in a sleep pattern can partially explain some overestimation of the “Sleep” state – the sleep medicine doctor tagged almost 90% of epochs in the analyzed recordings as “Sleep”, which naturally affects the algorithm.

The results of the algorithm’s evaluation are presented in Table 2. The achieved overall accuracy is 87% (Sleep: 91%, Wake: 52.4%), and the calculated Cohen’s kappa is 0.40, which is “fair to good” according to [23].

Table 2. Accuracy and Cohen’s kappa values

Overall accuracy	87%
Cohen’s kappa	0.40

4 Conclusion and Outlook

The section's "Results" outcomes have confirmed the appropriateness of the described in this manuscript approach. However, some improvements of the algorithm are still possible, in particular, to overcome the perceived overestimation of the "Sleep" state. For that, several methods may be considered:

- Selection of additional derived parameters with a strong correlation with "Sleep" or "Wake" state.
- The weighting of the used features increases the influence of the most relevant of them.
- Maintaining the set of selected parameters with modifying the processing algorithm, for instance, application of artificial intelligence.

Another aim of future work is to consolidate the algorithmic software part described in this document with the hardware system for measurement of respiratory and body movement signals like described in [14, 15]. This combination will provide a standalone contactless system for continuous sleep monitoring that can be used in both the in-home and medical centers.

Acknowledgments. We thank the Interdisciplinary Center for Sleep Medicine of Charité Clinic in Berlin and, in particular, Dr. rer. medic. Martin Glos for supporting the study.

This research was partially funded by the EU Interreg V-Program "Alpenrhein-Bodensee-Hochrhein": Project "IBH Living Lab Active and Assisted Living", grants ABH40, ABH41, and ABH66 and by the German Federal Ministry For Economic Affairs And Energy, ZiM project "Sleep Lab at Home" (SLaH) grant: ZF4825301AW9.

References

1. Mukherjee, S., et al.: An official american thoracic society statement: the importance of healthy sleep. Recommendations and future priorities. *Am. J. Respir. Crit. Care Med.* **191**, 1450–1458 (2015)
2. Lao, X.Q., et al.: Sleep quality, sleep duration, and the risk of coronary heart disease: a prospective cohort study with 60,586 adults. *J. Clin. Sleep Med.* **14**, 109–117 (2018)
3. Grandner, M., Mullington, J.M., Hashmi, S.D., Redeker, N.S., Watson, N.F., Morgenthaler, T.I.: Sleep duration and hypertension: analysis of 700,000 adults by age and sex. *J. Clin. Sleep Med.* **14**, 1031–1039 (2018)
4. Chaput, J.-P., et al.: Sleep duration and health in adults: an overview of systematic reviews. *Appl. Physiol. Nutr. Metab.* **45**, S218–S231 (2020)
5. Gaiduk, M., Seepold, R., Ortega, J.A., Martínez Madrid, N.: Comparison of sleep characteristics measurements: a case study with a population aged 65 and above. *Procedia Comput. Sci.* **176**, 2341–2349 (2020)
6. Gaiduk, M., et al.: A comparison of objective and subjective sleep quality measurement in a group of elderly persons in a home environment. In: Saponara, S., De Gloria, A. (eds.) *ApplePies 2020. LNEE*, vol. 738, pp. 286–291. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-66729-0_35

7. Schokman, A., et al.: Agreement between subjective and objective measures of sleep duration in a low-middle income country setting. *Sleep Health* **4**, 543–550 (2018)
8. Lauderdale, D.S., Knutson, K.L., Yan, L.L., Liu, K., Rathouz, P.J.: Self-reported and measured sleep duration: how similar are they? *Epidemiol. (Cambridge, Mass.)* **19**, 838–845 (2008)
9. Matthews, K.A., et al.: Similarities and differences in estimates of sleep duration by polysomnography, actigraphy, diary, and self-reported habitual sleep in a community sample. *Sleep Health* **4**, 96–103 (2018)
10. Short, M.A., Gradisar, M., Lack, L.C., Wright, H., Carskadon, M.A.: The discrepancy between actigraphic and sleep diary measures of sleep in adolescents. *Sleep Med.* **13**, 378–384 (2012)
11. Kryger, M.H., Roth, T., Dement, W.C.: *Principles and Practice of Sleep Medicine*. Elsevier, Philadelphia (2005)
12. Berry, R.B., Quan, S.F., Abreu A.R., et.al: The AASM manual for the scoring of sleep and associated events. Rules, terminology and technical specifications: Version 2.6. American Academy of Sleep Medicine, Darien, Illinois (2020)
13. Hirshkowitz, M.: Polysomnography challenges. *Sleep Med. Clin.* **11**, 403–411 (2016)
14. Gaiduk, M., Seepold, R., Martínez Madrid, N., Orcioni, S., Conti, M.: Recognizing breathing rate and movement while sleeping in home environment. In: Saponara, S., De Gloria, A. (eds.) *ApplePies 2019*. LNEE, vol. 627, pp. 333–339. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-37277-4_38
15. Gaiduk, M., Wehrle, D., Seepold, R., Ortega, J.A.: Non-obtrusive system for overnight respiration and heartbeat tracking. *Procedia Comput. Sci.* **176**, 2746–2755 (2020)
16. Gaiduk, M., Vunderl, B., Seepold, R., Ortega, J.A., Penzel, T.: Sensor-mesh-based system with application on sleep study. In: Rojas, I., Ortuño, F. (eds.) *IWBIBIO 2018*. LNCS, vol. 10814, pp. 371–382. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-78759-6_34
17. Pai, F.-Y., Huang, K.-I.: Applying the technology acceptance model to the introduction of healthcare information systems. *Technol. Forecast. Soc. Chang.* **78**, 650–660 (2011)
18. Kurihara, Y., Watanabe, K.: Sleep-stage decision algorithm by using heartbeat and body-movement signals. *IEEE Trans. Syst., Man, Cybern. A* **42**, 1450–1459 (2012)
19. Gaiduk, M., Penzel, T., Ortega, J.A., Seepold, R.: Automatic sleep stages classification using respiratory, heart rate and movement signals. *Physiol. Measur.* **39**, 124008 (2018)
20. Hosmer, D.W., Lemeshow, S., Sturdivant, R.X.: *Applied Logistic Regression*. Wiley, New York (2013)
21. Greene, W.H.: *Econometric analysis*. Pearson Education, Boston (2012)
22. Long, X., Foussier, J., Fonseca, P., Haakma, R., Aarts, R.M.: Analyzing respiratory effort amplitude for automated sleep stage classification. *Biomed. Signal Process. Control* **14**, 197–205 (2014)
23. Fleiss, J.L., Levin, B., Paik, M.C., Shewart, W.A., Wilks, S.S.: *Statistical Methods for Rates and Proportions*. John Wiley & Sons Inc., Hoboken (2003)



Comparison of a Medical-Grade and an Open ECG Biosensor Using a Soft Real-Time m-Health Platform

Miltos D. Grammatikakis^(✉) and Michael Androulakis

Hellenic Mediterranean University, 71410 Heraklion, Greece
mdgramma@cs.hmu.gr, tp1945@edu.hmu.gr

Abstract. Biosensor devices transform healthcare services from a physician, hospital- or clinic-centric system to one that directly involves patients. In this work, we evaluate, compare and analyze two complex biosensors, connected over Wi-Fi to a server. The first one is a complex industrial biosensor (called STMicro Bodygateway, or BGW) integrated with a multithreaded driver that runs on a single board computer (Odroid XU4). The driver supports data acquisition from multiple BGW sensors, capturing raw Bluetooth packets via rfcomm, processing them to retrieve associated biosignals, and transmitting extracted biometric data over Ethernet to a server for soft real-time analysis of arrhythmias. The second one is ProtoCentral's HeartyPatch, an open architecture that can be connected directly to the server over Wi-Fi. We consider power consumption and soft real-time signal acquisition, analysis, and visualization when both biosensors transmit ECG at 128 Hz to a low-cost Odroid XU3 board, which acts as the server. While our platform sustains soft real-time for either sensor, the open-source HeartyPatch solution shows a slightly better behavior, in terms of performance, power consumption, and data precision.

1 Introduction

Electrocardiography (ECG) examines the electrical signal taken from electrodes connected to the body to detect the heart rhythm. Using electrodes is more accurate compared to optical photo-plethysmogram (PPG) methods that illuminate the skin and essentially measure changes in light absorption. Duration, amplitude, and morphology of the main ECG spikes (QRS complex) is the main medical stress test (e.g., a short test on a treadmill or stationary bike, or 24-h Horter) for diagnosing heart diseases, such as cardiac arrhythmias. During an arrhythmia event, the heart may beat more slowly (bradycardia), or faster than normal (tachycardia), or with an irregular rhythm (e.g. with premature or extra heartbeats). Arrhythmias are caused by changes to heart tissue, imbalanced hormone level or electrolytes, exertion or stress, as a side effect of medication, or disorders of the electrical activity in the heart. They can lead to life-threatening situations, such as stroke, heart failure, or sudden cardiac arrest requiring immediate notification.

In this context, we consider a pervasive in-hospital use case, whereas a typical hospital server logs and analyzes patient data from a biosensor continually in soft real-time; in practice, the hospital server is often extended to a cloud-based analysis platform

monitored by care personnel, and more recently, artificial intelligence data acquisition/analysis algorithms. More specifically, we consider a continuous analysis of the ECG signal for detection of ventricular arrhythmias, together with automated visualization of the annotated waveform in soft real-time. We use filters from WFDB and OSEA, two open-source ECG analysis packages for obtaining the R-R metric for heart-rate variability. The R-R metric is defined as the time interval between consecutive apex points in the QRS complex which correspond to the depolarization of the right and left ventricles of the heart. R-R metric can also be applied to other clinical or physiological functions of the autonomous nervous system, such as detection of stress, sleepiness, or emotions.

In our experimental framework, we use two sensors: a) a typical medical-grade 1-lead Bluetooth-based pulse sensor with an ARM microcontroller supporting FreeRTOS (ST Micro BodyGateway, or BGW, similar to Preventice BodyGuardian (now Boston Scientific), and, b) a 1-lead ECG sensor, HeartyPatch from Protocentral (Karnataka, India), supporting FreeRTOS and Bluetooth or WiFi connectivity. We connect both sensors to the same low-cost single-board computer (Odroid XU3), acting as a Server. A comparison of specifications with other popular sensors is shown in Table 1. The main motivation for choosing BGW and HeartyPatch is their similarity, affordable cost, high resolution, and the relatively large range of supported pulse rates.

Table 1. A comparison of BGW and heartyPatch sensors with industrial ones

	Shimmer	Telemetry ePatch	Apple watch Series 6	BGW & preventice BodyGuardian	HeartyPatch
No. Leads	4	1	1	1	1
Sampling rate (Hz)	500 (recommended)	128, 256, 512, 1024	? (closed specs)	128 (BGW only), 256	128, 256, 512
Resolution (bits)	16	16	? (closed specs)	12	18
Certification	-	HIPAA	Fitness only	HIPAA	-
Sensors	ECG, ACC, EMG	ECG	ECG, 6-axis ACC, GPS, SpO2/VO2, Ambient Light	ECG, ACC, Bio-impedance	
Connectivity	BT, SD Card	USB	WiFi, BT, GSM/HSPA/LTE	BT	WiFi, BT, UART
Battery	Li-Ion Rechargeable 450 mAh	Li-Ion Rechargeable 450 mAh, 4.2 V	Rechargeable 304 mAh	Li-Ion Rechargeable 380 mAh	Li-Ion Rechargeable 380 mAh

Our framework running on the Server extends open-source ECG analysis packages (Harvard Physionet WFDB, OSEA, and WAVE) towards soft real-time monitoring and

analysis of arrhythmias (atrial and ventricular fibrillation) and visualization. Unlike our paper, most commercial products, such as AliveCor [1], LifeMonitor [2], NowCardio [3], and PhysioMem [4] only target real-time monitoring on smartphones. ECG analysis is performed offline by a physician, i.e., after the ECG signal from a patient-worn biometric device processed by the smartphone is transferred to a data center. Concerning real-time analysis, Apple Smart Watch Series 6 supports periodic analysis of a short ECG signal, obtained using PPG, specifically atrial fibrillation, and informs the user upon the occurrence of 5 out of 6 consecutive cases of atrial arrhythmia [5]. However, unlike our solution, this product is not medical-grade or open-source, is not server-based, and is not able to detect ventricular fibrillation, a more serious situation that frequently results in loss of consciousness and death due to interrupting blood supply.

When devices operate at 128 ECG pulses/s, we show that we can sustain soft real-time ECG analysis and visualization for both devices. Moreover, the open, HeartyPatch-based solution shows a slightly better behavior, in terms of accuracy (16- vs 12-bit causing fewer spurious annotations), performance (fewer spikes), and smaller projected power consumption that allows continuous ECG transmission using the same Li-Ion 380 mAh battery for more than 20 h versus approximately 14 h).

Next, in Sects. 3 and 4 we discuss the sensor architecture (BGW and HeartyPatch) and the application. Section 4 focuses on the experimental framework and results. Finally, Sect. 5 provides a summary and discusses future work.

2 Sensor Architecture and Configuration: BGW and HeartyPatch

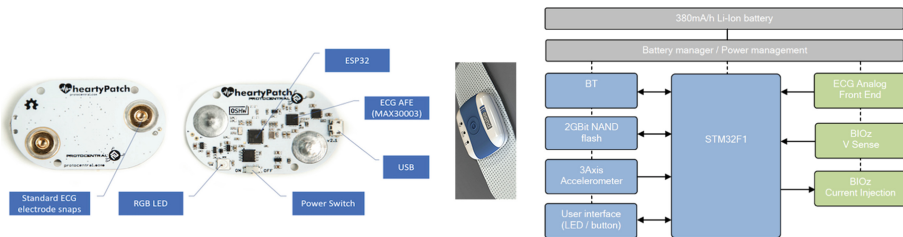


Fig. 1. HeartyPatch and BGW sensor hardware.

The ProtoCentral HeartyPatch (HP), shown in Fig. 1 (left), is an open-source, single-lead, wearable pulse sensor, capable of capturing ECG data by placing it on the patient's skin [6]. It features the popular two-core ESP32 SoC microcontroller and the Maxim MAX30003 analog front-end for capturing ECG and possibly obtaining Heart Rate Variability (HRV). It provides Wi-Fi/Bluetooth connectivity and RGB LED for status display. It is a low-cost device, intended for lifestyle (e.g., fitness or well-being applications), as well as for diagnostics and medical research (e.g., stress level). Despite its dual communication support, we make use of Wi-Fi and examine continuous ECG streaming over TCP, to compare the performance of the BGW firmware. Since the device is preloaded with a firmware version that works over BLE, we have flashed the official Wi-Fi capable

image, but shortened the packet structure (as explained below). The device sets the ECG data rate to 128 pulses/s and registers FreeRTOS logs. It also connects to our router and acts as a slave device (waiting for client connections). Detailed instructions for building the ESP32 environment with FreeRTOS and HeartyPatch firmware are per the official website [6].

The original ECG packet used in the firmware contains the following information.

```

struct RawPacket {

    static constexpr unsigned SamplesCount = 8; //amount of ECG samples contained in the payload

    uint16_t packetStart, //should be 0xFA0A
            payloadSize; //size of the Payload structure
    uint8_t version; //should be 3

    struct Payload {
        uint32_t id; //serial\sequence number
        timeval timestamp; //timestamp taken while assembling the packet
        uint32_t rr; //computed R-to-R value in milliseconds
        int32_t samples[SamplesCount]; //ECG samples; upper 18bits is the actual sample
    } payload;

    uint16_t packetStop; //should be 0x0BF0

};

```

The STMicro BodyGateway (BGW), shown in Fig. 1 (right), is a medical-grade, single-lead, wearable patch, widely used for remote monitoring of cardiac and respiratory functions. Its newer smaller version, Preventice Bodyguardian mini has been used to record data for over 750K patients per year in the USA [7]. The BGW device integrates different sensors (ECG, accelerometer, respiration rate, etc.) with an ARM STM32 F-series microcontroller and a real-time operating system (FreeRTOS). The system can be programmed to transmit vital physiological data at different frequencies using short cluster-structured packets of 256 bytes over an integrated Bluetooth interface (BT 3.0). BGW can be programmed to acquire, digitize, and either stream sensor data, or monitor (i.e., save in its internal 2 Gbit NVM memory and periodically transmit).

Based on STMicro specifications (covered by NDA in EU project DREAMS), we have written a multithreaded GNU/Linux driver that manages a complex BT-to-Wireless (IEEE 802.11) protocol stack [8]. The first thread configures the biosensor to send a specific biosignal, e.g., ECG at 128 pulses/s, the second one captures raw BT packets (via rfcomm protocol) and extracts data to a shared list, while the third one transmits data from the list to a connected Server over Ethernet. The **BGW driver** runs on an ODROID XU4 server (ARM v7a architecture) and **transmits only the necessary information**, i.e. up to 256 samples of ECG data per sec (with 12-bit precision). Therefore, we have modified the ECG packet in the HeartyPatch firmware to only include samples, without any overhead. This corresponds to a total of 8 samples of ECG data per packet (with 18-bit precision). A total of 16 HeartyPatch packets are transmitted per sec.

3 Soft Real-Time ECG Analysis and Timing Infrastructure at Server

Our embedded soft real-time application [9] relies on two open-source software libraries: a) the Harvard Physionet WaveForm DataBase (WFDB) [10] is used to smooth and standardize the ECG signal transmitted by the BGW driver to 200 samples/s according to ANSI/AAMI EC-13, and b) the EP Limited Open Source ECG Analysis (OSEA) to perform low- and high-pass QRS filtering (via easytest script) for heartbeat detection and classification to normal or abnormal beats [11–13]. To manage ECG annotation in soft real-time, we extend easytest functionality to avoid re-computation by applying a training signal only on the latest data; our framework theoretically achieves a positive predictivity close to 99.8% when using MIT/BIH and AHA arrhythmia databases. Other ECG analysis methods result in smaller predictivity rates [14]; deep learning techniques are promising, see Preventice study using BodyGuardian sensor (BGW successor) [15]. Finally, for viewing, annotation, and interactive analysis of ECG waveform in soft real-time (with asynchronous display) we use the Harvard Physionet WAVE software package [16]. It is based on a 32-bit XView open-source toolkit (a low-level XWindows client).

For evaluating soft real-time, we share performance statistics, such as the number of samples, latency, throughput, and packet loss, across different application processes using a dynamic shared memory timing infrastructure that supports simple (fast) atomic shared memory read/write operations. This infrastructure, based on POSIX shared memory, allows the definition of shared memory objects using the MAC id, process id, and/or network IP as part of their name.

4 Experimental Framework - Performance, Power Dissipation, and Precision

In our use case, we consider soft real-time ECG analysis at the server (Odroid XU3), when a single BGW or HeartyPatch pulse sensor is operated at 128 pulses/s; Odroid XU3 is a single-board computer based on ARM big.LITTLE multicore architecture. The big cluster consists of powerful ARM CortexA15 quad-cores clocked from 200 MHz to 2000 MHz at intervals of 100 MHz, while the LITTLE CPU cluster comprises low-power quad-core Cortex-A7 capable of operating at a cluster-wise frequency of 200 MHz to 1400 MHz at discrete intervals of 100 MHz. Our application runs on two big Cortex-A15 cores: a) Core 0: server receiving BGW/HeartyPatch data, and b) Core 1: animator analyzing and visualizing the annotated ECG.

The BGW device transmits ECG data via Bluetooth to an XU4. The Odroid XU4 device runs the Bluetooth-to-WiFi BGW Driver to transfer ECG signal data to the Server (Odroid XU3), via a 2.1 Gbit/sec router TP-Link Archer C5400. The HeartyPatch device also transmits ECG via the same router directly to the Server. Using our timing infrastructure, we collect measurements each time new ECG data is uploaded to the WAVE tool for visualization (via wave-remote). This includes the number of ECG samples processed, current time, and latency distribution to different ECG analysis subprocesses.

Figure 2(a) examines the average rate during visualization. Based on the number of samples, and the current timestamp that the data is visualized using WAVE (via

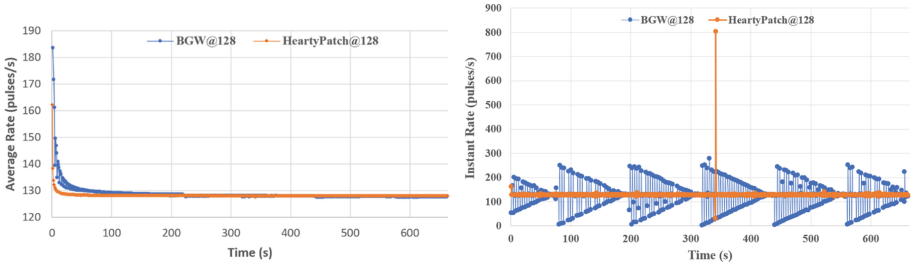


Fig. 2. a) Average processing rate, and b) Instant rate during server visualization when BGW and HeartyPatch operate at 128 pulses/s.

wave-remote), we compute the average processing rate of ECG data. From this graph, we observe that we can sustain soft real-time for both sensors when they operate at 128 pulses/s. Figure 2(b) shows the instant rate during visualization. We observe a large fluctuation around the average value for BGW, but not for HeartyPatch. This may occur, since BGW transmits ECG samples in bursts (up to 128 samples in a single burst), while the ECG data flow is more regular for HeartyPatch (with a maximum of 8 samples per burst). Notice that for BGW, missing packet information, e.g. due to missing an RTOS deadline, is subsequently transmitted in the next interval with a special (so-called, incomplete) packet. Incomplete packets are ~15% of the complete ones. For HeartyPatch, incomplete packets do not occur. However, it is possible that during some time interval, HeartyPatch packets are momentarily delayed, but this is rectified at the subsequent interval (see the peak in HeartyPatch instant rate at 340.1 s). In our experiments, we find that this occurs 2.4% of the time; this rise in the rate can be due to a prior system call or interrupt occurring in the Linux kernel at the server.

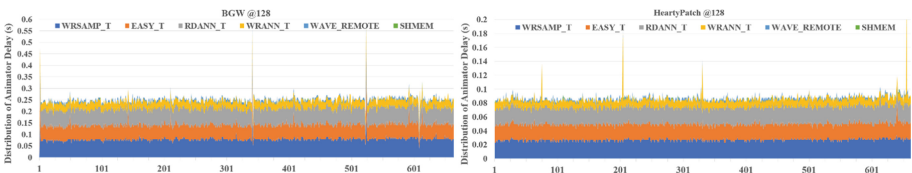


Fig. 3. Distribution of animation delays: a) BGW and b) HeartyPatch; both sensors operate at 128 pulses/s)

In Fig. 3 we examine the distribution of the different processing delays during animation when both pulse sensor devices operate at 128 pulses/s. The delays are normalized, i.e. they all assume the processing of 128 pulses. As shown, server delays for HeartyPatch are much shorter than those of BGW. The contribution to the different processing delays is similar for both sensors. This is mainly due to: a) wrsamp method used for conversion to std EC-13, b) easytest filtering used for heartbeat detection and classification, and c) wrann/rdann used for writing/reading to/from annotation files related to the latest data. Contribution from wave-remote, locking and shared memory constructs are marginal.

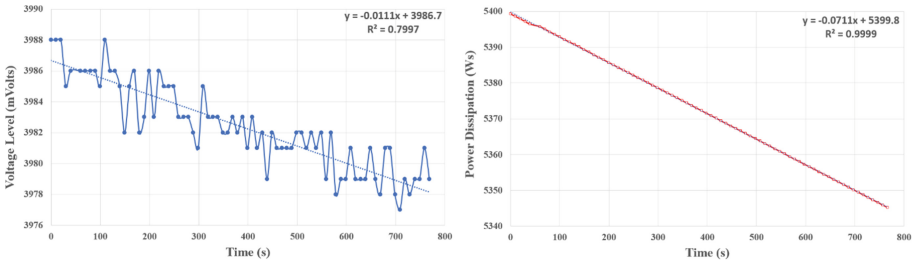


Fig. 4. Power dissipation during ECG transmission for both sensors: a) battery level (in mVolts) for BGW and b) power dissipation (in Watt secs) for HeartyPatch.

Both sensors use a 3.7 V rechargeable Li-Ion battery with an average capacity of 380 mA (equivalent to ~ 1.5 Wh). Figure 4 shows power dissipation during ECG transmission for BGW (on the left), and HeartyPatch (on the right). For BGW, battery depletion is obtained by programming the sensor to transmit its battery level every 10 s (along with ECG data). Using a least-squares trendline, we find that the BGW battery depletes to 3.4 V ideally after a maximum of 13.5 h of ECG transmission. For HeartyPatch, we have plugged the Wi-Fi enabled Odroid SmartPower 2 energy sensor [17] to the 5 V power supply of the board to measure its power consumption with a sample rate of 1 Hz (1 s). SmartPower captures the voltage (volt), current (Ampere), power (milliwatt), and energy consumed (kWh). Using a similar trendline model, we deduce that the HeartyPatch would support 21.1 h of ECG transmission before the battery is depleted. In comparison, continuous streaming of all data (ECG, ACC, respiration, battery, notifications) at the maximum sampling frequency of the BGW reduces the autonomy span to just 3 h.

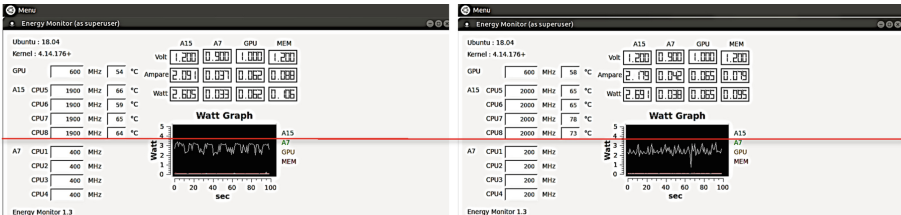


Fig. 5. Power consumption and annotated ECG at the server for a) BGW, and b) HeartyPatch.

Figure 5 compares the power consumption at the server (Odroid XU3) for a) BGW (on the left), and b) HeartyPatch (on the right). Over 90% of the energy consumed at the server is caused by our application processes (server and animator) running on two ARM Cortex-A15 cores. ARM Cortex-A7 (little cores) and GPU have an insignificant contribution to energy, while memory consumption is very small. According to the Watt graph (see red line), the server consumes less energy for HeartyPatch than for BGW (14% less dissipation). Instant power consumption is also more spikey for HeartyPatch. This may be related to the fact that for BGW, ECG data arrives at the server in large convoys (up to 128 12-bit values), and each one is transferred immediately to preserve

real-time. Convoys with HeartyPatch are limited to eight 18-bit values, packed in eight 32-bit unsigned values.

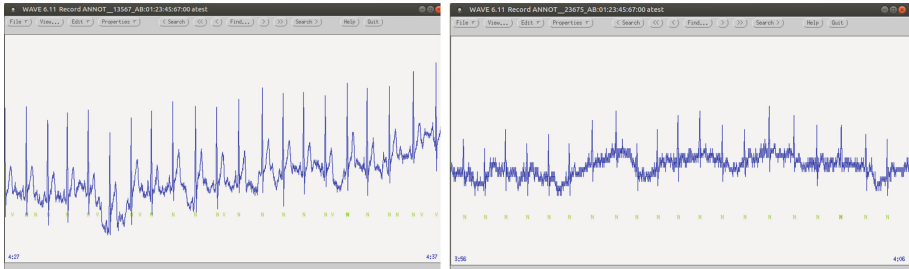


Fig. 6. Annotated ECG at the server for a) BGW, and b) HeartyPatch when operated with smaller packets.

Finally, Fig. 6 compares the annotated ECG at the server (Odroid XU3) for a) BGW (on the left), and b) HeartyPatch (on the right). Notice that the graph for HeartyPatch has a higher resolution as seen clearly from the images above (spikier). This is so since each HeartyPatch ECG sample is 18-bits compared to 12-bits precision for BGW. Occasionally, this may cause additional problems during the analysis, e.g. appearance of false arrhythmia notifications, as shown in the left figure. Notice that the 18-bit ECG value for HeartyPatch is next to the best available: 24-bit precision is offered from Coala Life Heart Monitor [18, 19].

5 Conclusions and Future Work

In this work, we compare a complex medical-grade biosensor (called STMicro Bodygateway, or BGW), with ProtoCentral’s HeartyPatch open-source electronic patch. As testbench, we use a soft real-time ECG monitoring, analysis, and visualization application that extends Harvard’s open-source WFDB and OSEA software packages. We consider power consumption and soft real-time signal acquisition, analysis, and visualization when both biosensors transmit ECG at 128 Hz to a low-cost server (Odroid XU3 board). While our platform achieves real-time performance for both sensors, the open, low-cost HeartyPatch solution is better, in terms of performance, signal precision, power consumption at the pulse sensor and server.

We have concentrated on a data rate of 128 pulses/s, since, in the current version of the HeartyPatch firmware, we have discovered problems when transmitting at higher data rates (256, and 512). The completely open (hardware/software) architecture, makes it simple to debug, optimize and extend its firmware in this direction. In addition, we hope to examine parallelization techniques, data compression, and socket options/flags to increase scalability, i.e., to support more sensor devices at higher rates. Finally, it is interesting to develop mixed-criticality scenarios that involve CPU/memory/network bandwidth management and evaluate overheads when implementing embedded network/system security mechanisms.

References

1. Alivecor (July 14, 2021 – Online). <https://www.alivecor.com/>
2. Lifemonitor (July 14, 2021 – Online). [http://www.equivital.co.uk/products/tnr/sense-and-tra
nsmit](http://www.equivital.co.uk/products/tnr/sense-and-tra
nsmit)
3. NowCardio (July 14, 2021 – Online). <https://contex-tech.com/medical/nowcardio>
4. Physiomem (July 14, 2021 – Online). [http://www.getemed.net/en/telemonitoring/physio
memr-pm-1000](http://www.getemed.net/en/telemonitoring/physio
memr-pm-1000)
5. Apple Watch, Series 5. (July 14, 2021 – Online). [https://www.apple.com/apple-watch-series-
5/health/](https://www.apple.com/apple-watch-series-
5/health/)
6. ProtoCentral: Heartypatch (July 14, 2021 – Online). <https://heartypatch.protocentral.com>
7. Preventice: BodyGuardian products (July 14, 2021 – Online). [https://www.preventicesolut
ions.com/healthcare-professionals](https://www.preventicesolut
ions.com/healthcare-professionals)
8. Grammatikakis, M.D., Koumarelis, A., Mouzakitis, A.: Software architecture of a user-level GNU/Linux driver for a complex E-Health biosensor. In: Saponara, S., De Gloria, A. (eds.) Applications in Electronics Pervading Industry, Environment and Society. ApplePies 2020, LNEE, vol. 738, pp. 1–7. Springer, Cham (2021). [https://doi.org/10.1007/978-3-030-372
77-4](https://doi.org/10.1007/978-3-030-372
77-4)
9. Grammatikakis, M.D., Koumarelis, A., Ntallaris, E. Validation of soft real-time in remote ECG analysis. In: Saponara, S., De Gloria, A. (eds.) Applications in Electronics Pervading Industry, Environment and Society. ApplePies 2020, LNEE, vol. 738, pp. 90–96. Springer, Cham (2021). <https://doi.org/10.1007/978-3-030-11973-7>
10. Physionet: WFDB (July 14, 2021 – Online). [https://archive.physionet.org/physiotools/wfdb.
shtml](https://archive.physionet.org/physiotools/wfdb.
shtml)
11. Hamilton, P.S., Patrick, S., Tompkins, W.J.: Quantitative investigation of QRS detection rules using the MIT/BIH arrhythmia database. *IEEE Trans. Biomed. Eng.* **12**, 1157–1165 (1986)
12. Tompkins, W.J.: A real-time QRS detection algorithm. *IEEE Trans. Biomed. Eng.* **3**, 230–236 (1985)
13. EP Limited: OSEA (July 14, 2021 – Online). <https://www.eplimited.com/confirmation.htm>
14. Pinto, J.R., Cardoso, J.S., Lourenço, A.: Evolution, current challenges, and future possibilities in ECG biometrics. *IEEE Access* **6**, 4746–4776 (2018)
15. Teplitzky, B.A., McRoberts, M., Ghanbari, H.: Deep learning for comprehensive ECG annotation. *Heart Rhythm J.* **17**(5), 881–888 (2020)
16. Physionet: WAVE (July 14, 2021 – Online). [https://archive.physionet.org/physiotools/wug/
wug.pdf](https://archive.physionet.org/physiotools/wug/
wug.pdf)
17. Hardkernel: Odroid SmartPower2 (July 14, 2021 – Online). [http://odroid.com/dokuwiki/doku.
php?id=en:acc:smartpower2](http://odroid.com/dokuwiki/doku.
php?id=en:acc:smartpower2)
18. Coala Heart Monitor: Coala Life (July 14, 2021 – Online). [https://www.coalalife.com/uk/cli
nicians](https://www.coalalife.com/uk/cli
nicians)
19. Coala Heart Monitor Technical Data Sheet: Coala Life (July 14, 2021 – Online). [https://cdn.
coalalife.com/uploads/2020/05/005013-Coala-Product-Note-Eng-ver-02.pdf](https://cdn.
coalalife.com/uploads/2020/05/005013-Coala-Product-Note-Eng-ver-02.pdf)



FPGA Implementation of a Configurable Vocal Feature Extraction Embedded System for Dysarthric Speech Recognition

Iacopo Casalini^(✉), Marco Marini, and Luca Fanucci

Department of Information Engineering, University of Pisa,
via G. Caruso 16, 561232 Pisa, Italy
iacopo.casalini@gmail.com

Abstract. In the last few years, users have been increasingly demanding for a hands-free interaction with their digital devices. This kind of technology is even more useful if used by people with disabilities, improving their quality of life. In particular, speech-impaired users (e.g. dysarthric speakers) represent a big challenge for an Automatic Speech Recognition (ASR) system because standard approaches are ineffective with them. Therefore, new speech analysis algorithms are implemented and generally tested on off-line datasets, but their performance can differ from a real case. Hence comes the need to easily validate their performance in a real scenario. The work presented in this paper shows an implementation of a highly configurable off-line embedded system for both MFCC and Mel Filterbanks extraction equipped with an on-board microphone. The results show that our system performs well in a real scenario case in terms of both power consumption and word error rate.

Keywords: Dysarthria · Automatic speech recognition · Speech analysis · MFCC · Filterbanks · FPGA

1 Introduction

Dysarthria is a motor speech disorder caused by a neurological injury that can manifest itself in different ways, most of which cause a slow and slurred spelling that results in a reduction in speech intelligibility.

The consequence is that commercial Automatic Speech Recognition (ASR) systems give bad performances for the transcription of dysarthric speech, as they are trained on unimpaired speech [1, 2].

An ASR system capable of making a correct transcription for dysarthric utterances would be very useful for these people, as it would allow them to interact with smart devices and domotics and even have a voice synthesizer speak in their place.

The first processing step in every ASR system is the vocal features extraction, which analyses and elaborates the voice signal in order to make it meaningful

for the acoustic model. The main features vectors extracted from signal voice are the Mel-Frequency Cepstral Coefficients (MFCCs) [9].

Previous research [3] has shown that the audio analysis can be tuned by changing a certain amount of settings (e.g. window size and shift size), which have been proved to increase the ASR performance for dysarthric speech up to 81%. These results can be confirmed by using a bigger Italian dysarthric database [4, 5] but the main issues about dysarthric speakers is that the recordings could be different from the real voice, being the dysarthria a degenerative disease. A mismatch between the results achieved using the recordings and the real case is then possible.

A possible solution is to test new ASR technologies with real people to validate recordings results. In order to do that, prototyping is the way but the software used to perform this kind of studies can rarely perform online and generally demands a great amount of resources.

The goal of this work is then to make prototyping process easier by implementing a hardware feature extraction device that works in real time and is fully configurable.

2 Data and Methods

The proposed system is composed of a hardware accelerator that takes the voice signal as input and extracts the features using the parameters set by the user in a previous settings procedure. Once extracted, the system streams the features over the serial peripheral to a desktop program that exploit an ASR system implemented by Kaldi toolkit, to interpret the features to a sequence of words. The settings procedure is leaded by the desktop program as well, thanks to a very user friendly Graphic User Interface (GUI).

2.1 Features Extraction and Configurability

The algorithm for MFCCs extraction is the following:

- Time domain: pre-emphasis of the audio signal (to emphasize high-frequency content, improving audio clarity [10]), followed by framing (splitting the input audio in small frames of fixed length and distance from one another) and windowing (in order to compute the Short-Time Fourier Transform of the frames mitigating the spectral leakage). Configurable settings in this step are:
 - Pre-emphasis coefficient;
 - Frame size and shift size (distance between the starting point of two consecutive frames);
 - Window type.
- Frequency domain: Short-Time Fourier transform of every frame, followed by the remapping of the STFT's frequency bins on the Mel scale to mimic the frequency resolution decay of the human ear [11]. Being the FFT size related to the number of samples in the frame, the only configurable setting in this step is the number of Mel bands.

- Quefrency domain: DCT transform of the Mel spectrum, keeping only a few of the lower order coefficients (the ones containing information on the phonemes articulation), followed by an optional liftering (filtering in quefrency domain) to reduce data variance [12]. Configurable settings in this step are:
 - Number of MFCC to keep (MFCC order);
 - Liftering coefficient.

All the parameters expressed in this paragraph are configurable in our system by the GUI.

2.2 System Implementation

The base behaviour of the system is shown in Fig. 1.

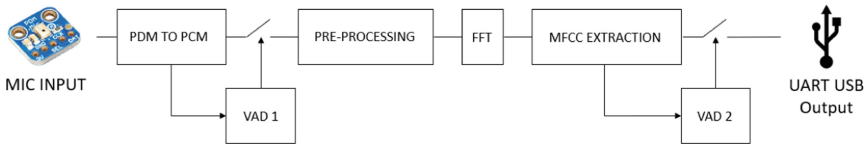


Fig. 1. System block diagram

The PDM input is fed to the PDM to PCM interface, whose output is then used to drive the hardware Voice Activity Detector (VAD) defined as VAD1 in the Figure. If vocal content is detected, the audio is transferred to the memory by a Direct Memory Access (DMA), pre-processed (pre-emphasis, framing and windowing) and then sent to the hardware FFT block, whose output is then taken back to memory, remapped on the Mel scale and DCT-transformed to get the MFCCs. These can be fed to a second software VAD, useful for a finer detection of false positives that are not needed in output. If no vocal content is detected, the system can behave in two different ways. The first is both DMA and CPU are still running (busy wait and no idle state) in order to be ready in case a vocal signal is detected; the second provides an idle state where both the input DMA and the CPU are in idle mode, so neither data transfers nor computation take place, keeping the power consumption low.

For this reason, the last two implementations of the device have been used for the final measurements and tests. **Hardware_1** behaves like the first case, while **hardware_2**, that represents the final version of the device, implements idle states for DMA and CPU, causing an almost complete deactivation of the system (only the PDM interface and the VAD1 are kept running), suspending the processing while waiting for an update of the input buffer from the DMA. This choice was made to try to lower the power consumption whenever possible.

Finally, a Graphic User Interface has been created to easily set all the parameters:

- Frame size (16 to 64 ms) and hop size ($\frac{1}{4}frame_size$ to $\frac{3}{4}frame_size$)

- Window type (Hann, Hamming and Povey)
- Number of Mel bands (19 to 27, default is 23)
- MFCC order (8 to 20, respectively the minimum [13] and maximum [14] value in literature)
- Pre-emphasis and liftering coefficients
- Software VAD configuration parameters
- Output features (MFCC or Filterbanks)

Figure 2 shows the GUI with the possibility to use a GMM model with all its parameters. Similarly, Fig. 3 shows the GUI with DNN configuration model settings. Notice that, only the Automatic Speech Recognition panel change between the GUIs.

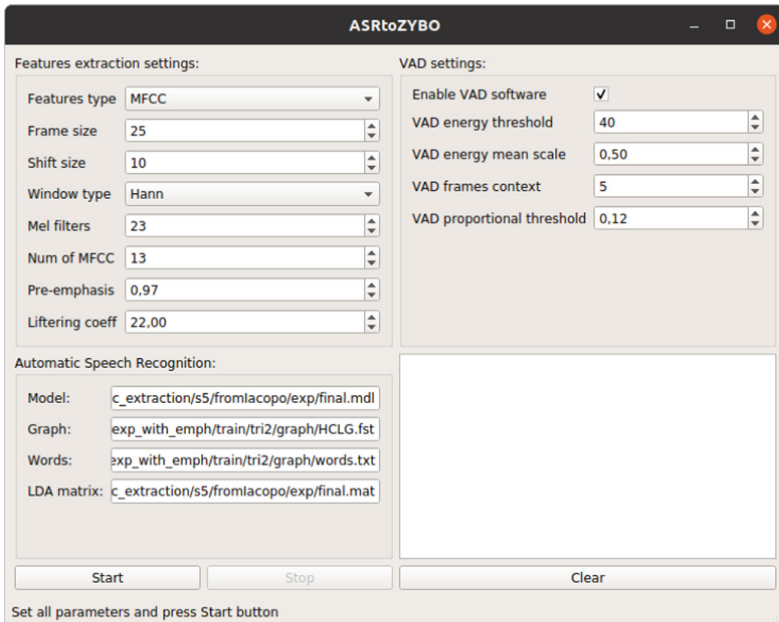


Fig. 2. GUI with GMM model settings

2.3 Materials

We used the Digilent Zybo FPGA as hardware accelerator and the Adafruit 3492 PDM microphone as input device because a system for PDM interfacing and Voice Activity Detection was already available for these platforms (Synapse [6], with its on-board VAD [7], implemented in VHDL language and based on energetic considerations, without any Fourier Transforms involved).

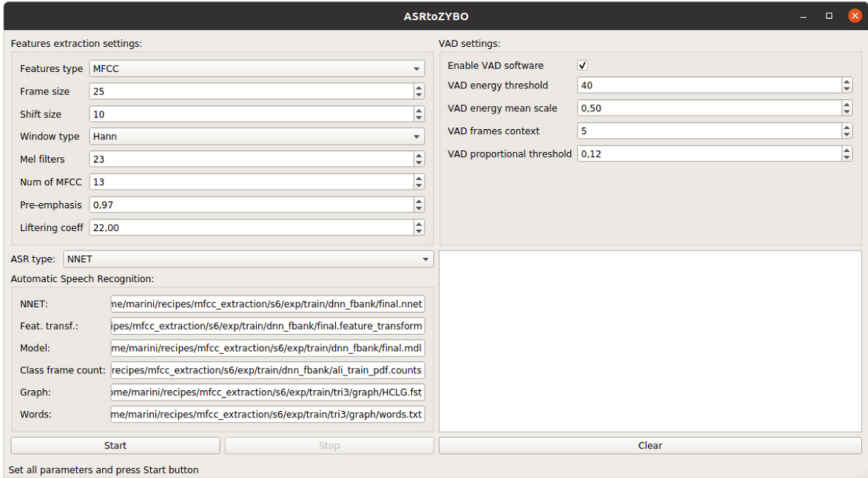


Fig. 3. GUI with DNN model settings

The hardware platform and source code for this project were built using respectively Xilinx Vivado 2020.2 and Xilinx Vitis 2020.2.

Kaldi toolkit [8] has then been used for the algorithm analysis and results verification.

2.4 Hardware Design

The input signal coming from the microphone is fed to the Synapse block, which converts it to a PCM signal that is analyzed by the hardware VAD (VAD1). The decision coming from VAD1 is then used to send two different interrupts to the CPU when it switches from 0 to 1 (no vocal signal detected \rightarrow vocal signal detected) or from 1 to 0 (vocal signal detected \rightarrow no vocal signal detected). In case of a 1 \rightarrow 0 transition, the interrupt handler puts both the DMA and the CPU in idle mode; in this state, the only interrupt signal the CPU can respond to is the one coming from the VAD when vocal content is detected again. In this case, the CPU wakes from the idle state and the DMA is activated so that it can read the samples from the input FIFO and transfer them to be elaborated. Since the asynchronous FIFO (a Xilinx IP) used for data input, when full, does not automatically erase the oldest sample to write the incoming one, a custom block has been written to take care of the problem when VAD1 is closed: in this condition, for every new incoming sample, the block reads the last one to keep the FIFO full and updated, ready to be read at the next VAD 0 \rightarrow 1 transition.

To prevent problems linked to a delayed opening of the VAD1, (especially in cases like sibilants at the beginning of a word or phrase) the incoming audio is also delayed by about 500 ms (value empirically tuned by testing). To take care of the early closing, instead, the approach is to put the CPU in idle state only if VAD1 output stays low for a certain period of time (empirically tuned to about

700 ms): in this way the early closing and the possible bouncing effect are both avoided.

The FFT hardware block, linked to another dedicated DMA instance, is a Xilinx IP.

For what concerns the FPGA resources utilization, the report from Vivado is available in Table 1.

Table 1. Vivado resources utilization report

Slice LUTs	Slice registers	F7 muxes	F8 muxes	Slice	LUT as logic
14107/17600	16005/35200	13/8800	2/4400	4326/4400	9618/17600
LUT as memory	Block RAM tile	DSPs	Bonded IOB	Bonded IOPADs	BUFGCTRL
4489/6000	22.5/60	29/80	4/100	130/130	3/32

2.5 Testing and Results

The testing procedure has two main objectives: measure the power consumption of the two implementations (hardware_1 vs hardware_2) in order to evaluate the effect of the idle state added in the second (and final) version, and compare the performance of an ASR system that analyses the features extracted by the two implementations. The parameters set used for all tests is: Frame size 25 ms, Hop size 10 ms, Window type Hann, Number of Mel bands 23, MFCC order 13, Pre-emphasis 0.97, Liftering coefficient 22.

The power consumption test has been done by measuring the power consumption of the embedded system during all its operation cycle. The measure was taken using a AVHzY CT-2 USB device¹, with one side connected to the PC and the other one powering the FPGA (with on-board microphone).

The Fig. 4 shows the measured power consumption of the wait for configuration step, and a couple of idle (when VAD1 is closed) and running (when VAD1 is open) cycles. The step between the first and second implementation has in fact led to a saving of about 5% in both situations. The hardware_2 implementation (orange line in the Figure) gives an average of 1.608 W for the running state and 1.572 W for the idle state, values that are to be referred to the already high consumption of the Zybo, which has a measured power consumption of 1.252 W when empty.

The effort spent for lowering the power consumption would have been useless if linked to a performance loss. For this reason, we have also compared the Word Error Rate (WER) of an ASR system fed by the features extracted by both the hardware configurations. We used an audio file played on a studio monitor at a fixed volume with the microphone in a fixed position, as input for the entire test. The audio file contains an male Italian utterance with 1785 words. Two ASR

¹ https://store.avhzy.com/index.php?route=product/product&product_id=50.

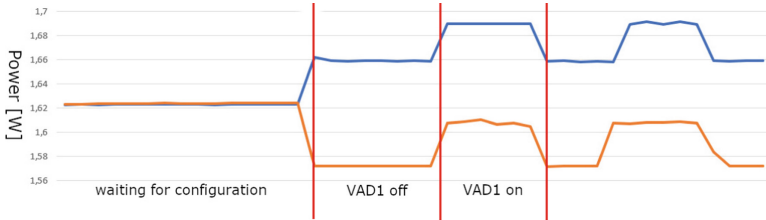


Fig. 4. Power consumption comparison between hardware.1 (blue line) and hardware.2 implementations (orange line). (Color figure online)

systems have been used for this test: one uses Gaussian Mixture Model (GMM) for the acoustic model, while the other one uses a Deep Neural Network (DNN). Both of them are implemented by Kaldi toolkit with the recipe described in [4].

Table 2 shows the results of WER test for both ASR systems and the power consumption for each hardware configurations. The results show not only the absence of a performance loss between the hardware configurations, but also an improvement of WER for the GMM model, while the DNN model seems to be quite unaffected by the difference between the two implementations. This behavior confirms that the DNN models are more reliable in presence of feature noise. Even though we cannot compare our ASR systems results with the state of the art, the very low WER values for both hardware implementations and the ASR models indicate a good user experience in terms of usability [2].

Table 2. WER results for different implementations of the feature extraction methods

Feature extraction method	Running consump.	Idle consump.	WER (GMM)	WER (DNN)
Hardware.1	1.690 W	1.660 W	2.297%	1.01%
Hardware.2	1.608 W	1.572 W	1.345%	1.01%

3 Conclusion

In this work, we introduce the implementation of a highly configurable off-line embedded system for both MFCC and Mel Filterbanks extraction. Our solution is equipped with an on-board PDM microphone and can be configured by using a GUI that also integrates two types of ASR systems (GMM and DNN based) that analyse the features extracted by the embedded system and generate the most likely transcription.

Two hardware implementation (hardware.1 and hardware.2) have been realized and compared in terms of power consumption and ASR performance (expressed by WER).

The results show that from the ASR performance point of view, both the implementations have a very low WER value, that means a good user experience in terms of usability. The second implementation is more efficient from the power point of view and the GMM system has a lower WER while, it seems that the hardware changes do not influence the performance of the DNN model.

References

1. Ballati, F., Corno, F., De Russis, L.: “Hey Siri, do you understand me?”: virtual assistants and dysarthria. In: *Intelligent Environments 2018*, pp. 557–566. IOS Press (2018)
2. Ballati, F., Corno, F., De Russis, L.: Assessing virtual assistant capabilities with Italian dysarthric speech. In: *Proceedings of the 20th International ACM SIGACCESS Conference on Computers and Accessibility* (2018)
3. Marini, M., Meoni, G., Mulfari, D., Vanello, N., Fanucci, L.: Enabling smart home voice control for italian people with dysarthria: preliminary analysis of frame rate effect on speech recognition. In: Saponara, S., De Gloria, A. (eds.) *ApplePies 2020*. LNEE, vol. 738, pp. 104–110. Springer, Cham (2021). https://doi.org/10.1007/978-3-030-66729-0_13
4. Marini, M., et al.: IDEA: an Italian dysarthric speech database. In: *2021 IEEE Spoken Language Technology Workshop (SLT)*. IEEE (2021)
5. Turrisi, R., et al.: EasyCall corpus: a dysarthric speech dataset. *arXiv preprint arXiv:2104.02542* (2021)
6. Ciarpi, G., Palla, A., Fanucci, L., Meoni, G., Pilato, L.: Fully Digital Low-Power Implementation of an Audio Front-End for Portable Applications. University of Pisa, Dept. of Information Engineering (DII) (2019)
7. Meoni, G., Pilato, L., Fanucci, L.: A low power Voice Activity Detector for portable applications. In: *PRIME 2018*, Prague, Czech Republic (2018)
8. Povey, D., et al.: The Kaldi speech recognition toolkit. In: *IEEE 2011 Workshop on Automatic Speech Recognition and Understanding* (2011)
9. Xu, M., Duan, L.-Y., Cai, J., Chia, L.-T., Xu, C., Tian, Q.: HMM-based audio keyword generation. In: Aizawa, K., Nakamura, Y., Satoh, S. (eds.) *PCM 2004*. LNCS, vol. 3333, pp. 566–574. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-30543-9_71
10. Vergin, R., O’Shaughnessy, D.: Pre-emphasis and speech recognition. In: *Proceedings 1995 Canadian Conference on Electrical and Computer Engineering*, vol. 2. IEEE (1995)
11. O’Shaughnessy, D.: *Speech Communication: Human and Machine*. Addison-Wesley, Boston (1987)
12. Gales, M., et al.: *The HTK Book*. Cambridge University Engineering Department (2006)
13. Sreenivasa, R.K., Koolagudi, S.G.: *Robust emotion recognition using spectral and prosodic features*. Springer Science & Business Media, Heidelberg (2013)
14. Mitrović, D., Zeppelzauer, M., Breiteneder, C.: Features for content-based audio retrieval. In: *Advances in Computers*, vol. 78, pp. 71–150. Elsevier (2010)



Classifying Simulated Driving Scenarios from Automated Cars

Marianna Cossu^(✉), Jorge Leonardo Quimi Villon, Francesco Bellotti,
Alessio Capello, Alessandro De Gloria, Luca Lazzaroni, and Riccardo Berta

Department of Electrical, Electronic and Telecommunication Engineering (DITEN), University
of Genoa, Via Opera Pia 11a, 16145 Genoa, Italy
marianna.cossu3250@gmail.com

Abstract. Detection of driving scenarios is getting ever more importance for assessment and control of automated driving functions. This paper investigates the performance of two versions of a high-end 3D convolutional network for scenario classification. The first one uses fully 3D kernels, the second one separates, in each constituting block, the 2D spatial convolution from the temporal convolution, $(2 + 1)D$. We made the tests on a synthetic dataset created by specifying scenarios in OpenScenario and running them in the CarLA 3D simulator. We focused our analysis on three main performance profiles: at different frame per second rates, different video clip lengths, and different weather conditions. Results show an overall robustness of the 3D predictors, and seem to suggest two different use cases: $(2 + 1)D$ looks more suited when the scenario changes quickly or a low latency is required, while the plain 3D solution is better for slow-changing scenarios and when FPS can be low.

Keywords: Deep learning · Video classification · Automated driving · CarLA driving simulator · Time-series · Convolutional neural network · Three-dimensional convolution · Spatio-temporal convolution

1 Introduction

Development of automatic driving functions (ADF) requires an accurate analysis of the operational design domain (ODD). This implies the recognition of the driving context, both for allowing an adaptation of the in-vehicle ADFs and, offline, for assessment and verification [1, 2].

Context awareness strongly benefits from Machine learning (ML) techniques that are now being widely applied in automotive (e.g., [3]). In this area, we are interested in exploring and benchmarking a set of state-of-the-art deep neural networks (DNNs), in order to understand and characterize the suitability of their features to the task of detecting driving scenarios. As a preliminary step, in the wide variety of possibilities given by the rapidly changing state of the art of neural network (NN) models, the goal of this paper is to focus on comparing two flavors of residual networks exploiting spatio-temporal convolutions directly since their lower layers [4]. For simplicity, we limit our

analysis to camera sensors. But a complete final system will need to fuse this information with other sources, such as lidars and radars.

Given the lack of publicly available datasets labeled with driving scenarios, we use a synthetic dataset [5] exploiting the Car Learning to Act (CarLA) simulator [6] running scenarios defined through the OpenScenario standard [7].

2 Background

2.1 Driving Scenario Dataset

In order to learn high-level meaningful information from data, supervised ML systems require datasets containing labeled samples for the target application. At present, while there is a variety of datasets in the automotive domain (e.g., KITTI [8], Prevention [9], NGSIM I-80 [10], US-101 [10], Apolloscape [11]), we are not aware of publicly available datasets labeled with driving scenarios.

On the other hand, [5] developed a tool-chain to automatically generate short video-clips, each one created by randomly assigning a set of parameters characterizing five types of simulated driving scenarios (namely: cut-in, following a lead vehicle, approaching a lead vehicle, free driving, lane change). The parametric descriptions of the scenarios are based on the Open Scenario format [7]. Therefore, we exploited one of such synthetic datasets, video clips of duration 8 s each, recorded at 22 frames per second (FPS).

Virtual datasets have a great generation controllability, at least for some factors (e.g., weather conditions), but still feature clear limits in terms of realism, likelihood, representativeness. Also, only simple, pseudo-random variations for the traffic scene among instances of the same scenario instance are considered, typically involving initial distance between main vehicles, speed of the ego vehicle, approximate traffic density. Moreover, videos are recorded in just two city maps, with only three types of vehicles. Despite these limitations, we think that the dataset can provide a significant preliminary test-bench for driving scenario recognition.

2.2 3D Convolutional Neural Networks

Extraction of information from videos through neural networks (NNs) has been the subject of several works presented in literature [12].

2D Convolutional Neural Networks (CNNs) (e.g. AlexNet [13]) are suitable for object detection but are not sufficient for processing videos, since they do not model temporal information and motion patterns. In order to model spatio-temporal information, a common approach consists in combining 2D CNNs with Recurrent Neural Networks (RNN), e.g., Long-Short Term Memory [14] (LSTM) and Gate Recurrent Unit [15] (GRU). A different approach – on which we focus in this article after some preliminary tests on 2D CNN and LSTM which did not yield relevant results – involves the direct use of convolutions for both the spatial and temporal dimensions. This three-dimensional architecture allows processing both spatial and temporal information from a sequence of images because time-dependent patterns can be recognized through the higher order convolution.

A recently proposed improvement over 3D CNN is $(2 + 1)$ D CNN [5], which separates 2D (spatial) convolution with the third, temporal dimension. $R(2 + 1)$ D is a convolutional residual block [5, 16, 17], which explicitly factorizes a 3D convolution into two separated and subsequent operations, a 2D spatial convolution and a 1D temporal convolution. From this kind of decomposition, two main advantages are obtained. This method duplicates the number of nonlinearities (with respect to 3D CNN) while having the same number of parameters, thus making the model able to represent more complex functions. This makes such models easier to optimize compared to 3D CNNs, generally leading to a lower training time and testing loss [5].

3 Experiment

Using the synthetic dataset mentioned in the Background section, we tested the performance of two versions of a 3D CNN that is shown in Fig. 1. Particularly, we focused our experiment on two main aspects:

- How much does the number of FPS in a clip affect the performance of the NNs? This factor is key for latency and memory footprint.
- What is the effect of weather and light conditions? How is the performance affected by varying the length of the input video clip?

In order to verify how rain and fog would affect the prediction, we used two datasets, created from the one above discussed (and having its same size): the first one is comprehensive of all the light and weather conditions (sun, night and fog & rain, see Fig. 1), while the second one contains day and night without fog and rain. RGB frames extracted from video clips were first resized from (1920×1080) pixels to (112×112) pixels. This implies a certain loss of information due to the frames downscaling, but saves the memory space and improves the timing performance.

The experiments were conducted on a Linux machine with one NVIDIA Quadro RTX 4000 GPU, with 8 GB V-RAM. Results are reported in terms of accuracy and precision [12].



Fig. 1. Different scenario samples from the left fog and rain, night and day

Each NNs input has thus size $(112p \times 112p \times 3 \times totalframes)$, where 3 is the number of input channels (red, green, blue), and $totalframes$ is the product of FPS times number of seconds (length of the clip).

For the $R(2 + 1)D$ and $R3D$ CNNs implementation, we took inspiration from [18]. However, since the training of such networks require High Performance Computing (HPC) machines, we halved the original number of residual blocks (150), going from 33 M total parameters to 14 M. We argue that this should not affect performance too much since we process short video clips (up to 8 s) and at a downscaled resolution (112×112) pixels. The input layer consists of 64 convolutional filters. The last layer of both CNNs is a linear one, returning the probability of all the possible scenarios (Fig. 2)..

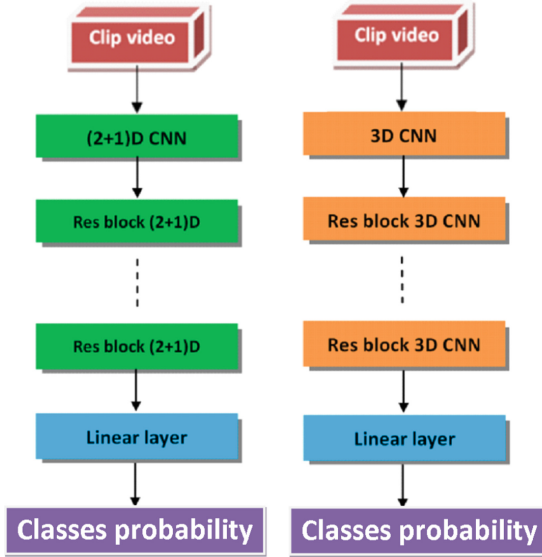


Fig. 2. Networks used during experiment. On the left we can see the residual network featuring $(2 + 1)D$ convolutions, and on the right the one using 3D convolutions.

4 Experimental Results

Considering the first analysis point (i.e., how much a change in the number of clip FPS affects the performance of the NNs), we tried to change the value of FPS and, correspondingly, the length of the input video clip between training and testing. Our tests (Fig. 3) clearly show that both CNNs reach the maximum accuracy and precision when the number of input frames is the same in training and testing. As an interesting consequence, tests can also be performed on a scene with a different duration than the training, but the *totalframes* value should be kept constant (e.g., if the test clip lasts twice as long as the training, its FPS value at inference time should be halved w.r.t. the training).

Then, we analyzed how the change of FPS value influences the prediction when the test videos are cut shorter (Table 1). In this case, the action of the scenario results incomplete due the shorter time span, and it is necessary to increase the FPS rate, in order to keep *totalframes* constant.

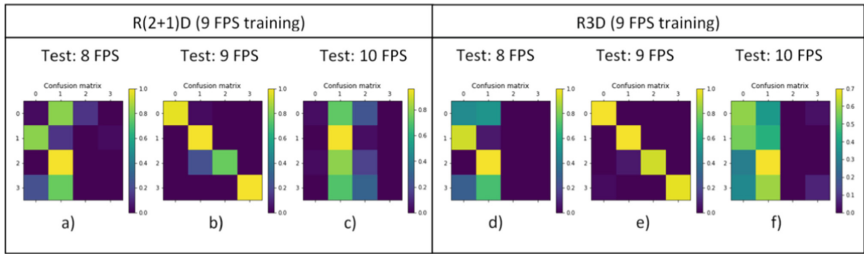


Fig. 3. Confusion matrices for CNNs trained with 9 FPS. Testing with a) R(2 + 1)D 8 FPS; b) R(2 + 1)D 9 FPS; c) R(2 + 1)D 10 FPS; d) R3D 8 FPS; e) R3D 9 FPS; f) R3D 10 FPS.

Overall, these experiments show that the R(2 + 1)D is more robust to the redundancy of the temporal information (increase of the testing FPS, with shorter videos), as the accuracy of the R3D starts decreasing earlier. Beyond a certain threshold, the accuracy of predictions drops (Table 1). This is probably caused by reducing both the video length and the time distance between frames, which makes the evolution of the action considerably slow, almost stationary, and limited in time, compared to the training phase.

Table 1. Test results with different FPSs in training, considering 2 s videos.

Train- FPS	total-frames	Test- FPS	R(2 + 1)D		R3D	
			Acc.	Prec.	Acc.	Prec.
1	8	4	0.81	0.86	0.96	0.96
2	16	8	0.97	0.98	0.84	0.90
3	24	12	0.99	0.99	0.89	0.91
4	32	16	0.86	0.90	0.85	0.90

Another set of experiments aimed at checking how light and weather conditions influence classification (Fig. 4). Particularly, we considered four different conditions: daytime, night, fog and rain, and all conditions. Results show a strong robustness of the classifiers. Going more in detail, R(2 + 1)D has overall a slightly worse accuracy compared to R3D, especially considering fog and rain scenarios which, as expected, result to be the most difficult ones to recognize. The fog and rain scenarios are particularly critical for the R(2 + 1)D configuration, whose results are on average at around 70% accuracy, with important fluctuations depending on the seconds of video used, probably due to the modest size of the test set (15 clips for each driving scenario types). R(2 + 1)D performs better with 1 s videos, which is an important outcome from a point of view of recognizing scenarios with low time latency. On the other hand, R3DCNN results to be quite more robust than R(2 + 1)D and suitable to recognize mid-long length scenarios.

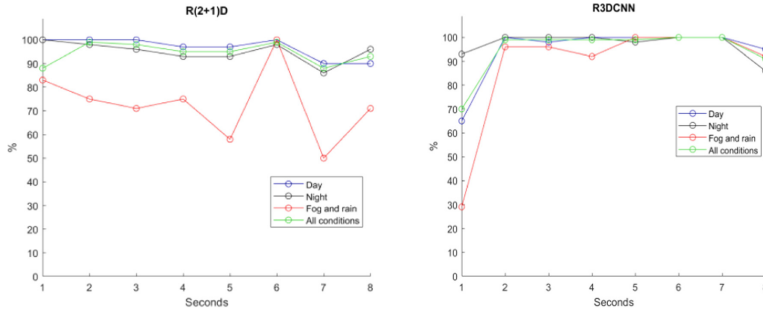


Fig. 4. R3D and R(2 + 1)D accuracy with different weather and light conditions and scenario lengths.

5 Conclusions and Future Work

As detection of driving scenarios is getting ever more relevance for ADF assessment and control, we have investigated the classification performance of two versions of a high-end 3D convolutional network in this task. First, we investigated the FPS rate, for which we identified different use cases for the CNNs. As the R(2 + 1)D scales better on FPS increase, it seems more suited for fast-changing scenarios, while R3D has a high accuracy even with a 1 FPS testing, and scales worse with high FPS values, so it looks more suited for slow-changing scenarios. The second question was about CNNs accuracy in different weather and light conditions. In this case, R3D performs significantly better than R(2 + 1)D, with the only exception of the 1 s clip test.

The third question compares the performance of the CNNs with different testing video durations. Considering the “all conditions” scenario (to avoid bias on a particular light or weather case), R3D has better or equal performance in terms of accuracy than the R(2 + 1)D, except for the 1 s case, as stated before.

These results seem to suggest two different use cases for the investigated CNNs:

- When the scenario changes quickly or a low latency is required, R(2 + 1)D can achieve better results.
- When the scenario changes slowly and FPS can be low, best choice is R3D.

As the preliminary results of our analysis are positive, our next goal is to train our CNNs with a more complex and bigger dataset to assess the CNNs with more challenging scenarios. Another goal is to increase the input resolution for instance using grayscale frames to compensate the increase in input size. A significant limit of the current implementations, that we would like to overcome, is the requirement that the product FPS times length of clips must be constant.

Finally, it will be interesting to analyze the performance of both CNNs with more complex inputs, like those coming from lidar and radar sensors, that can be simulated through the Carla simulator as well.

References

1. Weber, H., et al.: A framework for definition of logical scenarios for safety assurance of automated driving. *Traffic Inj. Prev.* **20**(sup1), S65–S70 (2019), <https://doi.org/10.1080/15389588.2019.1630827>
2. Bellotti, F., et al.: Managing big data for addressing research questions in a collaborative project on automated driving impact assessment. *Sensors* **20**, 6773 (2020)
3. Mozaffari, S., Al-Jarrah, O.Y., Dianati, M., Jennings, P., Mouzakitis, A.: Deep Learning-based vehicle behavior prediction for autonomous driving applications: a review. *IEEE Trans. Intell. Transp. Syst.* <https://doi.org/10.1109/ITITS.2020.3012034>
4. Tran, D., Wang, H., Torresani, L., Ray, J., LeCun, Y., Paluri, M.: A closer look at spatiotemporal convolutions for action recognition. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (2018). <https://doi.org/10.1109/CVPR.2018.00675>
5. Bonora, S., Motta, J.: Progettazione e realizzazione di una to. olchain per la generazione di dataset e il riconoscimento basato su deep learning di scenari per la guida autonoma. *Elettronica, Università di Genova, Tesi laurea magistrale Ing* (2021)
6. Carla community. Carla Simulator. Version 0.9.11. <https://carla.readthedocs.io/en/latest/>. Accessed 23 Apr 2021
7. OpenSCENARIO User Guide. https://releases.asam.net/OpenSCENARIO/1.0.0/ASAM_OpenSCENARIO_BS-1-2_User-Guide_V1-0-0.html, accessed on 22/04/2021
8. KITTI dataset. <http://www.cvlibs.net/datasets/kitti/>
9. Izquierdo, R., Quintanar, A., Parra, I., Fernández-Llorca, D., Sotelo, M.A.: The PREVENTION dataset: a novel benchmark for PREDiction of VEHicles iNTentIONS. In: 2019 IEEE Intelligent Transportation Systems Conference (ITSC), pp. 3114–3121, October 2019. <https://doi.org/10.1109/ITSC.2019.8917433>
10. NGSIM and US-101 datasets. <https://ops.fhwa.dot.gov/trafficanalysisitools/ngsim.htm>
11. apolloscape dataset. <http://apolloscape.auto/>
12. Géron, A.: *Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*. O'Reilly Media, Boston (2019)
13. Krizhevsky, A., Sutskever, I., Hinton, G.: ImageNet classification with deep convolutional neural networks. In: *InNIPS* (2012)
14. Hochreiter, S., Schmidhuber, J.: Long short-term memory. *Neural Comput.* **9**(8), 1735–1780 (1997)
15. Cho, K., et al.: Learning phrase representations using RNN encoder-decoder for statistical machine translation. [arXiv:1406.1078](https://arxiv.org/abs/1406.1078)
16. Feichtenhofer, C., Pinz, A., Wildes, R.P.: Spatiotemporal residual networks for video action recognition. In: *NIPS* (2016)
17. Qiu, Z., Yao, T., Mei, T.: Learning spatio-temporal representation with pseudo-3D residual networks. In: *ICCV* (2017)
18. R(2+1)D code. <https://github.com/irhum/R2Plus1D-PyTorch>



Mismatch Analysis of Parallel Li-Ion Batteries

Massimo Conti^(✉), Luca Stacchiotti, and Simone Orcioni

Department of Information Engineering, Università Politecnica delle Marche, via Brecce
Bianche 12, 60124 Ancona, Italy
m.conti@univpm.it

Abstract. Statistical variations of the electro-chemical-thermal characteristics of the single cells cause degradations of the performances of a battery pack. This paper presents a model and a simulation environment for the statistical analysis of the performances of a battery pack affected by variations of the parameters among the cells in the same battery pack.

1 Introduction

With the Green Deal, the European Union has set climate neutrality as a goal by 2050 through a transition towards a sustainable economy. Battery technology can facilitate the transition to a decarbonized society, through the integration of renewable energies with the electricity grid and zero-emission mobility [1].

Lithium batteries are among the most used energy storage system because of their excellent performance, which is related to their high specific energy, energy density, specific power, efficiency, and long life for either energy storage system or electrical vehicles [2]. The Battery Management System (BMS) uses advanced control algorithms for State of Charge (SoC) indication and State of Health (SoH) with electrical models that fully reflect the actual performance and cycle life characteristics of batteries [3–5]. One of the critical aspects of the use and management of lithium-ion battery packs is the statistical variations of the electro-chemical-thermal characteristics of the single cells. A battery pack consists of series and parallel connected cells. The effect of the mismatch among the cells causes degradation of the performances of the battery pack. In series connection, the cell charge active or passive equalization is carried out to mitigate the mismatches among the module cells and to maximize the charge throughput during charge/discharge. It is difficult, for the BMS to estimate the effect of cell mismatch in parallel connected battery pack, because the measurement of the current of each cell in parallel-connected battery packs is impractical due to the high cost of additional current sensor. Nevertheless, cell mismatch has negative effect in the performances and the life extension of the pack [6–9].

The mismatch effect on battery performances must be modeled to estimate the maximum mismatch allowed among the cells that must be placed in parallel.

This paper presents a simulation environment for the statistical analysis of the performances of a battery pack affected by variations of the parameters among the cells in the same battery pack.

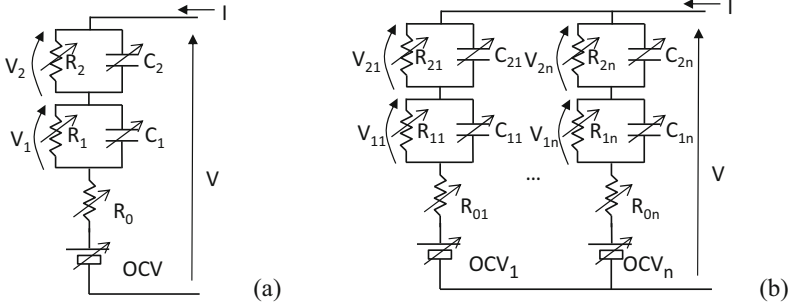


Fig. 1. Electric model of a single cell (a), and of parallel connected battery pack (b).

2 Battery Pack Statistical Model

The electric model of the single cell used is reported in Fig. 1(a). It consists of a voltage source OCV in series with a resistance R_i and two parallel connected resistors and capacitance R_1, C_1, R_2, C_2 . The 6 parameters have a nonlinear dependence with the SoC and temperature. The parameters have been estimated with experimental measurements using the HPPC technique, as reported in [10], on batteries with nominal capacity of 50 Ah, and nominal voltage of 3.2 V. In parallel configuration, battery pack voltage is the same for all its battery cells as shown in Fig. 1(b), whilst the battery pack current is the sum of all battery cell currents. The equations describing the circuit of Fig. 1(b) are therefore the following

$$\dot{v}_{j1} = \frac{i_j}{C_{j1}} - \frac{v_{j1}}{R_{j1}C_{j1}} \quad (1)$$

$$\dot{v}_{j2} = \frac{i_j}{C_{j2}} - \frac{v_{j2}}{R_{j2}C_{j2}} \quad (2)$$

$$V = OCV_j + i_j R_{ji} + v_{j1} + v_{j2} \quad (3)$$

$$SoC_j = SoC_{jinit} - \frac{1}{Q_j} \int_0^t i_j(\tau) d\tau \quad (4)$$

$$\sum_{j=1}^n i_j = I \quad (5)$$

A simulation environment widely used for cell modelling is Simulink [11, 12]. In our work, the electrothermal model has been implemented in SystemC-WMS. SystemC-WMS is an extension of SystemC for modelling heterogeneous systems [13, 14]. It has been already used in the simulation of battery pack in [15–17].

A model of the battery pack including parametric variations has been used to make a simulation of a battery pack with parallel connected cells. The system simulated consists of a battery pack of 9 parallel cells, considering the statistical variations of the 7 parameters ($Q, OCV, R_i, R_1, C_1, R_2, C_2$) of each cell, for a total amount of 48 independent

random variables. The statistical parameters of a single cell have been modeled by the following equation:

$$X = X_{deter}(1 + \sigma_X \psi) \tag{6}$$

where ψ is a gaussian random variable with zero mean value and unitary variance. Based on the analysis reported in [14] the following coefficients has been selected.

$$\sigma_{Ri} = \sigma_{R1} = \sigma_{C1} = \sigma_{R2} = \sigma_{C2} = \sigma_Q = 0.03, \sigma_{OCV} = 0.003 \tag{7}$$

As an example, Fig. 2 reports the experimental statistical variations of the OCV as a function of the SoC of 6 new Panasonic NCR18650B battery cells.

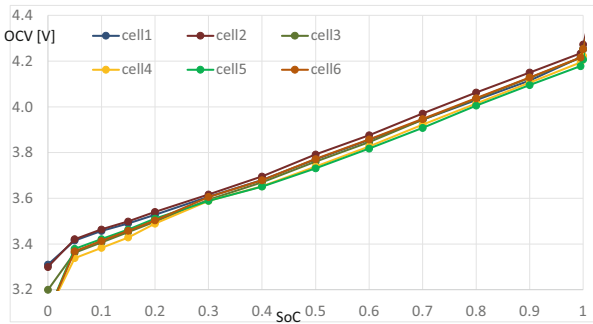


Fig. 2. Experimental statistical variations of the OCV as a function of SoC.

3 Statistical Simulations of Parallel Connected Cells

The Monte Carlo (MC) simulations of a battery pack of 9 parallel cells have been carried out considering the statistical variations of the cell parameters. The nominal capacity of the pack of 9 parallel cells is 450 Ah with a nominal voltage of 3.2 V.

The battery pack is charged with a CC CV technique with an initial constant load current of 90 A. When the voltage reaches 3.305 V the current is reduced to keep the voltage constant until the load current reaches 9 A, the 10% of the initial value. Then the battery pack is disconnected to the load. The duration of simulations is 25000 s, so that the duration of whole transient can be observed. The ambient and initial temperature of the cells have been fixed to 25 °C, the initial voltage is set to 3.24 V, corresponding to a nominal SoC_{jinit} of 25%, but due to the mismatch among the cell parameters the initial SoC_{jinit} is not the same for all the cells, as shown in Fig. 3b.

Figures 3, 4 and 5 show, as an example, one of the MC simulations. Figure 3a reports the current flowing in each of the 9 cells in case of mismatch, and in the case of identical cells in dashed line as a reference. The mismatch on the Q , OCV , R_i parameters cause a variation of about 10% among the cell currents. The average current over the 9 cells does not correspond to the value to the case of identical cells. Due to the complex dependence

of the mismatch parameters on the circuit, the cell that takes the maximum or minimum current is not always the same. At the end of the charge, when the cell is disconnected to the load, the average current is zero but a charge distribution among the cells is not negligible and it causes an additional internal power dissipation.

Figure 3b reports the state of charge of each of the 9 cells in case of mismatch and identical cells in dashed line as a reference. The initial state of charge is different for the different cells due to the mismatch on *OCV*. The *SoC* mismatch may be relevant (from 20% to 28% at the beginning and from 78% to 92% at the end of the charge. This effect is relevant on the performances of the parallel connected battery pack. To avoid a premature ageing due to overcharge of some of the cells, the excursion of the equivalent *SoC* must be reduced.

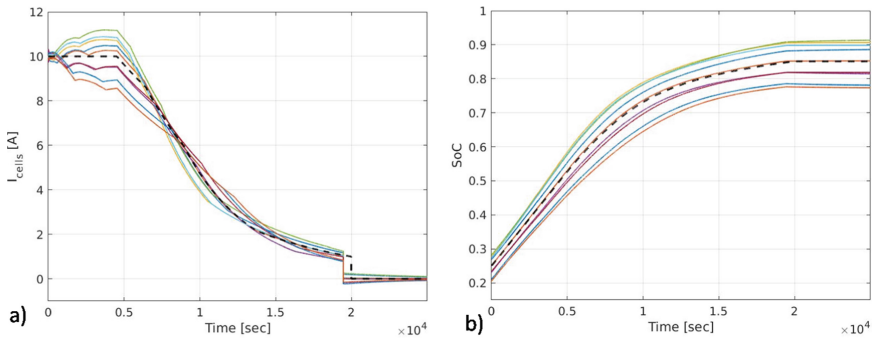


Fig. 3. Current (a) and state of charge (b) of each of the 9 cells and in case of mismatch and identical cells (dashed line).

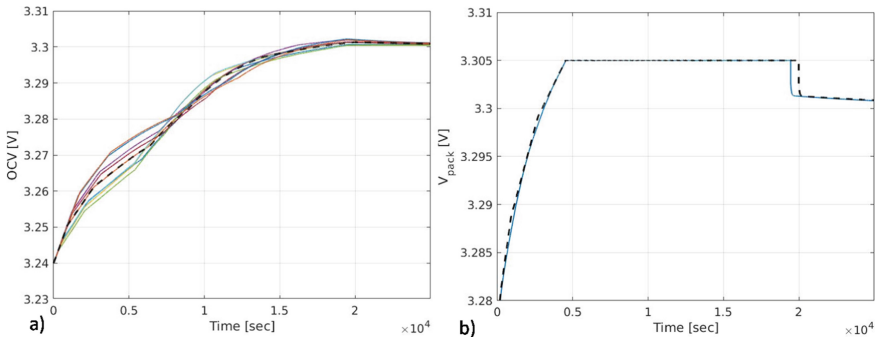


Fig. 4. OCV of the 9 cells (a) and voltage of the battery pack (b) in case of mismatch and identical cells (dashed line).

The same CC CV charge allows a charge of 85% in case of identical cells, but in case of mismatch one cell reaches 92%. Furthermore, the BMS cannot estimate the mismatch among the *SoC* only from the external current and voltage measurements.

Figure 4a reports the *OCV* of each of the 9 cells in case of mismatch and identical cells in dashed line as a reference. In stationary situation, that is at the beginning and at

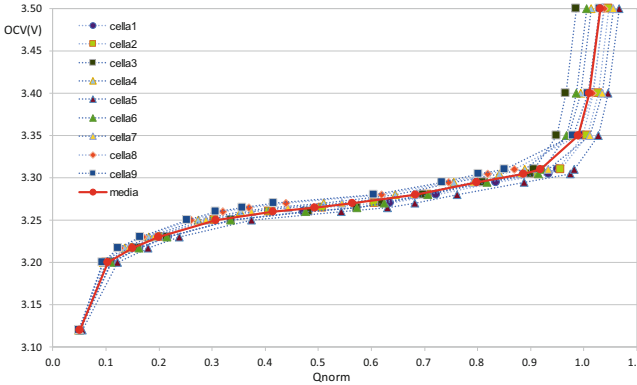


Fig. 5. OCV of the single cells and average value in stationary conditions in case of mismatch.

the end of the simulation, the voltage drop on the internal resistors is zero, therefore the *OCV* is the same for all the cells in parallel connection. The different currents in each cell cause different *OCV* in each cell. Figure 4b reports the voltage of the battery pack in case of mismatch and identical cells in dashed line as a reference. Due to the nonlinear relationship among *OCV* and *SoC*, the voltages of the battery pack in case of identical cells or in case of mismatch are similar but not identical in the *CC* part of the charge. In the *CV* phase the voltage is the same, since the voltage is fixed by the charging algorithm to 3.305 V. The time instant when charge stops, that is when the load current is 9 A, is different in case of mismatch and identical cells.

Finally, Fig. 5 reports the *OCV* as a function of the charge normalized to the nominal charge of each cell and the average value in stationary conditions in case of mismatch. The variations on the curves are due to the mismatch on *OCV* and *Q*.

The variation on the right side of the curves, when the *SoC* is close to 100% is mainly due to the mismatch on the charge *Q*. In stationary case with the load current is equal to zero, the voltage of the battery pack is equal to the *OCV*. Therefore, Fig. 5 shows the variations on the charge stored in the different cells for each value of the voltage of the battery pack.

4 Conclusions

This work studies the effect on a battery pack of mismatch among parallel connected cells. The BMS can estimate the effective current and *SoC* of the single cell only using additional high-cost current sensors. Current peaks, overcharge or over-discharge may cause cell damage or ageing acceleration. The maximum value of mismatch among the cell parameters that will be connected in parallel must be properly defined and the allowed voltage range of the battery pack must be reduced to avoid overcharge and over-discharge of some of the cells. This causes a reduction of the effective usable capacity of the battery pack. If usually the charge of a single cell is maintained between 10% to 90% of the nominal capacity, the charge of the battery pack in case of mismatch must be between 15% to 85% of the nominal capacity, that is an additional 10% of battery

capacity cannot be used. The effective usable capacity decreases increasing mismatch variance of the parameters, in particular of *OCV* and *Q* parameters.

References

1. European Commission: Commission staff working document on the evaluation of the Directive 2006/66/EC on batteries and accumulators and waste batteries and accumulators and repealing Directive 91/157/EEC (2019). https://ec.europa.eu/environment/pdf/waste/batteries/evaluation_report_batteries_directive.pdf
2. Horiba, T.: Lithium-Ion battery systems. Proc. IEEE **102**(6), 939–950 (2014). <https://doi.org/10.1109/JPROC.2014.2319832>
3. Byrne, R.H., Nguyen, T.A., Copp, D.A., Chalamala, B.R., Gyuk, I.: Energy management and optimization methods for grid energy storage systems. IEEE Access **6**, 13231–13260 (2018). <https://doi.org/10.1109/ACCESS.2017.2741578>
4. Hannan, M.A., Hoque, M.M., Hussain, A., Yusof, Y., Ker, P.J.: State-of-the-art and energy management system of lithium-ion batteries in electric vehicle applications: issues and recommendations. IEEE Access **6**, 19362–19378 (2018). <https://doi.org/10.1109/ACCESS.2018.2817655>
5. Xiong, R., Cao, J., Yu, Q., He, H., Sun, F.: Critical review on the battery state of charge estimation methods for electric vehicles. IEEE Access **6**, 1832–1843 (2018). <https://doi.org/10.1109/ACCESS.2017.2780258>
6. Maitreya, S., Jain, H., Paliwal, P.: Scalable and de-centralized battery management system for parallel operation of multiple battery packs. In: 2021 Innovations in Energy Management and Renewable Resources (52042), pp. 1–7 February 2021. <https://doi.org/10.1109/IEMRES2042.2021.9386861>
7. Kharisma, M.D., Ridwan, M., Ilmiawan, A.F., Ario Nurman, F., Rizal, S.: Modeling and simulation of lithium-ion battery pack using modified battery cell model. In: 2019 6th International Conference on Electric Vehicular Technology (ICEVT), pp. 25–30, November 2019. <https://doi.org/10.1109/ICEVT48285.2019.8994009>
8. Ye, M., Song, X., Xiong, R., Sun, F.: A novel dynamic performance analysis and evaluation model of series-parallel connected battery pack for electric vehicles. IEEE Access **7**, 14256–14265 (2019). <https://doi.org/10.1109/ACCESS.2019.2892394>
9. Gong, X., Xiong, R., Mi, C.C.: Study of the characteristics of battery packs in electric vehicles with parallel-connected lithium-ion battery cells. In: 2014 IEEE Applied Power Electronics Conference and Exposition - APEC 2014, pp. 3218–3224, March 2014. <https://doi.org/10.1109/APEC.2014.6803766>
10. Orcioni, S., Buccolini, L., Ricci, A., Conti, M.: Lithium-ion battery electro-thermal model, parameter estimation and simulation environment. Energies **10**(3) 375 (2017). <https://doi.org/10.3390/en10030375>
11. Makinejad, K., et al.: A lumped electro-thermal model for Li-ion cells in electric vehicle application. In: 28th International Electric Vehicle Symposium and Exhibition, EVS 2015, pp. 1–13 (2015)
12. Liu, L., Wang, L.Y., Chen, Z., Wang, C., Lin, F., Wang, H.: Integrated system identification and state-of-charge estimation of battery systems. IEEE Trans. Energy Convers. **28**(1), 12–23 (2013). <https://doi.org/10.1109/TEC.2012.2223700>
13. Orcioni, S., Biagetti, G., Conti, M.: SystemC-WMS: mixed-signal simulation based on wave exchanges. In: Vachoux, A. (ed.) Applications of Specification and Design Languages for SoCs, pp. 171–185. Springer, Dordrecht (2006). https://doi.org/10.1007/978-1-4020-4998-9_10

14. Biagetti, G., Giammarini, M., Ballicchia, M., Conti, M., Orcioni, S.: SystemC-WMS: wave mixed signal simulator for non-linear heterogeneous systems. *Int. J. Embed. Syst.* **6**(4), 277 (2014). <https://doi.org/10.1504/IJES.2014.064982>
15. Orcioni, S., Ricci, A., Buccolini, L., Scavongelli, C., Conti, M.: Effects of variability of the characteristics of single cell on the performance of a lithium-ion battery pack. In: 2017 13th Workshop on Intelligent Solutions in Embedded Systems (WISES), pp. 15–21 (2017). <https://doi.org/10.1109/WISES.2017.7986926>
16. Buccolini, L., Ricci, A., Scavongelli, C., DeMaso-Gentile, G., Orcioni, S., Conti, M.: Battery Management System (BMS) simulation environment for electric vehicles. In: 2016 IEEE 16th International Conference on Environment and Electrical Engineering (EEEIC), pp. 1–6, June 2016, <https://doi.org/10.1109/EEEIC.2016.7555475>
17. Scavongelli, C., Francesco, F., Orcioni, S., Conti, M.: Battery management system simulation using SystemC. In: 2015 12th International Workshop on Intelligent Solutions in Embedded Systems, WISES 2015, pp. 151–156 (2015)



Efficient Training and Hardware Co-design of Machine Learning Models

Mohammad Amir Mansoori^(✉) and Mario R. Casu

Politecnico di Torino, Turin, Italy
{mohammadamir.mansoori,mario.casu}@polito.it

Abstract. To implement a Machine Learning (ML) model in hardware (Hw), usually a first Design Space Exploration (DSE) optimizes the model hyper-parameters in search of the best ML performance, while a second DSE finds the configuration with the best Hw performance. Multiple iterations of these steps might be needed as the optimal ML model may not necessarily be implementable. To reduce the design-time and provide the designer with a single exploration environment, we propose a general framework based on Bayesian Optimization (BO) and High-Level Synthesis (HLS), which performs at once both DSEs generating efficient Pareto curves in the space of ML and Hw performance.

Keywords: Machine Learning (ML) · Hardware acceleration · High Level Synthesis (HLS) · FPGAs · Bayesian optimization

1 Introduction

Machine Learning (ML) techniques can be very effective in various edge applications (medical diagnosis, computer vision, robotics) but very often require hardware (Hw) accelerators for power- and cost-efficient inference. The usual design method consists of a Design Space Exploration (DSE) to fine-tune the hyper-parameters of an ML model, followed by another DSE that aims to optimize the Hw design for a given target. In this work we consider small-size FPGAs as Hw target and a high-level design approach using High-Level Synthesis (HLS).

This approach of separate DSE is shown in Fig. 1(a) and is useful for those applications where the ML accuracy has to be maximized and powerful Hw accelerators can be selected to meet the desired performance. However, being bounded to one specific small-size Hw architecture makes the design more challenging, calling for the joint optimization strategy shown in Fig. 1(b). The joint method avoids lengthy iterations that occur when the selected ML model is incompatible with the Hw constraints (e.g., it exceeds the available resources or cannot meet the timing requirements) and can obtain a better trade-off between ML performance and Hw performance.

For this reason, we propose a general framework for the joint optimization of training hyper-parameters and HLS-based hardware configurations based on

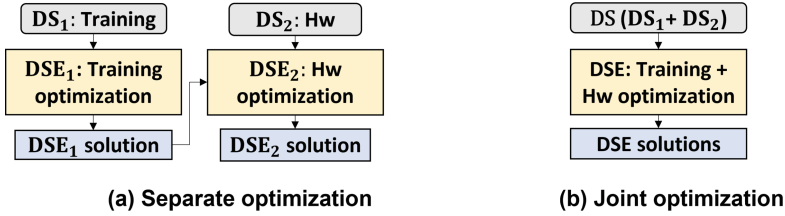


Fig. 1. Optimization of training hyper-parameters and hardware configurations: (a) traditional separate DSE, (b) more efficient joint DSE. DS_1 and DS_2 stand for Design Space of training and hardware design, respectively.

Multi-Objective Bayesian Optimization (BO) with constraints. This methodology allows for the optimization of multiple hardware parameters in HLS including clock frequency, data precision, and different levels of parallelism (unroll factor or array partitioning), together with the ML training hyper-parameters. BO optimizes black-box objective functions that are expensive to evaluate. It uses Gaussian Processes (GP) as the prior model to quantify the uncertainty of these functions. An acquisition function is then constructed based on the expected improvement of the prior model to decide where to sample the objective functions next [9]. In this work we use Predictive Entropy Search for Multi-Objective Optimization with Constraints (PESMOC) [7] for the acquisition function. In each BO iteration a new sample is suggested by the acquisition function and the objectives are evaluated for the new sample in order to update the Gaussian models and the acquisition functions in the next iteration. The main contributions of this work are as follows:

- An efficient method for the joint optimization of training hyper-parameters and HLS-based hardware configurations in the context of Machine Learning.
- Exploitation of the capabilities of BO for Multi-Objective optimization subject to positive constraints in the design of ML models in hardware.
- Inclusion of the clock frequency as an optimization target, which is currently not optimized in HLS tools for Xilinx devices (Vivado HLS).
- Evaluation of the proposed methodology for the efficient design of a Multi-Layer Perceptron (MLP) type of Neural Network in a Zynq FPGA.

2 Related Work

To find the optimum hyper-parameters of an ML model during training, more powerful approaches than simple Grid Search and Random Search are in use nowadays, such as Auto-Sklearn, HyperOpt, Auto-Keras, and Keras-Tuner. In [1], a review of commonly-used methods in automated ML has been presented. The algorithms used in these methods can be divided into Reinforcement Learning (RL), BO, and Evolutionary Algorithms (EAs). Although RL and BO share similar characteristics and have been proven effective in several works, one of the difficulties of RL is that the *policies* and *reward function* need customization for

each DSE problem, while the BO’s Gaussian Process is a natural fit for these optimization problems. EAs are computationally expensive, which led several authors to propose other methods to speed up the computations.

Regarding the DSE aimed at optimizing the hardware configurations, several works have focused on High Level Synthesis (HLS) as the hardware design tool and proposed different methodologies for the optimum selection of HLS *pragmas* [2,3]. In [4] Multi-Objective Bayesian Optimization is used to tune the HLS configurations. Although the generated Pareto Fronts are close to the actual optimal points, it could not consider *constraints* in addition to multiple *objectives* in the optimization process which increases the exploration time.

Recently, there has been a growing interest in the joint optimization of hardware and training parameters, specially in the context of Deep Neural Networks (DNNs) for which Hardware-aware Neural Architecture Search (HW-NAS) has been introduced. In the context of HW-NAS, numerous works have been presented for the co-optimization of network architecture and hardware-related parameters that are thoroughly described in a recent review paper [5]. One of the challenges in the recent HW-NAS methodologies is designing efficient algorithms to solve the expensive Multi-Objective Optimization (MOO) problem. According to [5], most of the previous works use single-objective optimization targeting network accuracy *constrained* to hardware-aware characteristics. Other approaches combine multiple objectives into one single function. Evolutionary algorithms are mostly used for MOO in HW-NAS which are computationally expensive. In [6] a two-step BO is used to reduce the complexity of MOO. Our work has features in common with previous HW-NAS works, but is not restricted to DNNs, uses BO as a more efficient exploration method for multi-objective functions, and uses HLS in the optimization loop to obtain an accurate estimation of the Hw performance for a given configuration.

Finding both optimum parameters for ML model training and hardware can be seen as a Multi-Objective Optimization problem with *Constraints* (MOOC). Recently, Garrido-Merchán et al. extended their Spearmint software for applying BO to MOOC and proved its efficiency in designing a DNN architecture in hardware [7]. The hardware area is selected as a constraint and is estimated by Aladdin, a pre-RTL performance estimator for ASIC accelerator design [8]. Unlike our work, FPGA is not considered as a Hw target and HLS is not used.

3 Proposed Methodology

Figure 2 illustrates our methodology. The BO takes as input the range of parameters for both training and Hw exploration and aims to optimize at once three objectives: (1) Training error on the ML dataset using floating-point precision in Python; (2) Prediction (i.e., inference) error after Hw implementation, which takes into account fixed-point quantization and is obtained with the HLS C-simulation tool from the Vivado suite; (3) Hw inference performance expressed as clock period times the number of clock cycles as obtained from the HLS tool¹. The constraints are the FPGA resources (BRAMs, DSPs, FFs, LUTs).

¹ We leave power optimization for future work.

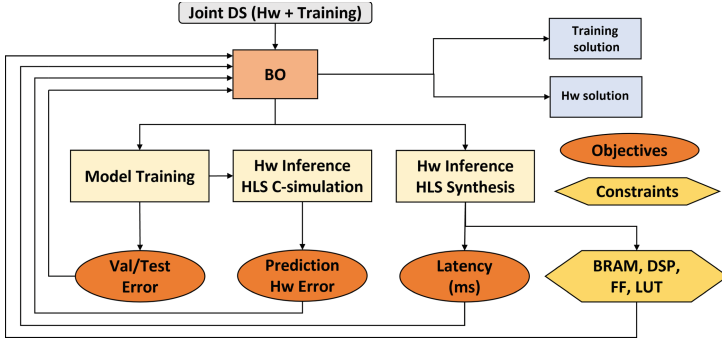


Fig. 2. Proposed methodology for the joint optimization of training parameters and HLS-based hardware configurations (BO = Bayesian Optimization).

The BO considers a Gaussian Process for each of the objectives and constraints. The training hyper-parameters and the Hw configuration parameters (HLS *pragmas*, data precision, clock frequency) are updated in each BO iteration based on the maximization of PESMOC acquisition function, which suggests a new sample in the design space to be evaluated by the objective functions. The non-dominated points of the design space are obtained at the end of the iterative process.

4 Results

For the evaluation, we used the MNIST dataset to train a Multi-Layer Perceptron (MLP) to be implemented in a Zynq7000 FPGA. We used hls4ml [10] to convert the MLP model to a synthesizable C++ code. The model hyper-parameters and the ranges of the HLS and Hw knobs are in Table 1.

Table 1. Ranges of parameters for the joint training/Hw optimization method.

Inputs	Clk (ns)	Hidden layers	Neurons	Precision #total	Precision #Integer	Reuse factor	Array partition	Learning rate	Regularization rate
Ranges	4-7	1-3	32-256 step = 32	12-16	4-6	1-4	2^x $x = [1-8]$	$1 \times 10^{(-x)}$ $x = [2-7]$	$1 \times 10^{(-x)}$ $x = [2-7]$

Figure 3 compares the evolution of the training error using a joint (Fig. 3(a)) and a separate (Fig. 3(b)) optimization approach. Figure 3(a) shows that both floating-point error during training and fixed-point error in hardware converge as the BO iterations progress. Since the separate method returns only the best training result, Fig. 3(b) shows only the fixed-point inference error and shows an immediately low error in all BO iterations (less than 5%). This is because the separate hardware design starts with a neural network already optimized in terms of training error, which only needs to be tailored to the hardware target.

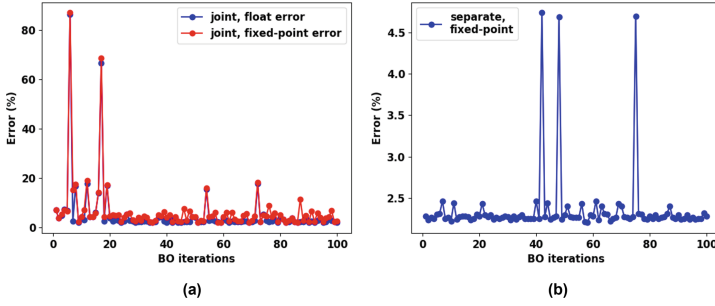


Fig. 3. Percentage of training error (float error) and hardware error (fixed-point error) in each BO iteration, (a) proposed joint optimization, (b) separate optimization.

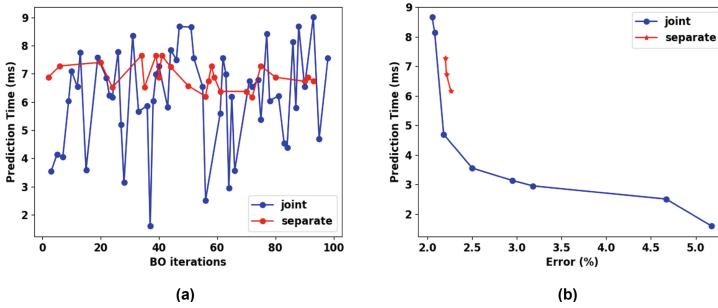


Fig. 4. Comparison of (a) prediction time and (b) Pareto fronts. (Color figure online)

This last optimization of the separate method, however, is constrained by the initial training. As a result, the BO cannot reach the same latency performance of the joint optimization. This is visible in Fig. 4(a), with the relatively high prediction time for the separate optimization method (red points).

The efficiency of the proposed methodology is apparent in Fig. 4(b), which compares the Pareto curves obtained by separate (red) and joint (blue) optimization methods. In the red curve, the training optimization is done by Keras-Tuner. Note that in this case Keras-Tuner suggests an initial MLP with three layers and a number of neurons that could not fit in the FPGA due to excessive BRAM usage, leading to a failure in the subsequent hardware BO. This required a second iteration to limit the neurons range from $[32 - 256]$ to $[32 - 128]$ in the training DSE, which returned a feasible three-layer MLP with 128 neurons in each layer, low error but relatively high prediction time. The subsequent Hw DSE returned only three (red) Pareto points. On the contrary, the joint method returns many more valid Pareto points (blue) because of its ample maneuverability in the combined training and hardware design spaces. Most importantly, the blue points dominate the red ones, as clearly shown in Fig. 4(b).

5 Conclusions and Future Work

We proposed a new strategy, based on Multi-Objective Bayesian Optimization subject to positive constraints, for the efficient design of Machine Learning models in FPGA-based hardware accelerators, which can simultaneously optimize the training hyper-parameters and HLS-based hardware configurations. It optimizes prediction error and latency, subject to FPGA resource constraints. The results show the efficiency of the Pareto sets obtained by the proposed method with near $2\times$ reduction of prediction time without an increase in the prediction error compared to the traditional separate optimization design. In the future, we will evaluate other ML models with this methodology, such as Support Vector Machine (SVM) and Random Forest (RF). Other techniques like Random Search and evolutionary algorithms can be compared with our approach in terms of computational time and efficiency of the Pareto fronts.

Acknowledgments. This work was supported by the EMERALD project funded by the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No. 764479.

References

1. Chen, Y., Song, Q., Hu, X.: Techniques for automated machine learning. *ACM SIGKDD Explor. Newsl.* **22**(2), 35–50 (2021)
2. Sohrabizadeh, A., et al.: AutoDSE: Enabling Software Programmers Design Efficient FPGA Accelerators. arXiv preprint [arXiv:2009.14381](https://arxiv.org/abs/2009.14381) (2020)
3. Zhao, J., et al.: Performance modeling and directives optimization for high-level synthesis on FPGA. *IEEE TCAD* **39**(7), 1428–1441 (2019)
4. Mehrabi, A., et al.: Bayesian optimization for efficient accelerator synthesis. *ACM TACO* **18**(1), 1–25 (2020)
5. Benmeziane, H., et al.: A Comprehensive Survey on Hardware-Aware Neural Architecture Search. arXiv preprint [arXiv:2101.09336](https://arxiv.org/abs/2101.09336) (2021)
6. Parsa, M., et al.: Pabo: pseudo agent-based multi-objective Bayesian hyperparameter optimization for efficient neural accelerator design. In: 2019 IEEE/ACM ICCAD (2019)
7. Garrido-Merchán, E.C., Hernández-Lobato, D.: Predictive entropy search for multi-objective Bayesian optimization with constraints. *Neurocomputing* **361**, 50–68 (2019)
8. Shao, Y.S., et al.: Aladdin: a pre-rtl, power-performance accelerator simulator enabling large design space exploration of customized architectures. In: 2014 ACM/IEEE 41st ISCA (2014)
9. Frazier, P.I.: A tutorial on Bayesian optimization. arXiv preprint [arXiv:1807.02811](https://arxiv.org/abs/1807.02811) (2018)
10. Fahim, F., et al.: An Open-Source Codesign Workflow to Empower Scientific Low-Power Machine Learning Devices. ArXiv preprint [arxiv:2103.05579](https://arxiv.org/abs/2103.05579) (2021)



Modeling the Line Interruption Issue in a Railway Network

Luca Fronda¹, Riccardo Berta^{1(✉)}, Paolo Cesario^{2(✉)}, Alessandro De Gloria¹,
and Francesco Bellotti¹

¹ Department of Electrical, Electronic and Telecommunication Engineering (DITEN),
University of Genoa, Via Opera Pia 11a, 16145 Genova, Italy
s4047957@studenti.unige.it, {alessandro.degloria,
francesco.bellotti}@unige.it

² Si Consulting, Srl - Via Gavotti 5/6, 16128 Genova, Italy
paolo.cesario@siconsulting.biz

Abstract. Rail line interruptions are rare but very costly events, as they require a complete re-definition not only of the timetable of the trains, but also of their path, with major variations at least in the hit area. To the best of our knowledge, the literature is rich of documentation on timetable re-scheduling in case of delays and/or disruption of train lines, but without considering path deviations.

The Flatland initiative has published a 2D railway-world toolkit that allows developers devise and experiment with different solutions to deal with the train re-scheduling issue particularly, but not exclusively, through reinforcement learning (RL) with multi-agent path finding (MAPF). While the approach looks very promising to deal with the complexity of the issue, the platform still has some limitations in terms of the modeling of a realistic railway network scenario. This article proposes the integration of some key features in order to make Flatland a valuable platform for training RL agents for supporting decision making in real-world train re-scheduling.

Keywords: Multi agent path finding · Reinforcement learning · Flatland · Railway interruption

1 Introduction

Rail line interruptions (e.g., due to bad weather or accidents) are rare but very costly events, as they require a re-definition not only of the timetable of the trains, but also of their path, with major variations in the hit area, and, possibly, in the whole national network as well.

In case of large disruptions, railway traffic controllers must apply fast and proper measures to guarantee the train services and minimize delay propagation to the rest of the network. Currently, contingency plans are used to assist traffic controllers in dealing with disrupted traffic [1]. Contingency plans are referred to a specific disruption scenario and to a specific infrastructure, so they are static and limited. In order to address

the specificity of the current disruption case, controllers often have to make adjustments to the original contingency plans. Hence, an efficient and robust decision support system would be a very useful instrument in this context.

The Timetable Rescheduling Problem (TRP) concerns the redefinition of the timetable in case of disruptions or delays. This involves the arrival and departure times of trains, the platforms at which the train should stop and the order of the different trains. To the best of our knowledge, a main limit of the literature on this problem is that it does not consider the possibility of train path deviation, which is anyways a significant opportunity in meshed railway networks, as is the case of several countries, especially in Europe.

Integer Programming (IP) [2], Mixed Integer Programming (MIP) [3] and Alternative Graph (AG) [4] are the three most commonly used models in literature for TRPs in railway networks. Other methods include discrete event simulation [5, 6] and simulation [7]. To find optimal solutions with these models, several techniques have been studied, such as Branch & bound (B&B), heuristic approaches, metaheuristic approaches (including Generative Algorithms (Gas), Simulated Annealing (SA), Tabu Search (TS), Ant Colony Optimization (ACO), Evolutionary Algorithms (EAs)), standard solvers (including CPLEX, GLPK, etc.), etc.

The Flatland initiative, supported by the Swiss and German federal railways, represents a significant novelty in the area. Flatland provides an open source 2D railway toolkit to develop and compare solutions for the problem of efficiently managing dense traffic on complex railway networks. A challenge is held yearly, that addresses the vehicle rescheduling problem by providing a simplistic grid world environment where researchers and practitioners can develop and test solutions based on any methodological approach, with a particular attention to (distributed or centralized) multi-agent reinforcement learning (RL) [8].

We consider this approach to be very promising, because of the ability, demonstrated by RL, of managing complex issues, by learning through (simulated) experience [9]. Moreover, the support to distributed multi-agent path finding opens new perspectives for new transportation services based on dynamic needs. However, the Flatland platform is still limited in terms of modeling a realistic railway network scenario, particularly as the current environment only allows assigning targets that trains have to reach as soon as possible, avoiding any crash. On the other hand, a realistic re-scheduling problem requires at least the definition of the original train schedules and paths, that the system should try to respect as much as possible, also in case of interruptions.

The remainder of this paper is organized as follows. Section 2 presents a critical analysis of the literature concerning the train rescheduling problem. Section 3 presents the proposed extension to the Flatland platform. Section 4 draws the conclusions on the ongoing work.

2 Literature Review

Literature presents various models to address TRP. The IP model relies on binary decision variables, like the priority of two trains, the sequence of trains, and on integer variables, like the arrival and departure times, and the delays. Constraints in the railway network

management are generally expressed by equations and inequalities. In [2], an IP model considering part of the Netherland railway is defined, with a time window of 2 h and 282 trains. The objective is to minimize the arrival time of the passengers.

In the MIP model, the arrival and departure times and delays are considered as a continuous variable. A MIP model of the Dutch railway network, considering a time window of 75 min, is presented in Cavone et al. [3], with the objective to minimize the average and maximum arrival delays of trains.

The AG is a generalization of the disjunctive graph and is used to model complex job shop scheduling problems (JSSPs). In [4], the Dutch national railway network is modeled, with a 1-hour time window and 679 trains. The objective of the model is to minimize the delay propagation to the entire network.

The discrete event models describe the railway system by the state of each train and the discrete events happening in the network at the given time. The dynamic process controls all the trains, determining the optimal velocities, orders, and connections for optimizing different objectives. In [5], the Madrid Regional Railway Network is simulated, exploiting real data provided by the Spanish national railway company, RENFE Operadora. In [6], the authors use a discrete event model to simulate three lines of the Indian railway. This work uses RL in order to minimize the propagation of the delays in the railway net.

A typical simulation model simulates the status of a railway system, but also forecasts the future status, also looking to resolve conflicts. Therefore, the rescheduling approaches can be integrated into the simulation model to support real-time dispatching. In [7], the authors propose a simulation of the Lucerne station area with 30 trains per hour.

Flatland is a two-dimensional simplified grid environment developed by the Swiss Federal Railway Company (SBB) and Deutsche Bahn (DB), that allows faster experimentation providing an easy interface. In the Swiss Railway, on a typical day of operations, more than 10,000 train runs are executed on a network of more than 13,000 switches and 32,000 signals. Almost 1.2 million passengers are transported on the railway network each day. Flatland is a high-performance simulator, which represents and simulates the railway infrastructure and the dynamics of train traffic. Flatland provides specific support to RL. [8] argues that the Flatland environment is a robust and valuable framework to investigate the rescheduling problem in railway networks.

However, Flatland presents some important simplifications with respect to a real railway environment, as stressed by [10]. Particularly relevant to the TRPs are the following:

- Trains cannot move backwards (which could happen in case of a sudden line interruption).
- Trains move with a constant velocity, based on the type of the train, but independent of the railway segment.
- Trains do not follow a timetable, with intermediate stations (to call at before reaching the target station) and arrival and departure times.

Table 1 synthesizes the features of the above cited papers. Mathematical models are the most used ones in literature. They require short computation times, but they grow up very quickly in terms of constraints as the network size and the number of trains

Table 1. Synthesis of the analyzed literature on TRP.

	Model	Model size	Time horizon	Solver	Computation time
[2]	IP	Part of Dutch railway	2 h	CPLex	30 s
[3]	MIP	Dutch railway	75 min	Two level heuristic	60 s
[4]	AG	Dutch railway	1 h	Branch and bound	Maximum of 90 s
[5]	Discrete event simulation	Madrid Regional Railway Network	20 h	Alternative greedy heuristic	~15.5 h
[6]	Discrete event simulation	Three Indian railway lines	72 h	Q-Learning (RL)	~11 min
[7]	Simulation	Lucerne station area	Not specified	Real-time rescheduling and train control approach	Not specified

increase. This is a limit for the scalability (e.g., a nation-wide network in an entire day). Moreover, mathematical models do not consider path deviations, which is necessary in case of disruptions in dense networks.

Simulation models are used in limited geographical areas, with limits in terms of scalability. In this case, attention should be given to [6], that uses a discrete event simulation in order to develop a RL algorithm. In this case, the computational time is limited and the network considered is larger with respect to the other simulation models. This hints in favor of RL as a promising tool to tackle the TRP.

3 Extending the Flatland Platform

From the literature analysis, we argue that the Flatland environment provides a significant opportunity for research, as it allows testing different ML solutions, particularly RL, that could be adequate to address, in a flexible and scalable way, the issue of train rescheduling in case of a rail line interruption, as it could learn from simulated experience.

However, as anticipated, even a realistic albeit simplified use case of train rescheduling requires some features that need to be integrated in Flatland. Particularly, we highlight the following:

- Need for allowing the user to specify generation of trains in the network also after the initial simulation time.
- Need for allowing the user to specify, for each train, such details as: stations to call at, expected arrival and departure schedule in each station, preferential routing. This information is mandatory for passenger's trains – apart from the case of line interruption, which may require the system to learn a different route.

- Need for adding the reverse train action, which is necessary for a train which is in a line segment that has been abruptly interrupted.
- Need for dynamic velocities, that vary depending on the type of train, on the line in which the train is traveling, and on the delay that the train has accumulated.
- Need for considering train run suppression (while keeping alive the train at a station until its end of daily service). Thus, trains should be considered as physical entities, not as independent source-sink services. This implies that the schedule involves the entire path of a day for a train. In this way, a train could avoid a run (i.e., a segment of its daily path), in order to avoid excessive total delay, but do a next one later in the day, starting from the station at which it was blocked.

After analyzing the Flatland source code, we argue that these additions should be addressable by properly extending the schedule generation (by subclassing *Schedule Generator*) and management modules (*rail_env*, *agent_utils*), and by writing proper RL reward/penalty functions, which last point is a standard use of Flatland.

Once implemented, these features should make Flatland environment able to deal with a real-world TRP. A simplified test case will be the one presented in Fig. 1. The use case involves a very simple double track railway network with five stations.

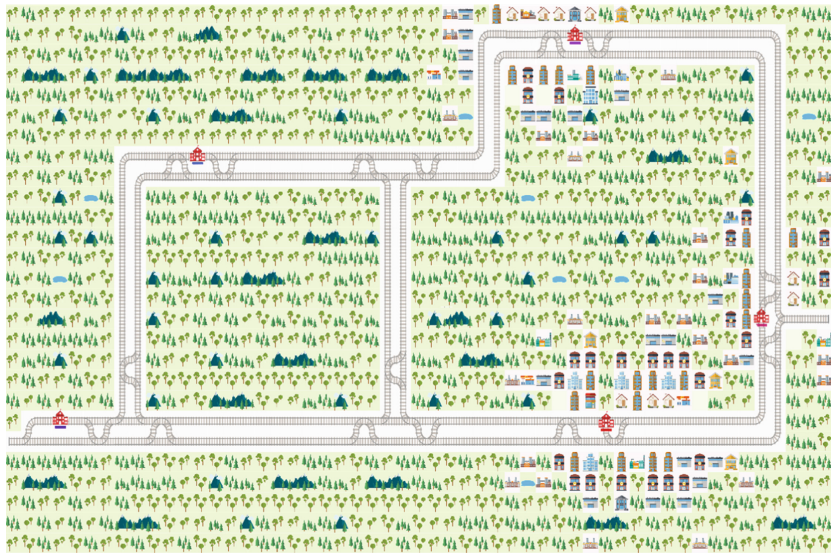


Fig. 1. A simple use case generated with the flatland environment, with five train stations (the red houses on the rail). The red cross represents a two-track disruption in the network.

The train agents should be trained on the whole network and daily activity, with random delays and interruptions, minimizing an overall cost function summing the daily delays and number of missed stations in case of deviation of a train from its ordinary path. It will be interesting to see whether additional information such as location of the

disruption should be provided by the environment to the agents in order to improve the overall performance.

4 Conclusions and Future Work

The problem of dealing with the TRP in case of disruption of a line is serious, but, to the best of our knowledge, has not yet been addressed in literature through a model considering the possibility of deviating trains from their original path in a meshed railway network. We argue that the problem could be tackled through a novel approach, based on RL, so that the train agents autonomously learn how to minimize a cost given by total delay and number of missed stations.

This paper has presented the extensions we are implementing on the open-source Flatland toolkit in order to make it able to better support decision making in real-world TRP cases. The classical research questions to be addressed in an upgraded model will concern the training time needed and accuracy achieved by RL path finding multiple agents, together with the ability of transferring learning in case of different railroad network configurations and, overall, interruption locations or delay states.

From the modeling point of view, a significant further extension will concern modeling train capacities and, overall, inter-station passenger flow demands.

Our focus is currently on a high-level railway network, considering only the connections among the cities, not the connections between a station and its railway depot(s). However, Flatland is a very flexible tool that can be used for this aspect as well. It is anyway of preliminary importance to see whether, once Flatland is extended, it is possible to train multiple RL agents to minimize the overall delay in case of an interruption in a railroad network.

Acknowledgments. This work was also supported by operative program Por FSE Regione Liguria 2014–2020 (Grant Agreement RLOF18ASSRIC).

References

1. Chu, F., Oetting, A.: Modeling capacity consumption considering disruption program characteristics and the transition phase to steady operations during disruptions. *J. Rail Transp. Plan. Manag.* **3**, 54–67 (2013)
2. Dollevoet, T., Huisman, D., Schmidt, M., Schöbel, A.: Delay management with rerouting of passengers. *Transp. Sci.* **46**(1), 74–89 (2012)
3. Cavone, G., Blenkers, L., van den Boom, T., Dotoli, M., Seatzu, C., De Schutter, B.: Railway disruption: a bi-level rescheduling algorithm. In: 2019 6th International Conference on Control, Decision and Information Technologies (CoDIT), pp. 54–59 (2019), <https://doi.org/10.1109/CoDIT.2019.8820380>
4. Kecman, P., Corman, F., D’Ariano, A., Goverde, R.M.P.: Rescheduling models for railway traffic management in large-scale networks. *Public Transp.* **5**(1/2), 95–123 (2013)
5. Almodóvar, M., García-Ródenas, R.: On-line reschedule optimization for passenger railways in case of emergencies. *Comput. Oper. Res.* **40**(3), 725–736 (2013)

6. Khadilkar, H.: A scalable reinforcement learning algorithm for scheduling railway lines. *IEEE Trans. Intell. Transp. Syst.* **20**(2), 727–736 (2019). <https://doi.org/10.1109/TITS.2018.2829165>
7. Lüthi, M., Medeossi, G., Nash, A.: Evaluation of an integrated realtime rescheduling and train control system for heavily used areas. In: *Proceedings of the IAROR International Seminar on Railway Operations Modelling and Analysis*, pp. 1–14 (2007)
8. Mohanty, S., et al.: Flatland-RL: multi-agent reinforcement learning on trains. arXiv preprint [arXiv:2012.05893](https://arxiv.org/abs/2012.05893) (2020)
9. Sutton, R.S., Barto, A.G.: *Reinforcement Learning: An Introduction*. MIT Press, Cambridge (2015)
10. Wälter, J.: Existing and novel approaches to the vehicle rescheduling problem (VRSP). Master's thesis, HSR Hochschule für Technik Rapperswil (2020)



The SENSIPLUS: A Single-Chip Fully Programmable Sensor Interface

Andrea Ria^{1,2}, Mattia Cicalini^{1,2}, Giuseppe Manfredini^{1,2}, Alessandro Catania¹, Massimo Piotto¹(✉), and Paolo Bruschi¹

¹ Dept. Ingegneria dell'Informazione, Università di Pisa, via Caruso 16, 56122 Pisa, Italy
{andrea.ria,mattia.cicalini,giuseppe.manfredini}@sensichips.com,
alessandro.catania@ing.unipi.it, {massimo.piotto,
paolo.bruschi}@unipi.it

² SENSICHIPS s.r.l., Via delle Valli 46, 04011 Aprilia, LT, Italy

Abstract. The SENSIPLUS, a recently introduced versatile sensor interface, is described. The SENSIPLUS is a single-chip solution that allows a wide variety of operations required by sensor systems, such as vector impedance, voltage and current measurements across a wide frequency range. Integration with standard embedded systems is facilitated by the presence of a configurable communication line. In this work, the capability of interfacing resistive sensors and performing impedance spectroscopy is demonstrated by means of experiments executed on external reference components.

1 Introduction

The pervasive introduction of electronic systems in an ever-growing number of different products has raised the interest on the development of smart sensor systems [1]. Collecting a large amount of data about physical and chemical quantities is a function that is mandatory for robots, wearable devices and monitoring networks that aspire to perform tasks that are really useful for improving the quality of life of modern citizens [2–4]. In order to reduce the impact on the whole system, both the size and the power consumption of sensors must be scaled down. This requirement also involves the electronic interface that should be coupled to sensors. One problem that has to be faced when designing sensor systems is that different sensors often require different readout systems [5, 6]. Using a dedicated electronic interface for each sensor results in low power and size optimization. Complex general purpose interfaces have been proposed over the years [7–10] where different sensors can be connected.

In last decade, SENSICHIPS s.r.l [11] has carried out the development of the SENSIPLUS programmable sensor interface, in cooperation with the University of Pisa. This interface performs temperature and humidity measurements by means of integrated sensors and the Electrochemical Impedance Spectroscopy (EIS) on external sensors connected using the two or four contact configurations. The SENSIPLUS can manage up to 12 external different sensors overcoming the limit of other interfaces [7–10].

Applications of the interface to different sensor systems and the development of Artificial Intelligence-based data processing protocols applied to measurements performed

with the SENSIPLUS platform have been proposed [12–14]. Recently, successful use of the SENSIPLUS for lithium batteries monitoring has been demonstrated [15].

In this work, we describe the SENSIPLUS interface, providing an overview of its main capabilities. Experimental results concerning vector impedance measurements of external reference components with the proposed interface are reported.

2 Device Description

A simplified block diagram of the SENSIPLUS interface is shown in Fig. 1. The most innovative part of the interface is represented by the Analog Front End (AFE), which includes (i) a Programmable Stimuli Generator (PSG) with either voltage or current output; (ii) a Trans-Impedance Amplifier (TIA), capable of setting the voltage at its output terminal and detecting the current that flows into it; (iii) a programmable Instrumentation Amplifier (IA) provided of input and output chopper modulators. Both the PSG and TIA have practically rail-to-rail outputs and are widely programmable. Signal generation is accomplished by means of highly efficient Switched Capacitors (SC) architectures, which also use Correlated Double Sampling (CDS) to reduce the impact of offset and flicker noise from the amplifiers.

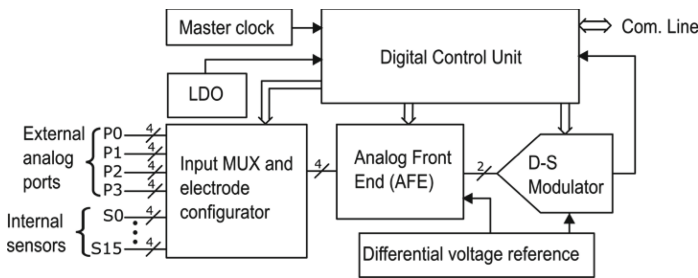


Fig. 1. Simplified block diagram of the SENSIPLUS versatile sensor interface.

The IA is based on an original architecture [16], having a bandwidth of 2 MHz and a gain programmable from 1 to 40 in four steps. Quasi rail-to-rail inputs (300 mV of margin from both rails) facilitate interfacing a wide variety of sensors. The AFE is accessible to four fully programmable terminals that can be configured to perform 2-wire and 4-wire complex impedance measurements, current and differential voltage measurements. The IA can be used to read an external voltage or to amplify and demodulate the output voltage of the TIA, operating a current reading.

The IA output voltage (differential) is read by a Delta-Sigma (DS) Analog-To-Digital (ADC) converter, based on a 2nd order discrete-time SC modulator. The reference voltage (differential) for the ADC and AFE is generated by an accurate programmable SC band-gap circuit, described in [17]. The current version of the SENSIPLUS has four analog ports (P0–P3), each one consisting of four wires. Selection of the port is performed by an analog multiplexer (MUX), which has also the function of configuring the AFE terminals according to the measurements to be performed.

All the functions of the chip are controlled by a digital finite state machine (Digital Control Unit - DCU in Fig. 1) which also includes the decimation filter of the ADC.

The DCU communicates with an external host through a versatile communication interface that can be programmed to implement the standard SPI, I2C protocols and a proprietary single-wire protocol (SENSIBUS).

The SENSIPLUS is fabricated using the UMC 0.18 μm CMOS process. All analog blocks are based on I/O MOSFETs, being able to be directly connected to the 3.3 V power supply. In this way, relatively wide ranges are allowed for the analog ports.

A photograph of the whole chip is shown in Fig. 2.

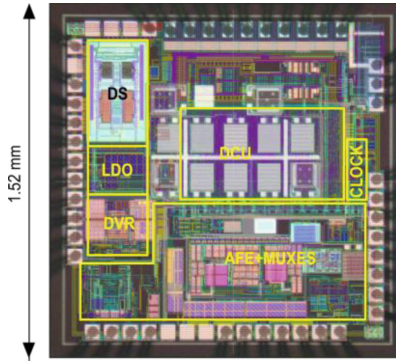


Fig. 2. Chip optical micrograph with the layout superimposed in order to show the blocks covered by planarization dummies. The main blocks are indicated (with reference to Fig. 1).

For the tests described in this work, the SENSIPLUS was bonded into a JLCC44 package allowing full access to the four analog ports and to diagnostic pads. A purposely built development board (7 cm \times 7 cm) was used to support the SENSIPLUS. A summary of performances is given in Table 1.

Table 1. Summary of specifications. Unless differently specified, all values refer to $V_{dd} = 3.3$ V.

Supply current (all on)	1.2 mA	PSG: DC component resolution	12 bit
Supply current (idle)	22 μA	PSG: maximum frequency	1 MHz
Power supply	1.8 V–3.3 V	PSG: minimum frequency:	40 Hz
Input voltage resolution (gain = 40)	4 μV	TIA: maximum current	10 mA
Maximum input voltage (gain = 1)	± 2.7 V	TIA: best current resolution	0.1 nA
Input offset voltage (max)	8 μV	ADC effective resolution	12 bit

3 Experimental Results

The experiments were performed by connecting the SENSIPLUS to a microcontroller through the communication line, configured as SPI. The microcontroller allowed sending commands from a personal computer (through an USB line) to the SPI.

In a first series of tests, the SENSIPLUS was configured for 2-wire resistance measurements. Resistors of different values (1% accuracy) were connected between the PSG and TIA terminals of port P0. The TIA was set to produce a $V_{dd}/2$ voltage, while the PSG was programmed to produce a 3 V, 78 kHz peak-to-peak sinusoidal waveform with a mean value of $V_{dd}/2$. The in-phase component of the current detected by the TIA was used to estimate the resistance value. The result is shown in Fig. 3(a). Note that a constant scale factor (nearly 0.88) distinguishes the measured value from the actual one. This is due to the unavoidable inaccuracy of the reference internal resistor used by the TIA. This error can be easily compensated by means of simple one-point calibration. The resolution in the resistance measurement is shown in Fig. 3(b), where the results of successive acquisitions are shown for two resistance values differing by 0.2%.

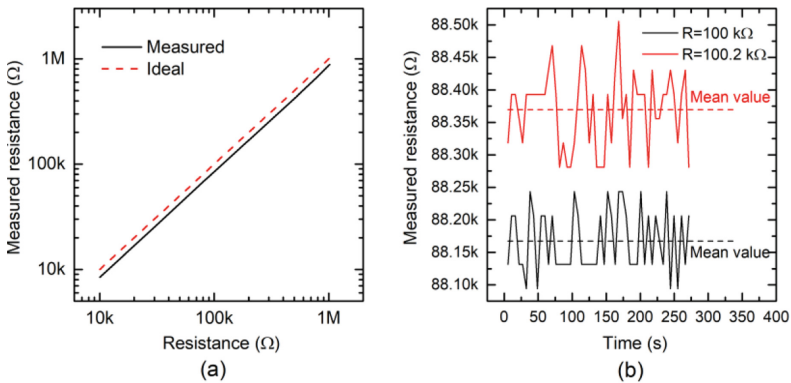


Fig. 3. (a) Measured resistance as a function of the input resistance; (b) sequence of measurements showing the equivalent resistance noise for two resistance values differing by 0.2%.

Examples of impedance spectroscopy measurements are shown in Fig. 4, where the magnitude and the phase acquired by the SENSIPLUS with 2-wire configuration are plotted as a function of frequency. In Fig. 4(a), the tested bipole is the parallel of a resistance (100 kΩ) and of a capacitor (560 pF). In Fig. 4(b) the same components are placed in series. Comparison with the calculated responses shows an optimum agreement, with only the scale factor error mentioned earlier. Finally, Fig. 5 shows the result of a C-V measurement performed with the SENSIPLUS on a commercial diode (Vishay, 1N4001). For this test, the magnitude of the sinusoidal stimulus (78 kHz) was reduced to 200 mV and the DC voltage was swept across the displayed range. The result is in agreement with the nominal curve reported in the device datasheet.

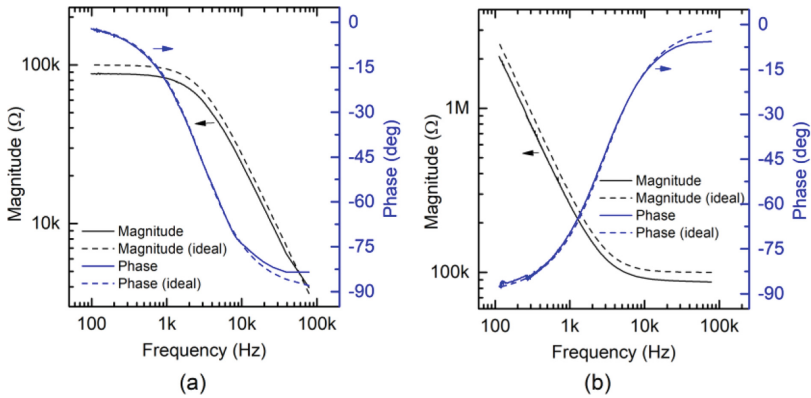


Fig. 4. Magnitude and phase of a bipole formed by the parallel (a) and series (b) of a capacitor (560 pF) and a resistor (100 kΩ) as a function of frequency measured by the SENSIPLUS platform. The experimental data (solid lines) are compared with the ideal curves (dashed lines).

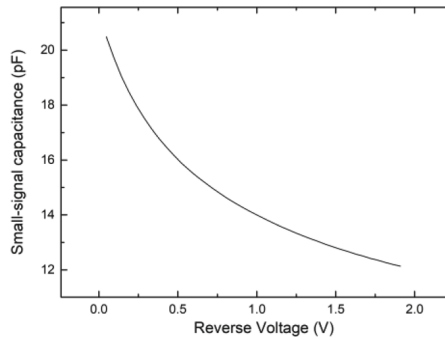


Fig. 5. C-V measurement performed on a commercial diode (Vishay, 1N4001).

4 Conclusions

The innovative SENSIPLUS general purpose sensor interface has been described. The interface is fully programmable and an input voltage resolution of $4 \mu\text{V}$ has been achieved. The use as an impedance analyser has been demonstrated considering both series and parallel configurations. The experimental results show a good agreement with the ideal behaviour.

References

1. Alioto, M. (ed.): Enabling the Internet of Things. Springer, Cham (2017). <https://doi.org/10.1007/978-3-319-51482-6>
2. Promphet, N., Ummartyotin, S., Ngeontae, W., Puthongkham, P., Rodthongkum, N.: Non-invasive wearable chemical sensors in real-life applications. *Anal. Chim. Acta* **1179** (2021)
3. Jo, S., Sung, D., Kim, S., Koo, J.: A review of wearable biosensors for sweat analysis. *Biomed. Eng. Lett.* **11**(2), 117–129 (2021). <https://doi.org/10.1007/s13534-021-00191-y>

4. Talebkhah, M., Sali, A., Marjani, M., Gordan, M., Hashim, S.J., Rokhani, F.Z.: IoT and big data applications in smart cities: recent advances, challenges, and critical issues. *IEEE Access* **9**, 55465–55484 (2021)
5. Crescentini, M., Bennati, M., Tartagni, M.: A high resolution interface for kelvin impedance sensing. *IEEE J. Solid-State Circuits* **49**, 2199–2212 (2012)
6. Jafari, H., Soleymani, L., Genov, R.: 16-Channel CMOS impedance spectroscopy DNA analyzer with dual-slope multiplying ADCs. *IEEE Trans. Biomed. Circuits Syst.* **6**, 468–478 (2012)
7. Van Helleputte, N., et al.: A multi-parameter signal-acquisition SoC for connected personal health applications. In: 2014 IEEE International Solid-State Circuits Conference (ISSCC) (2014)
8. Konijnenburg, M., et al.: A battery-powered efficient multi-sensor acquisition system with simultaneous ECG, BIO-Z, GSR, and PPG. In: 2016 IEEE International Solid-State Circuits Conference (ISSCC) (2016)
9. Kim, J., Ko, H.: Reconfigurable multiparameter biosignal acquisition SoC for low power wearable platform. *Sensors* **16**, 1–13 (2016)
10. Choi, S., et al.: A wide dynamic range multi-sensor ROIC for portable environmental monitoring systems with two-step self-optimization schemes. *IEEE Trans. Circuits Syst. I Regul. Pap.* **68**(6), 2432–2443 (2021)
11. www.sensichips.com. Accessed July 2021
12. Molinara, M., Ferdinandi, M., Cerro, G., Ferrigno, L., Massera, E.: An end-to-end indoor air monitoring system based on machine learning and SENSIPLUS platform. *IEEE Access* **8**, 72204–72215 (2020)
13. Bria, A., Cerro, G., Ferdinandi, M., Marrocco, C., Molinara, M.: An IoT-ready solution for automated recognition of water contaminants. *Pattern Recogn. Lett.* **135**, 188–195 (2020)
14. Bruschi, P., et al.: A novel integrated smart system for indoor air monitoring and gas recognition. In: 2018 IEEE International Conference on Smart Computing, pp. 470–475
15. Manfredini, G., et al.: An ASIC-based miniaturized system for online multi-measure and monitoring of lithium-ion batteries. *Batteries* **7**(3), 1–13 (2021)
16. Del Cesta, S., Ria, A., Piotto, M., Simmarano, R., Bruschi, P.: A compact current-mode instrumentation amplifier for general-purpose sensor interfaces. *AEU-Int. J. Electron. C.* **92**, 8–14 (2018)
17. Del Cesta, S., Ria, A., Simmarano, R., Piotto, M., Bruschi, P.: A compact programmable differential voltage reference with unbuffered 4 mA output current capability and $\pm 0.4\%$ untrimmed spread. In: 43rd IEEE European Solid State Circuits Conference, pp. 11–14 (2017)



DoS Detection on In-Vehicle Networks: Evaluation on an Experimental Embedded System Platform

Miltos D. Grammatikakis^(✉), Nikos Mouzakitis^(✉), Lefteris Kypraios^(✉),
and Nikos Papatheodorou^(✉)

Hellenic Mediterranean University, 71410 Heraklion, Greece

{mdgramma, nmouzakitis, lkypraios, ntpapatheodorou}@cs.hmu.gr

Abstract. Modern vehicles involve engine control units that communicate over multiple in-vehicle networks via a traditional Gateway. In this context, we develop an open, experimental distributed embedded platform that integrates CAN networks of different criticalities, populated with Raspberry Pi 3 nodes and an Odroid XU3 device that acts as the Gateway. During normal operation, a critical CAN (CAN2) emulates engine traffic (i.e., Korean car dataset). In contrast, a non-critical CAN (CAN1) sends packet requests related to the dashboard display (e.g., engine speed, RPM, temperature, airflow, etc.). Responses to these packets are forwarded back to CAN1, forming a request-response path. In our DoS attack scenario, a malicious CAN1 node broadcasts packet requests that are relayed by the Gateway towards CAN2. At the Gateway-level, we detect a DoS attack by monitoring perturbations of system metrics (Cortex-A15 power consumption, temperature gradients, and packet ID frequency) from pre-established thresholds using a sliding window-based cumulative sum approach. Also, we monitor variations of inter-arrival time in the request-response path at the periphery (CAN1). Our results on the experimental automotive platform indicate that frequency count at the Gateway and inter-arrival time at the network periphery are promising techniques for fast and accurate DoS detection using CUSUM. Furthermore, preliminary experimental results indicate that CUSUM is a more precise metric than entropy for detecting DoS.

1 Introduction

In modern vehicles, Engine Control Units (ECUs) exchange packets over multiple in-vehicle networks to monitor and control subsystems of different criticalities, such as throttle, ABS, lights control, window roll-up/down, dashboard display, and infotainment. CAN (Controller Area Network) is the most popular in-vehicle network for almost all vehicle models. Although the CAN data frame can be regulated, very few manufacturers currently provide detection and/or protection to security vulnerabilities. Hence, it is feasible to inject malware, e.g. over a physical CAN connection, to enable surveillance, or perform spoofing, replay, modification, man-in-the-middle, or denial-of-service (DoS) attacks [1–3].

In this paper, we focus on DoS detection when a non-critical node broadcasts packets that are relayed by a Gateway node to a critical CAN. By monitoring affected system metrics (e.g., packet frequency and latency, temperature, or energy), we can formulate this as a change-point detection problem. Hence, to evaluate the time interval when abrupt, irregular changes occur, we apply a sliding window cumulative sum algorithm (CUSUM) to compute an upper control limit (UCL) that represents deviation among expected and observed values. A DoS attack is signaled if UCL exceeds a pre-established threshold which represents the maximum statistical variance of the expected value for attack-free conditions. A lower control limit can also be used to detect a downward change, i.e., DoS recovery.

To evaluate our proposed methodology, we develop an open, distributed embedded platform prototype that integrates multiple CAN networks. Within this platform, we apply CUSUM towards DoS detection.

- at the Gateway, by monitoring perturbations of different system metrics: a) Cortex-A15 power consumption, b) temperature gradients and c) packet header (ID) frequencies, and
- at the non-critical CAN, by tracing round-trip time variations of inter-arrival packet delays. These variations refer to transactions in the request-response path, i.e., from the non-critical CAN to the critical CAN and back.

Results indicate that a) packet ID frequency at the Gateway and b) inter-arrival time at the network periphery are fast and accurate DoS detection techniques. It is also interesting that for specific ranges of malicious traffic, our proposed metrics can help estimate the relative extent of a DoS attack, and thus develop the most appropriate measures to contain the attack.

Different methods for detecting abnormal vehicle behavior have been proposed for DoS, fuzzy (random), or impersonation attacks [4]. A frequency-based anomaly detection system can identify DoS attacks on CAN bus by comparing the rate and time sequence of periodic data packets for each CAN ID versus a pre-established baseline [5–7]. Other techniques focus on detecting extraneous packets using correlation of the time intervals of response packets to remote frames [8], fingerprinting clock skews or voltage characteristics of different ECUs [9–11], machine learning (e.g., tree-based decision) [12], neural networks based on training [13, 14], Markov models [15], Petri Nets [16], transition matrices [17], hamming distance [18], cumulative sum [19], and entropy-based techniques [20, 21]. Alternatively, signal- and specification-based IDS validates predefined manufacturer specifications, e.g., data values in packet payloads, range, or inter-relation of packet IDs [22, 23]. Olufowobi et al. proposed a specification-based IDS using anomaly-based supervised learning algorithms that capture message timing and worst-case response time [24]. Analysis based on CAN logs indicates high accuracy and low false-positive rates compared to interval- and frequency-based detection.

In comparison with previous research, our work focuses on protecting a critical bus from non-critical access via a Gateway. To the best of our knowledge, our approach is the first one to consider DoS detection using power consumption and temperature metrics (computed at the Gateway) and inter-arrival time (estimated at the periphery). Moreover, unlike past research, our experimental framework does not focus on computation models

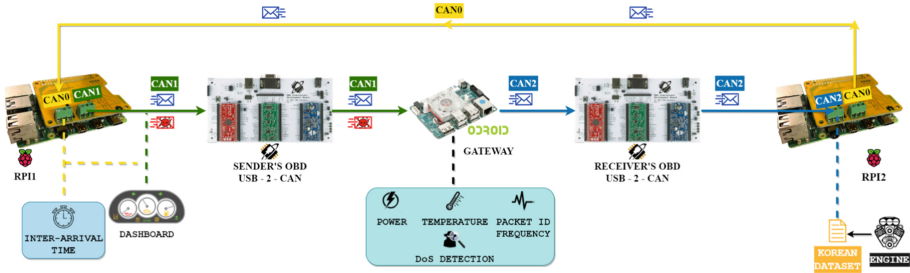


Fig. 1. Experimental framework of our open, distributed embedded platform

or spreadsheets analyzing pre-computed CAN logs. In contrast, as the first step towards application to real automotive systems, we use an open distributed embedded platform that carries actual automotive traffic in a critical subsystem, while non-critical subsystems communicate with it via a Gateway. This method allows us to compute detection time and accuracy at non-critical components outside the critical network, either at the periphery or at the Gateway. In contrast, when using models or spreadsheets, the implied attacker is already in the critical system. Hence, detection, in this case, must be deployed at each ECU node, causing unnecessary, possibly critical delays. Finally, as shown later, CUSUM supports fast and accurate detection (with almost zero false-positive and false-negative rates). Also, unlike popular neural network techniques, it provides a way to validate change-point phenomena in safety-critical situations. Finally, our experimental results indicate that CUSUM is a more precise metric than entropy for detecting DoS [20, 21].

Next, Sect. 2 details the experimental framework, including use case, configuration, detection metrics, and results. Section 3 discusses future work.

2 Experimental Framework: HW & SW

2.1 Platform Prototype, Use-Case & Configuration

As shown in Fig. 1, our distributed embedded platform interconnects multiple Raspberry Pi 3 nodes, (RPI1, RPI2) to an Odroid XU3 Gateway.

- Each Raspberry supports up to four independent CAN interfaces using the Linux `can-utils` software stack [25] and Industrialberry’s Canberry Dual v2.1 shield and drivers [26].
- Odroid XU3 uses Scantool OBD dev kits to support multiple (incoming or outgoing) UART-to-CAN interfaces [27]. CAN STN2120 microcontroller supporting ELM327 AT/ST commands [27]. For the software stack (~3K lines of code), we have extended a serial minicom terminal [28] to support POSIX threads to control incoming and outgoing Rx/Tx traffic. The operation follows a concurrent producer-consumer pattern, i.e., a receiver thread fills, while a sender empties a bounded buffer. The buffer size is selected to hold 100 CAN packets (header and data). Furthermore, using an I2C driver, we access internal CPU power/thermal metrics within our Gateway application via

the `sysfs` file system. This gives us the option to visualize using a Qt-based Energy Monitor; both metrics rise as more packets enter the Gateway.

Our use-case focuses on comparing the effectiveness and accuracy of detection when DoS attacks occur from the non-critical CAN1 towards the Gateway and critical CAN2. Therefore, it is critical to examine the operation of three components: RPI1 (dashboard sender), RPI2 (engine receiver), and Gateway (implementing DoS detection).

RPI Sender (RPI1 on non-critical CAN1). For DoS detection, the injection rate of malicious packets must be “sufficiently distinct” from the rate of legitimate ones. Hence, RPI1 (dashboard sender) uses a POSIX thread to transmit packet requests towards the engine subsystem. These requests travel via the Gateway (see red/blue packets in Fig. 1). By modifying an integer parameter, we can increase the number of malicious (red) packets transmitted per second. The number of legitimate (blue) packets always flows at a constant rate of 2 packets/sec, while the baud rate is set at 500Kb/sec. Since the injection rate of individual CAN packets can be controlled independently, we can use this rate without losing generality. At the same time, another thread on RPI1 listens for response packets on CAN0, with timing enabled using the Linux `can-utils` command: `can-dumpcan0 -t`. It enables computing the inter-arrival time on the request-response (round trip) path. Finally, a third thread detects DoS, see next Session on “DoS Detection at Gateway & Periphery”.

RPI Receiver (RPI2 on Critical CAN2). During normal operation, actual engine traffic is injected into CAN2, using an attack-free open Korean Hyundai YF Sonata engine dataset [14, 29]. This traffic is injected into CAN2 either via `canplay` or using a Viewtool Gingko UART-CAN [30] on an independent CAN (High/Low) interface on the OBD dev kit. The exact timing log of 988,987 packets makes this emulation very realistic. Notice that CAN2 operates at a fixed 500Kb/s baud rate. Simultaneously, RPI2 receives packet requests departing from Raspberry Pi device RPI1 that arrive via the Gateway. These requests are related to dashboard display meters (e.g., engine speed, RPM, temperature, airflow, etc.) and are answered back to RPI1 (on CAN0), forming a request-response path. Notice that in case of attack, packets along the request-response path (from RPI1 to RPI2 and back) are delayed not only due to engine traffic (Korean engine dataset) that has higher priority but also due to malicious packets arriving on the Gateway. Hence, we can examine DoS detection at the periphery, i.e., the RPI1 node in the non-critical CAN, by measuring the inter-arrival time in the request-response (round-trip) path.

DoS Detection at Gateway & Periphery. The Gateway, besides receiving and transmitting packets, evaluates three different DoS detection metrics: a) total Cortex-A15 power consumption, b) temperature gradient of each thermal zone associated with each Cortex-A15 tile, and c) packet ID frequency count. For all detection metrics, we use CUSUM with a sliding window of 10 packets; this is a compromise between the numbers of false negatives and false positives in detection and between the detector’s resolution and reliability. This improves sensitivity of our approach and allows us to distinguish attacks of smaller extent (i.e., far below 64 packets/sec). Upon DoS detection, we use

very fast (20nsec max) sleep/wake-up mechanisms of the OBD dev kit power module to throttle, sleep, or shut down traffic, alleviating the Gateway.

2.2 Results - DoS Detection at the Periphery

DoS detection metrics at the Gateway may be temporarily or permanently disabled. Then, to detect DoS, we monitor the inter-arrival delay of periodic CAN packets that originate from RPI1 (a.k.a. dashboard), flow to RPI2 (engine) via Odroid XU3 (Gateway) and return to RPI1. These “engine monitoring requests” are important in recently proposed automotive intelligent dashboard systems.

Figure 2 shows the inter-arrival time of packets arriving at RPI1 versus the injection rate of malicious packets. The latency is proportional to the injection rate of malicious packets. This indicates that by monitoring RTT, we can detect a DoS attack at the periphery (non-secure CAN). Finally, due to the almost linear relationship ($AnovaR^2 > 0.999$), RTT can provide direct insight into the relative extent of a DoS attack. This can help implement appropriate policies to stop the attack at its source.

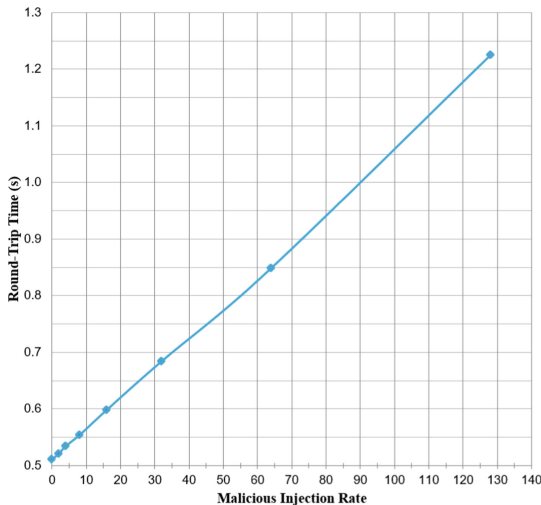


Fig. 2. Inter-arrival time (RPI1 to RPI2 and back) vs malicious injection rate

In Fig. 3, we show how to apply the proposed CUSUM detection metric to inter-arrival time. The allowable slack (CUSUMK parameter), which refers to the size of shift desired to detect, is set to one standard deviation [31–34]. The graph examines a “small” transition from M0 state (only legitimate packets) to M1 state, i.e. one malicious packet/sec. The CUSUM upper limit indicator allows DoS detection after only four samples or ~ 1.2 s; CUSUM calculation delay is insignificant since equations are rewritten to avoid re-computation. This upward change in statistical behavior is considered non-random as it exceeds the CUSUM Threshold (H) which is set to 5 times the standard deviation of our samples; this gives very high accuracy for DoS detection (99.99994%).

Notice, a tradeoff between detection accuracy and delay. For example, when the threshold is set to 3 times the standard deviation (99.973% confidence), the delay is reduced to 0.54 s. By comparison, this delay extends braking distance by 15 m on a car speeding at 100 km/h. In Fig. 3, we also compute the entropy metric [20, 21]. Notice from the graph that it is much harder to detect DoS using entropy metric, since there are several spurious peaks that appear the change-point.

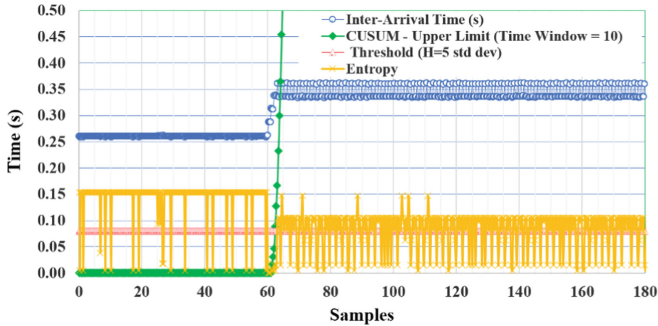


Fig. 3. CUSUM metric for predicting DoS attack using inter-arrival time.

2.3 DoS Detection at the Gateway

We now examine different system metrics that can be used to detect a DoS attack at the Gateway. Since our code (process and threads) on the Gateway runs solely on Cortex-A15,

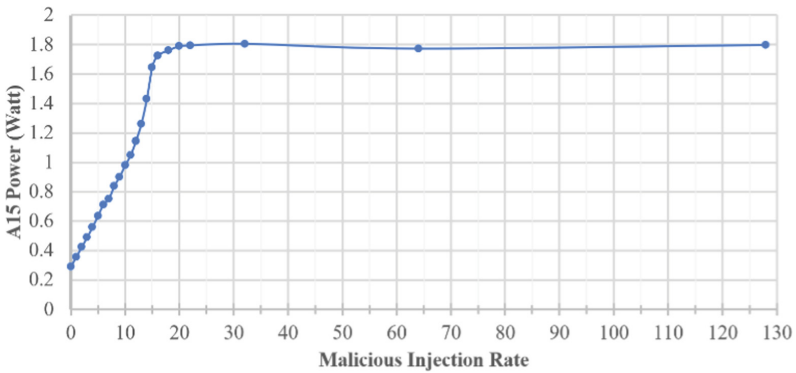


Fig. 4. Instant cortex-A15 power consumption vs. malicious injection rate

Figure 4 shows the total consumed power by Cortex-A15 tiles (averaged over multiple runs). Average power consumption on Cortex-A15 sharply increases between the rates of 1 to 16 malicious packets/sec. Then, it levels off, due to packet drops, since the

injection rate becomes too fast for the UART-to-CAN interface to handle. Hence, the power consumption metric can be used for detecting denial of service in this region, provided that a threshold value within this range of malicious injection rates is given. This threshold must be below the power saturation point which defines a system limit before packets arrive too fast to be accepted by the UART-to-CAN interface.

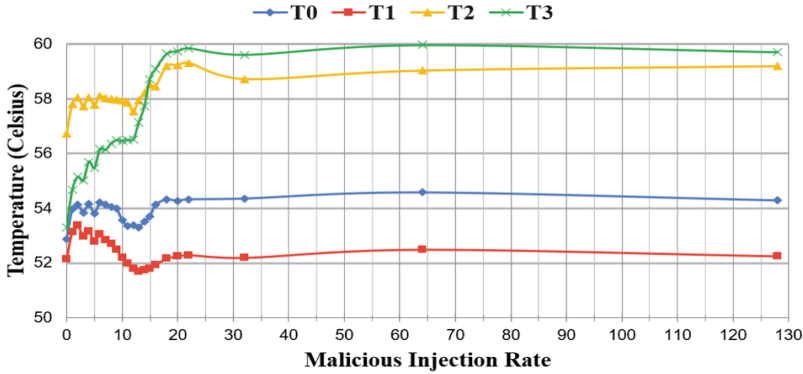


Fig. 5. Average temperature for each thermal zone vs malicious injection rate

In Fig. 5, we examine the average temperature of each of the four thermal zones that correspond to each Cortex-A15 tile versus the malicious injection rate. We monitor an abrupt increase in thermal zones 2 and 3, in the range of 1 to 18 malicious packets/sec. These zones correspond to the Cortex-A15 cores on which Gateway software (single producer and single consumer threads) is assigned to be executed. Similar to the power metric (Fig. 4), there are limits in the malicious rates that can be detected, due to saturation. However, there are also certain local minima in a plateau that exhibit irregularities. Hence, this metric is not so reliable in detecting denial of service and might only be used in conjunction with other metrics to improve detection.

In Fig. 6, we show the packet ID frequency count versus the malicious injection rate. Since the frequency count increases somewhat beyond the power consumption saturation rate, we have a slightly larger range of DoS detection. Furthermore, by examining the value of the frequency count, we can derive knowledge on the DoS attack size.

2.4 Prediction Accuracy of Gateway Metrics

Besides detecting the time of change in near real-time, we also consider statistical metrics related to the accuracy of detection due to modeling errors or operating conditions, e.g. Linux kernel bursts. For the false-positive case, we can consider a) the probability of false alarms, and b) the mean time between false alarms. Similarly, for the false-negative case, we can consider a) the probability of non-detection, and b) the mean time between non-detection events. In this context, we next compare the relative accuracy of our two prime DoS detection metrics at the Gateway, namely Cortex-A15 power consumption and packet ID frequency count. Assuming a time window of 10 samples, false-positive and false-negative outcomes are indicated by the overlapping curves in Figs. 7 and 8.

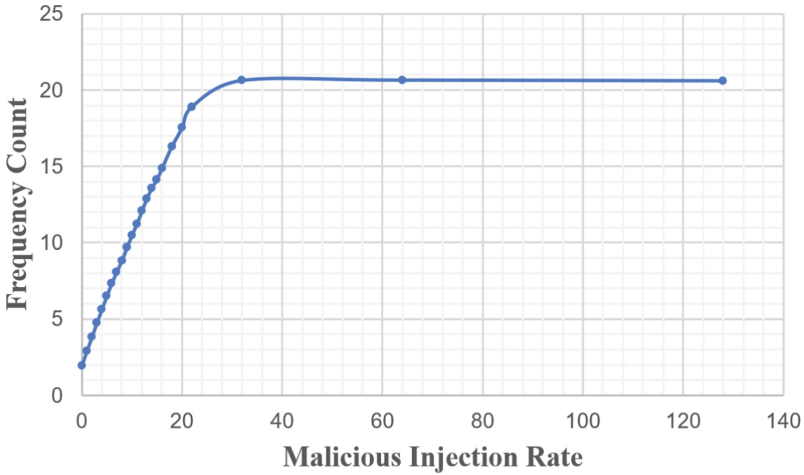


Fig. 6. Packet ID frequency count vs malicious injection rate

In Fig. 7, we consider the range below saturation for power consumption, i.e., below 16 malicious packets/sec. For the given 450-s interval, false-positive and false-negative detection rates are zero for a transition from M0 to M4 (i.e., from 0 to 4 malicious packets/sec), or M4 to M8. The rates are very low ($\sim 2\%$) when comparing M8 with M12. Due to power saturation (see Fig. 4), it is impossible to distinguish the extent of an attack above M16 (16 malicious packets/s); in fact, consumed power for the M16 case is very similar to that of M128.

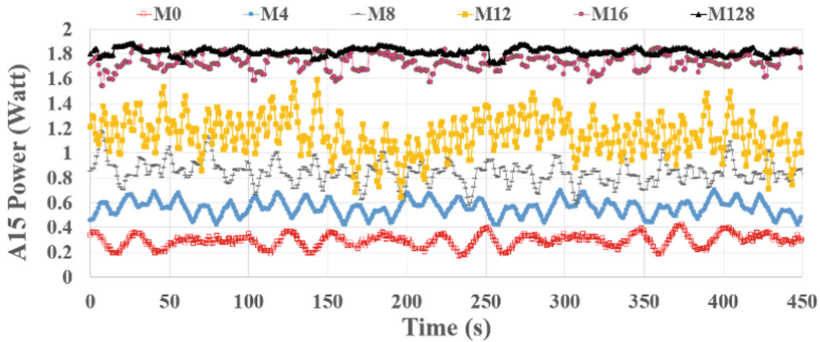


Fig. 7. DoS detection accuracy: power consumption vs. Time (for different malicious packet injection rates: M0 to M128)

In contrast, as shown in Fig. 8, the packet ID Frequency Count has a larger operating region, i.e., $[M0, M24]$ compared to $[M0, M16]$ for the power metric. It also offers a very good detection probability without any false positives or false negatives for all malicious rates below M24.

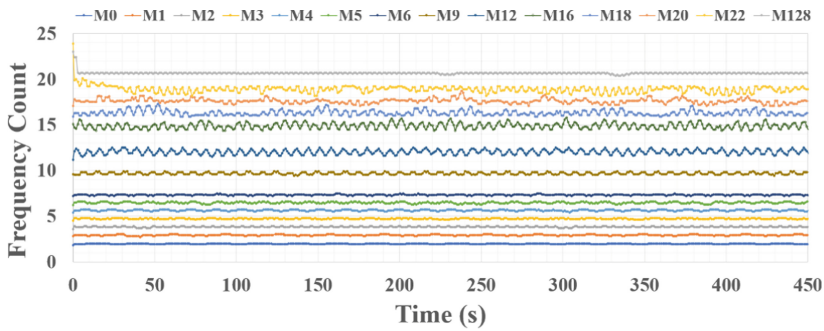


Fig. 8. DoS detection accuracy: packet ID frequency count vs. Time (for different malicious injection rates: M0 to M128)

3 Future Work

We have developed an open-source experimental automotive platform using ARM boards. The platform integrates CAN networks of different criticalities, populated with Raspberry Pi 3 nodes and an Odroid XU3 device that acts as the Gateway. During normal operation, a critical CAN (CAN2) emulates engine traffic (i.e., Korean car dataset). We detect a DoS attack by monitoring perturbations of system metrics (Cortex-A15 power consumption, temperature gradients, and packet ID frequency) from pre-established thresholds using a sliding window-based cumulative sum or entropy-based method. Our results indicate that frequency count at the Gateway and inter-arrival time at the network periphery are promising techniques for fast and accurate DoS detection using CUSUM. Furthermore, preliminary experimental results indicate that CUSUM is a more precise metric than entropy for detecting DoS.

We plan to extend our approach towards detecting spoofing and replay attacks on real vehicles. In this respect, we focus on lightweight cryptography and examine statistics sensitive to mean, variance, and composite statistics. It is also interesting to pursue the design of a universal CUSUM controller for DoS detection, experiment with multiple CAN interfaces, and develop scalable implementations of our Gateway packet transport mechanism using concurrent lock-free queues. The code, demo, and troubleshooting guide of our experimental platform will be available in SourceForge/GitHub (Q1/22).

Acknowledgment. The authors acknowledge support from EU Horizon 2020 project AVANGARD (Contract No. 869986).

References

1. Biron, Z.A., Dey, S., Pisu, P.: Real-time detection and estimation of denial of service attack in connected vehicle systems. *IEEE Trans. Intell. Transp. Syst.* **19**(12), 3893–3902 (2018)
2. Woo, S., Jo, H.J., Lee, D.H.: A practical wireless attack on the connected car and security protocol for in-vehicle CAN. *IEEE Trans. Intell. Transp. Systems* **16**(2), 993–1006 (2015)
3. Miller, C., Valasek, C.: Remote exploitation of an unaltered passenger vehicle. In: *BlackHat Conference 2015*, 19 June 2021. http://illmatics.com/car_hacking.pdf

4. Wu, W., et al.: A survey of intrusion detection for in-vehicle networks. *IEEE Trans. Intell. Transp. Syst.* **21**(3), 919–933 (2020)
5. Hoppe, T., Kiltz, S., Dittmann, J.: Applying intrusion detection to automotive IT: early insights and challenges. *J. Info Assur. Secur.* **4**(3), 226–235 (2009)
6. Song, H.M., Kim, H.R., Kim, H.K.: Intrusion detection system based on the analysis of time intervals of can messages for an in-vehicle network. In: *Proceedings of International Conference on Information Networks*, pp. 63–68 (2016)
7. Young, C., et al.: Automotive intrusion detection based on constant CAN message frequencies across vehicle driving modes. In: *Proceedings of ACM Workshop on Automotive Cybersecurity*, pp. 9–14 (2019)
8. Lee, H., Jeong, S.H., Kim, H.K.: OTIDS: a novel IDS for an in-vehicle network by using remote frame. In: *Proceedings of Conference on Privacy, Security and Trust*, pp. 5709–5757 (2017)
9. Halder, S., Conti, M., Das, S.K.: COIDS: a clock offset based intrusion detection system for controller area networks. In: *Proceedings Distributed Computing and Networking*, pp. 1–10 (2020)
10. Cho, K.-T., Shin, K.G.: Fingerprinting electronic control units for vehicle intrusion detection. In: *Proceedings of USENIX Security Symposium*, pp. 911–927 (2016)
11. Choi, W., et al.: Voltageids: low-level communication characteristics for automotive IDS. *IEEE Trans. Inf. Forensics Secur.* **13**(8), 2114–2129 (2018)
12. Weber, M., et al.: Embedded hybrid anomaly detection for automotive CAN communication. In: *Proceedings of Embedded Real-Time Software and System Congress* (2018)
13. Wasicek, A.R., et al.: Context-aware intrusion detection in automotive control system. In: *Proceedings of Embedded Security in Cars Conference* (2017)
14. Seo, E., Song, H.M., Kim, H.K.: GIDS: GAN based intrusion detection system for in-vehicle network. In: *Proceedings of Conference on Privacy, Security and Trust*, pp. 1–6 (2018)
15. Vasistha, D.K.: Detecting anomalies in controller area network for automobiles. MSc. thesis, Department Computer Engineering, University of Texas A&M (2017)
16. Rieke, R., et al.: Behavior analysis for safety and security in automotive systems. In: *Proceedings of Conference on Parallel, Distributed and Network Proceedings*, pp. 381–385 (2017)
17. Marchetti, M., Stabili, D.: Anomaly detection of CAN bus messages through analysis of ID sequences. In: *IEEE Intelligent Vehicles Symposium* (2017)
18. Stabili, D., Marchetti, M., Colajanni, M.: Detecting attacks to internal vehicle networks through hamming distance. In: *Proceedings of AEIT International Conference*, pp. 1–6, September 2017
19. Olufowobi, H., et al.: Anomaly detection approach using adaptive cumulative sum algorithm for controller area network. In: *Proceedings of ACM Workshop on Automotive Cybersecurity*, pp. 5–10 (2019)
20. Müter, M., Asaj, N.: Entropy-based anomaly detection for in-vehicle networks. In: *Proceedings of IEEE Intelligence on Vehicles Symposiu*, pp. 1110–1115 (2011)
21. Wu, W., et al.: Sliding window optimized information entropy analysis method for intrusion detection on in-vehicle networks. *IEEE Access* **6**, 45233–45245 (2018)
22. Larson, U.E., Nilsson, D.K., Jonsson, E.: An approach to specification-based attack detection for in-vehicle networks. In: *Proceedings of IEEE Intelligent Vehicles Symposium*, pp. 220–225 (2008)
23. Taylor, A., Japkowicz, N., Leblanc, S.: Frequency-based anomaly detection for CAN bus. In: *Proceedings of World Congress on Industrial Control Systems Security*, pp. 45–49 (2015)
24. Olufowobi, H., et al.: SAIDuCANT: specification-based automotive intrusion detection using controller area network timing. *IEEE Trans. Veh. Technol.* **69**(2), 1484–1494 (2020)

25. Kleine-Budde, M.: SocketCAN: the official CAN API of the Linux kernel. In: Proceedings of CAN Conference, pp. 5–17 (2012)
26. CanberryDual, Industrialberry, 19 June 2021. <http://www.industrialberry.com/canberry-v-2-1-isolated>
27. OBD dev kit, Scantool, 19 June 2021. <https://www.scantool.net/obd-development-kit>
28. Serial Console, Sourceforce, 19 June 2021. <https://sourceforge.net/projects/serialconsole>
29. Ginkgo UART-CAN, Viewtool, 19 June 2021. <http://www.viewtool.com/index.php/en/27-2016-07-29-02-13-53/44-ginkgo-usb-can-9>
30. Car-Hacking Dataset, HCRL, 19 June 2021. <https://sites.google.com/a/hksecurity.net/ocslab/Datasets>
31. CUSUM Charts, 19 June 2020. <https://www.spcforexcel.com/knowledge/variable-control-charts>
32. Page, E.S.: Continuous inspection schemes. *Biometrika* **41**(1), 100–115 (1954)
33. Poor, H.V., Hadjiliadis, O.: Quickest Detection. Cambridge University Press, Cambridge (2009)
34. Basseville, M., Nikiforov, I.V.: Detection of Abrupt Changes: Theory and Applications. Prentice Hall, Hoboken (1993)



Simulation Environment for Mixed AHB-NoC Architectures

Massimo Conti^(✉)

Department of Information Engineering, Università Politecnica Delle Marche,
via Breccie Bianche 12, 60124 Ancona, Italy
m.conti@univpm.it

Abstract. Communication architecture is crucial for performances and power constraints in modern multicore systems on chip (SoC). Network-on-Chip (NoC) is used to increase the bandwidth limitations of a traditional bus paradigm. In this work an interface between the AMBA-AHB bus and NoC has been designed. A NoC simulation environment has been modified integrating the AHB architectures in different NoC architectures. The performances of different heterogeneous (AHB-NoC) architectures have been compared.

1 Introduction

Complexity of SoC has recently increased significantly in terms of the number of processor cores and communication among the cores. Communication is often the bottleneck in modern multicore systems on chip. The traditional bus communication architecture has an intrinsic limit on bandwidth. The Network-on-Chip tries to overcome this limit. Furthermore, scalability and reusability of electronic design motivated the growth of the NoC paradigm.

A NoC consists of three types of modules: routers, links and interfaces. The messages are sent from an Intellectual Property (IP) source to a router and forwarded by this to other routers until they will arrive to the router connected to the destination IP. Routers are connected with each other by links forming a net of chosen topology, size and connection degree.

NoC architecture has many degrees of freedom. Topology for regular network can be chosen between a wide variety: mesh and torus, hypercubes, spidergon, hexagonal, binary tree, butterfly and Benes networks. Different architectures have been proposed by industry or academia: Spidergon STNoC, Mango, Aethereal, Arteris, Sonics, SoCbus and xPipes.

The design of the different parts of a NoC is a complex task [1–4]. A low power and low complexity link microarchitecture for NoC is presented in [1]. A configurable architecture for NoC routers, compliant with several NoC topologies is reported in [2]. Paper [3] presents the design of a Network Interface for on-chip communication infrastructure capable of advanced networking functionalities: store & forward transmission, error management, power management. The optimization of router buffers has a directly impact network throughput, and they dominate the on-chip router area [4].

Many works have focused on the design of NoC architectures, while the interfacing between NoC and IPs and the interface among NoC and Bus also needs to be considered. A widely used bus architecture is the Advanced Microcontroller Bus Architecture (AMBA) with different versions and protocols. The AMBA AXI-Stream protocol is a point-to-point protocol, connecting a single transmitter and a single receiver. In [5] an adapter from the AMBA-AXI to NoC architectures is designed at low level. Similarly, a network interface architecture to make the AMBA-AXI4 protocol compatible with a NoC is proposed in [6].

System-level design and transaction level modeling are the keys to fast SoC innovation with the capability to quickly try out different design alternatives, to confirm the best possible architecture in terms of speed and power dissipation [7–9].

A SystemC class library, named NOCEXPLORE, and relative tools for comparing different Network-on-Chip (NoC) architectures and investigating communication bottlenecks has been presented in [10] and free available on sourceforge.net. Different NoC architectures are compared by statistical analysis of packet delays and power dissipation in [11].

A system level design environment for the simulation of a hybrid NoC-Bus is not available up to now in the literature. In this work the NOCEXPLORE simulation environment has been modified integrating classical bus architectures into different NoC architectures. The simulation environment is able to compare the performances of different NoC, Bus, and hybrid NoC-Bus architectures. In this work a NoC-AMBA AHB interface has been developed and described at system level. A test example with different hybrid NoC-Bus architectures is developed. The performances have been compared.

2 AHB-NoC Architecture

The NOCEXPLORE simulation environment allows the simulation of different NoC architectures that can be defined using the a great number of parameters: quality of service, network size, topology, link type, link width, link delay, number of physical link per topological arc, flit per packet, packet per message, routing algorithm, arbitration scheme, switch structure, the Dynamic Power Management policy of each router, flow control, input port buffer length, input port buffer number, output port buffer length, output port buffer number.

A new SystemC library for the simulation of mixed NoC-AMBA AHB architecture has been integrated in NOCEXPLORE. Goal of the library is to allow a fast comparison between different mixed NoC-Bus architectures.

The interfaces between AMBA AHB bus and routers of the NoC has been developed. The new library allows to instantiate in the NoC the following additional blocks:

- AHB master, as traffic generators with configurable traffic;
- AHB slaves;
- AHB bus (arbiter, decoder, multiplexers);
- NoC to Slave AHB interface;
- NoC to Master AHB interface;
- NoC to AHB Bridge.

The interface behaves as a virtual slave or master in the AHB bus to which is connected. The interface is used to allow the masters, connected to the bus, to address all the other slaves or IPs connected to other parts of the NoC.

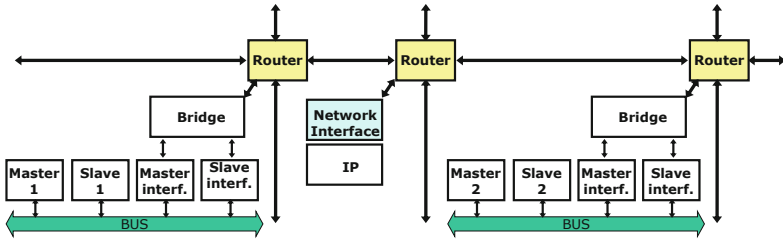


Fig. 1. Example of possible AMBA-NoC configurations.

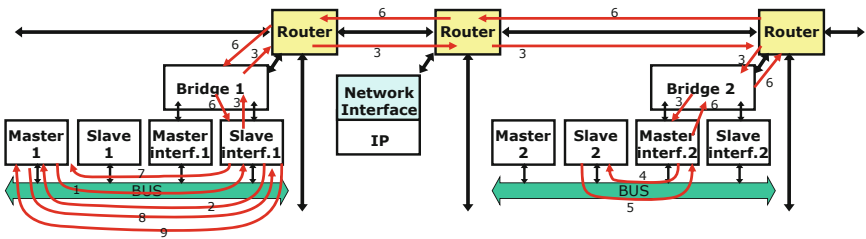


Fig. 2. Example of data transfer in the AMBA-NoC architecture.

The AHB bus is a synchronous system and the NoC is asynchronous, the interface is therefore a sync/async interface with buffers to store data. An example of possible interconnections is shown in Fig. 1.

The bridge acts as a multiplexer. The data coming from the router are sent to the NoC-Slave interface, if they are generated by an AHB master located in another part of the NoC. Conversely, the data are sent to the NoC-Master interface, if they are generated by an AHB slave located in another part of the NoC. Master1 and Master2 can address Slave1 and Slave2 directly or through the interfaces.

The sequence of operations during a read transfer from Master 1 to Slave 2 is the following, reported in Fig. 2.

- 1) Master 1 sends a read transfer request to Slave 2 through Slave interface 1.
- 2) Slave interface 1 stores the information, responds with “split” to keep the bus free, creates the packet for the NoC.
- 3) Slave interface 1 starts the transfer to Master interface 2 through the routers.
- 4) Master interface 2 converts the packet to the sequence of actions required by the AMBA AHB bus, and sends, as a master, a read transfer request to Slave 2.
- 5) Slave 2 responds sending the requested data.
- 6) Master interface 2 stores the information, creates the packet for the NoC and starts the transfer to Slave interface 1 through the routers.

- 7) Slave interface 1 deletes the “split” signal
- 8) Master 1 is able to ask for the data again
- 9) Slave interface 1 sends the data.

The delay time in a NoC is not deterministic depending on the NoC traffic, the number of links between source to destination, router buffers length and routing algorithms. Therefore, the AMBA AHB split option has been used to allow the Masters connected in same bus (only Master 1 and Slave 1 in this example) to use the bus during the data transfer request to a Slave (Slave 2) not directly connected to the bus.

3 Simulation Results

In the SystemC simulations performed the masters and IPs are random traffic generators sending data to the all the other IPs with variable packet length and traffic intensity. Figure 3 shows the four different AHB, NoC and heterogeneous AMBA-NoC architectures that have been simulated in the SystemC environment developed.

- (a) AMBA AHB bus with 2 Masters and 2 Slaves;
- (b) AMBA AHB bus with 4 Masters and 2 Slaves;
- (c) 4×4 Mesh NoC;
- (d) 4×4 Mesh NoC + 2 AMBA AHB BUS.

Three arbitration algorithms have been tested for the AMBA AHB bus:

- Priority: the masters have a fixed priority.
- Priority with waiting control: the masters have a fixed priority. When a master waits for a time longer than a fixed value, his priority becomes the highest.
- Shortest Job First: the shortest bus occupation request takes the bus.

Figures 4 and 5 report the results of the simulations. Figure 4 compares the normalized throughput for each IP of the different architectures as a function of the normalized traffic injected by the IPs of the network. The dots report the simulation results, continuous lines represent the ideal throughput. Different color are used for the different architectures. As an example, in case of 4 masters in the AHB, the bus saturation is reached when each IP generated 25% of its maximum traffic.

The throughput results of the simulation are lightly lower than the maximum ideal case, as expected. The arbitration algorithm has a limited influence on the throughput performances. In the case of a 4×4 mesh NoC, when the injected traffic increases, the throughput depends on the type of traffic (long packet or short packet). Long packets increase the probability of collision in the routers and reduces the performances.

Figure 5 reports the packet delay of the different architectures as a function of the normalized traffic injected by the IPs of the network. The dots report the simulation results, continuous lines represent interpolating curves. The delay of the bus is limited when the number of IPs is low. The average delay in a 4×4 NoC is higher and even worst in case of network congestion, since on average the packets must go across many routers and links going from source to destination.

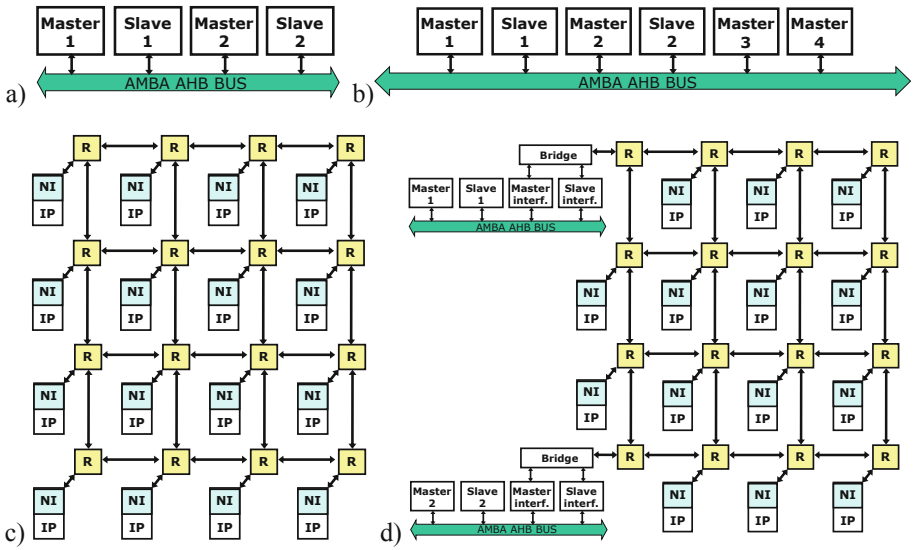


Fig. 3. Four different AHB, NoC and heterogeneous AMBA-NoC architectures simulated. (R: Router, NI: Network Interface)

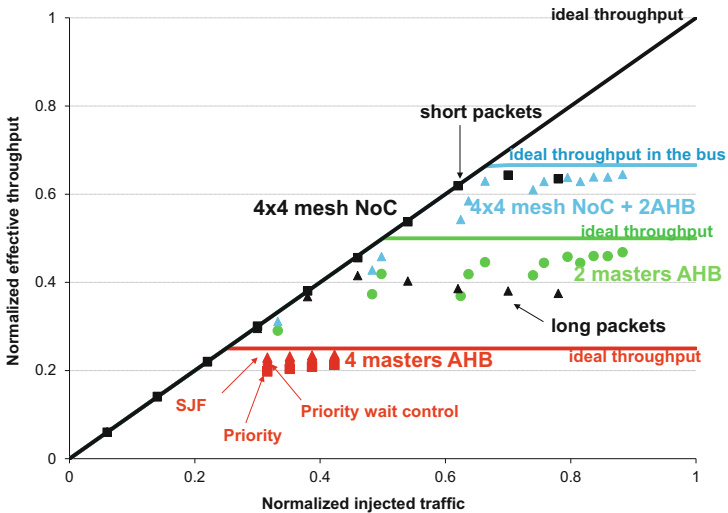


Fig. 4. Normalized throughput for each IP of the different architectures as a function of the normalized traffic injected by the IPs.

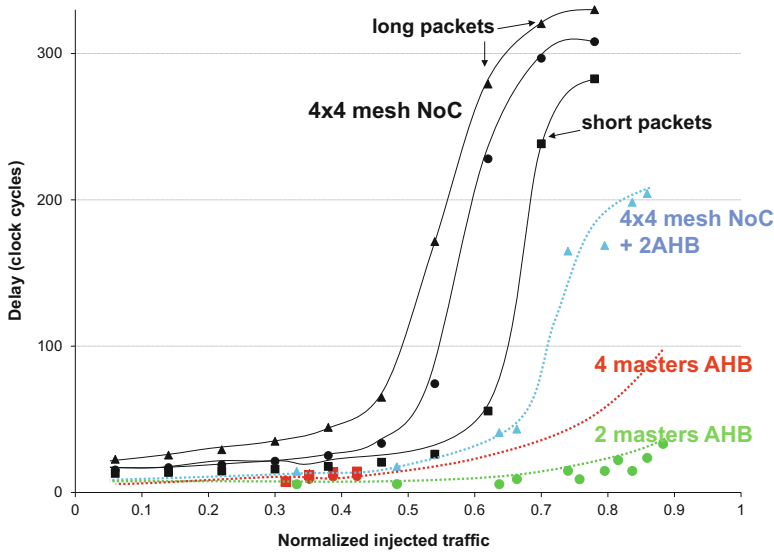


Fig. 5. Average packet delay of the different architectures as a function of the normalized traffic injected by the IPs.

4 Conclusions

The focus of this work is the integration of bus architecture on a system level NoC simulator developed in SystemC. The simulation environment allows to quickly compare the different hybrid NoC-Bus architectures. Furthermore it may be used to optimize different NoC parameters (e.g. network size, topology, routing algorithm, arbitration scheme), different hybrid NoC-Bus architectures, different NoC-Bus interfaces, the placement of the IPs in the hybrid NoC-Bus architecture.

In this work a specific NoC-AMBA AHB interface has been developed. A test example with different hybrid NoC-Bus architectures is developed.

The preliminary results reported allow to compare the performances of the different architectures. In the example reported the IP are simply random generators and each IP sends the packet to all the other IPs with equal probability. The simulation environment could be used to optimize the architecture for a real traffic in a specific application.

References

1. Vitullo, F., et al.: Low-complexity link microarchitecture for mesochronous communication in networks-on-chip. *IEEE Trans. Comput.* **57**(9), 1196–1201 (2008). <https://doi.org/10.1109/TC.2008.48>
2. Saponara, S., Fanucci, L.: Configurable network-on-chip router microcells. *Microprocess. Microsyst.* **45**, 141–150 (2016)
3. Saponara, S., et al.: Design of an NoC interface macrocell with hardware support of advanced networking functionalities. *IEEE Trans. Comput.* **63**(3), 609–621 (2014)

4. Jindal, N., Gupta, S., Ravipati, D.P., Panda, P.R., Sarangi, S.R.: Enhancing network-on-chip performance by reusing trace buffers. *IEEE Trans. Comput. Aided Des. Integr. Circuits Syst.* **29**(4), 922–935 (2020)
5. Tran, X.-T., Nguyen, T., Phan, H.-P., Bui, D.-H.: AXI-NoC: high-performance adaptation unit for ARM processors in network-on-chip architectures. *IEICE. Trans. Fundam. Electron. Commun. Comput. Sci.* **100**, 1650–1660 (2017)
6. Wang, B., Lu, Z.: Efficient support of AXI4 transaction ordering requirements in many-core architecture. *IEEE Access* **8**, 182663–1826678 (2020)
7. Vece, G.B., Conti, M.: Power estimation in embedded systems within a system C-based design context: the PKtool environment. In: *Proceedings of the 7th Workshop on Intelligent Solutions in Embedded Systems, WISES 2009* (2009)
8. Vece, G.B., Conti, M., Orcioni, S.: Transaction-level power analysis of VLSI digital systems. *Integr. VLSI J.* **50**, 116–126 (2015)
9. Giammarini, M., Orcioni, S., Conti, M.: Powersim: power estimation with system C: computational complexity estimate of a DSR front-end compliant to ETSI standard ES 202 212. In: Conti, M., Orcioni, S., Martínez Madrid, N., Seepold, R. (eds) *Solutions on Embedded Systems. Lecture Notes in Electrical Engineering*, vol. 81, 285–300. Springer, Dordrecht (2011). https://doi.org/10.1007/978-94-007-0638-5_20
10. Gigli, S., Conti, M.: NOCEXplore: a systemC platform for NoC analysis. In: Conti, M., Orcioni, S., Martínez Madrid, N., Seepold, R. (eds) *Solutions on Embedded Systems. Lecture Notes in Electrical Engineering*, vol. 81, 91–104. Springer, Dordrecht (2011). https://doi.org/10.1007/978-94-007-0638-5_7
11. Gigli, S., Conti, M.: A systemC platform for Network-On-Chip performance/power evaluation and comparison. In: *Proceedings of the 7th Workshop on Intelligent Solutions in Embedded Systems, WISES 2009* (2009)



Convolutional Neural Networks Based Tactile Object Recognition for Tactile Sensing System

Ali Ibrahim^{1,2(✉)}, Haydar Hajj Ali², Mohamad Hajj Hassan³, and Maurizio Valle²

¹ Department of Electrical and Electronic Engineering, Lebanese International University, Beirut, Lebanon

ali.ibrahim@liu.edu.lb

² Department of Electrical, Electronic and Telecommunications Engineering and Naval Architecture, University of Genova, Genoa, Italy

maurizio.valle@unige.it

³ American International University, As Salimiyah, Kuwait

Abstract. Recent advances have enabled machine learning methods to be integrated in many application domains to extract meaningful information from sensory data. Machine learning methods have been recently used in tactile sensing systems performing intelligent tasks with an effort to mimic human capabilities. This paper presents a convolutional neural network architecture for tactile object recognition in tactile sensing systems. The proposed architecture outperforms similar state of the art solutions by providing an average classification accuracy of 99.5%. This result paved the way towards the hardware implementation of such network to be integrated in the tactile sensing system.

1 Introduction

Enabling intelligent tasks by employing machine learning (ML) methods has been the focus of recent researches due to the success of such methods in many application domains [1–3]. ML methods have been used recently in tactile sensing systems to extract meaningful information trying to mimic human capabilities. Different ML algorithms have been engaged in the tactile sensing systems for a variety of tasks such as: touch modality classification, surface texture identification, slip and grasp detection, and object recognition.

A BioTac tactile sensor is used to generate a dataset with the size of 5300 instances composed of 53×10 trials each by scanning 53 household objects. A Convolutional Neural Network has been used in the classification with a batch size of one fifth for training i.e. approximately 1000 samples [4]. In [5], a CNN has been used in robotic domain to classify clothes when a tactile sensor is integrated in a robotic arm to sense the clothes grasping by analysing a large RGB pressure map. Textile properties classification has been addressed in [6]. The textile type, thickness, softness, washing method, smoothness, stretchiness, woolen, durability, and wind proof have been targeted. The VGG-19 pretrained on ImageNet has been adopted for the textile classification. Object structure bumpy, hard, soft, etc. and object state e.g. moving, static, etc. have been identified for a humanoid robot covered by tactile sensory array in [7]. K-Nearest Neighbor (KNN) classifier has been used to differentiate between 18 objects achieving an accuracy of

80%. In [8], a CNN network based on AlexNet has been used to recognize 22 objects from their pressure map obtained from a tactile sensor array of size 28×50 sensors. Transfer learning has been employed in [9] to classify three different touch modalities from an array of 4×4 piezoelectric sensors. This has been done by transforming the sensorial tactile data into images then adopting CNN models trained on ImageNet [6] for the classification.

This paper presents the implementation of a convolutional neural network for tactile object recognition. The dataset has been collected using an array of 160 piezoresistive tactile sensors. Experimental results demonstrate the effectiveness of the proposed approach by achieving a classification accuracy of 99.5% overcoming similar state of the art solutions.

2 The Tactile Sensing System

Figure 1(a) provides a block diagram of the tactile sensing system. The system involves a tactile sensor array, an electronic interface, and a graphical user interface on a computer.

- (1) A force Sensing Resistor (FSR) the MS9723 piezoresistive sensor is adopted as a sensor array [10]. The MS9723 is composed of 16 rows \times 10 columns forming a total of 160 sensors converting the applied pressure into electric resistance variation. When applying a force on the surface of the sensors the current increases by the mean of a decrease in resistance.
- (2) The snowboard system developed by Kitronyx is employed as an electronic interface for the system. It is considered as a resistive version of modern capacitive multi-touch sensing technology [11]. When the MS9723 is connected to the Snowboard, the forces applied to each sensor are detected and sent through a USB interface to the computer.
- (3) The graphical user interface (GUI) is based on the Snowforce 3 [12] which is a tool to communicate between the snowboard and the computer. It is in charge of setting the communication parameters, the visualization method, and the data logging period. This GUI has been used to visualize sensor data (heatmap images), to test the system, and also to save the collected data after each measurement.

The main task is to identify the object when they are on the surface of the tactile sensor array. The dataset has been collected using eleven different objects shown in Fig. 1: “bottle cap, eraser, gas lashes, highlighter cap, key, marble, rock, spray cover, tape, piece of wood, shaped screwdriver”. These objects have different weights and distinct texture properties (soft, hard, flexible). Each object has been placed on various positions of the sensor array for 50 times and pressed for a duration between 1 and 3 s to form a dataset with total size of 4000 samples.

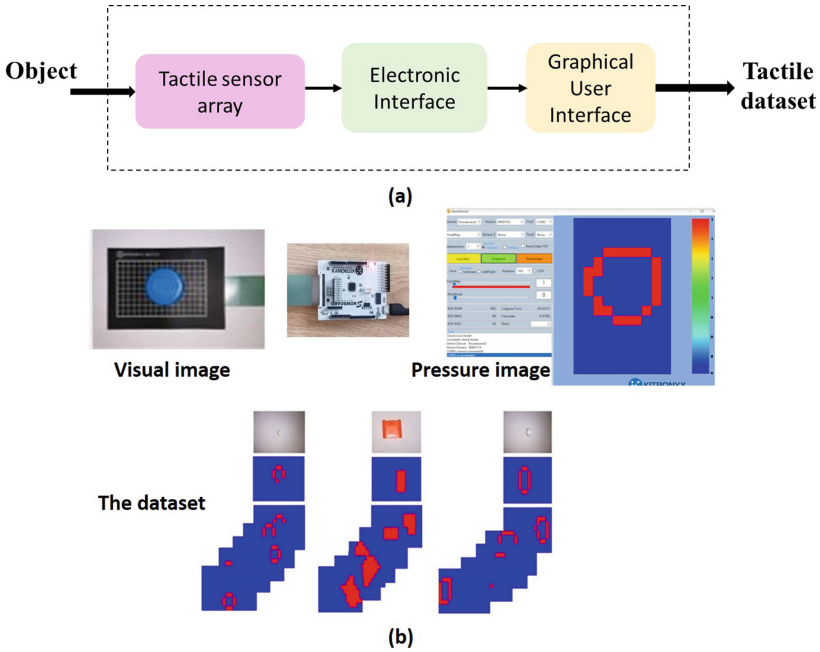


Fig. 1. (a) A block diagram of the tactile sensing system; (b) The dataset collection process. (Color figure online)

3 Network Architecture

As previously discussed, the objects have been placed and pressed on the sensor array to form pressure map figures for the eleven targeted classes. Hence, since we are dealing with figures, CNN is the direct choice since it is specifically designed to process input images. Matlab tool has been used for the implementation where Alexnet [13] has been adopted to perform the classification.

The network is shown in Fig. 2. It consists of 25 layers with an RGB image as input of size 256×256 . The Max Pooling layers come after the first two convolutional layers. The third, fourth and fifth convolutional layers are connected directly. The fifth convolutional layer is also followed by a Max Pooling layer, the output of which goes into a series of two fully connected layers. The second fully linked layer feeds 1000 class labels into a Softmax classifier. RELU function is applied after all the convolution and fully connected layers. The RELU of the first and second convolution layers is followed by a local normalization step before doing pooling. The learning rate of the considered network is equal to 0.0001 [13].

4 Experimental Results

Training the model from scratch required a large dataset to achieve high accuracy. For that reason, data augmentation techniques i.e. flipping, rotation, and translation in the x

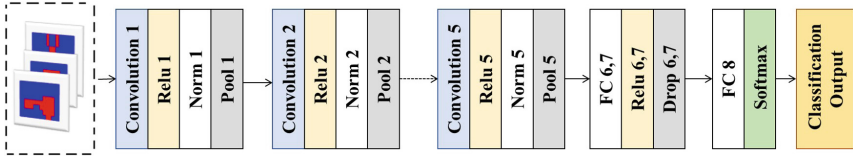


Fig. 2. The network architecture for tactile object recognition. (Color figure online)

and y axis, were applied to the dataset. Before the training, each image is resized to 256×256 pixels in order to fit the size of the network input layer. The dataset is composed 4000 images divided in 80% for training, 10% for validation, and 10% for testing. So, each object has 290 images resulting in 3200 for the training, and 36 images for each object in the testing and validation to form a total of 400 images for each. This process is repeated five times until all the folds are used, without having common elements across all folds for the validation and test sets as shown in Fig. 3(a). The training accuracy results for the 5 folds over 10 runs are illustrated in Fig. 3(b).

The performance of the implemented model is evaluated based on the recognition rates achieved by using the test set to classify 400 images into 11 classes as described above. The confusion matrix illustrated in Fig. 4 shows that the proposed network achieves an average accuracy of 99.5% overcoming similar state of the art solutions [8, 14].

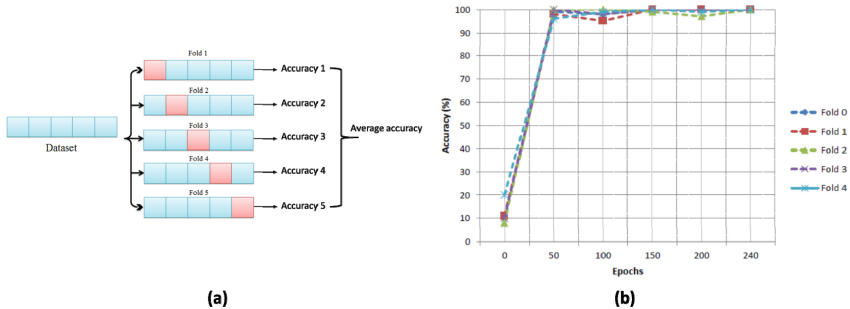


Fig. 3. (a) Dataset organization in 5-fold process; (b) Training results. (Color figure online)

bottle cap pic	55 9.1%	0 0.0%	0 0.0%	1 0.2%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	1 0.2%	0 0.0%	96.5% 3.5%
eraser pic	0 0.0%	54 8.9%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
gas lashes pic	0 0.0%	0 0.0%	55 9.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
highlighter cap pic	0 0.0%	0 0.0%	0 0.0%	54 8.9%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
key pic	0 0.0%	0 0.0%	0 0.0%	0 0.0%	55 9.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
marbel pic	0 0.0%	1 0.2%	0 0.0%	0 0.0%	0 0.0%	55 9.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	98.2% 1.8%
rock pic	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	55 9.1%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
shapped screwdriver pic	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	55 9.1%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
spray cover pic	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	55 9.1%	0 0.0%	0 0.0%	100% 0.0%
tape pic	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	54 8.9%	0 0.0%	100% 0.0%
wood pic	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	55 9.1%	100% 0.0%
	100% 0.0%	98.2% 1.8%	100% 0.0%	98.2% 1.8%	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	100% 0.0%	98.2% 1.8%	100% 0.0%	99.5% 0.5%
bottle cap pic												
eraser pic												
gas lashes pic												
highlighter cap pic												
key pic												
marbel pic												
rock pic												
shapped screwdriver pic												
spray cover pic												
tape pic												
wood pic												

Fig. 4. Confusion matrix for the testing results. (Color figure online)

5 Conclusions and Future Perspectives

This paper presented the implementation of a convolutional neural network for tactile object recognition. It presented the tactile sensing system together with the experimental setup carried out to collect a dataset from an array of 160 piezoresistive sensors. Experimental results demonstrate the effectiveness of the proposed approach by achieving a classification accuracy of 99.5% overcoming similar state of the art solutions. Future work will consist on embedding the proposed network into hardware devices e.g. STM32 microcontroller to enable the system of providing a classification on the edge. Moreover, optimization steps will be applied on the networks to reduce its size targeting energy efficient implementations. Such techniques may pave the way towards smart tactile sensing system for portable/wearable applications.

Acknowledgments. Authors would like to thank Eng. Mohamad Baalbaki and Eng. Fatima Saleh for their help in data collection.

References

1. Liu, Y., Bi, S., Shi, Z., Hanzo, L.: When machine learning meets big data: a wireless communication perspective. *IEEE Veh. Technol. Mag.* **15**(1), 63–72 (2020). <https://doi.org/10.1109/MVT.2019.2953857>

2. Shinde, P.P., Shah, S.: A review of machine learning and deep learning applications. In: 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBE), pp. 1–6 (2018). <https://doi.org/10.1109/ICCUBE.2018.8697857>
3. Franceschi, M., Nannarelli, A., Valle, M.: Tunable floating-point for embedded machine learning algorithms implementation. In: 2018 15th International Conference on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD), pp. 89–92 (2018)
4. Gao, Y., Hendricks, L.A., Kuchenbecker K.J., Darrell T.: Deep learning for tactile understanding from visual and haptic data. In: 2016 IEEE International Conference on Robotics and Automation (ICRA), pp. 536–543. Stockholm, Sweden (2016)
5. Yuan, W., Mo, Y., Wang, S., Adelson, E.: Active Clothing Material Perception using Tactile Sensing and Deep Learning. [arXiv:1711.00574](https://arxiv.org/abs/1711.00574) (2017)
6. ImageNet. <http://www.image-net.org>. Accessed 15 July 2021
7. Bhattacharjee, T., Rehg, J.M., Kemp, C.C.: Haptic classification and recognition of objects using a tactile sensing forearm. In: 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, pp. 4090–4097. Vilamoura-Algarve, Portugal (2012)
8. Gandarias, J.M., Garcia-Cerezo, A.J., Gomez-de Gabriel, J.M.: CNN-based methods for object recognition with high-resolution tactile sensors. *IEEE Sens. J.* **19**, 6872–6882 (2019)
9. Alameh, M., Ibrahim, A., Valle, M., Moser, G.: DCNN for tactile sensory data classification based on transfer learning. In: Proceedings of the 2019 15th Conference on Ph.D Research in Microelectronics and Electronics (PRIME). pp. 237–240. Lausanne, Switzerland, 15–18 July 2019
10. Kitronyx. https://www.kitronyx.com/store/p31/%5BMS9724%5D_FSR_Matrix_Array_Sensor_%2816x10_Rows_and_Columns_%2F_127mm_x_80mm_Active_Sensing_Area%29.html. Accessed 15 July 2021
11. Kitronyx. https://www.kitronyx.com/store/p68/Snowboard_2_Plus.html. Accessed 15 July 2021
12. Kitronyx. <http://sites.kitronyx.com/wiki/applications/snowforce-3>. Accessed 15 July 2021
13. Nayak, S.: Learnopencv. <https://learnopencv.com/understanding-alexnet/>. Accessed 15 July 2021
14. Alameh, M., Abbass, Y., Ibrahim, A., Valle, M.: Smart tactile sensing systems based on embedded CNN implementations. *Micromachines J.* **11**(1), 103 (2020). <https://doi.org/10.3390/mi11010103>



Design of V2X Communications Based on 5G NR: A Physical Layer Perspective

Fabiola Sapienza, Vincenzo Lottici^(✉), Filippo Giannetti, and Sergio Saponara

Department of Information Engineering, University of Pisa, 56122 Pisa, Italy
{fabiola.sapienza, vincenzo.lottici, filippo.giannetti, sergio.saponara}@unipi.it

Abstract. The call for high performance vehicular-to-everything (V2X) connectivity for automotive applications has spurred the considerable development of several standards, where the 5G New Radio (NR) is playing definitely an essential role. In this paper, after briefly reviewing the basics of the NR physical layer, we first evaluate by simulation the performance of NR-based V2X communication links for the NR modulation and coding scheme (MCS) options, and then, such outcomes are exploited to perform the link design for realistic wireless scenarios. Our results show considerable performance gains when the proper MCS is adopted, thus motivating the need for suitable resource allocation algorithms.

Keywords: 5G NR · NR-V2X · Packet error rate performance evaluation · Link design

1 Introduction

Future networks will have to provide broadband every-where and every-time connectivity to support a diverse range of services and applications including everything from voice, 3D video, virtual and augmented reality, safety, emergency, factory automation, self-driving cars, smart city. The demand for high performance connectivity has gained a considerable ever increasing interest also from the automotive area about what is called as Vehicular-to-Everything (V2X) communications. In the recent years, indeed, this has motivated the development of several yet different standards, such as IEEE 802.11p, 802.11bd, LTE-V2X [1]. However, the stringent requirements for advanced V2X applications has brought to give a significative emphasis to the newer standard based on 5G New Radio (NR), also known as NR-V2X [2]. 5G NR fully satisfies the required demanding features of the automotive applications, its components being designed to enable flexibility, ultra-lean factor, backward and forward compatibility, low-latency, ultra-reliability, greater coverage, low error performance, high data rates and the exploitation of new higher frequency bands such as millimeter wave (mm-Wave) [3].

This work was partially funded by the European Union's Horizon 2020 research and innovation programme "European Processor Initiative" under grant agreement No. 826647.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2022
S. Saponara and A. De Gloria (Eds.): ApplePies 2021, LNEE 866, pp. 286–292, 2022.
https://doi.org/10.1007/978-3-030-95498-7_40

Contributions

1. A brief overview of 5G NR PHY layer to emphasize some key features, as frequency bands, modulation and channel coding, multiple numerology, and time-frequency domain signalling, which are useful in the following sections.
2. The evaluation of the packet error rate (PER) performance for the Infrastructure to Vehicular (I2V) and Vehicular to Infrastructure (V2I) communication links employing the NR modulation and coding scheme (MCS) options, based on the simulation of the overall communication system over realistic wireless scenarios.
3. The design of the I2V and V2I links through combining the link budget equation with the PER performance obtained by the simulation runs, from which it can be deduced the need of adaptive resource allocation matched to the time-varying mobile environment.

2 Overview of 5G NR PHY Technology

According to the 3GPP specifications defined in Release 16, the design of the 5G NR physical layer (PHY) is based on up-to-date key techniques, [4,5], such as: *i*) efficient channel coding, *ii*) hybrid-ARQ (HARQ) soft combining, *iii*) error detection and signalling to the upper layers, *iv*) rate matching to the channel available resources, *v*) multicarrier modulation, *vi*) Multiple-Input Multiple-Output (MIMO) antennas, with the aim of enabling high system performance, notable flexibility and both backward and forward compatibility. Next, brief insights into some of the main PHY features are detailed.

RF Frequency Bands. The 5G spectrum can be basically classified into two ranges, the Frequency Range 1 (FR1) covering all the frequency bands lower than 6 GHz, and the Frequency Range 2 (FR2) including the new frequencies from 24.25 to 52.6 GHz, the so-called millimeter wave (mm-Wave).

Specifically, the low-frequency bands are the ones up to 2 GHz, which were previously used for the LTE transmissions. They offer wide coverage areas and 20 MHz wide bandwidths. The medium-frequency bands are in the range 3–6 GHz, with larger bandwidths till 100 MHz, thus allowing higher levels of capacity and data rates. The high-frequency bands are above 24 GHz with bandwidths up to 400 MHz. In view of the adverse propagation properties incurred in such frequency interval, however, they can not be used for wide coverage but only for ultra-small cells (hotspots), though giving remarkable throughput performance.

OFDM Modulation and Channel Coding. In order to enhance the spectral efficiency and the ease of implementation, Orthogonal Frequency Division Multiplexing (OFDM) is adopted as standard modulation format for both downlink (DL) and uplink (UL) transmissions, using either Time Division Duplexing (TDD) or Frequency Division Duplexing (FDD). Herein, the data stream is mapped onto different subcarriers belonging to the available bandwidth, with QPSK, 16-QAM, 64-QAM, and 256-QAM as possible modulation schemes. As alternative choice to the cyclic-prefix (CP) based OFDM, the DFT-spread-OFDM (DFT-s-OFDM) may be employed in the UL to reduce the peak-to-average power ratio (PAPR), which is useful when power efficiency is required.

As for the channel coding, the low density parity check (LDPC) and the cyclic redundancy check (CRC) aided polar codes are proposed for the data channel and the control channel, respectively. The LDPC code covers a wide range of coding rates and block sizes, and supports incremental redundancy (IR) HARQ techniques as well. On the other side, polar coding has been shown as the best performed short code with low algorithmic decoding complexity.

Numerology and Subcarrier Spacing. One of the key aspect of the 5G NR is the selection of a multiple numerology for the subcarrier spacing Δf , as given by

$$\Delta f = 2^\mu \cdot 15 \text{ kHz}, \quad \mu = 0, \dots, 4, \quad (1)$$

that allows the flexibility of addressing different spectrum, bandwidth, scenarios and services with varying requirements. Hence, the minimum spacing amounts to 15 kHz, the one employed for LTE, to ensure backward compatibility. According to the adopted NR numerology, each value of subcarrier spacing defined in (Eq. 1) corresponds to a given duration of the CP and the useful symbol interval as shown in Sect. 3. Further, (Eq. 1) is such that the duration in the time domain of one OFDM symbol includes a number of OFDM symbols at higher spacing obtained as a power of two.

Structure in Time and Frequency Domain. The transmission of NR physical signals in the time domain is based on 10 ms long frames, which are divided into ten 1 ms long subframes. The flexibility is reached by dividing each subframe into a variable number of slots depending on μ in (Eq. 1), while each slot contains a fixed number of 14 OFDM symbols. To achieve flexibility also in the frequency domain, the NR standard introduces the resource element as the smallest physical resource, one OFDM symbol upon one subcarrier, the resource block, 12 consecutive subcarriers upon one OFDM symbol, and the resource grids, each consisting (depending on μ) of a multiple of 12 consecutive subcarriers and some OFDM symbols.

3 V2X Communication Model

3.1 System Model

In the V2X communication system we are dealing with, a *packet*, usually corresponding to an IP packet or a layer 3 control signalling message, is taken as the basic information to be transmitted over the radio interface. At the transmitter side, packet-processing is first performed, then followed by frame processing. The packet-processing is based on: *i*) one-to-one mapping of each input packet to a radio link control (RLC) sub-layer protocol data unit (PDU), including a CRC sequence for error detection and the payload received from upper layers¹, *i*) coding each RLC PDU according to the schemes illustrated in Sect. 2 depending on the logical channel to be delivered, *ii*) rate matching by puncturing to select a specific set of encoded bits suitable to the available channel resources, and *iii*) proper interleaving. The frame processing section performs the

¹ If the input bit sequence is larger than the maximum LDPC block size (8448 or 3840), segmentation is performed and an additional CRC is attached to each block.

Table 1. MCS options and data rates for LDPC and Polar coding.

MCS	Modulation	Code rate	LDPC Code data rate (Mbps)	Polar Code data rate (Mbps)
6	QPSK	0.44	6.31	4.63
7	QPSK	0.51	7.36	5.40
10	16-QAM	0.33	9.57	8.10
13	16-QAM	0.48	13.67	10.79
21	64-QAM	0.65	27.34	16.19
27	64-QAM	0.92	38.27	32.38

Table 2. Channel model parameters.

Scenario	Power profile (dB)	Delay profile (ns)
Urban LOS	[0 -8 -10 -15]	[0 117 183 333]
Urban NLOS	[0 -3 -5 -10]	[0 267 400 533]

mapping of the coded and interleaved sequence to the symbols of the adopted QAM constellation and the IFFT-based OFDM modulation. Hence, for each input packet, a frame of OFDM symbols is obtained which conveys the information for the end-user. The resulting OFDM frame signal is thus transmitted over a mobile frequency-selective time-varying multipath channel. Then, the received signal undergoes OFDM demodulation, zero-forcing equalization, soft demapping, deinterleaving and eventually is decoded.

3.2 Simulation Setup

In accordance with the 5G NR Release 16 standard [4,5], we consider the following V2X communication systems, having carrier frequency of 5.9 GHz (sub-6 GHz) and bandwidth of 10 MHz.

1. **Infrastructure to Vehicular (I2V)** for DL. LDPC coding is employed to allow high throughput efficiency, the input packet being $12 \cdot 10^3$ bits long. The decoding is based on the suboptimal Normalized Min-Sum Algorithm, which reduces the computational complexity of the best known Belief Propagation Algorithm. As for the mobile multipath channel, it is assumed to be time-correlated according to the speed of the vehicle, which is taken as 13.8 m/s.
2. **Vehicular to Infrastructure (V2I)** for UL. Polar coding is employed to achieve high reliability for URLLC applications, the input packet being 300 bits long. The decoding is based on the CRC-Aided Successive Cancellation List Decoding (CA-SCL), whereas the mobile multipath channel is time-invariant within the frame interval, but randomly independent across consecutive frames (block fading).

In our simulations, according to the numerology illustrated in Sect. 2, the OFDM parameters are quantified as follows: *i*) subcarrier spacing, 15 kHz (I2V), 60 kHz (V2I); *ii*) symbol duration, $66.7 \mu\text{s}$ (I2V), $16.7 \mu\text{s}$ (V2I); *iii*) cyclic prefix, $4.69 \mu\text{s}$ (I2V), $1.17 \mu\text{s}$ (V2I); *iv*) FFT size, 667 (I2V), 167 (V2I); *v*) useful subcarriers, 624 (I2V), 132 (V2I). Further, the MCS options we address are shown in Table 1, while the communication scenarios of interest are Urban LOS and Urban NLOS [1], as described in Table 2.

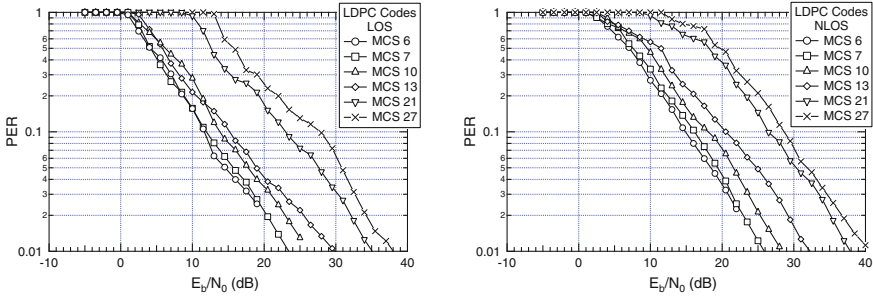


Fig. 1. PER performance for LDPC code in Urban LOS (left) and Urban NLOS (right) channels for various MCSs. (Color figure online)

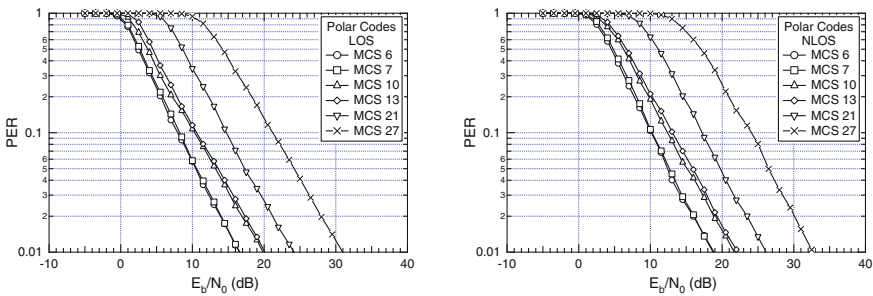


Fig. 2. PER performance for Polar code in Urban LOS (left) and Urban NLOS (right) channels for various MCSs. (Color figure online)

4 Performance Evaluation

In this section, the effectiveness of the V2X communication systems we are addressing is numerically tested for the scenarios described in Sect. 3.2. We take as performance metric the PER as a function of the received mean energy per bit-to-noise spectral density ratio E_b/N_0 . The simulation results for the LDPC code are shown in Fig. 1 and for the Polar code in Fig. 2. The figures at the left refer to the Urban LOS channel whereas those at the right to the Urban NLOS channel. As apparent from Figs. 1 and 2, better PER accuracy is reasonably offered for lower MCS orders, that is to say lower modulation formats and code rates. Additionally, for a given MCS and PER level, moving from the LOS channel to the NLOS one, the Polar codes show a better robustness to performance loss.

5 V2X Link Design and Conclusions

The parameters adopted to design the I2V and V2I communication links are as follows [6]: *i*) carrier frequency f_0 , 5.9GHz; *ii*) TX power P_T , 23 dBm (mobile), 49 dBm (fixed); *iii*) TX/RX antenna gain G_T , G_R , 0 dBi (mobile), 8 dBi (fixed); *iv*) TX/RX

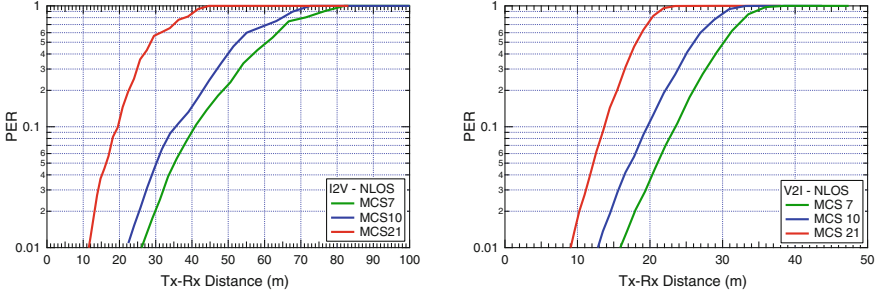


Fig. 3. PER performance vs. Tx-Rx distance for I2V (left) and V2I (right) for various MCSs. (Color figure online)

antenna height h_T, h_R , 1.5 m (mobile), 25 m (fixed); *v*) receiver noise figure F , 13 dB (mobile), 7 dB (fixed); *vi*) shadowing loss L_S , 5.2 dB; *vii*) path-loss exponent γ , 3; *viii*) reference distance d_0 , 1 m; *ix*) channel model, Urban NLOS. To the design aim, for a Tx-Rx distance d , the path-loss L_{PL} can be modelled as $(4\pi d/\lambda)^2$ for $d \leq d_0$, and $(4\pi d_0/\lambda)^2 \cdot (d/d_0)^\gamma$ for $d > d_0$, according to [6] and references therein. Hence, the total loss in the link (in dB) results as $L_{tot}(\text{dB}) = L_{PL}(\text{dB}) + L_S(\text{dB})$, where $L_S(\text{dB})$ denotes the additional contribution due to the shadowing. The resulting link budget equation is then $E_b/N_0 = (P_T G_T G_R) / [L_{tot} R_b k (F - 1) T_0]$, where R_b is the bit rate, k is the Boltzmann constant and $T_0 = 290 \text{ K}$ is the customary ambient temperature. Thus, for a given MCS, combining together the link budget equation with the PER performance evaluated as a function of the E_b/N_0 ratio in Sect. 4 specifies how the PER metric changes with the TX-RX distance d . Figure 3 gives such result for the I2V (left) and V2I (right) links for various MCSs (the leftmost, the middle and the rightmost curves refer to MCS7, MCS10 and MCS21, respectively), showing that i) the PER performance monotonically increases with d , even in a steep way especially for the higher MCS orders, and ii) the I2V link enables a larger coverage.

Concluding, this work evaluated the effectiveness of NR-based V2X communication systems taking the main MCSs adopted by the NR standard into account. Performance metrics obtained by simulation allowed also to design the I2V and V2I communication links over realistic scenarios. Assuming to maintain a given PER level, the main result of the link design is that coverage extension can be achieved only by adopting a lower MCS order. Hence, in a mobile environment where typically the TX-RX distance is time-varying, appropriate resource allocation algorithms are required to adaptively choose the MCS order which makes always available the best link quality.

References

1. Anwar, W., TraBl, A., Franchi, N., Fettweis, G.: On the reliability of NR-V2X and IEEE 802.11bd. In: 30th Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC). IEEE, Istanbul, Turkey, November 2019
2. Naik, G., Choudhury, B., Park, J.-M.: IEEE 802.11bd & 5G NR V2X: evolution of radio access technologies for V2X communications. *IEEE Access* 7, 70169–70184 (2019)

3. Ji, H., Park, S., Yeo, J., Kim, Y., Lee, J., Shim, B.: Ultra-reliable and low-latency communications in 5G downlink: physical layer aspects. *IEEE Wirel. Commun.* **25**(3), 124–130 (2018)
4. 3GPP: 5G NR, Multiplexing and channel coding, 3rd Generation Partnership Project (3GPP) Technical Specification (TS) **38.212**, version 16.2.0 Release, 16 July 2020
5. 3GPP: 5G NR, Physical channels and modulation, 3rd Generation Partnership Project (3GPP) Technical Specification (TS) **38.211**, version 16.2.0 Release, 16 July 2020
6. 5GCAR (Fifth Generation Communication Automotive Research and Innovation): Final 5G V2X Radio Design, Deliverable **D3.3**, version 1.1, November 2019



A Low Cost Compact Output Amplifier for Multichannel Muscle Stimulation

Massimo Ruo Roch^(✉) and Maurizio Martina

Dipartimento di Elettronica e Telecomunicazioni, Politecnico di Torino, Turin, Italy
massimo.ruoroch@polito.it

Abstract. Functional Electrical Stimulation (FES) is a technique widely used in different application fields. Among them, rehabilitation treatments, sports training activity and aesthetics. A critical component in FES is the output amplifier, which must be designed to accomplish both functional and safety specifications. In this paper, a novel power amplifier architecture with current controlled output is presented, specifically tailored to FES needs, but optimized to reduce cost and size, as needed in a multichannel electrical stimulator. The output amplifier topology, and the techniques adopted to satisfy both functional and safety specifications are herein detailed. Simulation and real circuit results are included, too.

Keywords: Functional Electrical Stimulator (FES) · Output amplifier · Rehabilitation

1 Introduction

Neuromuscular stimulation, also known as FES (Functional Electrical Stimulation) is a well known approach to rehabilitation [1,2]. Its application to spinal chord injury (SCI) treatment helps to strengthen muscles, reduce their atrophy, spasticity, and also bone demineralization [3]. The same technique find applications in sport training, where it is used to increase strength and/or cool-down after intense work-out sessions [4]. Last, the same approach is widely used in aesthetics and aesthetic medicine, where this kind of treatment is used to shape customer's body through an increase of muscles tone.

FES is obtained by applying conveniently shaped currents to the neuromuscular tissue of the muscle. These currents are usually transmitted to the muscle through skin electrodes, and, as a consequence, the impedance which must be driven is relatively high, ranging from 500 Ω to 2 k Ω [5].

The current used to obtain muscle contraction must be adjustable from few mA up to some tens of mA, due to the great variability in muscle sensitivity (depending on muscle extension, too), and the desired contraction level [6].

With the impedances stated above, voltage levels must vary from few Volts up to one hundred Volts. The output signal can be unipolar or bipolar. The

latter is preferred, as it avoids charge imbalance inside the tissue, or electrodes polarization effects, which can be both painful and harmful.

Last, special care must be adopted, from the point of view of safety, as electrodes are directly connected to the patient, and the risk of electrocution must be minimized. For this reason, regulations impose strong isolation of the outputs from the AC lines and from earth. Generation of excessive voltages and currents must be prevented by-design, too.

According to the requirements described above, a typical FES device is built-up by a low voltage signal generator, and a high voltage power amplifier. The first one can be a simple relaxation oscillator, or a microprocessor driving a digital-to-analog converter (DAC). The power amplifier is a discrete-based circuit, designed to generate voltage or current outputs. The second topology is typically preferred, as it better adapts to the unknown and varying impedance at the electrode-skin interface [7].

Several circuits and topologies are described in literature. Simplest ones are just pulse generators based on inductive charge [3,8,9]. A second family of circuits is represented by discrete components implementation of current mirrors. They can either be unipolar, followed by a four switch bridge to invert the current in the load [10], or fully symmetric ones [11]. Anyway, they do not fully accomplish system specification, lacking flexibility in signal generation (pulse generators), or missing safety characteristics. Moreover, they are not optimized for power consumption or size, and need to be improved.

To overcome these pitfalls, a novel circuit has been designed, with the following characteristics:

- Open-loop design, to guarantee stability with varying loads.
- Scalable design, adaptable to different environments (fixed/portable/battery-operated).
- Low cost and small size. As the design must be compatible to a multichannel design, cost and size optimization is mandatory.
- Safety by-design. As electrodes are directly attached to human body, fault conditions must guarantee that no injury can be caused to the patient.
- Arbitrary output waveform. This feature allows to apply the designed stimulator to different research fields, too, where experimental waveforms can be used.

2 Circuit Description

The solution proposed in this paper is based on a tight coupling of the low voltage signal generator with the output amplifier. This result is obtained using a mixed signal MCU driving a current amplifier built with discrete components. The resulting block architecture is visible in Fig. 1. The MPU, the DAC, the comparator and operational amplifier inside the dashed rectangle are included in the selected MCU, and are separately drawn just for the sake clarity.

The microprocessor contains the program able either to synthesize, or to store in a table, the samples of the signals to be generated. These samples are

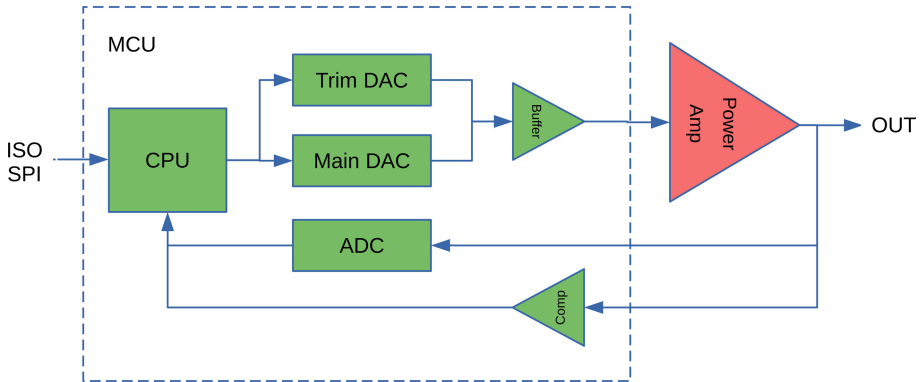


Fig. 1. Basic architecture of the system

sent to the integrated DAC, and converted to a low amplitude analog signal. Optionally, the microprocessor can receive configuration data from an isolated SPI interface.

The DAC output is sent to an operational amplifier, used to accomplish two different tasks. The first one is to have a higher current available to drive the input stage of the output amplifier, as described below. The second task is to sum a DC voltage, used to cancel offsets present in the output amplifier, due to components tolerances and temperature and/or aging induced drifts.

An ADC continuously monitors the output voltage, and the microprocessor averages it to extract DC components. The result is used to generate a further DC voltage, which is injected in the operational amplifier driving the power amplifier, to fully cancel output offset. This offset is not digitally added to the main DAC output, to avoid an output dynamic range reduction. Instead, it is sent to a secondary DAC, used to finely regulate output DC level.

Last, a solid state relay, not shown in the figure, connects the power amplifier output to the load, allowing for complete disconnection of the electrodes from electronics. For safety concerns, this switch is turned off by default.

2.1 Power Output Amplifier

The output amplifier is a fully symmetric, open-loop, transconductance amplifier, shown in Fig. 2. The topology is class B, to reduce power dissipation.

Due to design symmetry, the circuit analysis can start considering only the upper half, which generates currents flowing to the load.

The input voltage is first level shifted by the diode-connected transistor Q_1 , and then applied to the base of Q_3 . This device is configured as a current generator, through the introduction of the emitter resistance R_4 . Its collector current is approximately V_{IN}/R_4 .

Transistor Q_1 has a twofold usage. First, it adds its V_{BE} voltage to the input value, to reduce crossover distortion. In fact, Q_3 is slightly polarized, even with

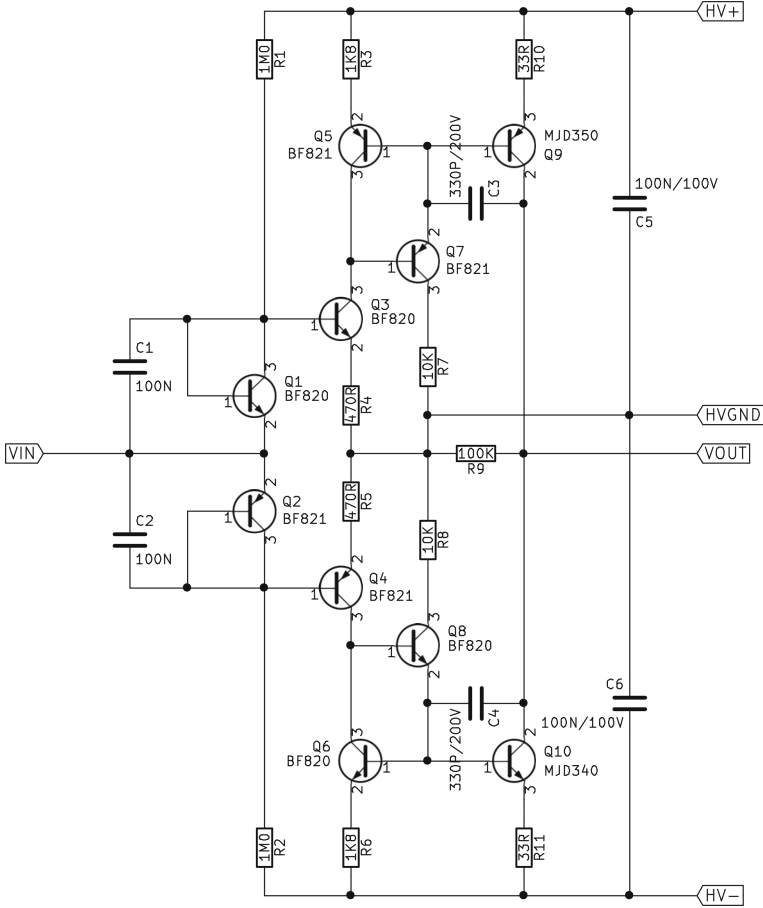


Fig. 2. Detailed schematic of the output power amplifier

an input voltage equal to zero. Second, the V_{BE} of Q_1 varies with temperature changes, and compensates thermal drift of $Q_3 V_{BE}$, improving overall thermal stability. A remarkable point is that only Q_3 temperature is relevant for output current change, as Q_5 and Q_9 mutually compensates. R_1 , biasing Q_1 , sets the quiescent current of the amplifier.

C_1 , shorting Q_1 , is used to improve large signal transient response, e.g., when a square wave is sent to the input.

Q_3 collector current feeds a current mirror composed by Q_5 , Q_7 , and Q_9 . Q_5 and Q_9 are the mirror pair, with current amplification given by the ratio R_3/R_{10} . Q_7 instead, is a current buffer, used to minimize the impact of the low current gain of the output power transistor Q_9 . Without Q_7 , the current ratio error would be unacceptable.

A Wilson topology would not be acceptable in this circuit, due to the asymmetries between weak and strong side of the mirror. This asymmetry is mandatory, to reduce power dissipation in the weak side.

Resistor R_7 is used to protect Q_7 in case of an open circuit on the output, which can by all means happen, as skin electrodes can be easily detached. In this case, without R_7 , the current through Q_7 will be $(H_{V+} - V_{BE})/R_{10}$, as Q_9 is no more a transistor but only a diode represented by its base-emitter junction. This current will create enough power dissipation to damage Q_7 . The introduction of R_7 , instead, limits the current in Q_7 to $(H_{V+} - V_{BE})/(R_{10} + R_7)$, decreasing it by three orders of magnitude.

As a drawback, the turn on time of the power output transistor Q_9 is slightly increased, as R_7 limits the maximum current available to turn it on to few mA's. Anyway, as it will be shown in Sect. 3, the resulting rise and fall time are more than sufficient for an effective muscle stimulation.

Capacitor C_3 , across the base-collector junction of Q_9 , is used to improve circuit stability in some rare conditions of a highly reactive load, again without a significant impact on overall amplifier bandwidth.

The same considerations applies for the lower half of the circuit, which sinks current from the load. Due to symmetry, the output current is given by the difference between Q_9 and Q_{10} collector currents.

2.2 Safety Countermeasures

Several different techniques are introduced, to guarantee circuit safety in any condition:

- Output amplifier fault. The output voltage of the current amplifier is continuously monitored, and if its DC value is out of range, the output is disconnected from the load through the solid state relay.
- Power supply rail failure. The power supply rails are monitored by two channels of the ADC, and the electrodes are disconnected if the read values are too far from the nominal one.
- Microcontroller firmware fault. A hardware watchdog monitors the correct execution of the application firmware, and, in case of a software malfunction, it resets the microcontroller, turning off the output solid state relay, too.
- Application software erroneous configuration. A hardware comparator is connected, through a low pass passive filter, to the output of the power amplifier. If the generation of a waveform with erroneous parameters is detected, then the comparator triggers an event, disabling the output switch.

3 Physical Implementation and Results

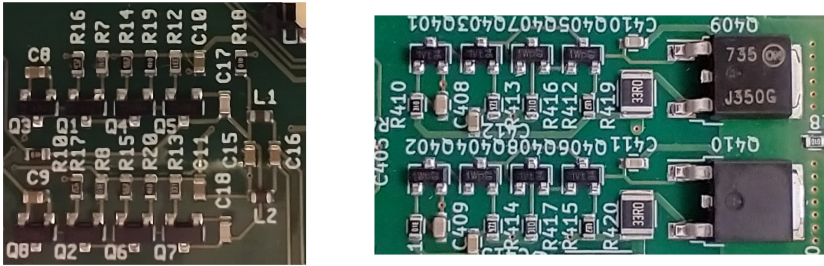


Fig. 3. Low power (left) and high power (right) version of the power amplifier

An important feature of the designed circuit is its scalability. In fact, it can be easily adapted to different current and voltage levels, just changing the BJTs, without any topology changes. To prove this feature, two different circuits have been implemented, one for a portable, single channel, battery operated device, and the second for an AC powered multichannel stimulator. Photos of the power amplifier section of the two circuits are represented in Fig. 3 and dimensions are 18 mm × 20 mm and 32 mm × 20 mm, respectively. The low power version has a 40 mA, 48 V output limit, while the high power one can reach 100 mA, 120 V output.

Simulated and measured waveforms for a large signal pulse are shown in Fig. 4. The oscilloscope photo shows a bipolar pulse of 110 mA peak-to-peak on a load of 1 kΩ. Pulse duration is 440 μs.

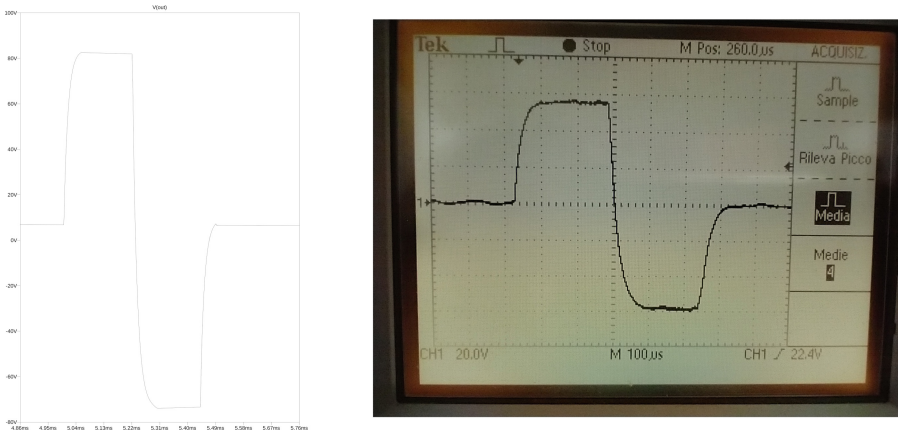


Fig. 4. Simulated (left) and measured (right) output waveforms

References

1. Badran, M., Moussa, M.: BioMEMS implants for neural regeneration after a spinal cord injury. In: 2005 International Conference on MEMS,NANO and Smart Systems, 2005, pp. 89–90 (2005). <https://doi.org/10.1109/ICMENS.2005.31>.
2. Brunetti, F., Garay, A., Moreno, J.C., Pons, J.L.: Enhancing functional electrical stimulation for emerging rehabilitation robotics in the framework of hyper project. In: 2011 IEEE International Conference on Rehabilitation, pp. 1–6 (2011). <https://doi.org/10.1109/ICORR.2011.5975370>
3. Cheng, K.E., Lu, Y., Tong, K.Y., Rad, A.B., Chow, D.H., Sutanto, D.: Development of a circuit for functional electrical stimulation. *IEEE Trans. Neural Syst. Rehabil. Eng.* **12**(1), 43–47 (2004). <https://doi.org/10.1109/TNSRE.2003.819936>.
4. Bélanger, M., Stein, R.B., Wheeler, G.D., Gordon, T., Leduc, B.: Electrical stimulation: can it increase muscle strength and reverse osteopenia in spinal cord injured individuals? *Arch. Phys. Med. Rehabil.* **81**(8), 1090–8 (2000). <https://doi.org/10.1053/apmr.2000.7170>. PMID: 10943761
5. Xu, Q., Huang, T., He, J., Wang, Y., Zhou, H.: A programmable multi-channel stimulator for array electrodes in transcutaneous electrical stimulation. In: Proceedings of the IEEE/ICME International Conference on Complex Medical Engineering, 2011, pp. 652–656 (2011). <https://doi.org/10.1109/ICME.2011.5876821>
6. Enoka, R.M., Amiridis, I.G., Duchateau, J.: Electrical stimulation of muscle: electrophysiology and rehabilitation. *Physiology* **35**(1), 40–56 (2020). <https://doi.org/10.1152/physiol.00015.2019>
7. Hsueh, Y., Chen, G.: Design of high voltage digital-to-analog converter for electrical stimulator. In: Proceedings of the IEEE Asia Pacific Conference on Circuits and Systems, 2012, pp. 77–80 (2012). <https://doi.org/10.1109/APCCAS.2012.6418975>
8. Velloso, B., Souza, M.N.: A programmable system of functional electrical stimulation (FES). In: Proceedings of the 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, 2007, pp. 2234–2237 (2007). <https://doi.org/10.1109/IEMBS.2007.4352769>
9. Chen, M., et al.: A self-adaptive foot-drop corrector using functional electrical stimulation (FES) modulated by tibialis anterior electromyography (EMG) dataset. *Med. Eng. Phys.* **35**(2), 195–204 (2013). <https://doi.org/10.1016/j.medengphy.2012.04.016>. Epub 2012 May 22. PMID: 22621781
10. Khosravani, S., Lahimgarzadeh, N., Maleki, A.: Developing a stimulator and an interface for FES-cycling rehabilitation system. In: Proceedings of the 2011 18th Iranian Conference of Biomedical Engineering (ICBME), pp. 175–180 (2011). <https://doi.org/10.1109/ICBME.2011.6168550>.
11. Yochum, M., Binczak, S., Bakir, T., Jacquir, S., Lepers, R.: A mixed FES/EMG system for real time analysis of muscular fatigue. In: Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology, 2010, pp. 4882–4885 (2010). <https://doi.org/10.1109/IEMBS.2010.5627264>



The Exploitation of Sustainable Composite Materials for the Manufacturing of High-Efficient Electric Cars

Jacopo Agnelli¹(✉), David Benedetti¹, Nicholas Fantuzzi², and Sergio Saponara³

¹ Carbon Dream Spa, Barberino Tavarnelle (FI), Sambuca, Italy
jacopo.agnelli@tiscali.it

² DICAM Department, University of Bologna, Bologna, Italy

³ DII Department, University of Pisa, Pisa, Italy

Abstract. To produce high-efficient electric cars one of the main innovations is the use of bio- composites, which are biodegradable materials that can be a valid alternative to fibers-reinforced plastics. They are nontoxic and recyclable because of their derivation from natural materials. The public demand for a sustainable environmental development has been the main driving factor of the present study. Innovation feeds the relationship between innovative companies and customers that confront each other on new challenges with demanding requirements. All these reasons have led to design a two-seats micro car with a bodywork in bio- composite material has been designed. The car has an electric propulsion and a high- end electronic equipment on board for a safe and sustainable urban mobility. The project includes also the construction of a car shelter, or garage, with photovoltaic tiles plus power electronics for recharging the full electric car battery and an off-grid system, in order to define a whole zero emissions mobility system.

1 Introduction

Industry nowadays is trying to take advantage from renewable energy sources using simultaneously recyclable materials or materials with a smaller global impact, in order to fight against climate changes. The main objective of this study will be to demonstrate the big step taken by presenting a feasible application of bio-composites combined with the use of photovoltaics and power electronics in automotive industry, particularly in the design of a lightweight electric car. To adopt composite materials in automotive industry needs increased competences in the field of biopolymer technologies. Such extent in competences is not industry-wide and in-depth analysis are needed according to competences especially in plastics industry, see Chadha (2011).

The automotive industrial sector, particularly in its life cycles and domestic market status, governs the interactions between technological-push and demanding-pull as indicated by Choi (2017). As a matter of fact, technology development occurs ahead of market demand in the US electric vehicle sector, while it isn't so in Europe, from that the importance of the technology-push factor and the role of the lead market in the electric vehicle sector. This is the reason why the development of new and innovative vehicles is

so important in the current worldwide market. Another way of pushing forward innovation is via sectorial characteristics. It has been observed by Faria and Andersen (2017) that sectorial characteristics are also important when eco-innovation is considered. A comprehensive collection of contributions related to traditional innovation theory which suggests that economic growth and technological change are strongly interlinked were presented by Leone and Belingheri (2017). Particularly, these studies were concentrated on the automotive sector and all its sub-fields. In this work, the use of bio-composites in vehicle bodies is presented linked with its repercussions both in the developed and Developing Countries and, in general, on its economic and environmental sociology. The model of the present micro-car using linen bio-composites is depicted in Fig. 1. Bio-composites or Natural Fiber Composites (NFCs) have been studied and applied in the industry since the beginning of the 20th century.

In England, during World War II, owing to the lack of aluminum, a special fiber based on reinforced linen was used with a phenol-formaldehyde resin (Gordon-Aerolite) for the production of the plies of the fuselage of the Spitfire aircraft. In 1941 H Ford had already produced a prototype using a composite based on hemp fibers but, unfortunately, that model never entered in production in consequence of economic limitations and difficult relations in the international commercial context of the II World War. In the 1950s, the first standard-built passenger car, produced with natural wool or cotton fiber reinforcements, was the Trabant produced from 1964 until the fall of the Berlin wall. During the Cold War the Trabant was the usual means of transport for families living beyond the Iron Curtain, under the Warsaw Pact and with a five-year stability plan. Today NFCs are used in the automotive sector, especially in order to save weight: we can now save up to 34% of the individual parts. The Lotus Eco Elise, presented at the end of 2008, weighed 32 kg less than the standard car thanks to bio-sustainable materials. Mercedes-Benz is also expanding the use of bio-compounds in its series production of the E class. The rest of the paper is organized as follows: Sect. 2 reviews the bio-composite material market and its use for vehicle design. Section 3 presents the new electric car design. Section 4 shows the electronic system design part. Conclusions are drawn in Sect. 5.



Fig. 1. Render and model of the present electric micro-car in linen bio-composite. (Color figure online)

2 The Bio-composites Market

Bio-technology clusters can be found in a number of locations both in Europe and US. The quality and performance of these clusters are identified by basic research funding, invention, regulatory variations between national business and innovation systems, as well as the nature of markets, the structure of industry and some qualitative aspects like ‘entrepreneurship’. These selected variables are the most important in explaining how biotechnology industry works in the most leading countries. According to Cooke (2001) “new economy” conventions have to be defined and their implications investigated for systemic innovation. As a result, bio-composites are employed in the present study in the avant-garde topic of electric vehicles for a sustainable and innovative mobility. According to Luncintel (2019) the bio-composites market looks promising thanks to the new opportunities in several engineering practices among which automotive is one of them, particularly NFC for automotive interiors, thanks to its aesthetics and growing concern to passenger safety. Bio-composites are estimated to reach an overall estimate of \$8.3 billion by 2024 and it is forecasted to grow to a CAGR of 7.5% from 2019 to 2024.

The expected growth of bio-composites market indicated by Luncintel (2019) from 2016 to 2030 is around 25%. NFC bio-composites are biodegradable materials, which can be a valid alternative to fiber-reinforced plastics. Plant fibers are not toxic and derive from natural materials which are readily available and can help to promote healthier and safer workplaces reducing the production of climate-altering carbon dioxide. It is well-known that composite materials in natural fiber, derived from hemp, cotton, pineapple, linen and jute are nowadays much used, even if only in the fashion and clothing sector. Interest in bio-composites has grown considerably in recent years thanks to their good mechanical properties, low cost, low density and high recyclability. In the enormous plethora of bio-fibers, are particularly suitable, in the field of structural reinforcement, linen, flax, cotton and hemp. According to Maximize Market Research (2019) these are the reference markets for bio-composites: building and construction, automotive, electrical and electronics components and consumer products. In general, all these fields observe an increase through the years. All vegetable fibers have an extremely complex molecular structure formed by a multiplicity of biopolymers (lignin, crystalline cellulose, pectin,) and a nano-structured architecture which gives these fibers mechanical properties usefully used as reinforcements in the transport sector and green mobility (Lau et al. 2009). The main properties of hemp, cotton and linen fibers are (see Table 1):

- high tensile strength and breaking deformation;
- Low thermal, electrical and acoustic conductivity;
- Electromagnetic transparency need of low energy for production;
- They are completely recyclable;
- They derive from renewable sources.

NFC can be impregnated both with thermoplastic and thermosetting matrices, becoming a valid alternative to glass fibers and aluminum, which instead require a lot of electricity for production. With a view to continuous innovation aimed at eco- sustainability, our industrial research intends to obtain solutions that allow

us to achieve higher technological performance, paying particular attention to the environment. The low cost of the material and the fact that the Developing Countries are the largest producers of many of these fibers are determining a growing interest in these bio-materials, applied also in aerospace and automotive.

Table 1. Comparison between mechanical properties and costs of traditional compounds and bio-composites (Cristaldi and Cicala 2011).

Fiber	Density	Tensile strength [MPa]	Tensile strength density, [MPa·m ³ /kg]	Unit cost \$/kg
Carbon	1880	1700–2400	0.90–1.28	220.00
Glass	2540	1400–2500	0.57–0.98	5.00
Broom	1250	400–800	0.32–0.64	0.50
Ramie	1560	800–1000	0.51–0.64	0.70
Cotton	1520	300–600	0.20–0.39	1.50
Jute	1450	400–600	0.28–0.41	0.30
Linen	1540	900–1200	0.58–0.80	1.50
Hemp	1480	400–700	0.27–0.47	1.30
Sisal	1450	500–600	0.34–0.41	0.75
Coconut	1150	100–200	0.09–0.17	0.50
Banana	1350	500–700	0.37–0.52	1.50
Pineapple	1440	400–1000	0.28–0.69	0.75

The NFC can be described starting from the basic concept of FRP (Fiber Reinforced Plastics, the carbon fiber used in many automotive applications, such as in Formula 1), i.e. the three-dimensional combination between a polymer resin and a reinforcing fiber, rather than exploiting materials of synthetic origin, deriving from the processing of oil. The bio-composites used in this project exploit a wider range of origin plants, besides cotton. This is a very important aspect in a global socio-economic vision, especially in the developing Countries, indicated by International Monetary Fund (2018), and it can represent a valid solution to the fourth economic macro-market failure: the underdevelopment of some areas of our planet. Table 2 presents a list of natural fibers and their main producers in relation with the GDP pro capita and the energy for their production.

The combination of natural fibers with polymers of renewable origin, is used to produce increasingly competitive materials, if compared to synthetic composites, even if their production very often requires more technological skills, and appropriate technological and industrial investments, including numerous process innovations, such as the use of a particular polyamide interface between bio-fiber and matrix and even more complex process phases (MAFF 2003). Rowell (1996) has reported that the so-called green bio-composites are therefore evolving by conglomerating natural fibers with natural polymers obtained from renewable sources such as plastics cellulose, polylactides,

Table 2. Comparison between countries main producers of natural fiber according to GDP pro capita (\$) and energy for production according to Int. Monetary Fund (2018), Agnelli et al. (2020).

Natural Fibre	Main producer Country	RANK 2017	GDP 2017 Pro capita \$	Energy for production MJ/ton
Ramie	China	72/188	8643	-
Jute	Bangladesh	145/188	1602	-
Linen	China	72/188	8643	-
Hemp	India	140/188	1983	4170
Sisal	Brazil	67/188	9895	2488
Coconut	Mozambique	184/188	429	-
Banana	India	140/188	1983	-
Flax	-	-	-	2752
Pineapple	Costa Rica	61/188	11685	-
Glass	-	-	-	31700
Carbon	-	-	-	355000

plastic starch derived from starch, poly-hydroxy-alkenoates (bacterial polyesters), soy or corn based plastics.

A non-complete classification is listed in Table 3. Many bio-composite materials use recycled materials or fibres derived from fast-growing plants such as hemp (*Cannabis Sativa*) (Oksman et al. 2003). They can therefore be recycled in a simple way or designed to be very quickly biodegradable. They also greatly reduce the need of products derived from the petrochemical industry or in any case from fossil fuels, with a relative lack of climate-altering CO₂, as they generally use natural binders and also favour the use of locally sourced products, as they generally use natural binders and also favour the use of products of local origin, reducing the cost of transport. Lastly, they can grant an increase in social well-being, becoming massively used in the production of urban electric vehicles because they are light and ecological. Bio-composites find different types of applications in the automotive field (Karus and Vogt (2004), Bledzki et al. (2006)) such as: bodysHELLS

Table 3. Classification of natural fibres adapted from Gurunathan et al. (2015).

Natural (Animal/Mineral)	Synthetic (Organic/Inorganic fibres)	Cellulose/Lignocellulose	
Silk	Glass	Bast	Flax
Wool	Carbon		Hemp Jute
			Ramie
		Leaf	Banana Sisal
			Pineapple
		Seed	Cotton
		Fruit	Coconut

of microcar; e-bikes; FEV; automobile interiors; frame structural elements. Table 4 lists examples of bio-composites automotive applications.

Table 4. Bio-composites different applications in the automotive field.

Car manufacturer	Model and application
Audi	A2, A3, A4, A6, A8, Road star, Coupe: Seat back, side and back door panel, boot lining, hat rack, space tire lining
BMW	3, 5 and 7 series and others: door panels, headliner panel, boot lining, Seat back
Daimler/Chrysler	A, C, E, S class: Door panels, windshield, dashboard, business table, pillar cover panel; A class, Travego bus: exterior under body protection trim; M class: instrumental panel (Now in S class: 27 parts manufactured from bio fibers, weight 43kg)
FCA	Punto, Brava, Marea, Alfa Romeo 146, 156
Ford	Mondeo CD 162, Focus: Door panels, B-pillar, boot liner
Opel	Astra, Vectra, Zafira: Headliner panel, door panels, pillar coverpanel, instrumental panel
Peugeot	New model 406
Renault	Clio
Rover	Rover 2000 and others: Insulation, rear storage shelf/panel
SEAT	Door panels, seat back
Volkswagen	Golf, Passat Variant, Bora: Door panel, seat back, boot lid finish panel, boot liner
Volvo	C70, V70
Mitsubishi	Space star: Door panels; colt: Instrumental panels.

3 Electric Car Design Case Study

First of all, innovation patterns can make a relevant difference with reference to the economic performance, primarily at the industrial level and consequently at the institutional level. Based on the direct experience of Carbon Dream, the know-how developed through the opportunity to work with and for the most important brands in the hyper-car and Formula1 businesses, as well as through a structured R&D, represents the knowledge offered by the Company to its traditional and new customers to base a service of concurrent engineering to approach not only the design and industrialization of new products, but even to access new markets. Such model can truly make the difference between traditional manufacturing companies and an innovation company. When a similar model is applied successfully by a relevant number of industrial structures, its beneficial impacts can be opportunely scaled up to the institutional level.

Second of all, the institutional underpinnings of the present research come from three different major organizational forms: Regional, National and European. The present project has been financed by POR–CREO–FESR project call which is a 50% co-financing between local Region and Europe, in addition, the project has received tax incentives from the MISE (Ministry of Economic Development). This is due to a strong pressure from the public for sustainable development and respect for the environment which are main driving factors for the economic growth of the Country. The present case study considers a design industrialization and manufacturing of a micro car, with two seats, with shell in bio-composite materials, with electric propulsion and high electronic equipment on board for safe and eco-sustainable urban mobility. The micro car will be made in two

versions, one with carbon-linen shell and the other in bio- composite linen and denim-like cotton. Samples of such finishing are depicted in Fig. 2 for linen natural composites. Obviously, such composites are highly user-adjustable and strongly versatile according to the market they are introduced in either Western or Eastern. This demonstrate on one hand the high flexibility of the working processes for the realization of composite shells, on the other hand the high performance achievable with these advanced bio-materials, in terms of both mechanical properties (strength, weight reduction) and aesthetic yield and customization for the end user. In both versions the microcar will have a high technological content in terms of electronics for digital traction control and vehicle dynamics, sensors, man-machine interface with the possibility of remote monitoring of the state of charge of the accumulators and connection with the vehicle network. The use of carbon fiber in the bio-composite gives a high added value in terms of both high structural strength with reduced weights (with benefits in terms of energy efficiency), aerodynamic and aesthetic performance. Mechanical performances of carbon fiber were listed in Table 1. However, Table 1 lists only the mechanical performances of the single fibre and not of the composite material given by the mixture of matrix and fibers (Fig. 3).

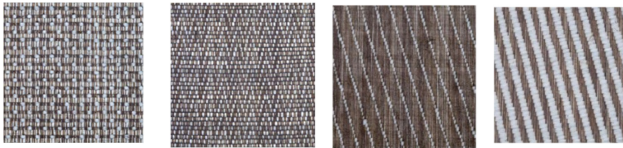


Fig. 2. Four samples of linen-based bio-composites (Fantuzzi et al. 2021). (Color figure online)



Fig. 3. Reels of natural fibres (left) and autoclaves (right). (Color figure online)

4 Electronic Systems for the Electric Car Design Case Study

The project also includes the design and production of a car shelter, or garage, with photovoltaic tiles to recharge the electric car battery with an off-grid system in order to define a whole zero emissions mobility system. The photovoltaic roof tiles used in the project are shown in Fig. 4. Lithium batteries of the latest generation and electronic systems with advanced algorithms for intelligent management of energy flows and monitoring of charge status and predictive diagnostics will be used for the energy accumulation of both the electric vehicle and the charging box. Taking advantage of an Internet-of-Things, it will be possible to implement a wireless network for vehicle to everything (V2X)

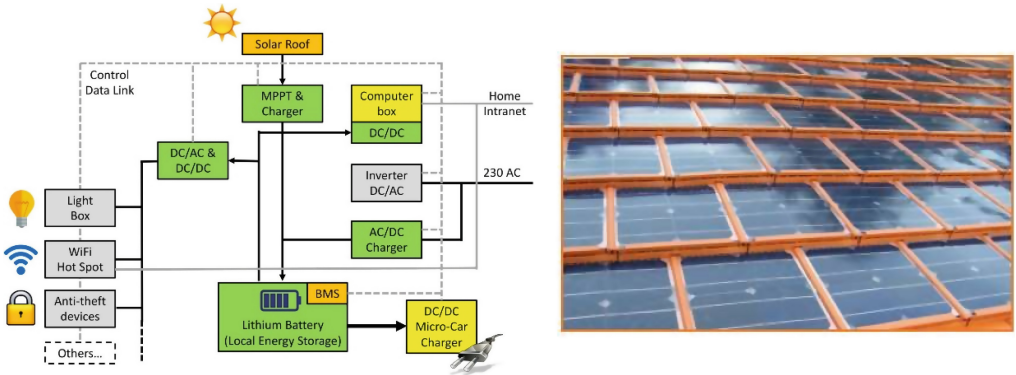


Fig. 4. The electric schematics of a garage system fitted with photovoltaic tiles (Benedetti et al. 2020). (Color figure online)

connectivity, among whose nodes there are the photovoltaic box, the mini-car and the remote control unit, to ensure remote monitoring and management of all elements of the network. The technologies developed in terms of off-the-box-shelving with photovoltaic cells, storage systems and intelligent energy management, traction and electrical actuation systems, Machine to Machine (M2M) and Human to Machine Interface (HMI) connectivity. Advanced materials made of carbon and bio-composite, have potential applications in terms of intelligent machines, mechatronics and advanced robotics, for digital industry scenario 4.0. This research, due to the relevance of the problems faced, assumes the characteristic of pure Public Good (complete indivisibility in the socio-economic benefits and not full exclusivity from the advantages deriving for each citizen) and of Merit Good in the economic meaning of the term as formulated by Musgrave in 1957 (good that the State and community must safeguard consumption beyond the “short-sighted” individual preferences of individual consumers). It will be recalled from the studies of Political Economy that this is a classic case of microeconomic failure or lack of equilibrium of the market, and that therefore the State, the Region and the EU must necessarily intervene; so, the regional intervention is well under co-financing with the EU. A simple schematics summarizing the present V2X model is shown in Fig. 4. All the system is controlled by the sunlight energy through solar roof tiles. This renewable energy is stored in a lithium battery storage which is kept for the micro-car. The IoT through control data link and a computer box interact with the home intranet and add-on devices such as WiFi hot spot, Light box, Anti-theft devices and so on.

5 Conclusions

This paper has shown an innovative circular, green, sustainable micro-mobility system which suits the needs of several markets world-wide. The use of bio-materials in modern automobiles is constantly increasing, as NFCs help to achieve precise weight targets safeguarding the structural properties necessary to passengers’ safety and therefore achieving better efficiencies in terms of energy consumption and FE vehicles autonomy. Furthermore, bio-composites offer enormous opportunities in the field of design

and are important both for the feelings to the human touch and aesthetic elegance, as well as in terms of car safety. All these applications have a green fingerprint and can be easily recycled once their function is over. When working in a high end market with a structured R&D, companies are able to provide industrial access to new applications for green products and markets to customers.

In fact, thanks to a greater respect for the environment, consumers and suppliers of the automotive industry and Original Equipment Manufacturers (OEMs) are nowadays increasingly looking for green materials alternative to those based on oil, that is a non-renewable source. The modern consumer is not short-sighted, and his decisions are increasingly eco-sustainable. At last, the use of NFC can finally give an important growth opportunity of GDP and GDP-pro-capita to many Developing Countries which still base their economy mainly on agriculture.

Acknowledgment. The research MI.CA.EL.A. of Carbon Dream Spa described in this paper was financially supported by RS - POR CREO 2014–2020 N. 2 adopted by decree no. 7165 of 24/05/2017 of the Tuscany Region.

References

- Fantuzzi, N., Baccocchi, M., Benedetti, D., Agnelli, J.: The Use of Sustainable Composites for the Manufacturing of Electric Cars. *C Composites/Elsevier* (2021)
- Agnelli, J., Benedetti, D., Gagliardi, A., Dini, P., Saponara, S.: Design of an off-grid photovoltaic carport for a full electric vehicle recharging. In: *IEEE International Conference on Environment and Electrical Engineering* (2020)
- Benedetti, D., Agnelli, J., Gagliardi, A., Dini, P., Saponara, S.: Design of a digital dashboard on low-cost embedded platform in a fully electric vehicle. In: *2020 IEEE International Conference on Environment and Electrical Engineering* (2020)
- Bledzki, A.K., Faruk, O., Sperber, V.E.: Cars from Bio-Fibres. *Macromol. Mater. Eng.* **291**(5), 449–457 (2006)
- Chadha, A.: Overcoming competence lock-in for the development of radical eco-innovations: the case of biopolymer technology. *Ind. Innov.* **18**(3), 335–350 (2011)
- Choi, H.: Technology-push and demand-pull factors in emerging sectors: evidence from the electric vehicle market. *Ind. Innov.* (2017). <https://doi.org/10.1080/13662716.2017.1346502>
- Cooke, P.: New economy innovation systems: biotechnology in Europe and the USA. *Ind. Innov.* **8**(3), 267–289 (2001)
- Cristaldi, G., Cicala, G.: Sviluppo di materiali compositi rinforzati con fibre naturali per l'ingegneria civile. PhD Thesis. University of Catania (2011)
- Faria, L.G.D., Andersen, M.M.: Sectoral dynamics and technological convergence: an evolutionary analysis of eco-innovation in the automotive sector. *Ind. Innov.* (2017). <https://doi.org/10.1080/13662716.2017.1319801>
- Gurunathan, T., Mohanty, S., Nayak, S.K.: A review of the recent developments in biocomposites based on natural fibres and their application perspectives. *Compos. Part A Appl. Sci. Manuf.* **77**, 1–25 (2015)
- International Monetary Fund. *World Economic Outlook Database* (2018)
- Karus, M., Vogt, D.: European hemp industry: Cultivation, processing and product lines. *Euphytica* **140**, 7–12 (2004)
- Lau, K.-T., Cheung, K.H.-Y., Hui D.: Preface. *Comp. Part B Eng.* **40**, 591–593 (2009)

- Leone, M.I., Belingheri, P.: The relevance of innovation for ethics, responsibility and sustainability. *Ind. Innov.* **24**(5), 437–445 (2017)
- Lucintel. Bio-Composites Market Report: Trends, Forecast and Competitive Analysis (2019). <https://www.lucintel.com/bio-composites-market.aspx>
- MAFF. The use of natural fibers in nonwoven structures for applications as automotive component substrates. R&D Report (2003)
- Maximize Market Research. Biocomposite Market – Global Industry Analysis By Fiber, By Plastic, By End Use, By Region for Forecast Period (2019–2026). Report ID 13055 (2019). <https://www.maximizemarketresearch.com/market-report/biocomposite-market/13055/>
- Oksman, K., Skrifvars, M., Selin, J.-F.: Natural fibres as reinforcement in polylactic acid (PLA) composites. *Comp. Sci. Tech.* **63**(9), 1317–1324 (2003)
- Rowell, R.M.: Opportunities for composites from agro-based resources. In: Rowell, R.M., et al. (eds.) *Cap. 7*. CRC press- Lewis Publishers, Boca Raton (1996)



Developing a Synthetic Dataset for Driving Scenarios

Jacopo Motta, Francesco Bellotti, Riccardo Berta, Alessio Capello, Marianna Cossu, Alessandro De Gloria, Luca Lazzaroni^(✉), and Stefano Bonora

Department of Electrical, Electronic and Telecommunication Engineering (DITEN),
University of Genoa, Via Opera Pia 11a, 16145 Genoa, Italy

{franz, berta, adg}@elios.unige.it, luca.lazzaroni@edu.unige.it

Abstract. Driving scenarios detection is an important aspect of the development of automated driving functions (ADF). Given the lack of publicly available datasets with driving scenario labels, we designed a toolchain for generating synthetic video datasets of driving scenarios, based on the OpenSCENARIO format, a well-established, public and vendor-independent standard. The experience reported in this paper shows the feasibility of a full end-to-end implementation of a workflow allowing designers to quickly create datasets for pre-training machine learning models. Video clips are recorded through a driving simulator which runs different sessions implementing variations of a pre-defined set of driving scenarios. The user specifies through a configuration file each parameter value range (e.g., vehicle speed, distance, weather conditions) that represent the intended variability within each scenario. Preliminary results show effectiveness of the approach and indicate directions on how to improve the system and reduce the need for human intervention in post-production.

Keywords: Driving scenarios · Synthetic datasets · Automated driving · CARLA driving simulator · Deep learning · Video classification

1 Introduction

Introduction of on-road automatic driving functions (ADF) requires that all possible situations are correctly handled within the operational design domain (ODD) of the vehicle. A promising approach is to formally characterize a set of scenarios that provide an abstraction for behavior of subjects (e.g., vehicles and pedestrians) in various driving situations [1]. A driving scenario describes the maneuvers of multiple entities in a given time window.

Detection of driving scenarios is important not only in real-time, during a car ride, when this high-level information can for instance set an operational mode of vehicle control, but also off-line, when vehicular signals are post-processed, for instance in order to assess effectiveness of automated driving [2].

Machine learning (ML) is now playing a key role in the implementation of ADF for automotive control (e.g., [3]), and can be used also for detecting driving scenarios. In order to learn high-level meaningful information from data, supervised ML systems

require datasets containing labeled samples. At present, we are not aware of publicly available datasets labeled with driving scenarios. This spurred us to create a toolchain to allow automotive system designers and ADF impact analysts to quickly create datasets usable for new research solutions. High quality real-world sensor datasets are fundamentals (e.g., [4]). However, they are very expensive and take a long time to create, and some of them are proprietary or shared in limited consortia. Virtual simulations (e.g., [5]) could offer a useful complement for various research cases. For example, they could support the set-up of ML models for early feasibility test and the pre-training of real-world automotive systems, that would then be tuned on smaller-size real-world datasets exploiting transfer learning [6]. Virtual datasets have still clear limits in terms of realism, likelihood, representativity, but they have a great generation controllability, at least for some factors (e.g., weather conditions), which is particularly relevant for corner cases, dramatically expensive to record in the real world.

In this context, we are interested in investigating feasibility of a toolchain for generating datasets of virtual-reality driving scenarios. Particularly, we intended to develop a workflow, usable by typical target users such as traffic data analysts or ML engineers, resumable in the following steps:

- Scenario instance set specification, in which the user fills a configuration file specifying the variation factors inside a chosen driving scenario type.
- Scenario generation, in which the system processes the user input to generate the scenario file.
- Scenario reproduction, in which each scenario instance is played and recorded in a driving simulator.
- Scenario check and utilization.

The remainder of the paper is organized as follows. The next session gives an overview of the fundamental building blocks of the toolchain. Section 3 shows the developed system architecture, while Sect. 4 present and discusses our initial achievements. Section 5 draws the conclusions on the work done.

2 Fundamental Building Blocks of the Toolchain

This section describes the three building blocks we have chosen as fundamental parts of the proposed toolchain: the OpenSCENARIO data model, the CARLA driving simulator and the Scenario Runner and Scenario Generation modules.

In order to create a dataset of driving scenario instances, we decided to describe each of them through the above introduced OpenSCENARIO, that defines a data model for the description of scenarios (and a corresponding XML-based file format, namely XOSC) [7]. The key element in a scenario is the *Storyboard*, that specifies the instructions for the dynamic behavior and interaction of *Entities* (objects that may change their location over time, such as *Vehicles* and *Pedestrians*). The *Storyboard* involves an initialization and one or more *Story* elements, further subdivided in *Acts*, that define conditional groups of *Actions*. *Acts* contain groups of *Maneuvers* that are activated and deactivated by *Triggers*. *Actions* serve to create or modify dynamic elements of a scenario, e.g., change in lateral

dynamics of a vehicle, change of the time of day, change simulator settings. *Actions* are singular elements which may be combined to create meaningful behavior in a scenario.

The second key component of the toolchain is the simulation tool. The simulation environment models behavior and physics of the involved participants. Our choice went on Car Learning to Act (CARLA), an open-source simulator developed (relying on the Unreal engine) for automated driving research, which supports OpenSCENARIO, offers a good environment realism and features a large community of users and developers [8]. The environment inside the simulator is composed of 3D models of static and dynamic objects. A basic controller has been implemented, regulating non-player vehicles (NPVs) behavior: follow the lane, respect traffic lights, speed limits and make decisions at intersections. Vehicles and pedestrians can detect and avoid each other. To increase visual variety, appearance of NPVs, when added to the simulation, is randomized. A variety of weather conditions and lighting regimes is available.

The third basic component of the system architecture is given by the Scenario Runner and Scenario Generation modules. Scenario Runner allows the definition and execution of automotive scenarios inside the CARLA simulator [9]. Scenarios can be defined through a Python file or following the OpenSCENARIO standard, allowing to reproduce inside the simulator scenarios defined in XOSC format. We thus needed a method to quickly produce a series of parametric scenarios in XOSC in order to build a varied dataset. To this end, we exploited the Python *scenariogeneration* package, a collection of libraries for generating OpenSCENARIO (.xosc) files [10].

3 System Architecture

Figure 1 gives an overview of the developed system architecture. For each targeted scenario type, the end user specifies the relevant parameters configuration in a json file (an example is provided in Table 1). The Scenario Generator processes this input to produce the corresponding set of .xosc files, then run in CARLA through the Scenario Runner. We developed a video recorder plug-in for CARLA, in order to record each clip, which is finally saved in a directory structure corresponding to the actual scenario label. We implemented the following scenario types:

- Free ride, in which the ego vehicle freely proceeds along its lane, without interacting directly with other vehicles.
- Following a lead vehicle, in which the ego vehicle imitates the maneuvers of the adversary, at a given distance.
- Lane change, in which the ego vehicle overtakes the adversary
- Approaching a static object, in which the ego vehicle approaches the adversary, which is stationary in the middle of the lane
- Cut-in, in which the ego vehicle proceeds regularly in its lane and its way is cut by an adversary coming at higher speed, from an adjacent lane.

The preparation of the .xosc files to be played by the Scenario Runner is performed in the Scenario Generation module through pyosex, a python wrapper for OpenSCENARIO. The module generates as many scenario instances as the permutations of the

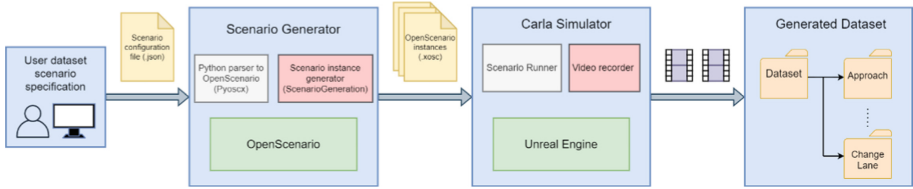


Fig. 1. Overview of the developed toolchain

Table 1. Example of a scenario type configuration file parameters

Parameter	Values	Description
Map	Town 01...06	CARLA map to be used in the scenario
Rain	True	Boolean to activate a rain scenario
Hour	15	Time of the day
Speed	10, 15, 20	Initial speed of the ego vehicle
Random pos	True	Activate random position of the ego vehicle
Initial dist	50, 80	Initial distance between ego and adversary
NPVs	4	Number of influent vehicles (NPVs, they represent random traffic surrounding the scenario)
Check lane	Right, left	Check on the adjacent lanes
Iterations	20	Number of random repetitions for each combination of parameter values

parameter values multiplied by the value of the repetition parameter. The randomization is related to the initial position of the ego vehicle and of the NPVs, which are context vehicles not involved in the scenario, representing traffic. Initial positions are generated exploiting the spawn points available in CARLA maps. However, spawn points are arbitrarily positioned in a map, by its author. This implies that NPV traffic generation does not exactly correspond to the parameter value set by the user, since some vehicles may never appear inside the scenario. We set a rule that spawn points chosen for NPVs generation should be close to the initial position of the ego vehicle, while avoiding conflicts with the nominal target scenario (e.g., avoid that an NPV is generated before the ego vehicle in the freeride scenario). But this is not sufficient to guarantee the actual presence of all the requested NPVs in each instantiated scenario. While some variability beyond the user specifications can be accepted, a more appropriate approach would involve the exploitation of waypoints, which are positioned at regular positions in a map and contain information about their context. We have explored this possibility when setting the initial position of the adversary vehicle, if present in a scenario, which must be accurate, in order to guarantee the correctness of a scenario. In this case, we developed an algorithm which follows the waypoints on the same lane and direction of the ego vehicle until a distance is reached on the map equal to the initial distance

parameter set by the user in the configuration file. The algorithm may be extended and employed also for the generation of NPVs.

A scenario instance may be perturbed also by presence of traffic lights. We thus inserted a trigger so that when the ego vehicle approaches an intersection, the traffic light switches to green. However, the direction taken by a vehicle is random. This, again, may perturb a scenario instance, and represents a limitation to be addressed. Another major limitation is given by the availability of three types of vehicles only.

4 Results

We have tested the toolchain by creating a set of synthetic datasets, whose main features are reported in Table 2. The last column reports the extremely short time needed to produce each dataset end-to-end, from user specifications to availability on disk. Inspecting the produced clips, we noticed some issues, mainly due to the limitations discussed in the previous section. A certain number of the produced clips must be discarded after a manual inspection. The removal percentage ranges from 10% to 50% and is typically related to the presence of intersections (at which, in the current implementation, we still let the vehicles take a random direction) and to not fully appropriate description of lanes in some of the CARLA maps. Also in this case, a better utilization of waypoints and investigation of the CARLA vehicle behavior logic should allow a much finer control over the produced clips.

As a final validation step, we have prepared a deep learning neural network classifier for recognizing the labeled clips. The network has three main modules: an initial time-distributed 2D convolutional step, which processes ten $112 * 112$ RGB frames per clip in order to detect relevant spatial patterns, followed by a gated recurrent unit (GRU) / long short-term memory (LSTM) layer, which processes the temporal relationships, and a final dense network for determining the classification [11]. The total number of trainable parameters is 5.5M, of which the vast majority (about 4.5M) concerns the initial convolutional network. The requested training time on Google Colab graphics processing units (GPUs) is about 20 min. Results (on a dataset of 4 s clips partitioned into an 80%–20% training and testing set) show a recognition rate of up to 95%, but

Table 2. Characterization of the generated datasets. Videos are 8 s long. 4, 2 and 1 s versions are also produced. The fifth scenario is cut-in.

ID	Map	Weather/time	Nr. scenario types	Nr. videos	Prod. time (hrs)
D1	Single	Sunny	4	800	6
D2	Multi	Sunny	4	800	6
D3	Single	Night	4	400	4
D4	Single	Rain	4	400	4
D5	Single	Fog	4	800	6
D6	Single	Sunny	5	400	4

there is a strong dependency on the training seed. Performance degrades with weather conditions, particularly fog (60%), and length of the clip (70%, 1 s).

5 Conclusions and Future Work

We designed a toolchain for the generation of synthetic video datasets of driving scenarios, based on the OpenSCENARIO format. The present implementation relies on CARLA and features five types of driving scenarios. While CARLA provides readings from other sensors as well, such as radars and lidars, we preferred focusing on the camera output. The main goal, in fact, was to have full end-to-end system implementation, including the test case of training and testing a neural network for scenario recognition, in order to have a preliminary assessment (considering the generation of simple scenario instances) of the feasibility of a workflow allowing designers to quickly create datasets for pre-training ML models.

Our intention is to increase the variability of the clips within the single scenarios, through a better control of the generation points and of the behavior of all the vehicles involved in a scenario instance, particularly increasing the level of traffic realism.

References

1. Weber, H., et al.: A framework for definition of logical scenarios for safety assurance of automated driving. *Traffic Inj. Prev.* **20**(sup1), S65–S70 (2019). <https://doi.org/10.1080/15389588.2019.1630827>
2. Bellotti, F., et al.: Managing big data for addressing research questions in a collaborative project on automated driving impact assessment. *Sensors* **2020**(20), 6773 (2020). <https://doi.org/10.3390/s20236773>
3. Mozaffari, S., Al-Jarrah, O.Y., Dianati, M., Jennings, P., Mouzakitis, A.: Deep learning-based vehicle behavior prediction for autonomous driving applications: a review. *IEEE Trans. Intell. Transp. Syst.* (2020). <https://doi.org/10.1109/TITS.2020.3012034>
4. Izquierdo, R., Quintanar, A., Parra, I., Fernández-Llorca, D., Sotelo, M.A.: The PREVENTION dataset: a novel benchmark for PREDiction of VEHicles iNTentIONS. In 2019 IEEE Intelligent Transportation Systems Conference (ITSC), pp. 3114–3121 (2019). <https://doi.org/10.1109/ITSC.2019.8917433>
5. Li, X., Wang, Y., Yan, K., Wang, F., Deng, F., Wang, F.-Y.: ParallelEye-CS: a new dataset of synthetic images for testing the visual intelligence of intelligent vehicles. *IEEE Trans. Veh. Technol.* **68**(10), 9619–9631 (2019). <https://doi.org/10.1109/TVT.2019.2936227>
6. Torrey, L., Shavlik, J.: Transfer learning. In: Soria Olivas, E., Martin Guerrero, J.D., Martinez-Sober, M., Magdalena-Benedito, J.R., Serrano Lopez, A.J. (eds.), *Handbook of Research on Machine Learning Applications and Trends: Algorithms, Methods, and Techniques*, IGI Global, Hershey (2010)
7. ASAM: OpenSCENARIO User Guide. <https://www.asam.net/standards/detail/openscenario/>. Accessed 22 Apr 2021
8. CARLA community CARLA Simulator, Version 0.9.11. <https://carla.readthedocs.io/en/latest/>. Accessed 22 Apr 2021

9. ScenarioRunner for CARLA. https://github.com/carla-simulator/scenario_runner. Accessed 21 Apr 2021
10. Scenariogeneration. <https://github.com/pyoscx/scenariogeneration>. Accessed 2 Apr 2021
11. Ferlet, P.: Training a neural network with an image sequence – example with a video as input. <https://medium.com/smileinnovation/training-neural-network-with-image-sequence-an-example-with-video-as-input-c3407f7a0b0f>. Accessed:5 Apr 2021



Smart On-Board Surveillance Module for Safe Autonomous Train Operations

G. Mezzina^{1,2}(✉), M. Barbareschi³, Salvatore De Simone³,
Alessandro Di Benedetto^{1,2}, G. Narracci^{1,2}, C. L. Saragaglia^{1,2}, D. Serra³,
and Daniela De Venuto^{1,2}

¹ Department of Electrical and Information Engineering, Politecnico di Bari, 70124 Bari, Italy
{giovanni.mezzina, alessandro.dibenedetto, giuseppe.narracci,
cataldoluciano.saragaglia, daniela.devenuto}@poliba.it

² CINI - Consorzio Interuniversitario Nazionale per l'Informatica, 00185 Roma, Italy

³ Rete Ferroviaria Italiana SpA, Ricerca e Sviluppo – Sviluppo Sistemi, 80021 Afragola (NA),
Italy

{m.barbareschi, sa.desimone, d.serra}@rfi.it

Abstract. This paper proposes the hardware implementation of a novel in-cabin module to realize a smart surveillance on autonomous train operations (ATO). The proposed smart surveillance module (SSM) consists of a 10-layers PCB carrier card holding a central computing core based on an Ultrazed-EG system on module (SoM). To interface cabin equipment, the SSM includes several common communication buses in the railway context, favoring its inclusion in already existing apparatus. The SSM also provides a multi-piggyback slot, whose layout is designed to allow the independent housing of two different communication boards (Profibus or Continuous Signal Repetition), when included in a redundant architecture. These functionalities have been condensed in a single Eurocard board with a height of only 4 horizontal pitches. To improve the fault detection, the SSM has been also supplied with several diagnostic interfaces concerning power management, debugging and diagnostic and SoM temperature monitoring.

1 Introduction

In recent years, the European Union (EU) pushed the railway sector towards the development of flexible, smart and real-time operating traffic management and decision support systems. In this context, the partnership Shift2Rail Joint Undertaking proposed a new technological revolution with the main objective of coordinating any innovative research to realize a common and interoperable European railway area [1]. The objective has been partially addressed through the introduction of the European Rail Traffic Management System/European Train Control System (ERTMS/ETCS). Through a standard definition, it guaranteed a transition, as smooth as possible, between European national borders. However, ERTMS/ETCS did not solve all the critical issues found by the Shift2Rail analysis and that concern the timing, logistics and efficient use of resources and fuels on the same route, but traveled by different operators, at different times and/or days. To bridge this gap, the EU railway industry started investing in systems that, making extensive use

of automation. In this framework, interoperable Automatic Train Operation (ATO) over European Train Control System (ETCS), briefly ATOoverETCS. To carry out automatic operations with high safety level, such as autonomous driving, dedicated surveillance systems are needed [2].

In this respect, we propose the design and the implementation on printed circuit board (PCB) of a novel and compact smart surveillance module (SSM) for the ATOs control. The proposed PCB interfaces and monitors most of the on-board subsystems currently housed in the train cabin, computes their responses and acts to preserve the overall system safety.

The paper is organized as follows. Section 2 outlines the framework in which the SSM fits. Section 3 describes the SSM architecture, and its PCB implementation. Section 4 provides experimental results, and, finally, Sect. 5 concludes the paper.

2 The SSM in ATOoverETCS

Figure 1 shows the overall system in which the proposed SSM is intended to be used. Specifically, Fig. 1 reports typical built-in train system equipment (e.g., ERMTS/ETCS and so on) and ATO on-board (ATO-OB) related apparatus (excluding the here-proposed SSM in red).

Currently, the operations of ATO-OB are carried out under the supervision of the ETCS that continuously calculates a safe maximum speed for each train, with cab signaling for the on-board systems. Alongside this supervision, systems that evaluate the reliability of each on-board subcomponents are also needed to prevent potentially hazardous events [3, 4]. Several algorithmic approaches have been proposed to improve the safety by analyzing data from each system subcomponent [5–7], but diagnostic interfaces result in cumbersome separated racks to be added to cabin equipment [4].

The SSM proposed in this paper partially bridges this gap proposing a compact solution with several roles: (i) operating vital control over all the other subsystems; (ii) running part of the ETCS logic when unavailable (following voluntary isolation of the on-board subsystem); (iii) implementing the communication master node with all the other modules of the platform; (iv) managing the communication with external devices; (v) running its own application logic, aiming to improve the train system safety.

To provide an overview of a possible architecture implementation of the SSM, Fig. 1 shows an example of interconnections map between the SSM and other train system devices such as Multifunction Vehicle Bus (MVB), Balise Transmission Module (BTM) and so on. The 2oo2 architecture [8] composed of two SSMs in inter-processor communication (IPC), and reported on the right in Fig. 1, constitutes the normal section only. Typically, within the dedicated rack, the normal section is replicated with a backup one, realizing a 1oo2 architecture.

Due to the limited space in cabin racks, it was required that all functionalities be included in a single PCB, rather than in several dedicated boards as in current practice. For this aim, the SSM has been designed with limited footprint and reconfigurable layout. In the same context, the SSM design also addressed the problem of lack of forced air cooling in the rack, by designing the mechanics of the PCB for housing a Conduction Cooling Assembly (CCA).

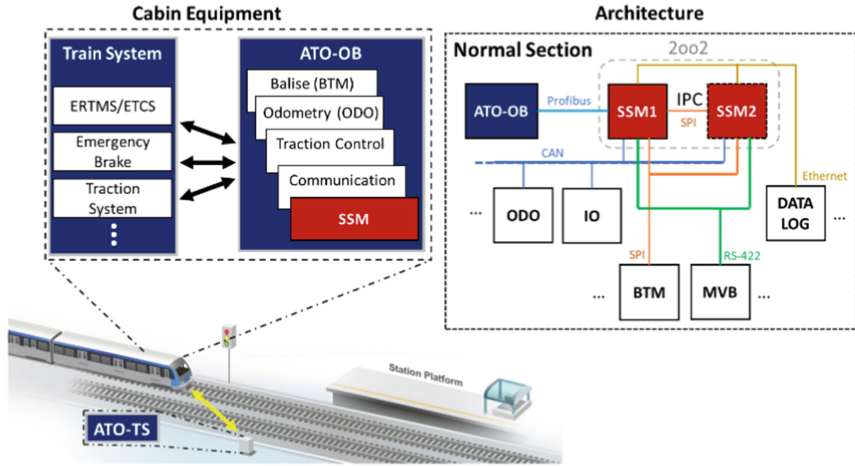


Fig. 1. SSM in ATOOverETCS with normal section architecture example

Since the SSM constitute the vital control core of the section, it has been designed to embed all the railway used communication protocols to interface the cabin equipment as per Fig. 1 [8].

To compute the responses and formalize the requests from/to other subsystems with a minimum latency, as well as to instantiate most of the communication drivers, the SSM has been design to house a processing core based on a highly performant system on module (SoM). Due to the parallel computational capabilities and the presence of several IPs for iocols, an Ultrazed-EG by AVNET[®] has been chosen.

The SSM has been also designed to embed, on the same board, two different configurations for communication with external apparatus without substantial hardware/software modifications. Indeed, the SSM should be able to instantiate a Profibus communication with ATO-OB and a Continuous Signal Repetition (CSR) interface for national signaling system.

The requested high safety level for the SSM board also prompted to the design of a fault detection section. For the purpose, the SSM has been supplied with diagnostic interface to check and provide information about the power rail status on the carrier card and on the SoM, as well as to control system temperature.

3 The Smart Surveillance Module (SSM)

3.1 The SSM Architecture

Figure 2 shows the block diagram view of the here-proposed SSM divided per main section blocks.

Processing Unit. As introduced in Sect. 2, the SSM computation core exploits a Ultrazed-EG (AES-ZU3EG) by AVNET[®] based on the Zynq Ultrascale+ by Xilinx [9–12]. This computation core runs the SSM application layer and manages all the communication drivers to interface other cabin submodules. The SoM provides access to

different pin banks, which allow the I/O access to signal lines from Processing System (PS) and signal lines from the FPGA Programmable Logic (PL).

Communication Interface. To realize the interconnections map provided in Fig. 1, the SSM embeds dedicated sections to SPI, CAN, RS-485, RS-422 and Ethernet interfaces [13]. All the communication lines are connected to the backplane through isolated transceivers (as per EN 50124-1 [13]) realizing the physical layer (PHY). Specifically, the SSM embeds six SPI interfaces for external communications with equipment like the Balise Transmission Module, and for the Inter-Processor Communication (IPC) mechanism, when the SSM is placed in a 2oo2 voting architecture (Fig. 1).

Two RS-485 and two RS-422 isolated transceivers have been also embedded in the architecture to interface the MVB, while a CANBus transceiver is provided for several inter-communications with odometry boards, watchdogs, and so on. SPI, RS-485 and RS-422 drivers are fully integrated on PL, while CAN driver exploits PS IPs. Six general purpose I/O lines are also provided to add vitality check lines for other systems. Finally, for the communication with the data log board an Ethernet section has been provided. It consists of 5 interfaces, 1 Gigabit Ethernet and 4 Fast Ethernet (100 Mbps).

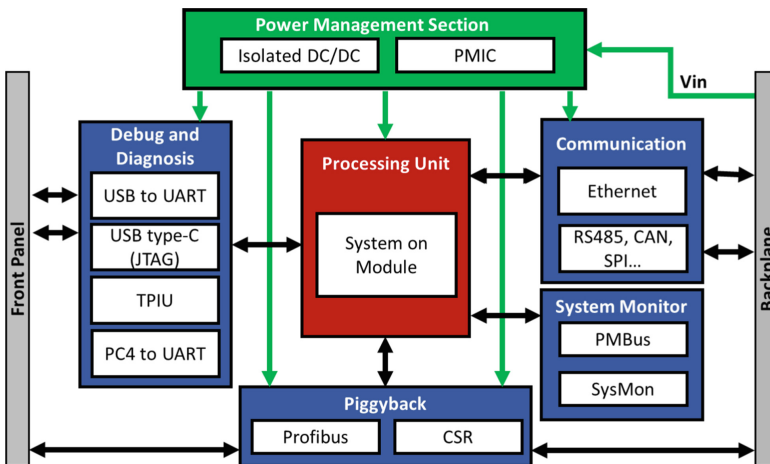


Fig. 2. SSM architecture overview

The Gigabit Ethernet media access control and the PHY layers are located on the SoM and are connected via the RGMII interface. The 4 Fast Ethernet interfaces are obtained through a Gigabit Ethernet Switch, connected to the SoM through an RGMII interface implemented on the PS of the UltraScale+ and programmable via I2C protocol.

Piggyback Section. According to Subset-058 [14], a surveillance module must communicate with the onboard ATO core via Profibus. For this reason, the SSM includes a piggyback slot designed to hold the insertion of a Profibus DP Master device. At the same time, the slot must also be able to permit the insertion of a piggyback board dedicated to the train driving control, the CSR, and widely used in the Italian signaling system. The two piggyback modules must be interchangeable.

Power Management. The SSM main power source is provided via the backplane and typically consists of +24 V generated from an external module that oversees supply all the boards included in the rack. The first stage of the designed SSM power delivery network consists of an isolated DC/DC converter that step-down the +24 V to +12 V. Starting from this isolated +12 V, used as SoM working voltage, all the voltages useful for the SSM internal functioning (i.e., +5 V, +3.3 V, +1.8 V, +0.85 V, +1.8 V linear) are derived. For this purpose, a power management IC (PMIC) has been exploited. The implemented PMIC supports 5 power rails, of which 4 independent switching regulators and one linear regulator. This device is programmable in terms of channels characteristics and soft-start conditions, through an I²C/PMBus. The UltraZed-EG SoM provides PMIC with voltage sense feedback to adjust voltages on each channel [15].

System Monitor. To facilitate fault detection, the SSM provides the user with two main interfaces: the SysMon and the PMBus [16]. The first one allows monitoring the output of on-SoM analog-to-digital converters (ADC), as well as on-chip sensors that are used to sample external voltages and to monitor FPGA on-die operating conditions, such as temperature and supply voltage levels. While the PMBus is used to monitor, in real-time, the current absorption on the PMIC placed on the carrier card, as well as on the two PMICs placed on the AVNET[®] SoM. Since UltraScale+ contains embedded processor systems with multiple ARM cores, the SSM has been also designed to provide a parallel Trace Port Interface Unit (TPIU) to export trace data.

3.2 SSM Hardware Implementation

The above described SSM architecture has been implemented on a 10-layers PCB, shown in Fig. 3. For the sake of clarity, Fig. 3 also shows labels for each main section.

PCB Footprint. The realized PCB complies with the long Eurocard standard, condensing the above-reported functionalities in a small footprint of 100 mm × 220 mm and a height of 4 HP. All the chosen COTS components comply with the standard for railway application and ensure an on-top space of at least 3 mm to allow the insertion of CCA for passive dissipation in still air (typical cabin rack environment).

Different Operation Modes. The SSM is designed to work both when inserted into the rack and when removed for testing and maintenance. For instance, the debugging section provides a USB-C at the front panel for the JTAG interface, and, at the same time, PC4 access onboard is also provided. The Debug also includes an 8 pin DIP slide switch and several LEDs connected to PS/PL, to provide test and debug functionalities.

Piggyback Shared Layout. The multi-piggyback slot is designed to house two independent modules with minimum hardware variation. The proper combination of a set of zero-ohm resistors and solder joints has been used to create the custom configurations. The lines toward PL are shared and the software running on the SoM is capable of recognizing the inserted piggyback. It selects between a dual-port memory configuration when the Profibus piggyback is inserted, or general digital I/O and a local SPI when CSR is housed. To ensure a minimum hardware variation between the two modules, a customized DB-15 receptacle layout has been also realized and provided on the front panel.

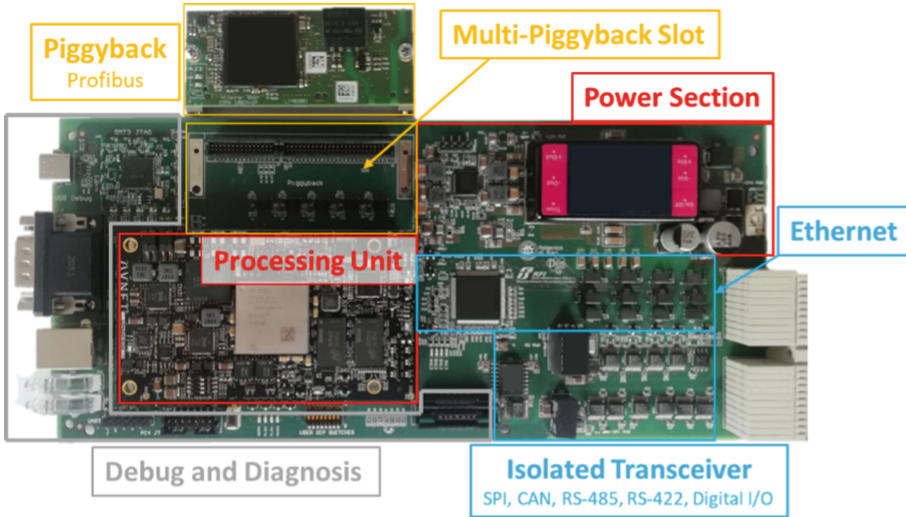


Fig. 3. PCB implementation of SSM

4 Experimental Results

4.1 Transmission Test

For testing purposes, the Ultrazed-EG SoM, which constitutes the processing core of the SSM, has been programmed through Vivado 2018.2 and related Vivado SDK to instantiate all the needed drivers (SPI, CAN, RS-485, RS-422, Ethernet and Profibus) on PS and PL [17, 18]. All the instantiated driver IPs exploit standard library for the messages exchange, but interface a dedicated middleware that allow a safe validation of the message.

In this preliminary test, once the SSM is turned on, a simple test message is created and sent via all the isolated communication branches. The signals from the backplane connector are transmitted to an oscilloscope with decoding function.

Figure 4a shows an example of the SSM transmission of an increasing counter value for an SPI and a CANbus interface. The decoded message is displayed in the table on the top of the oscilloscope screen.

4.2 FPGA Resources and Power Consumption

The Vivado block design related to the drivers initialization exploits Advanced eXtensible Interface (AXI) based IPs on the PL. A resource utilization analysis returned that all the communication interfaces of the SSM occupy 15.54% (10968/70560) of the total available Look-Up Tables (LUTs), only 2.5% (720/28800) of LUTRAM (LUTs used as RAM), 10.20% (14406/141120) of the available Flip-Flops, the 3.73% (8/216) of the block RAM (BRAM), 12/196 (6.12%) buffers for global clock distribution (BUFG), and the 56.66% (102/180) I/O pins. Specifically, 38 lines are dedicated to the DPM for the Profibus/CSR interface, 24 lines both for SPIs and Ethernet connections, 6 lines for the

general purpose digital I/O, 2 for CANbus and 4 for both RS-485 and RS-422. Figure 4b shows a device view of the implemented communication section.

Figure 4c shows the UltraScale+ on-chip power consumption when all the communication branches are active with a typical temporal sequence. Results show how only the communication involvement consumes a total of 2.406 W of which the 13% (318 mW) concerned static power dissipation, while the 87% (2.088W) are due to dynamic power consumption. Details on the consumption is shown in Fig. 3c.

Finally, a test on the overall SSM power consumption in the same condition, realized via PMBus monitoring for 3 h, shown a total power consumption of $9.46\text{ W} \pm 0.64\text{ W}$ with a global temperature increment of $+4^\circ$ on the PMIC, $+6^\circ$ on the Ethernet Switch and $+19^\circ$ on the UltraScale + [15]. These results stress the need for a dedicated CCA.

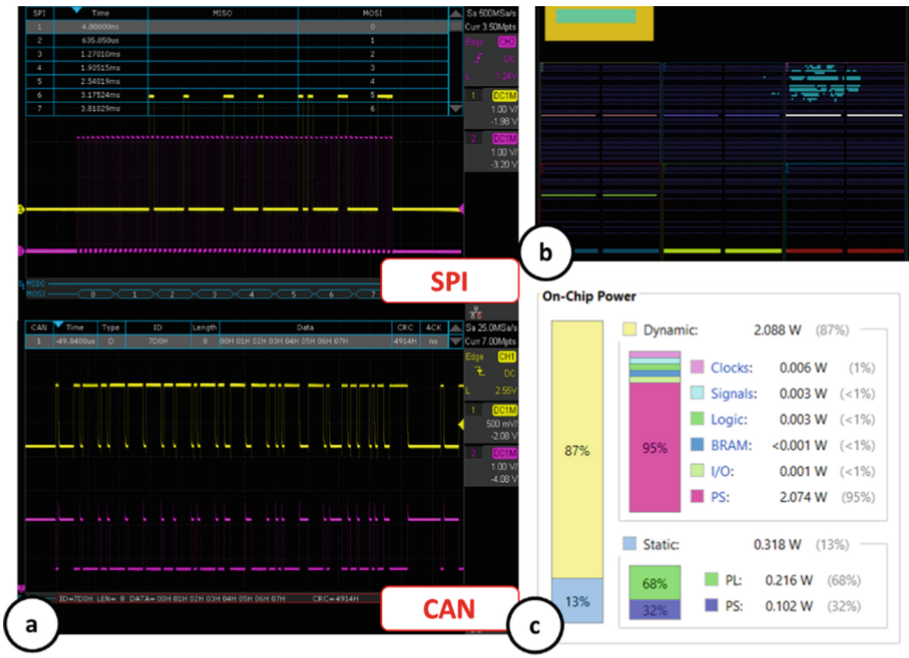


Fig. 4. (a) Example of interface test (SPI, CAN); (b) on-chip used area for drivers implementation; (c) power consumption on UltraScale+ FPGA for drivers implementation.

5 Conclusions

In this paper, the PCB implementation of an SSM for the ATO over ETCS has been presented. To realize the surveillance, the proposed module is supplied with a number of interfaces to monitor the onboard subsystems housed in the train cabin, computes their responses, and acts to preserve the overall system safety. For this purpose, six SPIs, one CAN, two RS-485, two RS-422, 5 Ethernet interfaces have been implemented. To allow

the maximum reconfigurability, SSM embedded a multi-piggyback slot, whose shared PCB layout that through a set of zero ohm resistors and solder joints is able to house two different communication boards, the Profibus and the CSR. The SSM integrates all the communication and computing capabilities, typically distributed over several dedicated boards, in a single Eurocard board with a height compatible with a CCA insertion. Experimental results on the realized PCB showed that the fully communication section with drivers occupies a low number of resources, only 15.54% LUTs, 2.5% LUTRAM, 10.20% of the available Flip-Flops, the 3.73% of BRAM. Moreover, it occupies the 56.66% of the total I/O pins. Experimental tests on the SSM power consumption (PMBus monitoring on 3 h) showed a power consumption of 9.46 W of which 2.406 W related to the Ultrascale+ dissipation corroborating the need for a conduction cooling system, realized via the predisposition of the SSM board to CCA holding.

References

1. Ďuračik, M., Kršák, E., Meško, M., Ružbarský, J.: Software architecture of Automatic Train Operation. In: 2019 IEEE 15th International Scientific Conference on Informatics, pp. 000051–000054 (2019)
2. Hwang, J., Jo, H.: RAMS management and assessment of railway signaling system through RAM and safety activities. In: 2008 International Conference on Control, Automation and Systems, pp. 892–895 (2008)
3. Wang, J., Wang, J., Roberts, C., Chen, L.: Parallel monitoring for the next generation of train control systems. *IEEE Trans. Intell. Transp. Syst.* **16**(1), 330–338 (2015). <https://doi.org/10.1109/TITS.2014.2332160>
4. Cai, G., Zhao, J., Song, Q., Zhou, M.: System architecture of a train sensor network for automatic train safety monitoring. *Comput. Indust. Eng.* **127**, 1183–1192 (2019). ISSN 0360–8352
5. Meng, L.H., et al.: Evaluation of reliability of urban rail train traction inverter system. *J. China Railway Soc.* **36**(9), 34–38 (2014)
6. Huo, Z., Zhang, Y., Francq, P., Shu, L., Huang, J.: Incipient fault diagnosis of roller bearing using optimized wavelet transform based multi-speed vibration signatures. *IEEE Access* **5**, 19442–19456 (2017). <https://doi.org/10.1109/ACCESS.2017.2661967>
7. Pugel, A.E., et al.: Use of the surgical safety checklist to improve communication and reduce complications. *J. Infect. Public Health* **8**, 3, 219–225 (2015)
8. Bertieri, D., Ceccarelli, A., Zoppi, T., Mungiello, I., Barbareschi, M., Bondavalli, A.: Development and validation of a safe communication protocol compliant to railway standards. *J. Braz. Comput. Soc.* **27**(1), 1–26 (2021). <https://doi.org/10.1186/s13173-021-00106-w>
9. Akbariavaz, K., Fazel, S.S., Khosravi, M.: Fully FPGA-based implementation of a modified control strategy for power electronic transformer in railway traction applications. *IET Power Electron.* **14**(11), 1904–1919 (2021). <https://doi.org/10.1049/pel2.12158>
10. De Venuto, D., Annese, V.F., Mezzina, G.: An embedded system remotely driving mechanical devices by P300 brain activity. In: Design, Automation and Test in Europe Conference and Exhibition (DATE), pp. 1014–1019 (2017). <https://doi.org/10.23919/DATE.2017.7927139>
11. De Venuto, D., Ohletz, M.J.: On-chip test for mixed-signal ASICs using two-mode comparators with bias-programmable reference voltages. *J. Electron. Test.* **17**, 243–253 (2001). <https://doi.org/10.1023/A:1013377811693>

12. De Venuto, D., Annese, V.F., Mezzina, G., Ruta, M., Di Sciascio, E.: Brain-computer interface using P300: a gaming approach for neurocognitive impairment diagnosis. In: 2016 IEEE International High Level Design Validation and Test Workshop (HLDVT), pp. 93–99 (2016). <https://doi.org/10.1109/HLDVT.2016.7748261>
13. CEI EN 50124- Railway Applications - Insulation Coordination
14. UNISIG - Subset-058. FFFIS STM Application Layer
15. Blagojevic, M., Kayal, M., Gervais, M., De Venuto, D.: SOI Hall-sensor front end for energy measurement. *IEEE Sens. J.* **6**(4), 1016–1021 (2006). <https://doi.org/10.1109/JSEN.2006.877996>
16. White, R.V., Durant, D.: Understanding and using PMBus/spl trade/data formats. In: Twenty-First Annual IEEE Applied Power Electronics Conference and Exposition, APEC 2006. IEEE (2006)
17. De Venuto, D., Annese, V.F., Mezzina, G., Defazio, G.: FPGA-Based embedded cyber-physical platform to assess gait and postural stability in parkinson's disease. *IEEE Trans. Compon. Pack. Manuf. Technol.* **8**(7), 1167–1179 (2018). <https://doi.org/10.1109/TCPMT.2018.2810103>
18. De Venuto, D., Rabaey, J.: RFID transceiver for wireless powering brain implanted microelectrodes and backscattered neural data collection. *Microelectron. J.* **45**(12), 1585–1594 (2014). <https://doi.org/10.1016/j.mejo.2014.08.007>

Author Index

A

Abid, Usman, 139
Addabbo, Tommaso, 180
Agnelli, Jacopo, 300
Akkad, Ghattas, 76
Ali, Haydar Hajj, 280
Androulakis, Michael, 212
Apicella, Tommaso, 86
Armenise, M. N., 45
Ayoubi, Rafic, 76

B

Baggio, Federico, 155
Baldanzi, Luca, 31
Barbareschi, M., 317
Baronti, Federico, 107, 120, 126
Bassoli, Marco, 53
Bellotti, Francesco, 229, 249, 310
Benedetti, David, 300
Berta, Riccardo, 229, 249, 310
Berzi, Lorenzo, 24
Biagioni, Ian, 126
Bianchi, Valentina, 53
Bigongiari, Franco, 194, 200
Billé, Fulvio, 173
Boni, Andrea, 53
Boni, Enrico, 18, 24
Bonora, Stefano, 310
Brunelli, Davide, 139, 155
Brunetti, G., 45
Bruschi, Paolo, 256

C

Canepa, Alessio, 86
Capello, Alessio, 229, 310
Cardarilli, Gian Carlo, 39
Careri, Maria, 53
Carlioni, Andrea, 120
Carrato, Sergio, 69, 93, 173
Casalini, Iacopo, 221
Casino, Fran, 114
Casu, Mario R., 243
Catania, Alessandro, 256
Cervetto, Marcos, 69
Cesario, Paolo, 249
Cicalini, Mattia, 256
Cicuttin, Andres, 69, 93
Cimatti, Alessandro, 149
Ciminelli, C., 45
Cococcioni, Marco, 61
Comino, Corrado, 187
Constà, Stefano, 120
Conti, Massimo, 206, 236, 273
Cossu, Marianna, 229, 310
Crespo, Maria Liz, 69, 93
Crocetti, Luca, 31

D

De Gloria, Alessandro, 229, 249, 310
De Munari, Ilaria, 53
De Simone, Salavatore, 317
De Venuto, Daniela, 317
Di Benedetto, Alessandro, 317
Di Credico, Gioia, 173
Di Matteo, Stefano, 31

Di Nunzio, Luca, 39
 Di Rienzo, Roberto, 107, 120, 126
 Donati, Massimiliano, 100

E

Elhanashi, Abdussalam, 1, 10
 ElHassan, Bachar, 76

F

Falaschi, Francesco, 31
 Fantuzzi, Nicholas, 300
 Fanucci, Luca, 31, 100, 187, 194, 200, 221
 Fazzolari, Rocco, 39
 Florian, Werner, 69
 Fort, Ada, 180
 Fortunati, Simone, 53
 Fronza, Luca, 249

G

Gaiduk, Maksym, 206
 García, Luis G., 69
 García, Luis Guillermo, 93
 Gastaldo, Paolo, 86
 Gemma, Luca, 139
 Giannetti, Filippo, 286
 Giannetto, Marco, 53
 Gianoncelli, Alessandra, 173
 Giardino, Daniele, 39
 Giorgio, A., 45
 Giuffrida, Gianluca, 100
 Grammatikakis, Miltos D., 212, 262
 Grasso, Francesco, 24
 Guadagno, Matteo, 39
 Guzzi, Francesco, 173

H

Hassan, Mohamad Hajj, 280

I

Ibrahim, Ali, 280
 Inaty, Elie, 76

K

Kourousias, George, 173
 Kypraios, Lefteris, 262

L

Lazzaroni, Luca, 229, 310
 López-Aguilar, Pablo, 132
 Lottici, Vincenzo, 286

M

Madrid, Natividad Martínez, 206
 Manfredini, Giuseppe, 256
 Mansoori, Mohammad Amir, 243

Mansour, Ali, 76
 Marchi, Edgardo, 69
 Marini, Marco, 221
 Marocco, G., 45
 Martina, Maurizio, 293
 Martínez-Ballesté, Antoni, 114
 Maya-López, Armando, 114
 Mezzina, G., 317
 Molina, Romina Soledad, 93
 Morales, Iván René, 93
 Moretti, Riccardo, 180
 Motta, Jacopo, 310
 Mouzakititis, Nikos, 262
 Mugnaini, Marco, 180

N

Nannipieri, Pietro, 31
 Narracci, G., 317

O

Orcioni, Simone, 206, 236
 Ortega, Juan Antonio, 206
 Osman, Anas, 139

P

Pacini, Tommaso, 187
 Palopoli, Luigi, 164
 Panicacci, Silvia, 100
 Papatheodorou, Nikos, 262
 Pasquali, Manlio, 120
 Passerone, Roberto, 149, 164
 Penzel, Thomas, 206
 Perez, Hector, 93
 Perotto, Matteo, 139
 Piotto, Massimo, 256
 Prastowo, Tadeus, 164
 Pugi, Luca, 18, 24

R

Ragusa, Edoardo, 86
 Ramponi, Giovanni, 93
 Rapuano, Emilio, 187
 Re, Marco, 39
 Ria, Andrea, 256
 Romani, Andrea, 194, 200
 Roncella, Roberto, 107, 120, 126
 Rossi, Federico, 61
 Ruffaldi, Emanuele, 61
 Ruo Roch, Massimo, 293

S

Saletti, Roberto, 107, 120, 126
 Sapienza, Fabiola, 286
 Saponara, Sergio, 1, 10, 31, 61, 286, 300
 Saragaglia, C. L., 317

Savi, Raffaele, [24](#)
Seepold, Ralf, [206](#)
Serra, D., [317](#)
Shah, Ayub, [164](#)
Simonte, Gianluca, [126](#)
Solanas, Agusti, [114](#), [132](#)
Spanò, Sergio, [39](#)
Stacchiotti, Luca, [236](#)

T

Tierno, Antonio, [149](#)
Torrise, Alessandro, [155](#)
Turri, Giuliano, [149](#)

V

Valinoti, Bruno, [69](#)
Valle, Maurizio, [280](#)
Venturi, Alessio, [18](#)
Verani, Alessandro, [107](#)
Vignoli, Valerio, [180](#)
Villon, Jorge Leonardo Quimi, [229](#)

W

Wang, Hongjun, [1](#)
Weber, Lucas, [206](#)

Z

Zhang, Deliang, [1](#)
Zheng, Qinghe, [1](#)