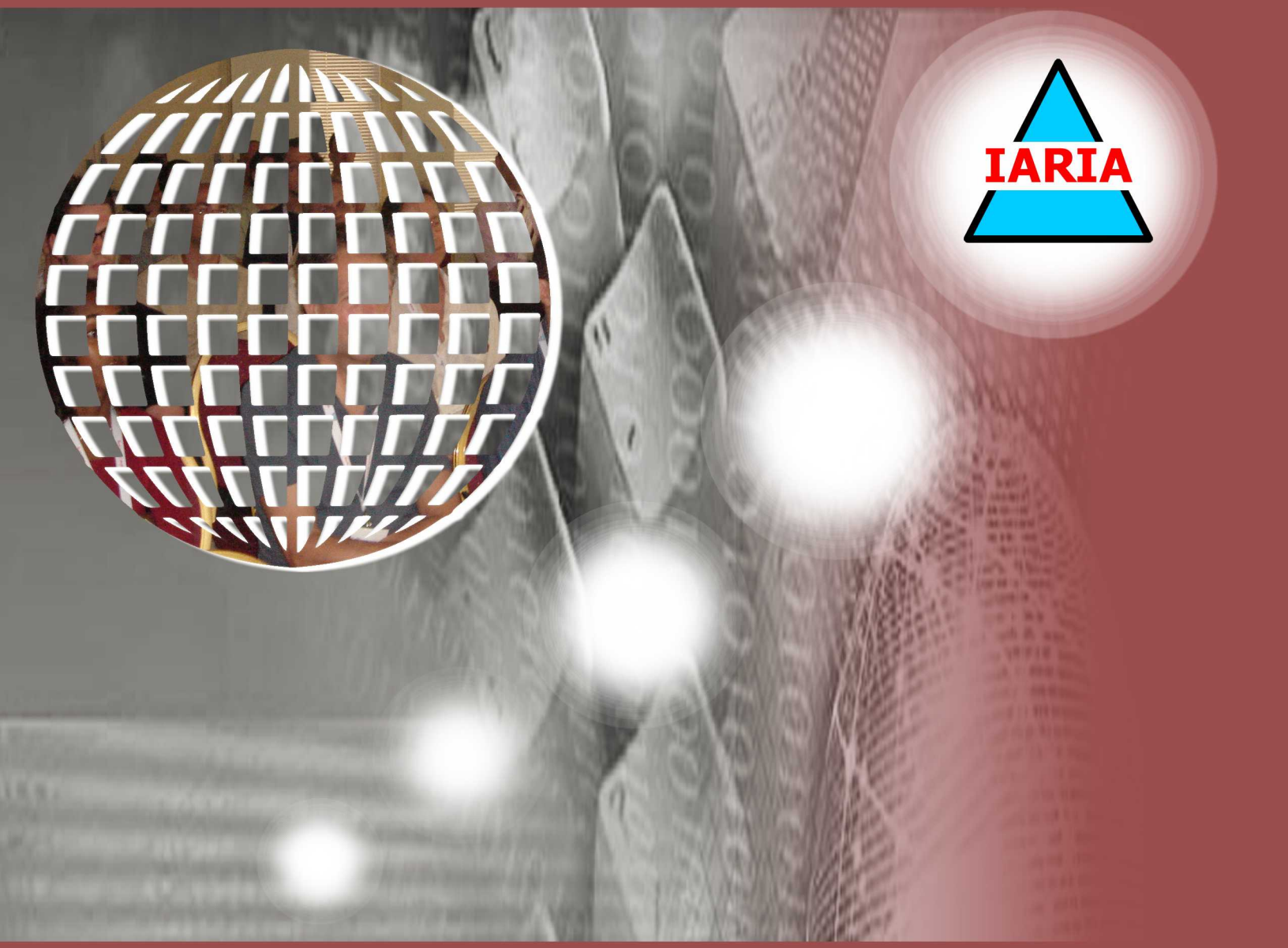
International Journal on Advances in Telecommunications



2013 vol. 6 nr. 3&4

The *International Journal on Advances in Telecommunications* is published by IARIA.
ISSN: 1942-2601
journals site: http://www.iariajournals.org
contact: petre@iaria.org

Sponsored by IARIA
www.iaria.org

International Journal on Advances in Telecommunications
Volume 6, Number 3 & 4, 2013

Ilija Basicevic, University of Novi Sad, Serbia
Carlos Becker Westphall, Federal University of Santa Catarina, Brazil
Mark Bentum, University of Twente, The Netherlands
David Bernstein, Huawei Technologies, Ltd., USA
Eugen Borgoci, University "Politehnica"of Bucharest (UPB), Romania
Fernando Boronat Seguí, Universidad Politecnica de Valencia, Spain
Christos Bouras, University of Patras, Greece
David Boyle, Tyndall National Institute, University College Cork, Ireland
Martin Brandl, Danube University Krems, Austria
Julien Broisin, IRIT, France
Dumitru Burdescu, University of Craiova, Romania
Andi Buzo, University "Politehnica" of Bucharest (UPB), Romania
Shkelzen Cakaj, Telecom of Kosovo / Prishtina University, Kosovo
Enzo Alberto Candreva, DEIS-University of Bologna, Italy
Rodrigo Capobianco Guido, University of Sao Paulo, Brazil
Hakima Chaouchi, Telecom SudParis, France
Emmanuel Chaput, IRIT-CNRS, France
Silviu Ciochina, Universitatea Politehnica din Bucuresti, Romania
José Coimbra, Universidade do Algarve, Portugal
Hugo Coll Ferri, Polytechnic University of Valencia, Spain
Denis Collange, Orange Labs, France
Noel Crespi, Institut TELECOM SudParis-Evry, France
Leonardo Dagui de Oliveira, Escola Politécnica da Universidade de São Paulo, Brazil
Gerard Damm, Alcatel-Lucent, USA
Danco Davcev, University for Information Science & Technology "St. Paul the Apostle", Ohrid, Macedonia
Francescantonio Della Rosa, Tampere University of Technology, Finland
Chérif Diallo, Consultant Sécurité des Systèmes d'Information, France
Klaus Drechsler, Fraunhofer Institute for Computer Graphics Research IGD, Germany
Jawad Drissi, Cameron University , USA
António Manuel Duarte Nogueira, University of Aveiro / Institute of Telecommunications, Portugal
Alban Duverdier, CNES (French Space Agency) Paris, France
Nicholas Evans, EURECOM, France
Fabrizio Falchi, ISTI - CNR, Italy
Mário F. S. Ferreira, University of Aveiro, Portugal
Bruno Filipe Marques, Polytechnic Institute of Viseu, Portugal
Robert Forster, Edgemount Solutions, USA
John-Austen Francisco, Rutgers, the State University of New Jersey, USA
Kaori Fujinami, Tokyo University of Agriculture and Technology, Japan
Shauneen Furlong , University of Ottawa, Canada / Liverpool John Moores University, UK
Ana-Belén García-Hernando, Universidad Politécnica de Madrid, Spain
Bezalel Gavish, Southern Methodist University, USA
Christos K. Georgiadis, University of Macedonia, Greece
Mariusz Glabowski, Poznan University of Technology, Poland
Katie Goeman, Hogeschool-Universiteit Brussel, Belgium
Hock Guan Goh, Universiti Tunku Abdul Rahman, Malaysia
Pedro Gonçalves, ESTGA - Universidade de Aveiro, Portugal

Valerie Gouet-Brunet, Conservatoire National des Arts et Métiers (CNAM), Paris
Christos Grecos, University of West of Scotland, UK
Stefanos Gritzalis, University of the Aegean, Greece
William I. Grosky, University of Michigan-Dearborn, USA
Vic Grout, Glyndwr University, UK
Xiang Gui, Massey University, New Zealand
Huaqun Guo, Institute for Infocomm Research, A*STAR, Singapore
Song Guo, University of Aizu, Japan
Ching-Hsien (Robert) Hsu, Chung Hua University, Taiwan
Javier Ibanez-Guzman, Renault S.A., France
Lamiaa Fattouh Ibrahim, King Abdul Aziz University, Saudi Arabia
Theodoros Iliou, University of the Aegean, Greece
Mohsen Jahanshahi, Islamic Azad University, Iran
Antonio Jara, University of Murcia, Spain
Carlos Juiz, Universitat de les Illes Balears, Spain
Adrian Kacso, Universität Siegen, Germany
György Kálmán, ABB AS, Norway
Eleni Kaplani, Technological Educational Institute of Patras, Greece
Behrouz Khoshnevis, University of Toronto, Canada
Ki Hong Kim, ETRI: Electronics and Telecommunications Research Institute, Korea
Atsushi Koike, Seikei University, Japan
Ousmane Kone, UPPA - University of Bordeaux, France
Dragana Krstic, University of Nis, Serbia
Archana Kumar, Delhi Institute of Technology & Management, Haryana, India
Romain Laborde, University Paul Sabatier (Toulouse III), France
Massimiliano Laddomada, Texas A&M University-Texarkana, USA
Thomas D. Lagkas, University of Western Macedonia, Greece
Wen-Hsing Lai, National Kaohsiung First University of Science and Technology, Taiwan
Zhihua Lai, Ranplan Wireless Network Design Ltd., UK
Jong-Hyouk Lee, INRIA, France
Wolfgang Leister, Norsk Regnesentral, Norway
Elizabeth I. Leonard, Naval Research Laboratory - Washington DC, USA
Jia-Chin Lin, National Central University, Taiwan
Chi (Harold) Liu, IBM Research - China, China
Haibin Liu, China Aerospace Science and Technology Corporation, China
Diogo Lobato Acatauassu Nunes, Federal University of Pará, Brazil
Andreas Loeffler, Friedrich-Alexander-University of Erlangen-Nuremberg, Germany
Michael D. Logothetis, University of Patras, Greece
Renata Lopes Rosa, University of São Paulo, Brazil
Hongli Luo, Indiana University Purdue University Fort Wayne, USA
Christian Maciocco, Intel Corporation, USA
Dario Maggiorini, University of Milano, Italy
Maryam Tayefeh Mahmoudi, Research Institute for ICT, Iran
Krešimir Malarić, University of Zagreb, Croatia
Zoubir Mammeri, IRIT - Paul Sabatier University - Toulouse, France
Herwig Mannaert, University of Antwerp, Belgium

Mariusz Skrocki, Orange Labs Poland / Telekomunikacja Polska S.A., Poland
Leonel Sousa, INESC-ID/IST, TU-Lisbon, Portugal
Liana Stanescu, University of Craiova, Romania
Cosmin Stoica Spahiu, University of Craiova, Romania
Radu Stoleru, Texas A&M University, USA
Young-Joo Suh, POSTECH (Pohang University of Science and Technology), Korea
Hailong Sun, Beihang University, China
Jani Suomalainen, VTT Technical Research Centre of Finland, Finland
Fatma Tansu, Eastern Mediterranean University, Cyprus
Ioan Toma, STI Innsbruck/University Innsbruck, Austria
Božo Tomas, HT Mostar, Bosnia and Herzegovina
Piotr Tyczka, Poznan University of Technology, Poland
John Vardakas, University of Patras, Greece
Andreas Veglis, Aristotle University of Thessaloniki, Greece
Luís Veiga, Instituto Superior Técnico / INESC-ID Lisboa, Portugal
Calin Vladeanu, "Politehnica" University of Bucharest, Romania
Natalija Vlajic, York University - Toronto, Canada
Benno Volk, ETH Zurich, Switzerland
Krzysztof Walczak, Poznan University of Economics, Poland
Krzysztof Walkowiak, Wroclaw University of Technology, Poland
Yang Wang, Georgia State University, USA
Yean-Fu Wen, National Taipei University, Taiwan, R.O.C.
Bernd E. Wolfinger, University of Hamburg, Germany
Riaan Wolhuter, Universiteit Stellenbosch University, South Africa
Yulei Wu, Chinese Academy of Sciences, China
Mudasser F. Wyne, National University, USA
Gaoxi Xiao, Nanyang Technological University, Singapore
Bashir Yahya, University of Versailles, France
Abdulrahman Yarali, Murray State University, USA
Mehmet Erkan Yüksel, Istanbul University, Turkey
Pooneh Bagheri Zadeh, Staffordshire University, UK
Giannis Zaoudis, University of Patras, Greece
Liaoyuan Zeng, University of Electronic Science and Technology of China, China
Rong Zhao , Detecon International GmbH, Germany
Zhiwen Zhu, Communications Research Centre, Canada
Martin Zimmermann, University of Applied Sciences Offenburg, Germany
Piotr Zwierzykowski, Poznan University of Technology, Poland

International Journal on Advances in Telecommunications
Volume 6, Numbers 3 & 4, 2013

CONTENTS

Tulsi Pawan Fowdur, University of Mauritius, Mauritius
Yogesh Beeharry, University of Mauritius, Mauritius
K.M. Sunjiv Soyjaudah, University of Mauritius, Mauritius

pages: 153 - 163

High Quality Video Conferencing: Region of Interest Encoding and Joint Video/Audio Analysis

Christopher Bulla, Institut für Nachrichtentechnik, RWTH Aachen University, Germany
Christian Feldmann, Institut für Nachrichtentechnik, RWTH Aachen University, Germany
Magnus Schäfer, Institute of Communication Systems and Data Processing, RWTH Aachen University, Germany
Florian Heese, Institute of Communication Systems and Data Processing, RWTH Aachen University, Germany
Thomas Schlien, Institute of Communication Systems and Data Processing, RWTH Aachen University, Germany
Martin Schink, MainConcept GmbH, Germany

Pilot service and Efficient OEO-based Remote Terminal Providing a Higher Power Budget of an Asymmetric 10/1G-EPON

Kwang-Ok Kim[1,2], Sang-Soo Lee[1]
[1]Department of Optical Internet Research
Electronics and Telecommunications Research Institute
218 Gajeong-ro, Yuseong-gu, Daejeon, 305-700, Korea
E-mail: {kwangok, jhlee, soolee}@etri.re.kr

Youn-Seon Jang[2]
[2]Department of Electronics Engineering
Chungnam National University
99 Daehak-ro, Yuseong-gu, Daejeon, 305-764, Korea
E-mail: jangys@cnu.ac.kr

***Abstract*—This paper proposes the design of an efficient optical-electrical-optical (OEO)-based remote terminal (RT) that can provide the higher power budget required for a long-reach transmission in an asymmetric 10 Gbit/s Ethernet passive optical network (10/1G-EPON). The current 10/1G-EPON specification supports a maximum physical distance of only 20km in a 32-way split due to a power budget limitation. However, many service providers prefer a transmission reach of over 40km in a 64-way split for an efficient access network design. In this paper, the proposed OEO-based RT provides quad-port architecture for a cost-effective design, supports a high power budget of 58 dB through 3R signal regeneration, and offers over a 50 km reach and 128-way split per port with no modification of a legacy 10/1G-EPON system. In addition, it can satisfy a packet loss rate (PLR) of 10^{-10} in the downstream and upstream paths. The current 1G-EPON RT included within this proposed 10/1G-EPON RT is being the pilot service at a large residential apartment over 2,000 subscribers.**

Keywords-10Gbit/s EPON, Remote Terminal, Long Distance EPON, Reach Extender, Giga internet service

I. INTRODUCTION

A 10 Gbit/s Ethernet passive optical network (10G-EPON) is one of the fastest access technologies for providing next-generation ultra-broadband services to subscribers. In the current fiber-to-the-home (FTTH) optical access system, 1 Gbit/s EPON (1G-EPON) is being extensively utilized, particularly in Asian nations, including Japan, South Korea, and China. However, with the recent growth of user traffic, a 10G-EPON is expected to provide end users with a more comfortable online environment in the near future [1], [2].

The 10G-EPON specification was ratified as the IEEE 802.3av standard in 2009, and supports two configuration modes: symmetric mode, operating at a 10 Gbit/s data rate in both directions; and asymmetric mode, operating at a 10 Gbit/s in the downstream direction and 1 Gbit/s in the upstream direction [3]. Additionally, to reduce the costs for laying fibers and equipment, 1G-EPON and 10G-EPON use the same outside plant. In particular, an asymmetric 10G-EPON (i.e., 10/1G-EPON) can be easily applied to the single family unit (SFU) market as a cost-effective next-generation solution, as its upstream transmission is identical to that of 1G-EPON, and its downstream transmission relies on the maturity of 10Gbit/s Ethernet devices.

The current 10/1G-EPON is defined into three classes of power budget: PRX10, PRX20, and PRX30. For compatibility with the PX10 and PX20 power budgets defined for a 1G-EPON, a 10/1G-EPON should mainly use the PRX10 and PRX20 power budgets. These power budgets support channel insertion losses of 20 and 24 dB, respectively. Therefore, a legacy 10/1G-EPON can support a physical distance of only 20 km for single mode fiber (SMF) in a 1:32 split ratio [4], [5].

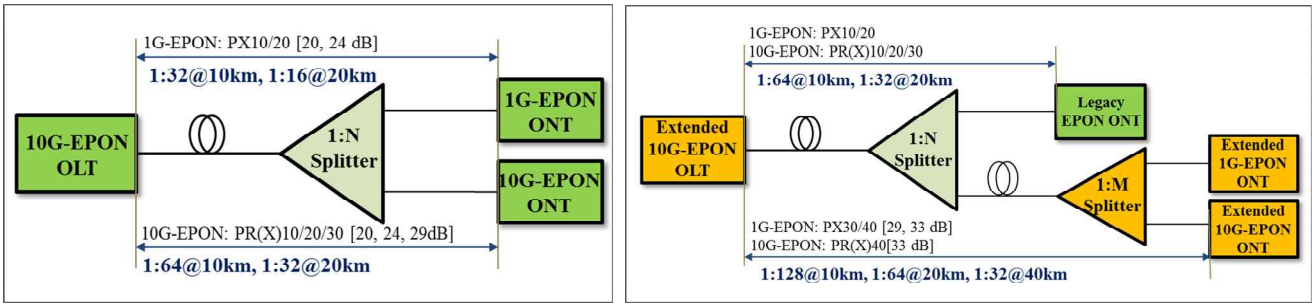
However, many network operators worldwide have placed an increased emphasis on combining an optical access network with a metro network by consolidating their central offices (COs) through a long-distance EPON solution. This combination results in a considerable reduction in capital expenditure (CAPEX) and operating expenditure (OPEX) budgets. In particular, EPON service providers require the high power budget to support long distances of 40 km and a high 64-way split. In addition, they hope to discover a solution satisfying the following key questions: how to leverage the EPON architecture in rural areas, how to further increase subscriber density in their COs, how to decrease the connection cost per subscriber, and how to serve more people at a larger distance from the COs using IEEE802.3 EPON equipment [6].

To satisfy these requirements, we suggest the cost-effective 3R-type optical-electrical-optical (OEO)-based remote terminal (RT) that can provide a higher power budget of a 58 dB in a legacy 10/1G-EPON without modification. We also demonstrate the performance of the OEO-based RT using a commercialized EPON optical line terminal (OLT) and 64 optical network terminals (ONTs) [1].

The remainder of this paper is organized as follows. In Section II, we briefly review related work, while in section III we describe the detailed structure and design scheme of the proposed OEO-based RT. In Section IV, we show experimental results proving the effectiveness of our method and provide an analysis of its performance. We also present the pilot service site for Giga internet service in S. Korea. Finally, we present a brief summary of our work in Section V.

II. RELATED WORK

An extended EPON solution helps with network evolution, and reduces network levels and nodes from an increased high power budget. It can also provide significant



(a) (b)

Figure 1. EPON link structure: (a) 10G-EPON link structure ratified by the IEEE802.3av standard, and (b) the 10G-EPON link structure suggested by the IEEE P802.3bk extended EPON working group.

cost savings by reducing the amount of electronic equipment and real estate required at a local exchange. Moreover, it can support service to small towns, suburbs, and rural areas [7].

To achieve these purposes, the IEEE P802.3bk extended EPON working group is standardizing a new definition of the power budgets or reach extender solutions that can support a higher power budget [12]. Recently, several methods were suggested by the extended EPON working group. The first method is to define new power budget classes (i.e., PX30, PX40, PR(X)40) through an increase in the receiver sensitivity and launch power of the transmitter. The second method is optionally to use an extender box providing optical amplification (OA) or OEO for passing data streams. The final method is to decrease optical distribution network (ODN) loss through an improved splitter design [8-9].

As a first option, the current extended EPON working group is focusing mainly on the physical media dependent (PMD) development of new power budget classes for application in the multiple dwelling unit (MDU) market, as many operators would prefer a completely passive solution. However, much of the market demand for high split ratios of over 1:64 also requires a 40 km reach through a high power budget of over 33 dB [13].

Among the methods described above, only the use of an extender box can easily satisfy this requirement. In an extended EPON working group, the line cost and power consumption per subscriber are also primary considerations when designing a high power budget solution for a 10G-EPON [10].

Figure 1 shows the 10G-EPON link structure ratified by the IEEE802.3av standard, and suggested by the IEEE P802.3bk extended EPON working group [3]. This 10G-EPON can support a maximum transmission reach of 10 km in a 1:64 split ratio when using a PR(X)30 with a high power budget of 29 dB under worst-case ODN design scenarios without any problems, as shown in Figure 1-(a). An extended PMD solution is provided through the insertion of an optical amplifier within the transceiver, and can support a power budget of 33 dB using the newly defined PR(X)40 PMD, as shown in Figure 1-(b). A 10G-EPON using a PR(X)40 PMD can support a long distance of up to 40 km in a 1:32 split ratio without an extender box in the remote node, and can also support a high split ratio of 1:128 at the distance of 10 km and a very high split ratio of 1:256 within a very short distance of 2 km [11].

However, an efficient 10/1G-EPON extender box solution supporting a cost-effective design, low power consumption, and a power budget of about 58 dB using the already developed legacy PRX30 PMD has yet to be reported. Therefore, to support a physical distance of over 40 km and a greater than 1:64 split ratio under the worst-case ODN design scenarios without any problems, a 10/1G-EPON must apply a remote terminal as an extender box utilizing an active device. Active in-field components are also acceptable to many operators.

In this paper, our proposed OEO-based 10/1G-EPON RT can efficiently provide a high power budget of 58 dB using the following functions: bit-level 3R-based signal retiming, remote management through a simple network management protocol (SNMP) agent and an embedded ONT, and upstream burst-mode to continuous-mode signal conversion to support the wavelength division multiplexing (WDM)-enabled reach extension.

III. PROPOSED OEO-BASED REMOTE TERMINAL FOR ASYMMETRIC EXTENDED 10/1G-EPON

Figure 2 illustrates the 10/1G-EPON link structure applied to the proposed OEO-based 10/1G-EPON RT in the remote office to support a long distance and high split ratio. A extended 10/1G-EPON system utilizing the 10/1G-EPON RT can provide a physical reach of over 60 km using an existing PRX30 PMD in the feeder fiber, and can support a 1:128 high split ratio for a 10 km reach, respectively, using the legacy PRX30 PMD supporting the link budget of 29 dB under the worst-case ODN design scenarios without any problems.

That is, when considering an optical fiber loss of 0.4 dB/km, the extended 10/1G-EPON applied to our 10/1G-EPON RT can easily support a high split ratio of 1:128 at a 80 km reach from a central office to end users. This is possible to make the flexible access network configuration. The 10/1G-EPON RT mainly provides wavelength conversion and a signal retiming function based on 3R-based signal regeneration between a 10/1G-EPON OLT and 10/1G-EPON ONTs or 1G-EPON ONTs at the remote office.

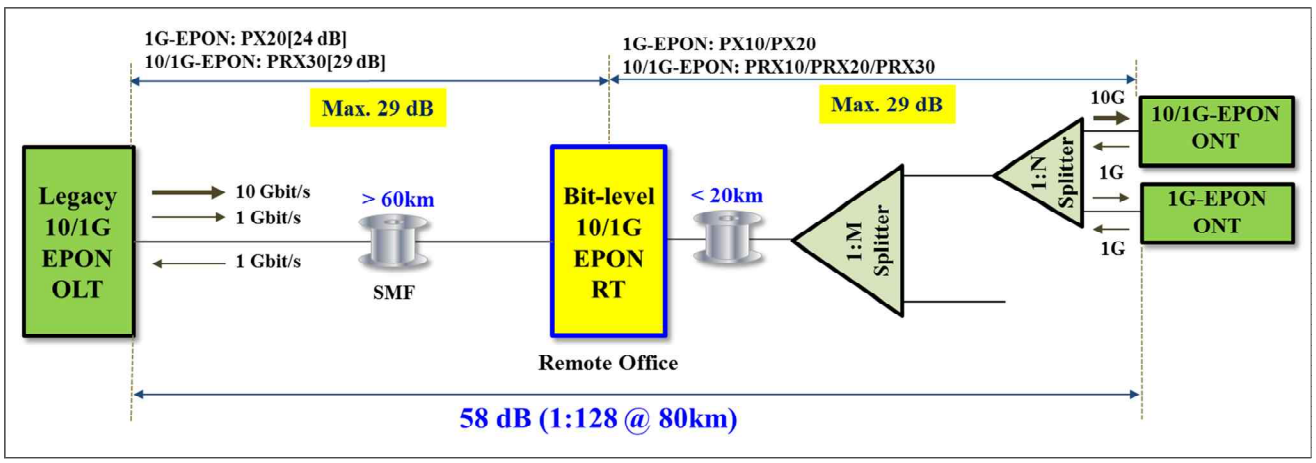


Figure 2. The extended asymmetric 10/1G-EPON link structure utilizing the proposed bit-level OEO-based RT.

In addition, it provides an optional upstream burst-to-continuous signal conversion to apply continuous mode WDM solutions in the trunk fiber. In particular, because the 10/1G-EPON RT is necessary for electrical power, it requires a remote management function and low-cost, low-power design. The extended asymmetric 10/1G-EPON provides the transmission rate of 11 Gbit/s at a downstream direction and 1 Gbit/s at an upstream path the allows to access of the asymmetric 10/1G-EPON ONTs and symmetric 1G-EPON ONTs as shown in Fig. 2.

Figure 3 shows the design architecture and quad-port prototype of the proposed 10/1G-EPON RT. The proposed 10/1G-EPON RT is composed of a 3-port edge WDM filter for interconnection with the 10/1G-EPON OLT equipment, a single low-cost FPGA(Field Programmable Gated Array) for the retiming of a 1 Gbit/s downstream signal, an embedded ONT for SNMP packet transmission to the 10/1G-EPON OLT, a CPU processor providing SNMP for remote management, a burst-mode clock and data recovery (BCDR) device for retiming of the burst-mode upstream signals, a crosspoint switch (CS) device for electrical signal division, a 1G-EPON ONU transceiver for receiving and transmitting a 1 Gbit/s optical signal, a 10/1G-EPON ONU transceiver for receiving a 10 Gbit/s downstream optical signal, and a 10/1G-EPON OLT optical transceiver for interconnection with the ODN [15].

The 10/1G-EPON RT divides a 10 Gbit/s wavelength into 1 Gbit/s wavelength channel using a 3-port edge WDM filter. These wavelength channels are then inserted into each EPON ONU transceiver, and an optical signal is then converted into an electrical signal. These electrical signals are retimed by an FPGA and 10 Gbit/s CDR in the electrical domain. The retimed signals then are retransmitted to the optical domain using a 10/1G-EPON OLT transceiver. In

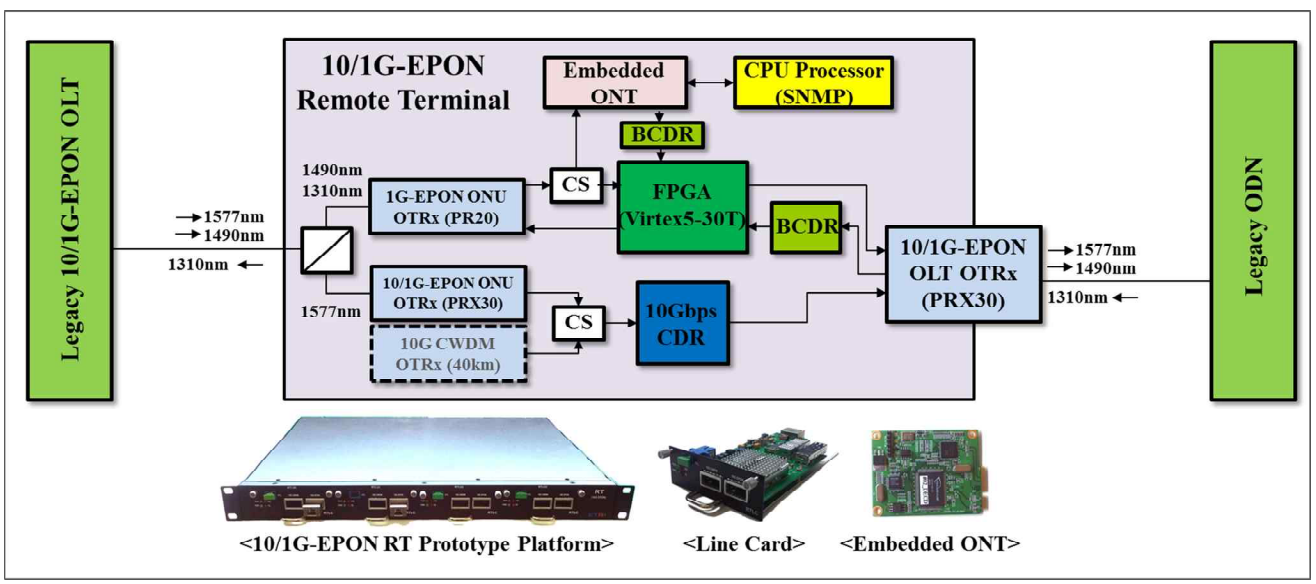


Figure 3. The design architecture and prototype of the proposed 10/1G-EPON RT.

contrast, the signal retiming for a 1 Gbit/s signal in the upstream is performed using a BCDR device, which then converts the retimed upstream burst-mode signal into a continuous-mode signal through the FPGA.

The proposed 10/1G-EPON RT is designed using a quad-port architecture, and provides a signal retiming function through a Virtex-5 FPGA (i.e., XC5VLX-30T) for lower design costs and power consumption. However, the signal retiming function for a 10 Gbit/s downstream signal is performed using a commercialized low-cost 10 Gbit/s CDR device. Each port of the 10/1G-EPON RT was also designed as a plug-and-play type. That is, the 10/1G-EPON RT can support a total transmission rate of 44 Gbit/s in the downstream and 4 Gbit/s in the upstream by using the cheaper solutions.

To provide remote management of the 10/1G-EPON RT, an embedded ONT is activated using a 10/1G-EPON OLT system through a CS device, BCDR device, and FPGA. An embedded ONT is provided using a compact-type of commercialized 1G-EPON ONU MAC as shown in Figure 3. That is, the 10/1G-EPON RT is connected with a 1 Gbit/s data channel of the 10/1G-EPON OLT, and only port 0 of the 10/1G-EPON RT is supported.

An embedded ONT receives a downstream signal through a CS device, and transmits an upstream signal to the 10/1G-EPON OLT using a BCDR device and FPGA, as shown in Figure 3. This upstream signal is merged with the upstream signal of port 0. The CPU processor gathers and manages the status of the installed optical transceivers based on the SFF-8472 and FPGA through a local CPU interface. The CPU processor is also connected directly with a user network interface (UNI) port of an embedded ONT at the electronic domain using a CS device without an external optical tap (e.g., an optical splitter), which is unlikely to have been used in previous methods. The proposed 10/1G-EPON RT can use optionally as a hybrid-type remote terminal by replacing the 10/1G-EPON and 1G-EPON ONU optical modules with 10 Gbit/s and 1 Gbit/s coarse WDM (CWDM) optical modules.

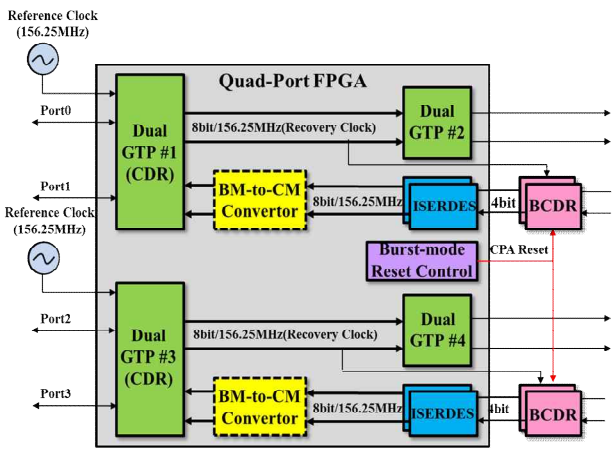


Figure 4. A Internal structure of the FPGA based on a quad-port for the signal retiming of 1G-EPON.

Figure 4 illustrates the internal architecture of the signal retiming logic within the FPGA based on a quad-port used to recover the 1G-EPON signals. Because the 10/1G-EPON OLT transceiver does not require a burst-mode reset signal, the 10/1G-EPON RT can easily provide a signal retiming function using the CDR within the FPGA for a downstream signal and the BCDR device for an upstream signal.

In the downstream direction, the FPGA provides a 3R signal retiming using a recovery clock extracted from the CDR, which is included in the dual gigabit transfer protocol (GTP). The dual GTP extracts a 156.25 MHz recovery clock and 8-bit data from a 1.25 Gbit/s continuous-mode downstream signal using an external 156.25 MHz reference clock. This recovery clock is then used as a reference clock source necessary for an external BCDR device. On the other hand, in the upstream direction, the 10/1G-EPON RT performs a signal recovery using the BCDR device and transmits this recovered signal to an input serializer and deserializer (ISERDES) as shown in Figure 4. This burst-mode parallel is transmitted the continuous-mode 1.25 Gbit/s to EPON OLT through the GTP block and BM-to-CM convertor [14].

Also, a burst-mode reset signal for the burst-mode upstream signal detection at a BCDR device is generated by a loss of signal (LOS) output from the 10/1G-EPON OLT transceiver. The BCDR device aligns with the input data within the 12-bit start of the preamble, and changes a 1-bit serial signal into a 4-bit parallel signal to provide the lower clock speed at the data transmission to the FPGA. The burst-to-continuous convertor within the FPGA optionally changes a burst-mode signal into a continuous-mode signal through the insertion of a particular pattern (e.g., h'55 or 8B/10B idle) during a laser off time.

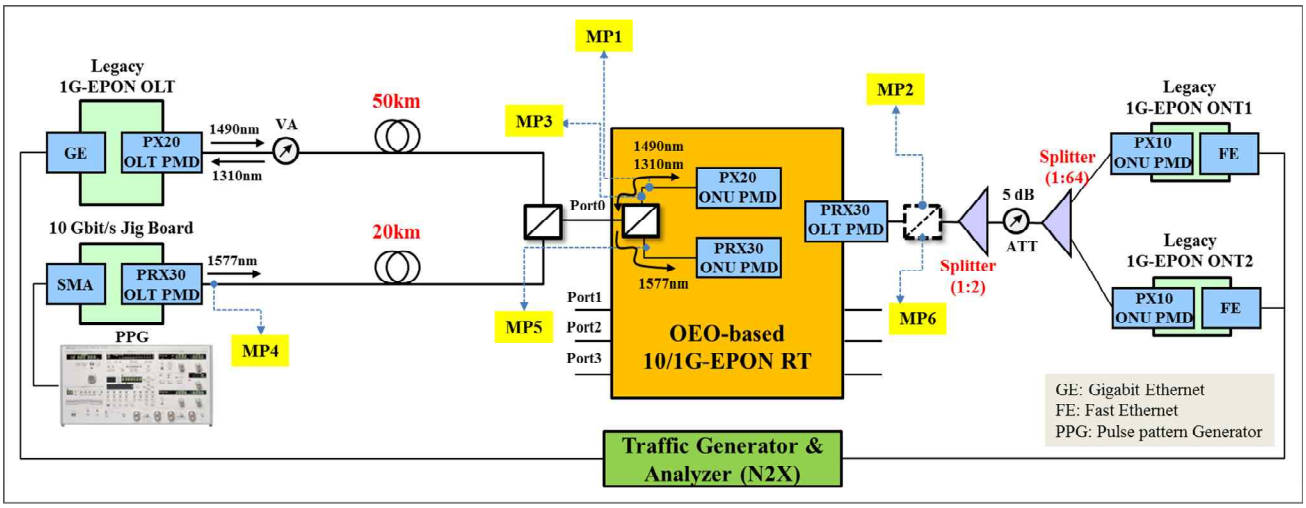
In this paper, our proposed OEO-type 10/1G-EPON RT can provide low-power consumption and a cost-efficient design through a quad-port structure using a single cheaper FPGA and legacy cheaper EPON optical transceivers.

IV. Experimental Setup and Results

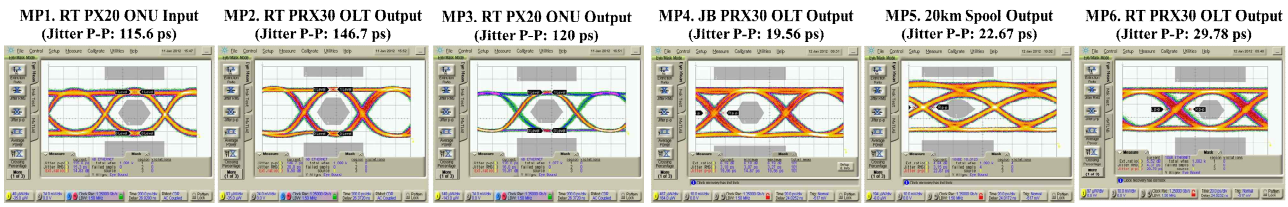
A. Experimental setup

The experimental setup for a performance measurement of the proposed OEO-based 10/1G-EPON RT is shown in Figure 5. As there is no commercialized 10/1G-EPON systems at present, we used the existing 1G-EPON system and a 10 Gbit/s jig board for the performance test of the 10/1G-EPON RT. The legacy 1G-EPON system generates downstream and upstream signals with a line rate of 1.25 Gbit/s, while the 10 Gbit/s jig board transmits a downstream signal with a line rate of 10.3125 Gbit/s using a pulse pattern generator (PPG).

For the 1 Gbit/s path link configuration, we used a single legacy 1G-EPON OLT and two 1G-EPON ONTs, and connected the trunk fiber using a 50 km SMF between the 1G-EPON OLT and 10/1G-EPON RT. In addition, we configured a fixed 5 dB attenuator and 128-way split as the drop fiber between the 10/1G-EPON RT and 1G-EPON ONUs. For the 10 Gbit/s path link configuration, we also connected the trunk fiber using a 20 km SMF between the 10



(a)



(b) (c)

Figure 5. Experimental setup for the performance measurement and optical eye-diagrams observed at MPs of the proposed OEO-based 10/1G-EPON RT: (a) experimental setup link, (b) 1.25 Gbit/s optical diagrams, and (c) 10.3125 Gbit/s optical diagrams.

Gbit/s jig board and 10/1G-EPON RT, as the 10/1G-EPON PMDs used in this experimental setup are unable to provide a transmission reach of 40 km owing to a dispersion problem.

The 1G-EPON used optical modules supporting both the IEEE 802.3ahTM-2004-PX20 and PX10, while the 10/1G-EPON RT used a 1G-EPON ONU optical module compliant with an IEEE 802.3ahTM-2008-PX20 supporting the link budget of 28 dB, and a 10/1G-EPON OLT/ONU optical module compliant with an IEEE802.3avTM-10/1GBASE-PRX30.

The 1.25 Gbit/s and 10.3125 Gbit/s optical signals generated were merged using a 3-port edge WDM filter, and then separated again into 1.25 Gbit/s and 10.3125 Gbit/s optical signals using a 3-port edge WDM filter in the 10/1G-EPON RT. The 10/1G-EPON RT transmits the retimed 1.25 Gbit/s and 10.3125 Gbit/s optical signals to the optical splitter via the 10/1G-EPON OLT optical module. The optical power budget in the trunk fiber is adjusted using a variable attenuator (VA) value. In this experimental setup, the insertion losses in the trunk and drop fibers are about -12.8 dB and -26.8 dB, respectively.

B. Optical eye digrams and analysis

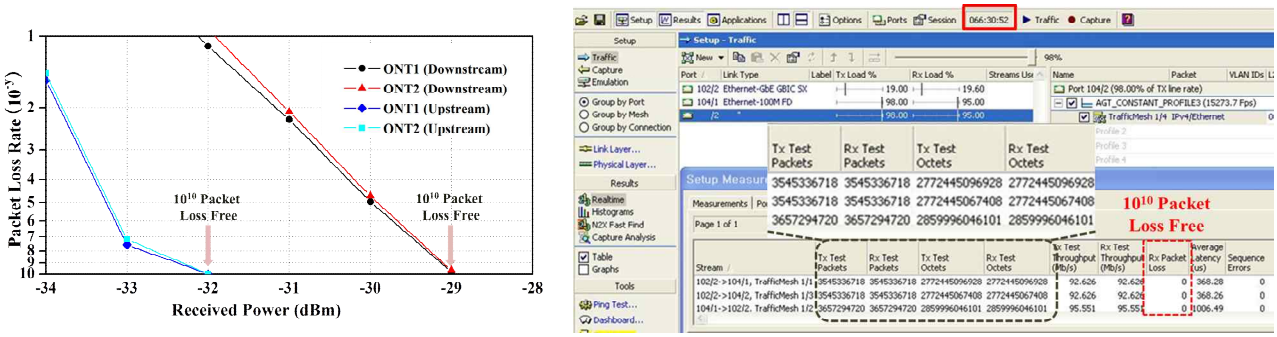
Figure 5 shows also the optical eye diagrams of each measurement point (MP) in the experimental setup link for the proposed 10/1G-EPON RT. Our 10/1G-EPON RT performs signal retiming using a recovery clock extracted through a 1 Gbit/s CDR within the FPGA and 10 Gbit/s CDR device, and this retimed signal is again recovered by the 1G-EPON ONUs.

In a 1.25 Gbit/s path, the downstream optical signal measured at MP1 is received by the 1G-EPON ONU optical module installed in the 10/1G-EPON RT through a 50 km SMF and two 3-port edge WDM filters. In an optical eye diagram measured at MP2 and MP3, shown in Fig. 5-(b), we can see a clear eye-pattern satisfying the optical eye mask adapted from the IEEE Gigabit Ethernet standard through a bit-level 3R-based signal regeneration. We can confirm that a jitter of about 31 ps is added by the signal recovery.

In a 10.3125 Gbit/s path, MP5 shows the results of a transmission dispersion problem caused by the 20 km SMF. We can also confirm that the output optical eye diagram measured at MP6 satisfies the optical eye mask adapted from the IEEE 10.3125 Gbit/s Ethernet standard by the 3R signal retiming, as shown in Fig. 5-(c), although a slight amount of jitter is added. This means that is possible to support a transmission reach of over 20 km through our proposed 10/1G-EPON RT, including a legacy 10G-EPON PMD in a 10 Gbit/s channel.

C. Packet transmission results and analysis

Using a commercially available router tester (i.e., Agilent N2X), the performance of the proposed 10/1G-EPON RT was evaluated in terms of packet loss rate (PLR) through



(a) (b)

Figure 6. The packet transmission results measured at 1.25 Gbit/s EPON link with two ONTs: (a) PLR results, and (b) result of a long-term test.

Ethernet frames with random lengths ranging from 64 to 1518 bytes at a downstream and upstream path. We transmitted 10^{10} frames to two ONTs for the PLR test.

Figure 6 shows the packet transmission results measured at the 1.25 Gbit/s path of the proposed 10/1G-EPON RT using a legacy 1G-EPON OLT/ONTs according to the VA value. Because a 1-Gb/s EPON OLT and ONU transceivers use an avalanche photodiode (APD) and a positive intrinsic negative photo-detector (PIN-PD), respectively, we can show that the upstream link budget increases by about 3 dB more than that of the downstream.

Our experimental results confirm that the 1G-EPON system supports a physical distance of 50 km in a 1:128 split ratio using the proposed OEO-based 10/1G-EPON RT, and satisfies a loss-free service up to -29 dBm in a downstream path, and -32 dBm in an upstream path as shown in Fig. 6-(a). This means that the 10/1G-EPON RT is able to provide transmission service at a distance of about 50 km with a 128-way split at PLR of the minimum 10^{-10}, when we take into account a budget loss of a 0.4 dB/km in an optical fiber.

Figure 6-(b) shows the result of a long-term packet transmission test of the proposed 10/1G-EPON RT. Over a 66-hour period, we transmit a packet load of 19 % (i.e., 190 Mbit/s) from an EPON OLT to each EPON ONT, and assign a packet load of 98 % (i.e., 98 Mbit/s) at each EPON ONT to measure the upstream PLRs, as a commercialized EPON ONT supports a Fast Ethernet port. From Fig. 6-(b), we can confirm the possibility of 10^{10} packet loss-free service in the downstream and upstream paths.

D. BMT environment and analysis

The 1G-EPON RT technologies included in the proposed 10/1G-EPON RT are commercialized since last year and we have performed the benchmark test (BMT) to multiple system operators (MSOs) (i.e., CJ HelloVision) in S. Korea on October 2012.

Figure 7 shows the experimental setup environment for the BMT of commercialized 16-port 1G-EPON RT platform with an existing EPON OLT and 64 ONU. For BMT, we configured a physical distance of 50 km with a 64-way split using a commercialized EPON OLT (i.e., Ubiquoss U9024a), 64 ONTs (i.e., Ubiquoss C504L), and the 16-port 1G-EPON RT platform. We applied a 50 km service profile map applying a window discovery size of 40,800 (i.e., 40,800 x 16-byte = 652 μs) and a maximum propagation delay time of 15,625 (i.e., 250 μs) to the EPON OLT.

As shown in Fig. 7, we configured a trunk fiber using two 20-km SMFs, a fixed attenuator of 10 dB and a VA to adjust the downstream signal power to -27 dBm in front of the 1G-EPON RT. For the drop section, we used a 10-km SMF, followed by a 1:4 optical splitter, and four 1:16 optical splitters. The loss budgets of the feeder and drop sections are

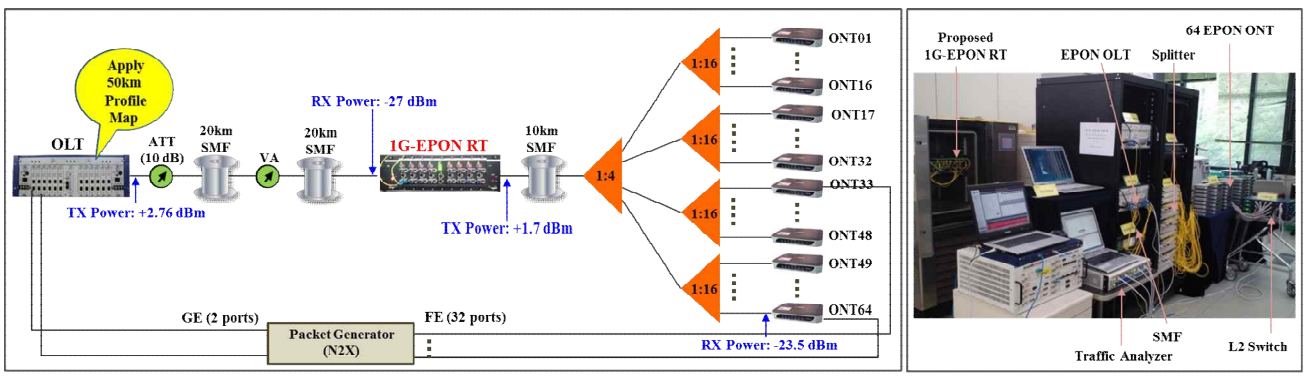


Figure 7. An experimental setup environment for the BMT with 1G-EPON RT platform.

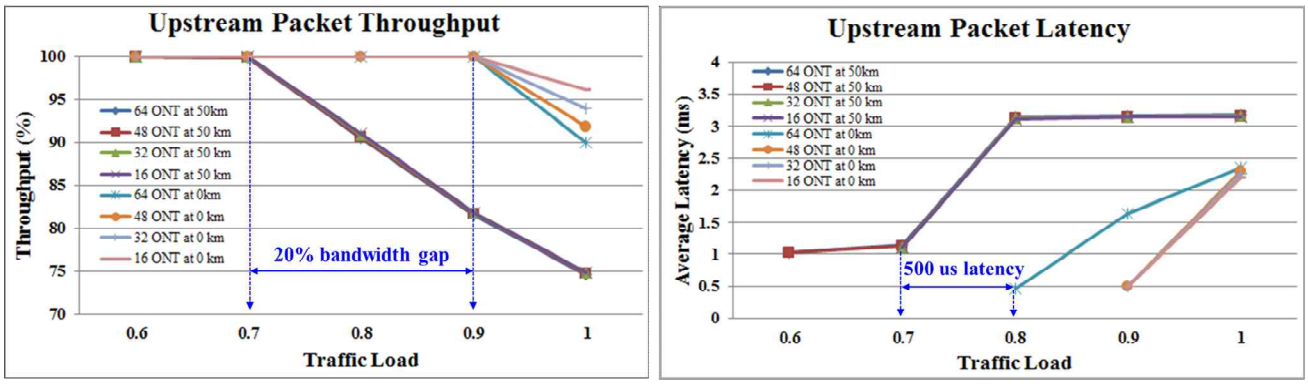


Figure 8. Upstream packet throughput and latency according to the traffic load at a single ONT.

29.8 and 25.2 dB, respectively.

Because the downstream throughput does not degrade depending on the transmission reach, at the only upstream path, we measured the maximum packet throughput that can be transmitted from one of the ONT and the average latency according to the traffic load in a single ONT.

For the measurement of the maximum packet throughput, we transmitted the packet with random lengths ranging from 64 to 1,518 bytes to only a single ONT. As shown in Fig. 8, although an EPON link at a back-to-back link (i.e., 0 km transmission reach) supports the maximum upstream packet throughput of about 920 Mbps, however, the extended EPON link of 50 km transmission reach utilizing the 1G-EPON RT platform provides a maximum packet throughput of about 700 Mbps regardless of the number of registered ONTs.

This is because of an an increased window discovery size (i.e., 391 μs) than existing window discovery (i.e., 261 μs) and an arrival time delay of the gate message (i.e., 300 μs) in order to allow 50-km reach. A maximum average latency is 3.2 ms at a traffic load of above 80 % (i.e., 800 Mbit/s) , as the input packet rate is faster than the bandwidth allocation time. Therefore, when extending an EPON with a 1G-EPON RT, to support over a 50-km transmission reach, a greater

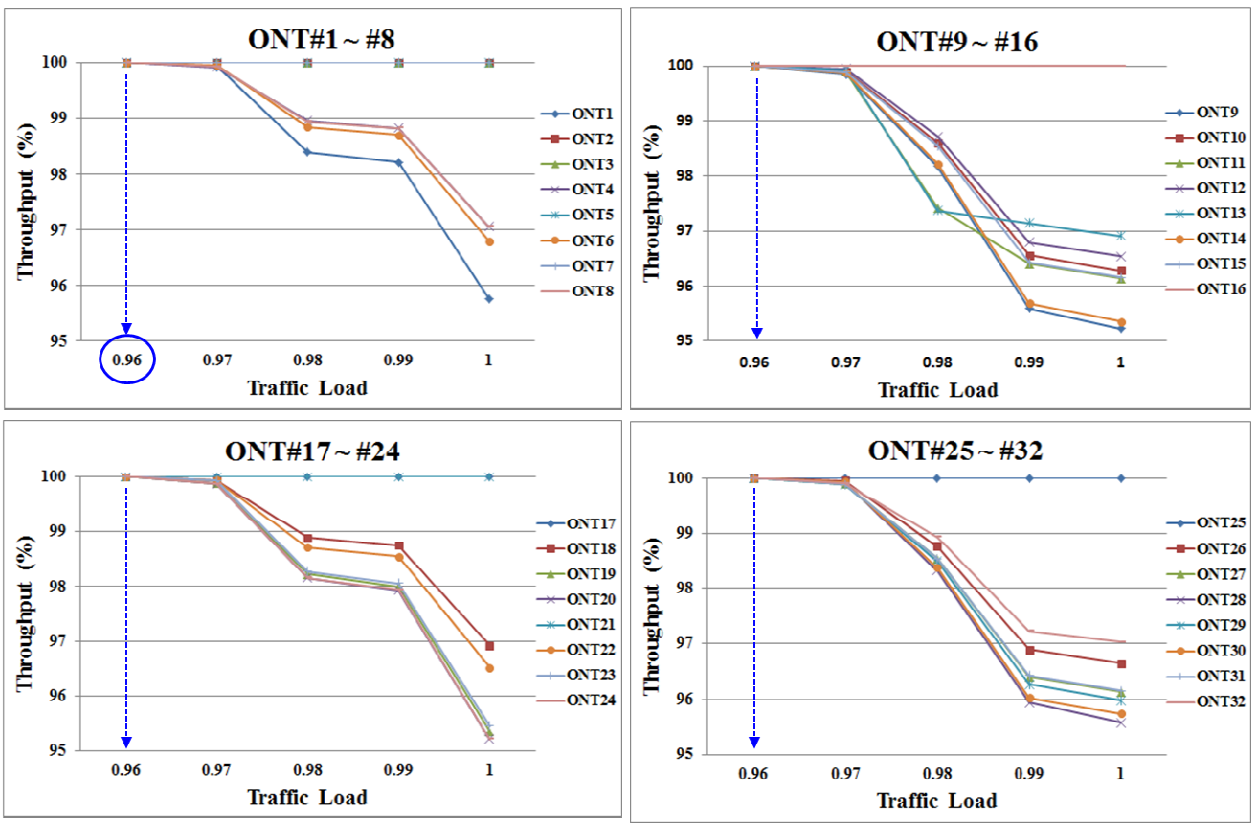


Figure 9. The downstream packet throughput according to the traffic load of a 32 ONTs.

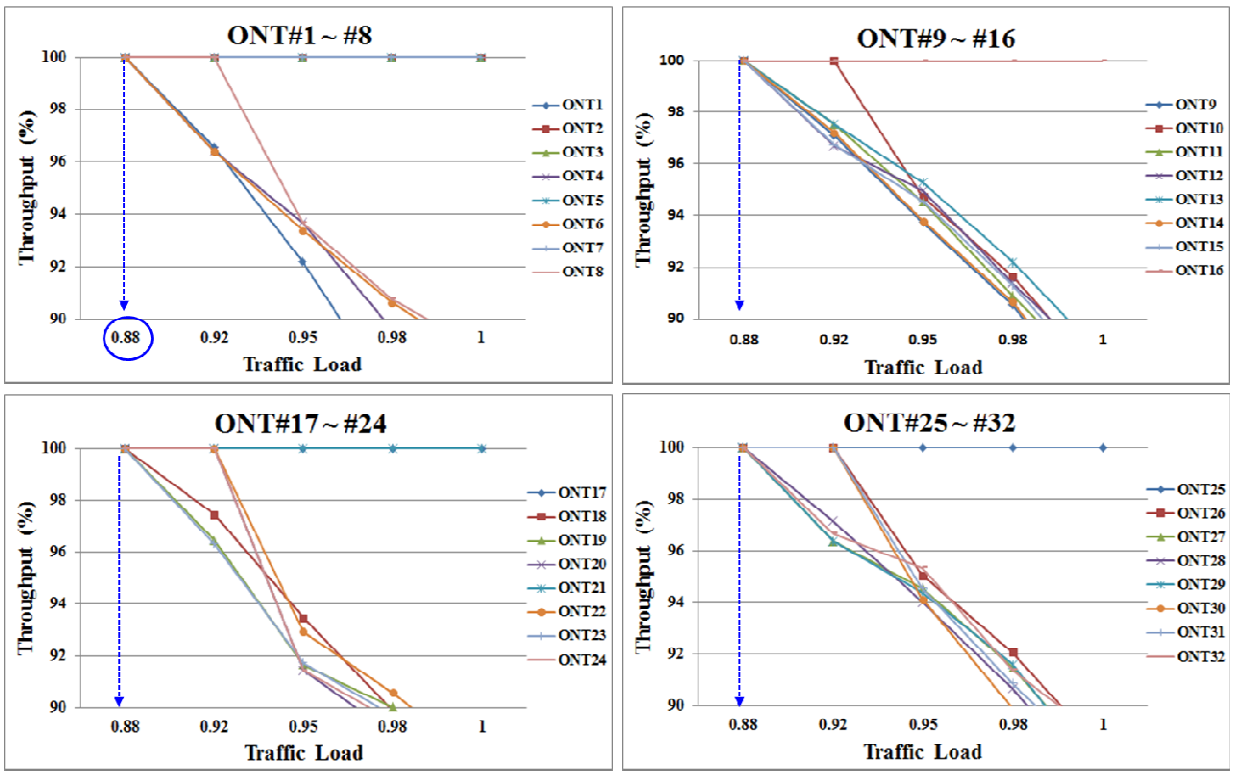


Figure 10. The upstream packet throughput according to the traffic load of a 32 ONTs.

byuffer size should be considered at the ONTs to cope with the increased latency of Gate message.

For performance analysis in actual environment, we transmitted 10^9 packets to 32 ONTs among 64 ONTs. Here, the downstream and upstream bandwidths per ONT were assigned about 30 and 28 Mbps, respectively.

Figures 9 and 10 show the results of packet throughput measured at the downstream and upstream paths according to the traffic load with 32 ONTs.

From the experimental results, we can confirm that the 1G-EPON RT platform satisfies the packet throughput of 100 % when traffic load is 96 % (960 Mbit/s) at the downstream path as shown in Fig. 9. On the other hand, at the upstream path, it can confirm that the 1G-EPON RT platform satisfies the packet throughput of 100 % when traffic load is 88 % (880 Mbit/s) as shown in Fig. 10. This is

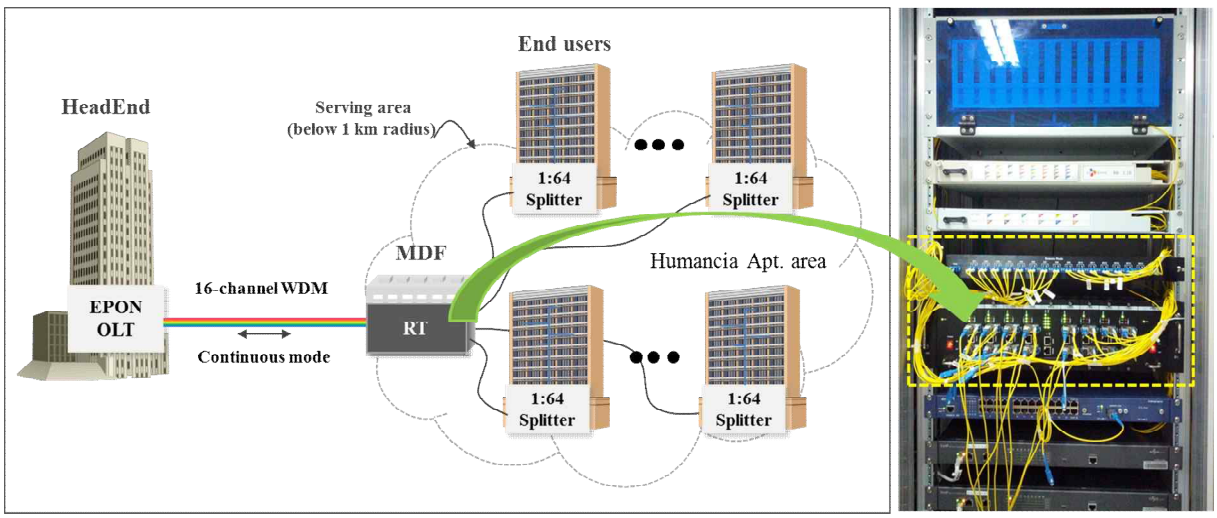


Figure 11. The results of long-term packet transmission and pilot service site.

because the packets are distribute to 32 ONTs. Therefore, the input packet rate at each ONT is lower than the asigned bandwidth allocation time from EPON OLT.

From the experimental results, we confirm that it provides loss-free service during a 1.6 x 10^{10} and 1.48 x 10^{10} packet transmission at the downstream and upstream path, respectively, when the received power at 1G-EPON RT platform is -27 dBm. Consequently, the extended EPON link based on the 1G-EPON RT platform can provide the maximum downstream bandwidths of 960 Mbps and upstream bandwidth of 880 Mbps similar to legacy EPON performance under the condition of packet loss-free.

E. Pilot services and link structure

The currently 16-port 1G-EPON RT had been installed with a DWDM solution at the feeder section to provide a Giga internet service at Gimpo area by Korean cable operator (i.e., CJ HelloVision) as shown in Fig. 11. The 1G-EPON RT had commercial service at the main distribution frame (MDF) within Humancia apartment to accommodate 2,234 subscribers, respectively, on December, 2012.

Each port of 1G-EPON RT platform is connected to end users through a 1:64 optical splitter as shown in Fig. 11. Therefore, cable operator can support a bandwidth of 16 Gbps per optical fiber, and reduce significantly the leased cost of the optical fiber using the designed a 16-port hybrid-type 1G-EPON RT platform.

F. Technical comparisons

Table I shows the results of a technology comparison of 10/1G-EPON using the proposed OEO-based 10/1G-EPON RT, a 10G-EPON standardized by IEEE802.3av-2009, and an extended 10/1G-EPON suggested by IEEE P802.3bk working group. As the table shows, with the exception of an active component used in a remote node, our proposed 10/1G-EPON RT can provide greater efficiency with respect to power budget, distance and user accommodation, cost per subscriber, and long-distance trunk fiber costs. However, to make up for the weak point above, we adopted quad-port architecture at the proposed 10/1G-EPON-RT.

TABLE I. TECHNOLOGY COMPARISON

Items	***IEEE802.3av 10/1G-EPON***	***IEEE P802.3bk 10/1G-EPON***	***RT-based 10/1G-EPON***
Power Budget	Max. 29 dB	Max. 35 dB	Max. 58 dB
Reach & Split Ratio	1:32 & 20km	1:32 & 40km	1:128 & 80km[Note.1]
BW per User	300Mbps	300Mbps	100Mbps
Upstream Mode	Burst	Burst	Burst & Continuous
WDM application	No	No	Yes
Remote Node	Passive	Passive	Active
Cost per User	Middle (100% / 32)	High (130% / 32)	Low (180% / 128)
Cost of trunk fiber	1	1	1/8[Note.2]

Note.1. Using a PRX30 PMD type at the EPON ONU.

Note.2. When a CWDM solution is applied to the trunk fiber.

V. CONCLUSION

We proposed and experimentally demonstrated an efficient bit-level OEO-based 10/1G-EPON RT based on quad-port architecture to overcome the physical limitations of a legacy asymmetric 10/1G-EPON system. We also confirmed that our proposed 10/1G-EPON RT can achieve a high power budget of 58 dB through 3R-based signal retiming using an existing 10/1G-EPON PMD without new PMD solution, and can be a cost-effective solution for an extended asymmetric 10/1G-EPON system.

Our experimental results verified that the proposed 10/1G-EPON RT can provide a distance of more than 40 km in a 1:64 split ratio, which many service providers desire. If 10/1G-EPON PMDs use the PRX30 power budget class, the 10/1G-EPON RT can be expected to support a reach of over 80 km in a 1:128 split ratio with no modifications of the legacy 10/1G-EPON standard.

In our future work, we will perform a feasibility test with either a commercial 10/1G-EPON OLT/ONTs, and apply WDM multiplexing to a trunk fiber with an asymmetric 10/1G-EPON wavelength converter performing wavelength conversion between the EPON and WDM.

ACKNOWLEDGMENT

This research was supported by the KCC (Korea Communications Commission), Korea, under the R&D program supervised by the KCA (Korea Communications Agency) - (KCA-10913-05002).

REFERENCES

[1] K.O. Kim, J.H. Lee, S.S. Lee, and Y.S. Jang, “Efficient OEO-based Remote Terminal Providing a Higher Power Budget of an Asymmetric 10/1G-EPON,” ACCESS2012, pp. 47-52, June 2012.

[2] C. Chen, Z. Chair, and B. Velmurugan, “10G EPON: Next Generation Ethernet Passive Optical Networks,” Proc. of OFC/NFOEC2009, March 2009, pp. 1-3.

[3] IEEE Std 802.3av, “Carrier Sense Multiple Access With Collision Detection (CSMA/CD) Access Method and Physical Layer Specification, Amendment 1: Physical Layer Specifications and Management Parameters for 10 Gbit/s Passive Optical Networks,” 2009.

[4] IEEE Std 802.3ah, “Carrier Sense Multiple Access with Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications,” 2004.

[5] IEEE Std 802.3ah, “Carrier Sense Multiple Access With Collision Detection (CSMA/CD) Access Method and Physical Layer Specifications, ” 2008.

[6] M. Hajduczenia, “Call For Interest consensus building meeting: Extended EPON PMDs,” IEEE802.3 Extended EPON Study Group, online presentations, available for download at: http://grouper.ieee.org/groups/802/3/EXTND_EPON/public/1107/index.html, July 2011.

[7] FTTH Council, “The Adavantage of Fiber,” 3rd ed., Spring 2009, pp. Sec2:1-32.

[8] B. Wang, J. Xu, and Z. Fu, “Proposal for Extended EPON PMD,” IEEE802.3 Extended EPON study Group, http://grouper.ieee.org/groups/802/3/EXTND_EPON/public/1111/index.html, Nov. 2011.

[9] D. Piehler, Extending EPON link budget without new PMD definitions," IEEE802.3 Extended EPON Study Group, http://grouper.ieee.org/groups/802/3/EXTND_EPON/public/1111/index.html, Nov. 2011.

[10] E. Mallette, "Five Criteria Board Market Potential," IEEE802.3 Extended EPON Study Group, http://grouper.ieee.org/groups/802/3/EXTND_EPON/public/1111/index.html, Nov. 2011.

[11] M. Hajduczenia, "Project Objectives…how far, how long..," IEEE802.3 Extended EPON study Group, http://grouper.ieee.org/groups/802/3/EXTND_EPON/public/1111/index.html, Nov. 2011.

[12] IEEE P802.3bk D1.2 "Draft Standard for Ethernet Amendment 2: Physical Layer Specifications and Management Parameters for Extended Ethernet Passive Optical Networks," Oct. 2012.

[13] Y. Ma, X. Zhang, and Z. Fu, "Proposed for Extended Box of EPON," IEEE 802.3bk Task Force, 802.3 Plenary meeting, San Diego, CA, USA, July 2012.

[14] K.O. Kim, J.H. Lee, S.S. Lee, J.H. Lee, and Y.S. Jang, "Low-cost, low power, high-capapcity 3R OEO-type reach extender for a long-reach TDMAPON,"ETRI Journal, vol. 34, pp. 352-360, 2012.

[15] K.O. Kim, S.S. Lee, J.H. Lee, and Y.S. Jang, "Design of a cost-effective hybrid-type PBEx providing a high power budget in an asymmetric 10G-EPON," ETRI Journal, vol. 34, pp. 838-846, 2012.

A Reliability and Survivability Analysis of US Local Telecommunication Switches

Andrew Snow, Aimee Shyirambere, and Julio Arauz
School of Information and Telecommunication Systems
Ohio University
Athens, Ohio USA
e-mail: snow@ohio.edu

Gary Weckman
Department of Industrial and Systems Engineering
Ohio University
Athens, Ohio USA
e-mail: weckmang@ohio.edu

***Abstract*— This paper presents a comprehensive analysis of the reliability and survivability of US local telecommunication switches over a 14-year study period (from 1996 to 2009). Using local switch outage empirical data, the causes, failure trends and impacts have been identified, analyzed and assessed. A total of 12,860 switch outages were investigated for which very significant reliability growth was identified over the study period. Outages were also studied temporally, from time of day, day of week, and month of year perspectives. Additionally, 2,623 of the outages were found to come from only 156 unique switches, each of which experienced eight or more outages over the study period. The data were separated into two categories, for comparison: more frequently failing switches and less frequently failing switches. Major findings are that scheduled maintenance activities and hardware failures are the major causes of outages in local telecommunication switches; there are significant causality differences between more frequently and less frequently failing switches; and there are considerable differences in switch characteristics between the more and less frequently failing switches. Additionally, the manufacturers of the more frequently out switches are identified.**

Keywords-network reliability; reliability trends; public switched telephone network; local telecommunication switches

I. INTRODUCTION

Early in 2013, research on local communication switches suffering frequent outages was reported in [1]. This work is expanded here, not just to switches with frequent outages, but to all local switches experiencing outages over a 14 year period. Additionally, in this paper, information on frequently out switches is expanded to include manufacturer. In this section, the research purpose, importance, and scope of this work are discussed.

A. Research Purpose

As an extension and expansion of previous research, this study analyzes the reliability and survivability of US local switches using data on Public Switched Telephone Network (PSTN) local switch outages. In addition to that, this study determines if a minority of switches account for significant reliability or survivability deficits in this important component of the PSTN. This study also examines local telecommunication switch outage trends and causes. Finally, the national policy regarding local switch outage collection regulations is briefly assessed.

In this study, the PSTN is considered as a single repairable system. A repairable system is defined as a system that passes from an operating mode to a failed mode, and then returns to operating mode after a certain period of time [2]. In fact, the system is returned to the operating mode by means other than replacing the entire system. As local switches serve as access nodes for users, it is important to understand switch outages in their natural operating setting.

B. Research Importance and Scope

Analyzing the reliability and survivability of local switches is of great importance because it can help in monitoring and improving the efficiency of the entire PSTN switching system since local switches form a large percentage of all the PSTN switches. Additionally, mobile and fixed wireless systems will always benefit from the reliability and survivability of the PSTN switching system because they greatly depend on it. This is especially true when wireless subscribers want to communicate with landline subscribers or call wireless subscribers in a different geographical area. The reliability assessment of local switches, which also includes determining the nature and trend of failure events, can help designers (switch manufacturers) and operators (service providers) in taking corrective or preventive actions where needed. Also, investigating the causes of outages in PSTN local switches can help in improving wireless switches, as they are very similar to wireline switches (i.e., same vendors and similar models). Definitely, the reliability of the PSTN is crucial, as it is the heart of landline and mobile voice communications.

The PSTN is a complex system composed of three main systems, namely the switching system, the signaling system, and the transmission system. The switching system controls and routes voice or data signals throughout the network. The signaling system enables switches to cooperate in call initiation, maintenance, and termination. Finally, the transmission system ensures physical connections between switches. These three systems enable end-to-end connections among PSTN subscribers. The signaling and transmission systems are not included in this study.

The switching system can also be subdivided into subsystems that include, the local exchange switching subsystem (local switches), the tandem switching subsystem (tandem switches), and the international gateway exchange subsystem (access switches) (Fig. 1). Only the local exchange switching subsystem is investigated in this study. Local switches include standalone, host, and remote local

switches. As an exception, some tandem switches also have access lines, and although they are included in this study, they represent a very small percentage of all reported switches or more outages. This study investigates reported local telecommunication switch outages in the U.S. of at least 2 minutes in duration for a 14-year period (1996-2009) and considers only totally failed switches rather than partially failed switches. Data below 2 minutes and partially failed switches are of interest, but not reported by the carriers.

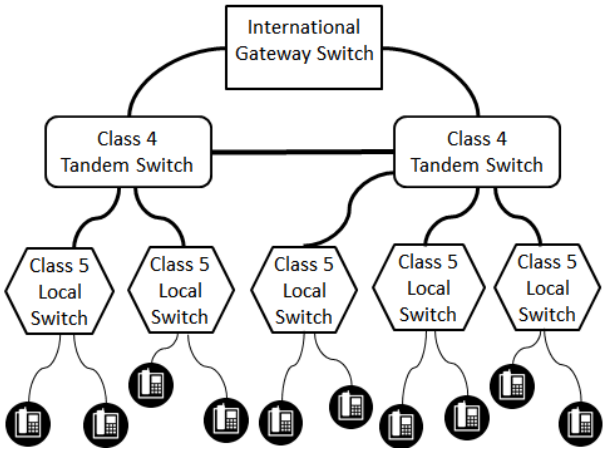


Figure 1. The PSTN switching system.

Again, the local exchange switching subsystem is analyzed as a whole, where all outages of individual local switches are pooled to analyze the overall reliability and survivability of the PSTN local exchange switching subsystem. Availability, safety, security, maintainability are other system dependability attributes not assessed in this study. Also, the calculated impact of outages includes only access lines but not blocked calls, as blocked call data were not reported.

In industry, local switches are called by different names, including "class 5 local exchange switch", "local switch", "central office switch.

II. LITERATURE REVIEW

In this section, a review of previous relevant research is presented in order to assess areas of focus of past researches, and establish a relationship between those studies and this research. The information presented was gathered from different sources, both electronic and non-electronic; including, but not limited to, research papers, books, conference papers and journals.

A. Past Studies on Local Telecom Switch Outages

There have been very few published research papers focusing on the reliability or survivability of US local telecommunication switches. In a previous research study, Snow analyzed local switch outages from 1991 to 1995, finding significant reliability growth [3]. Also, a few results regarding frequently failing switches were reported, but not studied in [4]. Kuhn studied local telecommunication switch outage records reported by telephone companies to the Federal Communication Systems (FCC), consisting of outages affecting at least 30,000 customers for at least 30 minutes. From those records, covering the period April 1992 through March 1994, Kuhn reported that the principal causes of PSTN large-scale outages were human error, acts of nature and traffic overloads [5].

Later, Snow investigated the effectiveness of the FCC outage reporting threshold, consisting of all reported outages affecting at least 30,000 users for at least 30 minutes. He used over 18,000 local telecommunication switch outages above and below the FCC reporting threshold and reported that the FCC outage reporting threshold was not optimal [6]. Again, Snow and Agarwal investigated over 19,000 local switch outages that occurred over the period 1993 through 2002 in order to explore an optimal outage reporting threshold that could reduce the number of outage reports and at the same time allow enough insight into network survivability. They found that "PRODUCT" thresholds such as lost line-hour (lines out times outage duration) are more optimal than AND (lines and duration) thresholds for assessing the survivability of telecommunication networks [7].

B. PSTN Overview

The PSTN, which is a collection of interconnected voice-oriented public telephone networks, was originally designed to support circuit-switched landline (or fixed line) voice communication. However, it is also used as the backbone network of mobile (or wireless) voice communication. Some of its elements are also used for Internet based network technologies such as Voice over Internet Protocol (VoIP). PSTN subscribers are connected to the PSTN network through local loops, which physically connect users' homes to central or local offices switches (also known as class 5 switches or end offices switches) [8]. In fact, local class 5 switches are the access and delivery points of voice communication to and from landline subscribers, and they receive numerous software upgrades during their lifetime in order to meet user and network ever-changing requirements.

Currently, most of the PSTN core system uses digital switching and transmission whereas many local loops still use analog mechanisms. It is clear that today's PSTN is transitioning to a packet switching, IP-based network, but this transition will not happen overnight. It will take many years to transition the entire PSTN into an all-IP-based network, especially the local loops [9]. In fact, although most of the interoffice transport network has been replaced by IP technology, the majority of the PSTN customers are still connected to the PSTN through local circuit-switched networks [9]. A recent report found that, "[in] December 2010, there were 117 million end-user switched access lines in service and 32 million interconnected VoIP subscriptions in the United States…" [10].

In fact, many network operators (services providers) want to provide a smooth migration from the legacy PSTN to a Next Generation Network (NGN), so they decided to consider changes to local loop networks only when

expanding them or when replacing a failed system. "VoIP began another evolution of the PSTN architecture. The PSTN is a large infrastructure that will likely take some time to completely migrate to the next generation of technologies; but this migration process is underway" [11].

In the cellular system, Mobile Telephone Switching Offices (MTSOs), also known as Mobile Switching Centers (MSCs), are very similar to PSTN Central Offices (COs), and they function like class 5 switches [12]. MSCs are connected to the PSTN switching centers through Gateway Mobile Switching Centers (GMSCs), which are a type of MSC, as shown in Fig. 2. The main function of an MSC is to connect mobile telephones to landline telephones or two mobile telephones in the same area, and an "MTSO is supposed to appear as a seamless extension of the public switched telephone network from the customer's perspective" [12]. In fact, both the PSTN and the cellular system use circuit switches made by the same equipment manufacturers, which are very similar.

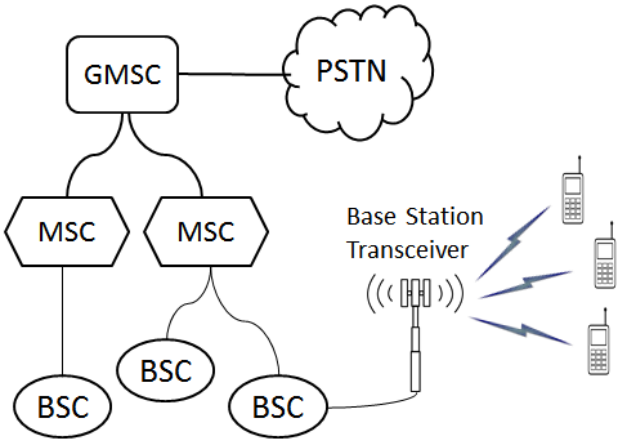


Figure 2. PSTN and mobile network.

As noted from the FCC website, the number of local switches has been slightly decreasing with time. This is partly due to the current migration from the legacy PSTN infrastructure to the more cost effective NGN infrastructure. In fact, at the end of their lifecycle, traditional PSTN Class 4 and Class 5 switches are transformed to media gateways or replaced by VoIP soft switches [13, 14, 15]. Media gateways (also called access servers) interconnect the traditional PSTN to VoIP networks, and they can originate or terminate landline phone calls. Soft switches function as other telephone switches except that they are software installed on servers, and they deal with IP to IP phone calls only. Fig. 3 depicts an interconnection between the traditional PSTN and a VoIP network.

C. *Reliability and Survivability Theories*

The reliability of a system is the probability of that system performing and maintaining its designated functions at an adequate level of performance, under specified circumstances and for a specified period of time [2]. Thresholds are used in specifying adequate levels of performance for repairable systems in order to differentiate operating and failed states [2].

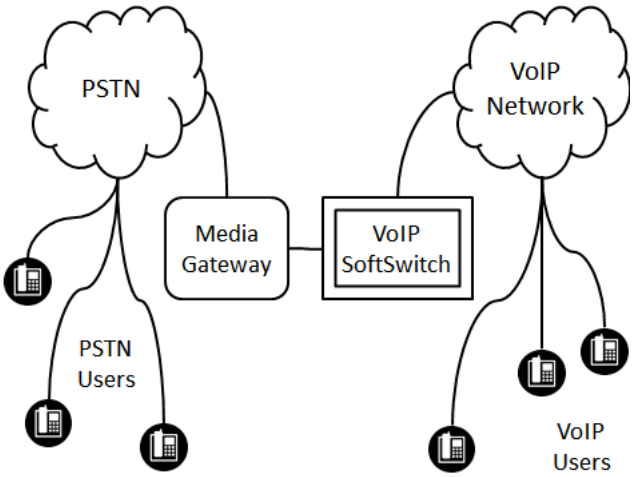


Figure 3. PSTN and VoIP network.

As mentioned before, a repairable system is defined as a system that passes from an operating mode to a failed mode, and back again after a certain period of time by means other than replacing the entire system [2]. A repairable system is also a counting failure process where successive inter-arrival failure times (time-between-failures) will tend to become larger for an improving system or become smaller for a deteriorating system [16]. The reliability of local switches can be assessed by analyzing the nature of failures experienced by those switches.

Point processes have been chosen for modeling the failure times since the time to repair or replace the system (local switches in our case) was negligible compared to its operating time. When the failure rate is constant over time, it can be modeled as a homogeneous Poisson process (HPP), which means that there is no improvement. On the other hand, when the failure rate varies over time, a non-homogeneous Poisson process (NHPP) is a better fit to model either an improving or deteriorating system [2].

The Laplace trend test, a reliability trend test, determines if there is a significant change in the pattern of successive failures of a repairable system over time. In fact, "[the] Laplace test, also known as the centroid test, is a measure that compares the centroid of observed arrival times with the midpoint of the period of observation. This measure approximates the standardized normal random variable (e.g., z-score)" [17]. The reliability trend test for repairable systems assumes the null hypothesis (H_o) to be HPP (no trend) and the alternative hypothesis (H_a) to be NHPP (there is a trend). If the null hypothesis can be rejected at a specified significance level, then it can be concluded that the system is either improving or deteriorating over the timeframe of interest [16]. The Laplace score U is given by the formula [17]:

$$U = \left[\left(\frac{\sum_i^n t_i}{n}\right) - \frac{T}{2}\right] / T\sqrt{1/12n} \quad (1)$$

where:

t_i is the time (e.g., number of days) from a fixed start point to the time of each event (outage).

n is the number of outage events (if t_n = T, then *n-1* is used instead of *n* in the formula).

T is the time from the start point to the end of the study period.

A positive score implies an upward or increasing trend (i.e., the system is deteriorating), a negative score implies a downward or decreasing trend (i.e., the system is improving), and a null score implies a constant trend (i.e., no change). Furthermore, "[when] the score is greater than (less than) +1.96 (-1.96), we are at least 95% confident that there is a significant trend upward (downward)" [17].

As an example of reliability assessment using the Laplace trend test, consider the failure arrival times given in Table I, where U-scores were calculated using the aforementioned formula. We can see from the table that the first set of sample failure arrival times represents an increasing trend (i.e., reliability deterioration) where the Laplace score U equals +3.46. The second set of sample failure arrival times represents a constant trend (i.e., no change in reliability) because the Laplace score U, which equals +1.79, is neither greater than +1.96 nor less than -1.96. The third set of sample failure arrival times represents a decreasing trend (i.e., reliability growth) where the Laplace score U equals -2.07. Visual representations of the increasing trend, the constant trend and the decreasing trend are given in Figs. 4, 5 and 6 respectively.

TABLE I. LAPLACE SCORE & FAILURE TREND EXAMPLES

Failure Trend Examples	*Increasing*	*Constant*	*Decreasing*
Failure Arrival Times	1	1	1
	8	3.5	1.1
	10	6	1.3
	11.5	8.5	1.8
	12.5	10.5	2.6
	13	13	3.5
Laplace Score (U)	**3.46**	**1.79**	**-2.07**

Survivability is defined as "the capability of a system to fulfill its mission, in a timely manner, in the presence of attacks, failures, or accidents." Typically, for a system to survive, it must automatically react to (and recover from) a harmful incident well before the root cause has been identified [18]. The survivability of local switches can be determined by analyzing the frequency and impact of failures experienced by those switches.

A system is delivering correct service when it is adequately fulfilling its functions. A service failure (or simply a failure) occurs when the system is not adequately implementing its functions, and the period during which the system is delivering incorrect service is called a service outage (or simply an outage) [19]. Failures are incidents that are likely to disturb the system and cause it to deliver incorrect service. They may be caused by deficiencies in the system or by external components to which the system is attached. Failures may be due to such things as software/hardware design errors, human errors, traffic overload, and natural disasters [18].

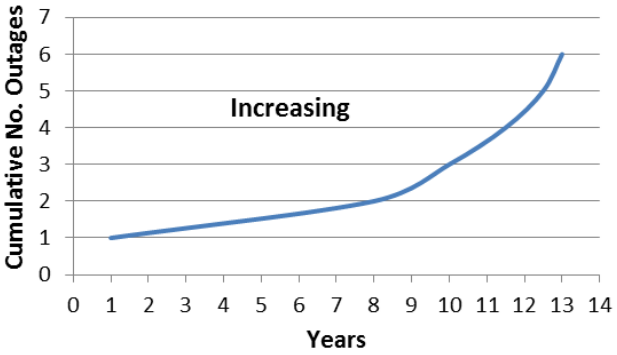


Figure 4. Example of reliability deterioration .

Figure 5. An example of a constant failure trend.

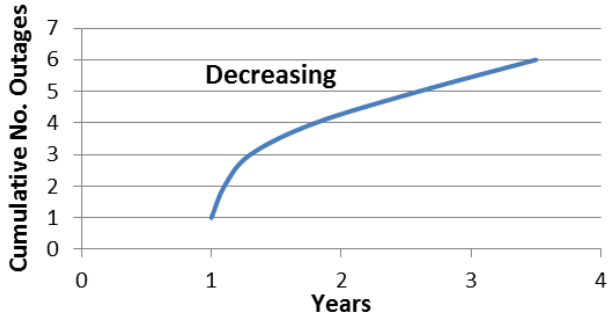


Figure 6. Example of reliability growth (decreasing failure trend).

Lost line hours, LLH, is an outage impact metric (or communication loss metric) that can be used to assess system survivability. LLH is the product of the number of access lines served by the and the duration of the outage [6]. For example, if a 3,000 access lines switch went down for 20 minutes, the LLH would be 1,000, meaning equivalency with 1,000 lines out for an hour. In this research, the LLH metric has been chosen to assess the survivability of local switches because it is simple and intuitive.

The metric also takes into account the size of the failed switch and the duration of the outage. It does not however account for blocked calls, but that data is not available from the switch outage reports.

D. The FCC Reporting Systems

The FCC, established by the Communications Act of 1934, is an independent agency of the US government. It regulates interstate and international communications by radio, television, wire, satellite and cable, and its jurisdiction covers the 50 states, the District of Columbia, and other US territories [20].

Until 2009, all US Local Exchange Carriers (LECs) reported each telecommunication switch outage of two minutes or more to the ARMIS (Automated Reporting Management Information System) section of the FCC website. Those public reports are part of FCC Report 43-05, the ARMIS Service Quality Report [21]. Note that the two-minute ARMIS reporting limit is used as the reliability threshold in this study. Again note that only totally failed switches were reported. The switch population required to report by LECs account for over 90% of the landline telephone access lines in the US [3].

ARMIS was initiated in 1987 to collect financial and operational data from the largest carriers in the US. Later in 1991, additional ARMIS reports were added to collect service quality (i.e., the FCC Report 43-05) and network infrastructure information from all US LECs subject to price cap regulations. However, after 2009, the FCC stopped collecting Report 43-05 as stated in the ARMIS Forbearance Order, where

> The Commission granted conditional forbearance from carrier obligations to file ARMIS Reports 43-05 and 43-06 provided that the carriers committed to file the data voluntarily for 24 months after September 6, 2008. The 24 months ended on September 6, 2010; and carriers do not file Reports 43-05 and 43-06 for reporting year 2010 and subsequent years [21].

The reports submitted to ARMIS contain information on each failed switch, including the date, time of outage occurrence, number of lines supported, outage duration, and outage cause. Additionally reported items include whether the switch is located in a Metropolitan Statistical Area (MSA = urban area) or not (Non-MSA = rural area), COSA (Company Operating company Study Area), and the switch CLLI code (Common Language Location Identifier). CLLI codes are unique identifiers for individual local switches. Carriers reported outage cause using one of fifteen different cause codes created by the FCC. In this study, cause code 1 is a scheduled maintenance outage; while cause codes 2 through 15 are considered failures, resulting in an outage. All 15 cause codes are defined in Section V.

III. RESEARCH GOALS AND OBJECTIVES

For a system to be properly improved it is necessary to know its past state and performance. Hence, from this perspective, monitoring the performance and assessing the reliability of local switches during past years will help understand what caused the outages to occur, and thus take corrective/preventive actions to alter future trends. Preliminary data exploration reveals a small number of switches or more outages an appreciable percent of all switch outages. The principal goal of this research is to compare the outage causality and switch characteristics between more frequently failing local switches and less frequently failing local switches. As its objective, this research focuses on addressing the following questions concerning local switches performance over a study period of 14 years:

A. Research Questions

1. What are the major causes of local telecommunication switch outages?
2. Is the reliability of local switches improving, constant, or deteriorating?
3. Are there individual switches that experience outages/failures more so than others?
4. Are there similarities/dissimilarities between switches failing more often and those that do not? In terms of:
    a. LLH
    b. Rural versus Urban
    c. Outage causes
    d. Outage duration
    e. Time of Day (TOD)
    f. Day of Week (DOW)
    g. Month of Year (MOY)
    h. Outage/failure trends
5. Are there switch manufacturers that account for outages/failures more so than others?

IV. METHODOLOGY

A. Research Data

The data used in this study were drawn from the FCC's ARMIS website (http://transition.fcc.gov/wcb/armis/) where US LECs reported switches that experienced a downtime of two minutes or more, and the data cover a 14-year period (from 1996 to 2009). The data include different information on the failed switches such as date, week day and time when the outages occurred, number of lines supported by the switches, duration and cause of the outages, MSA, COSA, and CLLI codes. Fifty-nine records were removed due to errors in recording the data, which left a total of 12,860 records.

In order to conduct research on frequently failing local switches, ARMIS data has been augmented by data from TelcoData.us, where additional information on US telecommunication switches is provided, such as switch models/manufacturers, and switch locations [22]. The Local Calling Guide website also provides information on US local switches, and was used to verify data from TelcoData.us and match these data to data from ARMIS [23].

B. Data Analysis Methods

The frequency of outage causes has been analyzed to assess the reliability of local switches, and the impact of those outage events has been analyzed to assess switch survivability. In order to assess differences that might exist in local telecommunication switch outages, the data have been separated into two categories: more frequently failing switches (8 or more outages over the study period) and less frequently failing switches (7 or less outages over the study period). The reliability difference between categories has been investigated by comparing failure trends and Laplace scores. Likewise, the survivability difference has been investigated by comparing impact trends and their respective LLH values.

In this study two measures of central tendency have been used to compare the two categories of data: the mean and the median. The t-Test (Two-Sample Assuming Unequal Variances) has been used to determine whether any difference between the means in the two local switch categories is statistically significant. We assumed that the two data sets came from distributions with unequal variances because one data set is considerably larger than the other one. Acts of god include such circumstances as wind, flooding, and earthquake.

For the t-Test, the null hypothesis (H_o) is that the means are the same, while the alternative hypothesis (H_a) is that the means are different.

V. RESEARCH RESULTS ANALYSIS

A. Causes of Local Telecommunication Switch Outages

As mentioned earlier, the cause of each reported outage was classified by carriers using 15 different cause codes. An abbreviated definition of each cause code together with the total number of reported outages for each are shown in Table II. The outage distribution is also shown in Fig. 7. From this table, two causes account for over 50% of the local switch outages:

- *Scheduled* outages (cause code 1)
- *Random hardware failure* (cause code 8)

Scheduled outages are planned for short duration, and therefore have little impact on the PSTN users because they are scheduled during hours and days of low traffic on the PSTN network. On the contrary, failures resulting in outages are unpredictable, occurring at any time and any day, impacting PSTN users in many cases. For that reason, it is also important to investigate the major causes of failures in PSTN local switches.

By discounting scheduled outages (cause code 1) and looking at failed switches (cause codes 2 to 15), we note from Table III that the main sources of failures in local switches account for almost 2/3 of all failures:

- *Hardware failures* (cause code 8)
- *Software design errors* (cause code 6)
- *Acts of God* (cause code (9)
- *External power failures* (cause code 12)

Tables III and IV give detailed information on the causes of local telecommunication switch outages; where Table IV consolidates cause codes to provide insights into major causal categories.

In this paper, major cause codes have been categorized as follows:

- *Scheduled Outages* (Cause Code 1): scheduled or planned maintenance activities.
- *Human Procedural Errors* (Cause Codes 2 to 5): installation/non-installation and maintenance/non-maintenance related errors made by the operating company technicians, or other errors made by system vendors or other vendors.
- *Design Errors* (Cause Codes 6 and 7): errors made by system vendors in designing the software or
- *Hardware Failures* (Cause Code 8): other hardware failures except design errors.
- *External Circumstances* (Cause Codes 9 to 14): other events, but external to the switch, which cause the switch to fail.
- *Others/Unknown* (Cause Code 15): all other events, different from the above, that cause the switch to fail.

TABLE II. OUTAGE DISTRIBUTION BY CAUSE

Cause Code	Description	No. Outages	%
1	Scheduled	3,885	30.2%
2	Procedural error (Telco install./maint.)	446	3.5%
3	Procedural error (Telco non-install./non-maint.)	376	2.9%
4	Procedural error (System vendor procedural error)	315	2.4%
5	Procedural error (Other vendor procedural error)	257	2.0%
6	Software design	1,078	8.4%
7	Hardware design	136	1.1%
8	Hardware failure	2,951	22.9%
9	Acts of God	935	7.3%
10	Traffic overload	17	0.1%
11	Environmental	83	0.6%
12	External power failure	896	7.0%
13	Massive line outage, cable cut, other	660	5.1%
14	Remote (Loss of facilities between host and remote)	309	2.4%
15	Other/unknown	516	4.0%
	Total	**12,860**	**100%**

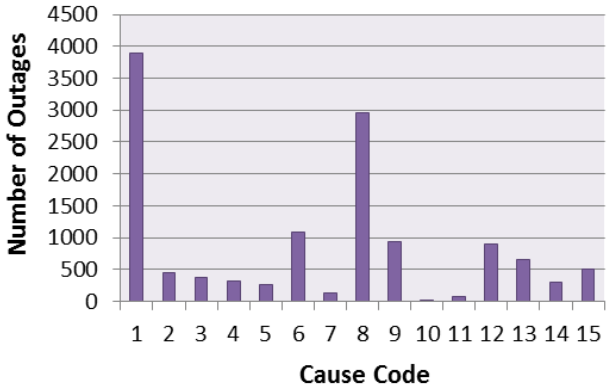


Figure 7. Local switch outage distribution.

TABLE III. FAILURE DISTRIBUTION BY CAUSE

Cause Code	Description	No. Outages	%
2	Procedural error (Telco install./maintenance)	446	5.0%
3	Procedural error (Telco non-install./non-maintenance)	376	4.2%
4	Procedural error (System vendor procedural error)	315	3.5%
5	Procedural error (Other vendor procedural error)	257	2.9%
6	Software design	1,078	12.0%
7	Hardware design	136	1.5%
8	Hardware failure	2,951	32.9%
9	Acts of God	935	10.4%
10	Traffic overload	17	0.2%
11	Environmental	83	0.9%
12	External power failure	896	10.0%
13	Massive line outage, cable cut, other	660	7.4%
14	Remote (Loss of facilities between host and remote)	309	3.4%
15	Other/unknown	516	5.7%
	Total	8,975	100%

TABLE IV. LOCAL SWITCH FAILURE DISTRIBUTION BY CATEGORY

Cause Code Category	No. Outages	%
Human Procedural Errors (2-5)	1,394	15.5%
Design Errors (6-7)	1,214	13.5%
Hardware Failures (8)	2,951	32.9%
External Circumstances (9-14)	2,900	32.3%
Others/Unknown (15)	516	5.7%
Total	8,975	100%

The percentage distribution of outages among the different cause code categories is shown in Table IV. It can be seen that scheduled outages (about 30% of all outages), hardware failures (about 23% of all outages) and external circumstances (also about 23% of all outages) are the main cause categories of local telecommunication switch outages. Table IV indicates that after hardware and external circumstances, procedural and design errors account for 29% of failures.

B. *Reliability Trends of Local Switches*

The outage rate of local switches can give us insight into reliability trends. As mentioned before, the variation of the outage rate over time implies a NHPP, which allows us to determine whether the system has been improving or deteriorating. Again, Laplace scores greater than 1.96 and less than -1.96 indicate strong statistical evidence of an increasing or decreasing trend, respectively.

A cumulative time series graph of all outages occurring over the 14-year study period is shown in Fig. 8.

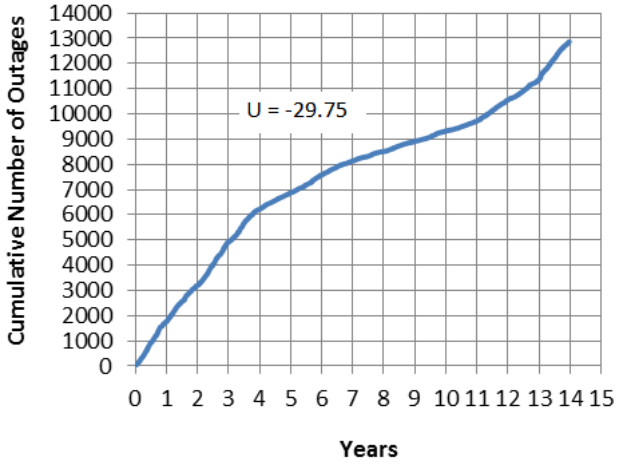


Figure 8. Switch reliability trend (all cause codes).

The figure, as well as the accompanying Laplace score (U = -29.75), show that, overall, there has been exceptionally strong reliability growth over the study period. However, there also seems to have been periods of reliability growth and periods of reliability deterioration along the study period. For example, the reliability was relatively constant from the start of the study period until the end of year 4; then it improved from year 5 until the end of year 11; subsequently, the trend started bending upwards towards the end of the study period (from year 12), which means that the reliability was starting to deteriorate.

By separating failures from scheduled outages we can gain a clearer picture of local switch reliability improvement. The reliability trend, considering only the scheduled outages, is shown in Fig. 9. From the figure, as well as from the accompanying Laplace score (U = -58.41), we conclude

remarkably strong reliability growth, especially from year 5 (i.e., 2000). These switches have "A" and "B" processors, one of which is primary and the other a live backup. Early in the study period vendors/carriers took the entire switch down for software/feature changes, resulting in short outages. Towards the end of the study period, one processor was taken down at a time for upgrade, and switch continued operating with many fewer outages..

On the other hand, from Fig. 10, which shows failures (all outages other than those that were scheduled), we notice a statistically significant reliability decrease (U = 2.81).

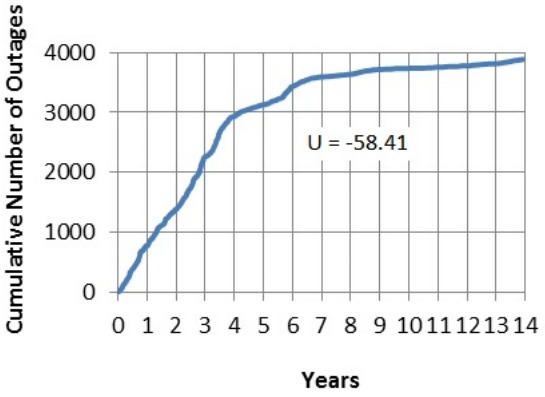


Figure 9. Switches reliability trend (Scheduled: cause code 1).

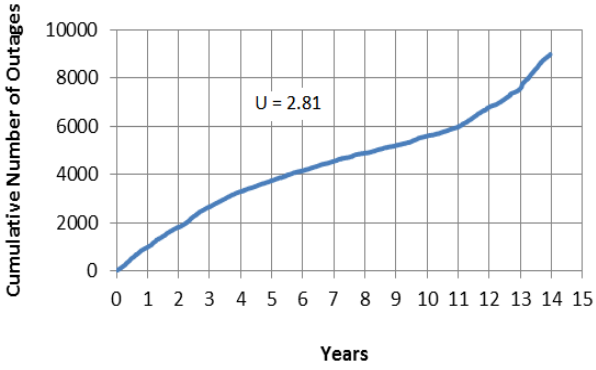


Figure 10. Switch reliability trend (Failures: cause codes 2-15).

This means that the switches have been failing more and more often due to increasing failures for the last three years of the study period. Therefore, by combining the two trends, one from scheduled outages and the other one from failures that resulted in outages, we get a trend that shows an overall reliability improvement in local switches from 1996 to 2009, but interesting insights are gained by separating scheduled outages and failure-induced outages.

C. *More Frequently Failing Switches Analysis*

There are many switches that failed more than once during the 14-year study period. A logarithmic chart of the number of outages encountered by failed switches during the study period is shown in Fig. 11. A total of 6,132 local switches have been responsible for 12,860 outages. The failed switches can be divided into two categories according to the number of outages that each unique switch encountered during the study period. In fact, 5,976 unique switches (about 97% of all switches) experienced 7 outages or less (which made about 80% of all outages); while 156 unique switches (about 3% of all switches) experienced 8 or more outages (which made about 20% of all outages) during the study period. The choice to focus on 8 or more outages was influenced by the fact that the curve from 1 outage to 8 outages in Fig. 11 was smooth and started to become irregular from 9 outages and up.

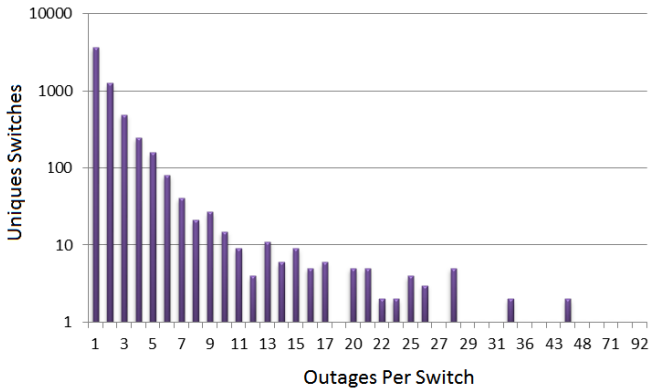


Figure 11. Unique local switch outage frequency (logarithmic scale).

Indeed, the more frequently failing switches account for only 3% of all individual switches failing over the study period. However, those 3% of the individual switches caused 20% of all outages that occurred.. From a survivability perspective, even if the more frequently failing switches caused 20% of all outages, they are responsible for only 7% of all LLH. Additionally, these more frequently failing switches are responsible for 9% of all lines down during the study period, and those lines out account for 22% of the total outage duration (more details are shown later in Table V).

As Fig. 11 uses a logarithmic scale in presenting the number of switches, instances of 1 switch having numerous outages is not indicated. Those instances are in Table V, where it is seen that remarkably, five different switches experienced 92, 75, 71, 60, and 48 outages. Additionally, two different switches both experienced 47 outages.

D. *Frequently and Less-Frequently Failing SwitchComparisons*

As mentioned earlier, the local switches have been divided into two categories, one of less frequently failing switches and another one of more frequently failing switches. Comparisons are made between impact (LLH),

switch location (rural or urban), outage cause codes, outages duration, time of day, day of week and month of year, and outage trends.

Each switch that fails has an impact on customers connected to that switch. A large switch (i.e., that has many lines connected to it) can go down for few minutes and still have the same impact as a small switch that goes down for many minutes. That is why the LLH metric is the best means to assess the impact of many different outages pooled together because it gives only one value, which is the product of the number of lines supported by the failed switch and the duration of the outage.

TABLE V. UNIQUE SWITCH OUTAGE FREQUENCY OVER 14 YEARS

No. Outages	No. Unique Switches	No. Outages	No. Unique Switches
1	3,701	22	2
2	1,259	23	2
3	489	25	4
4	247	26	3
5	158	27	1
6	81	28	5
7	41	29	1
8	21	30	1
9	27	31	1
10	15	32	2
11	9	36	1
12	4	40	1
13	11	43	1
14	6	47	2
15	9	48	1
16	5	60	1
17	6	71	1
18	1	75	1
20	5	92	1
21	5		

Before comparing the two categories of switches that experienced outages during the study period, let us have a look at the total LLH per year that resulted from all outages that occurred during the study period. We can see from Fig. 12 that there have been higher survivability deficits in 2001, 2005, and 2007. The figure also shows that the majority of outages were due to cause codes other than cause code 1, which is scheduled outage.

The high survivability decrease in 2001 is related to the 9/11 attacks in New York City since 77% (about 33,400,000 LLH) of the 2001 total LLH resulted from five switches located in New York, on September 11, 2001. Furthermore, the high survivability decrease in 2005 appears to be due to the 2005 Atlantic hurricane season, "the most devastating hurricane season the country has experienced in modern times," [25] as 97% of the 2005 LLH occurred during the hurricane season, which begins June 1st and ends November 30th. The 2005 Atlantic hurricane season's strongest hurricanes include hurricanes Wilma and Katrina. Similarly, the survivability decrease in 2007 appears to be due to the 2007 Atlantic hurricane season as 69% (about 47,000,000 LLH) of the 2007 LLH resulted from outages occurring during hurricane season. The 2007 Atlantic hurricane season's strongest hurricane was hurricane Dean.

A LLH comparison for each local switch category during the study period is given in Fig. 13. It is seen that the more frequently failing switches account for a very small portion of the impact on the PSTN.

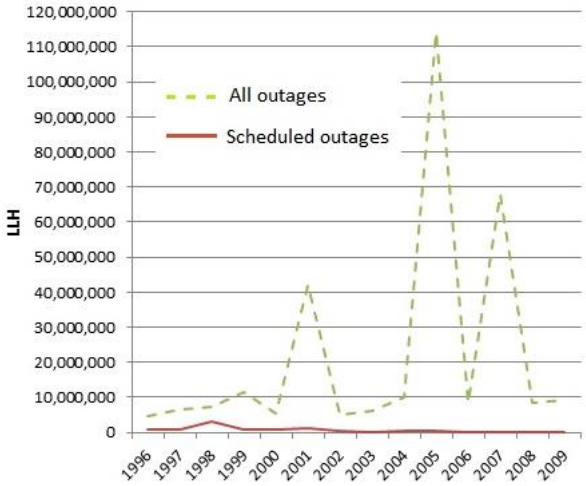


Figure 12. LLH per year for all versus scheduled outages.

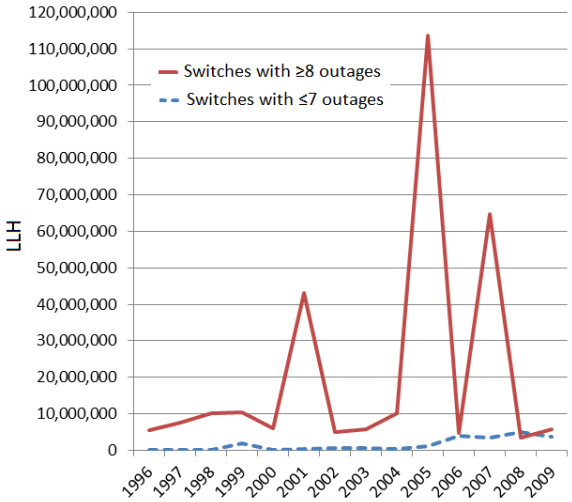


Figure 13. LLH per year for frequent/infrequectly out switches.

As for the 8 or more outages category (shown more clearly by Fig. 14), there have been two major survivability deficits along the 14-year study period. The first one occurred in 1999 where the LLH count reached almost 2,000,000 lines hours. The second and longest survivability decrease started in 2005 and continued until 2008 where it started to increase. There has been a slight survivability increase in 2007 but it was nothing compared to the survivability decrease that preceded in 2006 and the one that

followed in 2008. During this deterioration period, the highest LLH count reached almost 5,000,000 lines hours. The reason for the peaks is not discernible from the data.

A comparison of the number of outages/year experienced by switches in both categories, taking into account whether they are located in urban or rural areas, is shown in Figs. 15 and 16. In both categories rural switches suffered outages more often than urban switches.

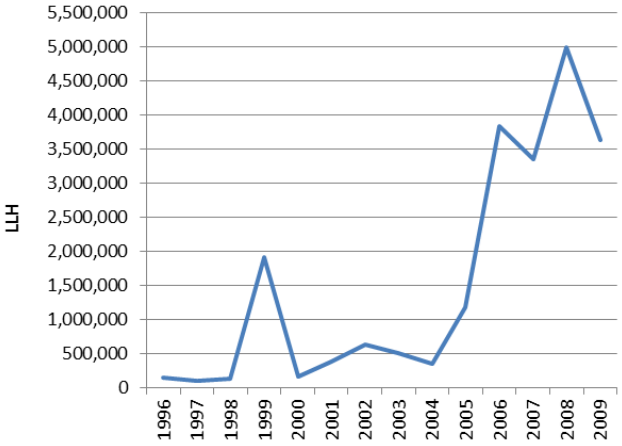


Figure 14. LLH per year for switches out ≥8 times.

From Fig. 15 we can see switches or more outages 7 outages or less had most of its outages at the start of the study period, from both rural and urban switches. On the other hand, switches or more outages 8 or more outages had most of its outages towards the end of the study period, mostly from rural switches (Fig. 16).

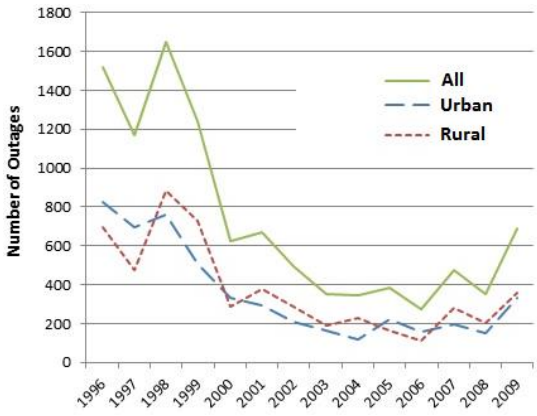


Figure 15. Urban vs rural switch outages per year (≤ 7 outages per switch)

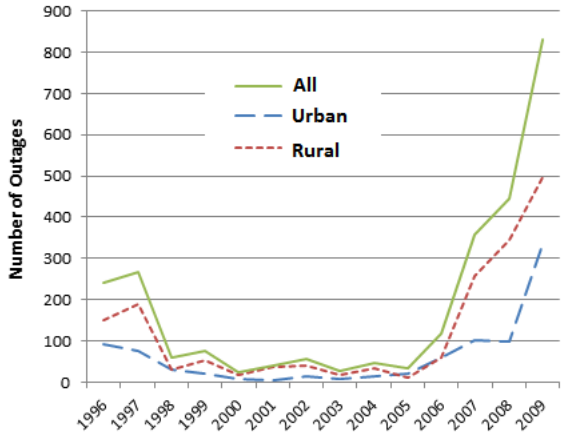


Figure 16. Urban vs. rural switch outages per year (≥8 outages per switch)

A comparison of the cause code frequency in both categories is shown in Fig. 17. Note that the switches that experienced 7 outages or less suffered considerably from scheduled outages (cause code 1) more often than the switches that experienced 8 or more outages. On the other hand, the switches that experienced 8 or more outages considerably suffered from hardware failures (cause code 8) and acts of God (cause code 9) more often than the switches in the other category.

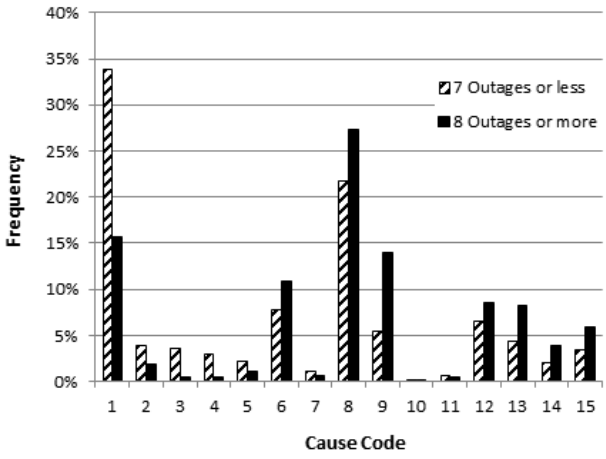


Figure 17. Outage cause code frequency.

By categorizing cause codes, we see from Fig. 18 that switches experiencing 8 or more outages suffered from design errors, random hardware failures, external circumstances and other/unknown causes more often than the switches experiencing 7 outages or less. Those failures might be the consequence of insufficient maintenance activities for the more frequently failing switches since most are located in rural areas. It is most interesting that the less frequently failing switches suffered a higher percentage of outages due to human procedural errors (Fig. 18).

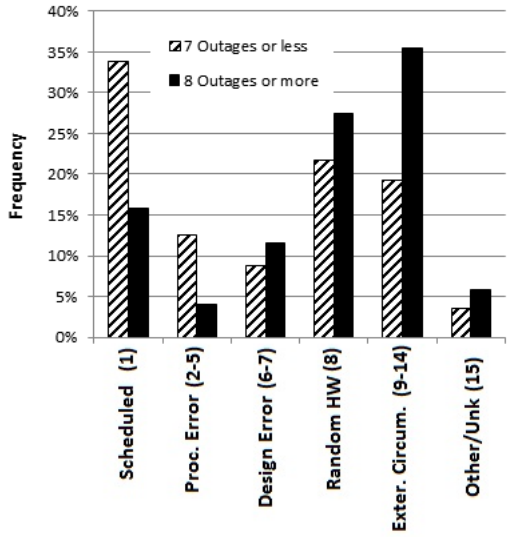


Figure 18. Causal Category frequency distribution.

Outages can be examined for all causes, scheduled causes, and failures. For each of these cases, the two outage frequency categories can also be compared.

1) Outages from all Causes

A numerical comparison between the two outage frequency categories, for all outage causes, is given in Table VI. Again, note that even if the switches experiencing 8 or more outages represent only 3% of all unique switches that failed during the 14-year study period, they are responsible for 20% of all outages that occurred. Additionally, those more frequently failing switches are responsible for 9% of all lines that went down during the study period, which corresponds to 22% of the total down time, in addition to 7% of the total LLH. Although the average duration values are roughly equal in both categories (3.6 and 3.2 hours per outage), the average LLH values are considerably different (8,129 and 28,890 LLH per outage) because the more frequently failing switches are in most cases smaller switches.

2) Scheduled Outages

A numerical comparison between the two outage frequency categories for scheduled outages is given in Table VII. In this case, the more frequently failing switches represent 3% of all unique switches that went down due to scheduled maintenance activities, which made 11% of all scheduled outages that occurred during the 14-year study period. Those more frequently failing switches are also responsible for 7% of all lines that went down due to scheduled maintenance activities during the study period, which corresponds to 21% of the total down time and hence 9% of the LLH due to scheduled maintenance activities. Also notice that average duration and average LLH for both categories are approximately equal.

TABLE VI. OUTAGE CHARACTERISTICS COMPARISON

All Cause Codes	All Outages	≥ 8 Outages	≤ 7 Outages	% ≥ 8	%≤ 7
No. of Outages	12,860	2,623	10,237	20%	80%
No. of Switches	6,132	156	5,976	3%	97%
Tot. Switch Lines Out	123.2 M	10.5 M	112.7 M	9%	91%
Tot. Dur. (Hours)	42,391	9,443	32,947	22%	78%
Total LLH	317.1 M	21.3 M	295.8 M	7%	93%
Avg. Switch Size (Lines)	9,584	4,040	11,005		
Avg. Dur. (Hours)	3.3	3.6	3.2		
Avg. LLH	24,655	8,129	28,890		
Median TOD	10:14am	11:38am	9:51am		
Mean TOD	10:55am	11:41am	10:43am		
Median DOW	4.13	4.36	4.11		
Mean DOW	4.27	4.37	4.25		
Median MOY	6.89	6.98	6.89		
Mean MOY	6.92	6.90	6.93		
Med. MSA No=Rural	No	No	No		

TABLE VII. SCHEDULED OUTAGE COMPARISON

Cause Code 1	All Outages	≥ 8 Outages	≤ 7 Outages	% ≥ 8	% ≤ 7
No. of Outages	3,885	413	3,472	11%	89%
No. of Switches	2,470	86	2,384	3%	97%
Tot. Switch Lines Out	57.6	4.0 M	53.6 M	7%	93%
Tot. Dur. (Hours)	998	211	787	21%	79%
Total LLH	9.2 M	0.9 M	8.3 M	9%	91%
Avg. Switch Size(Lines)	14,830	9,605	15,451		
Avg. Dur. (Hours)	0.3	0.5	0.2		
Avg. LLH	2,385	2,093	2,420		
Median TOD	5:47am,	4:37am	6:00am		
Mean TOD	10:38am	10:32am	10:39am		
Median DOW	4.18	5.03	4.16		
Mean DOW	4.40	4.68	4.36		
Median MOY	6.92	6.56	6.97		
Mean MOY	7.04	6.89	7.06		
Med. MSA (No=Rural)	No	No	Yes		

3) Outages Due to Failures

A numerical comparison between the two outage frequency categories, due to failures (cause codes 2-15) is in Table VIII. In this case, the more frequently failing switches still represent only 3% of all unique failed switches resulting in outages, but they caused 25% of all failures occurring. Those more frequently failing switches are also responsible for 10% of all lines that went down, which corresponds to 22% of the total down time, and hence 7% of the total LLH. Also notice that even if the average duration in both categories are roughly equal, the average LLH per switch are very different (9,257 vs. 42,475 LLH).

The t-Test results are given in Table IX. Again, a significance level of 0.05 has been used. Significant temporal differences between frequently out and less frequently out switches are apparent.

E. Local Switch Outage and Failure Trends

The outage trends of local switches in both categories during the 14-year study period, taking into consideration all cause codes, are shown in Figs. 19 and 20.

TABLE VIII. FAILURE CHARACTERISTICS COMPARISON

Cause: Codes 2-15	All Outages	≥ 8 Outages	≤ 7 Outages	% ≥ 8	% ≤ 7
Number Outages	8,975	2,210	6,765	25%	75%
Number Switches	4,517	154	4,363	3%	97%
Tot. Switch Lines	65.6 M	6.6 M	59.0 M	10%	90%
Tot. Dur. (Hours)	41,393	9,232	32,160	22%	78%
Total LLH	307.8 M	20.5 M	287.3 M	7%	93%
Avg. Switch Size	7,313	3,000	8,723		
Avg. Dur. (Hours)	4.6	4.2	4.8		
Avg LLH	34,295	9,257	42,475		
Median TOD	10:57am	12:00pm	10:34am		
Mean TOD	11:02am	11:54am	10:45am		
Median DOW	4.09	4.29	4.06		
Mean DOW	4.22	4.32	4.19		
Median MOY	6.89	7.08	6.82		
Mean MOY	6.87	6.90	6.86		
Med. MSA (N=Rural)	N	N	N		

TABLE IX. TEMPORAL T-TEST RESULTS

t-Test (MEANS)	≥ 8 Outages	≤ 7 Outages	Difference	P-value (one-tail)
TOD All Codes	11:41 am	10:44 am	YES	0.000
TOD Scheduled	10:32 am	10:40 am	NO	0.400
TOD Failures	11:54 am	10:46 am	YES	0.000
DOW All	4.37	4.25	YES	0.001
DOW Scheduled	4.68	4.36	YES	0.000
DOW Failures	4.32	4.19	YES	0.003
MOY All	6.90	6.93	NO	0.335
MOY Scheduled	6.89	7.06	NO	0.146
MOY Failures	6.90	6.86	NO	0.322

For the switches that experienced 7 outages or less, we can see from Fig. 19, as well as from the accompanying Laplace score (U = -48.66), that there has been steady reliability growth from the start of the study period until the end of the study period.

Figure 19. Outage trend for switches with ≤ 7 outages.

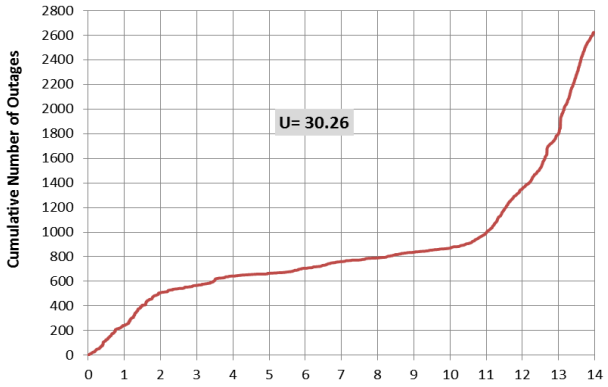


Figure 20. Outage trend for local switches or more outages ≥8 outages.

However, for the switches that experienced 8 or more outages, we can see from Fig. 20, as well as from the accompanying Laplace score (U = 30.26), that the reliability has exhibited dramatic deterioration. However, there seem to have been periods of strong reliability growth and deterioration during the study period. In order to gain more insight into the causes of the outage trend for switches that experienced 8 or more outages along the study period, the outage trend has been divided into three regions as follows:

- 0 ≤ Region I < 2 years
- 2 years ≤ Region II < 11 years
- 11 years ≤ Region III < 14 years

The outage trend of more frequently failing switches divided into regions and considering all the cause codes is shown in Fig. 21. From the figure we can see that the reliability was constant in Region I (U = -0.29) from 1996 to 1997; it then slightly decreased in Region II (U = 3.22) from 1998 to 2006; and finally things got worse in Region III where the reliability sharply decreased (U = 11.84) from 2007 to 2009.

The number of outages per cause code that occurred in each region is shown in Table X. We can see from the table that most of the outages occurred in Region III (62% of all outages in all regions). We can also note from the table that most of the outages in Region I resulted from cause code 1 (scheduled outages) and cause code 6 (software design); most of the outages in Region II resulted from cause codes 1 and 8 (hardware failures), and most of the outages in Region III resulted from cause codes 8 and 9 (acts of God).

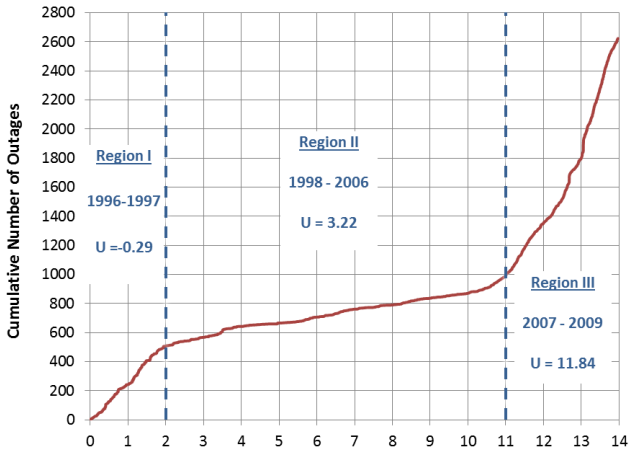


Figure 21. Trend analysis for local switches or more outages ≥8 outages.

By combining the cause codes in categories we can see from Table XI that the major causes of outages in Region I are scheduled outages and design errors; the major causes of outages in Region II are external circumstances, hardware failures and scheduled outages; and the major causes of outages in Region III are external circumstances and hardware failures.

TABLE X. OUTAGE FREQUENCY BY CAUSE CODE FOR MORE FREQUENTLY FAILING SWITCHES IN REGIONS

Cause Code	All		Reg. I		Reg. II		Reg. III	
	No.	%	No.	%	No.……	%	No.	%
1	413	15.7	253	49.8	95	19.5	65	4.0
2	49	1.9	5	1.0	24	4.9	20	1.2
3	12	0.5	1	0.2	5	1.0	6	0.4
4	14	0.5	1	0.2	7	1.4	6	0.4
5	30	1.1	1	0.2	8	1.6	21	1.3
6	284	10.8	213	41.9	52	10.7	19	1.2
7	19	0.7	3	0.6	3	0.6	13	0.8
8	718	27.4	26	5.1	112	23.0	580	35.6
9	368	14.0	0	0.0	21	4.3	347	21.3
10	4	0.2	0	0.0	1	0.2	3	0.2
11	14	0.5	0	0.0	3	0.6	11	0.7%
12	224	8.5	4	0.8	41	8.4	179	11.0%
13	215	8.2	1	0.2	61	12.6	153	9.4%
14	104	4.0	0	0.0	41	8.4	63	3.9%
15	155	5.9	0	0.0	12	2.5	143	8.8%
Total	**2,623**	**100**	**508**	**100**	**486**	**100**	**1,629**	**100%**
Pcnt..	**100%**		**19.4%**		**18.5%**		**62.1%**	

Note that the number of scheduled outages considerably decreased from Region I to Region III. Perhaps that is the reason why the number of outages resulting from hardware failures and external circumstances considerably increased from Region I to Region III, as the switches received maintenance.

TABLE XI. OUTAGE FREQUENCY BY CAUSE CATEGORY FOR MORE FREQUENTLY FAILING SWITCHES IN REGIONS

Cause Category	All		Region I		Region II		Region III	
	No.	%	No.	%	No.	%	No.	%
Sched. (1)	413	15.7	253	49.8	95	19.5	65	4.0
Proc. Err (2-5)	105	4.0	8	1.6	44	9.1	53	3.3
Design Err.(6-7)	303	11.6	216	42.5	55	11.3	32	2.0
Hdw (8)	718	27.4	26	5.1	112	23.0	580	35.0
Ext. Circum. (9-14)	929	35.4	5	1.0	168	34.6	756	46.4
Other (15)	155	5.9	0	0.0	12	2.%	143	8.8
Total	**2,623**	**100**	**508**	**100**	**486**	**100**	**1,629**	**100**

The outage trends of local switches in both categories during the 14-year study period, for scheduled outages , are shown in Figs. 22 and 23. For the switches that experienced 7 outages or less, we can see from Fig. 22, as well as from the accompanying Laplace score (U = -56.03), that the reliability strongly improved during the study period. This means that there have been less and less scheduled maintenance activities in those switches along the study period.

For the switches that experienced 8 or more outages, we can see from Fig. 23, as well as from the accompanying Laplace score (U = -16.67), that the reliability in those

switches also improved during the study period. In fact, the reliability has been constant from the start of the study period until year 11, but with a high outage rate from the start to year 2 (around 125 outages per year) and then a low outage rate from year 3 to year 11 (around 11 outages per year). The overall reliability growth was moderate because the switches seemed to have had a short period of reliability deterioration at the end of the study period (from year 12 to the end). This means that there have been more scheduled maintenance activities during the last few years of the study period.

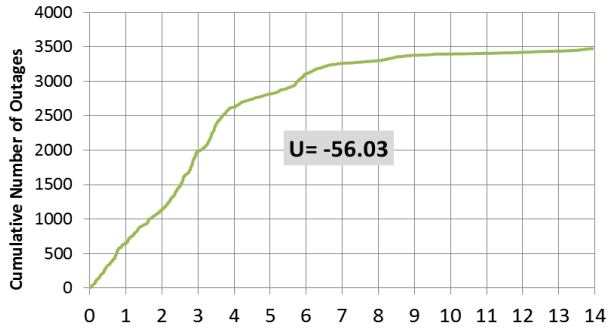


Figure 22. Scheduled outage trend for switches with ≤ 7 outages.

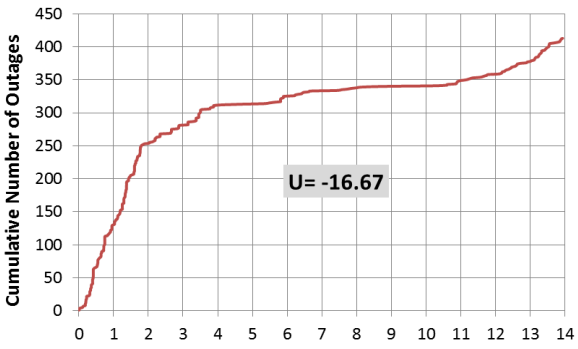


Figure 23. Scheduled outage trend for switches with ≥8 outages.

The failure trends of local switches in both categories during the 14-year study period, for failures (cause codes 2-15) are shown in Figs. 24 and 25. For the switches that experienced 7 outages or less, we can see from Fig. 24, as well as from the accompanying Laplace score (U = -19.72), that there has been steady reliability growth from the start of the study period until the end of the study period, with exception of last year.

In general, the more frequently failing switches showed a serious reliability deterioration, for cause codes 2-15 over the study period as shown by the Laplace score (U = 40.17). For the switches that experienced 8 or more outages, we can see from Fig. 25 that the reliability seemed to improve from the start of the study period to year 11, and then it extremely deteriorated from year 12 until the end of the study period.

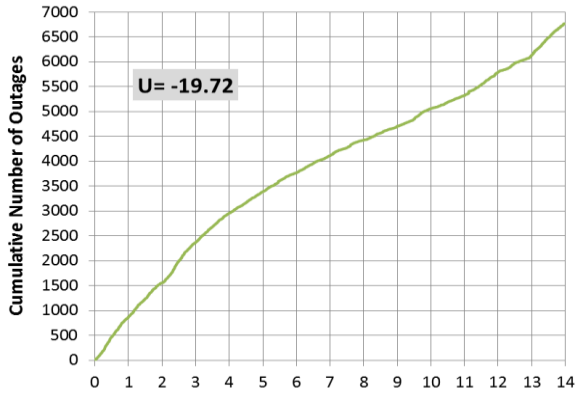


Figure 24. Failure trend for switches with ≤7 outages.

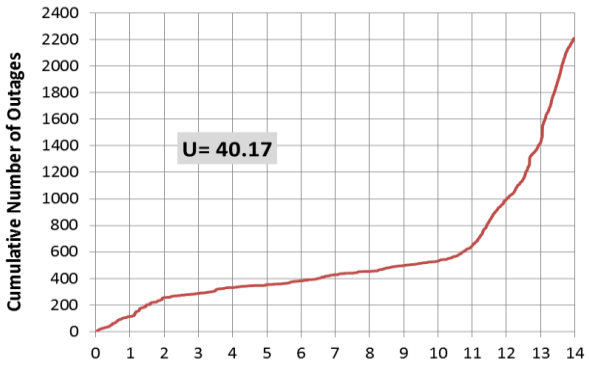


Figure 25. Failure trend for switches with≥ 8 outages.

F. *Manufacturers of More Frequently Failing Local Telecommunication Switches*

This section discusses the manufacturers of local switches or more outages 8 or more outages during the 14-year study period. As the percentages of switches by manufacturer in the total switch population are unknown, this presentation in this ection is of interest, but not a conclusive assessment.

The 156 unique switches that compose the 8 or more outages were classified among 25 different switch models, then aggregated into outages per manufacturer. There are indeed some switch models that account for the majority of outages. Here however, the 156 unique switches with 8 or more outages are classified among seven different switch manufacturers. The Northern Telecom switches are the ones that failed the most (about 38% of all outages) as shown in Table XII. The Automatic Electric and Ericsson switches come at the second place with roughly equal percentages of about 19% each. Lucent technology switches represent a very low percentage of the frequently out switches. Northern Telcom and Lucent switches are the most prolific in the switch population. The age of switches was not reported, however, we would expect older switches to have more random hardware outages.

TABLE XII. SWITCH MANUFACTURER OUTAGE FREQUENCY FOR MORE FREQUENTLY FAILING SWITCHES

Switch Manufacturer	Outage Frequency (No. of Outages)		No. Unique Switches	
Northern Telecom	999	38.1%	57	36.5%
Automatic Electric	505	19.3%	25	16.0%
Ericsson	493	18.8%	35	22.4%
Stromberg Carlson	264	10.1%	16	10.3%
Siemens	193	7.4%	9	5.8%
Unknown	130	5.0%	11	7.1%
Lucent Technologies	39	1.5%	3	1.9%
Total	**2623**	**100 %**	**156**	**100%**

VI. RESEARCH FINDINGS AND CONCLUSION

A. Research Findings

The research findings are presented for each research question, in the order presented earlier:

1) What are the major causes of local telecommunication switch outages?

In general, the major causes of local telecommunication switch outages between 1996 and 2009 were scheduled maintenance activities (i.e., scheduled outages) and hardware failures. When considering only failure induced outages, most local switch failures resulted from hardware failures, software design errors, acts of God and external power failures.

2) Is the reliability of local switches improving, constant, or deteriorating?

In general, the reliability of local switches has been improving over the study period. However, significant deterioration towards the end of the study period is very apparent. Also, when scheduled outages are discounted, the outage trend due to switch failures dramatically increased.

3) Are there individual switches that experience outages/failures more so than others?

Only 3% of all individual switches experiencing outages were responsible for 20% of all outages and 22% of the total downtime. These more frequently failing switches encountered eight or more outages each during the study period.

4) Are there similarities/dissimilarities between switches failing more often and those that do not?

a) Impact of Outages

Over the 14 years study period, scheduled outages had little apparent impact. Three years in particular (2001, 2005, and 2007) account for the lion's share of impact due to switch failures. The year 2001 corresponds to the 9-11 disaster where five switches were out for over a month, while the years 2005 and 2007 correspond to heavy hurricane induced outages. Switches having 7 or less outages are several times larger than switches suffering 8 or more outages. The less frequently failing switches account for 3% of the switches, 25% of the outages, but only 7% of the lost line hours.

b) Rural versus Urban

In both categories, rural switches failed more often than urban switches. However, most of the outages in the category of less frequently failing switches occurred at the start of the study period; whereas, most of the outages in the category of more frequently failing switches occurred towards the end of the study period. Additionally, most of outages in both categories, no matter where the switches are located, resulted from causes other than scheduled maintenance activities. Interestingly, switches failing more often were about three times smaller than switches failing less often. About 52% of less frequently out switches were rural, while about 66% of more frequently failing switches were rural.

c) Outage causes

For all failing switches, random hardware accounts for 33%, Design error 14% and Acts of God/Power outages 20%. The less frequently failing switches suffered considerably more often from scheduled outages; whereas, the more frequently failing switches suffered considerably more often from hardware failures and acts of God (i.e., natural disasters). The more frequently failing switches also often suffered from software design errors, external power failures, massive line outages, loss of facilities between host and remote, and other/unknown causes. There is an impression that the more frequently failing switches suffered from random hardware failures and external circumstances because they were not frequently maintained (i.e., no many scheduled outages); on the other hand, the less frequently failing switches suffered from human procedural errors maybe because they encountered a lot of scheduled outages.

d) Outage duration

The average switch outage duration was 3.3 hours. Even though the average duration per outage in both categories is approximately the same (considering all cause codes), the average impact (LLH) per outage is considerably different. This implies that the more frequently failing switches are smaller than the less frequently failing switches. Perhaps larger switches receive better maintenance responses, but this is not discernible from the data.

e) Time of Day

The differences between the mean times of day in both categories are statistically significant except when

considering the scheduled outages only.. No additional insights as to why were discernible from the data.

f) Day of Week

The differences between the mean days of week in both categories are also statistically significant. No additional insights as to why were discernible from the data.

g) Month of Year

The differences between the mean months of year in both categories are not statistically significant.

h) Outage/failure trends

For all switches, failure trends exhibited a dramatic increase, when scheduled outages were discounted. The reliability of the less frequently failing switches has been steadily improving during the study period; whereas, the reliability of the more frequently failing switches has been deteriorating during the study period, especially towards the end. No additional insights as to why were discernible from the data.

5) Are there switch models or manufacturers that account for outages/failures more so than others?

In the category of more frequently failing switches, the top three manufacturers of switches that experienced more outages than others are the Northern Telecom, Automatic Electric and Ericsson, although the percentage of those switches in the total switch population is not known. However, Northern Telecom switches accounted for about 38% of the more frequently failing switch outages while Lucent Technologies accounted for only 1.5%.

B. Conclusions and Limitations

This research reveals that there are significant reliability deficits in more frequently failing local telecommunication switches, especially towards the end of the study period (i.e., 2009). In fact, the last three years of the study period show significant reliability deterioration. This research also shows that there are significant differences in the causes of outages, impacts of outages, and switch characteristics between the two categories (the more frequently failing switches and the less frequently failing switches). In fact, only 3% of the failing switches accounts for 20% of all switch outages and for 7% of the lost line hours. Limitations include that the data is non-experimental, reported by a number of different carries who applied many cause codes. The consistency between reporting companies is unknown, as is the accuracy of the reports. Although scheduled outages reduced dramatically, it is not known if those outages induce subsequent failures, such as introducing software/hardware bugs.

The limitations of this work include the fact that outage reports are made by different carriers, and by different people at individual carriers. The danger is lack of consistency in reporting. A counterpoint is that the reporting rules were constant over the study period. The fact the FCC has stopped collecting reports on switch outages since 2009 is unfortunate, as future trends of local telecommunication switch reliability and survivability cannot be assessed from publically available data. However, the trends reveal opportunities for corrective/preventive actions on the part of switch manufacturers and service providers.

REFERENCES

[1] Snow, A. P., Shyirambere, A., Weckman G., Arauz, J., "A reliability and survivability analysis of US local telecommunication switches that sxperience frequent outages," The Twelfth International Conference on Networks (ICN 2013), IARIA.

[2] Leemis, L. M., Reliability: Probabilistic Models and Statistical Methods. Englewood Cliffs: Prentice-Hall, 1995.

[3] Snow, A P. "The Reliability of Telecommunication Switches," Six International Conference on Telecommunications Systems: Modeling and Analysis, pp. 288-295, March 1997.

[4] Snow, A. P., "Internet implications of telephone access," IEEE Computer, Vol. 32, No. 9, pp. 108-110, September 1999.

[5] Kuhn, R., "Sources of failure in the public switched telephone network," IEEE Computer, Vol. 30, No. 4, April 1997.

[6] Snow, A. P., "Assessing pain below a regulatory outage reporting threshold," Telecommunications Policy, Vol. 28, pp. 523–536, Elsevier, 2004.

[7] Snow, A. P., and Agarwal, S., "Towards an optimal network survivability threshold," Ninth IFIP/IEEE International Symposium on Integrated Network Management, pp. 761 – 774, 2005.

[8] Davidson, J., Peters, J., Bhatia, M., Kalidindi, S., and Mukherjee, S., Voice over IP Fundamentals, Second Edition. Cisco Press, 2006.

[9] Gillan, J., and Malfara, D. The Transition to an All-IP Network: A Primer on the Architectural Components of IP Interconnection. National Regulatory Research Institute, May 2012.

[10] FCC. Local Telephone Competition: Status as of December 31, 2010. Industry Analysis and Technology Division, Wireline Competition Burea, October 2011.

[11] Dryburgh, L., and Hewett, J., Signaling System No. 7 (SS7/C7): Protocol, Architecture, and Services. Cisco, 2004.

[12] Bedell, P., Cellular/PCS Management, New York: McGraw-Hill, 2000.

[13] Dialexia., White Paper: Softswitch Technology and the Migration to Full Convergence, 2005, November, Retrieved October 2012, from Dialexia: http://www.dialexia.com/index.php/en/technology/white-papers/86-white-paper-softswitch-technology-and-the-migration-to-full-convergence

[14] Alcatel-Lucent, STRATEGIC WHITE PAPER: How to Effectively Transition to VoIP and IMS - Big Bang or Phased Approach?, 2008, Retrieved October 2012, from Alcatel-Lucent: http://www.webtorials.com/main/resource/papers/lucent/paper92/VoIP-and-IMS.pdf

[15] ITALTEL, and CISCO, NGN for PSTN Transformation, Tech. Overview, May 2009, Retrieved Oct.2012,from http://www.cisco.com/web/partners/pr67/pr36/docs/NPT_Cisco_Italtel_technical_flyer_May09.pdf

[16] Wang, P., and Coit, D. W. , "Repairable systems reliability tend tests and evaluation," Reliability and Maintainability Symposium Proceedings Annual, pp. 416-421, January 2005.

[17] Adams, T. C., THE LAPLACE TEST, Retrieved September 2012, from National Aeronautics and Space Administration, KSC:

http://kscsma.ksc.nasa.gov/Reliability/Documents/Laplace_Test.pdf

[18] Ellison, R., Fisher, D., Linger, R., Lipson, H., Longstaff, T., and Mead, N., Survivable Network Systems: An Emerging Discipline, Pittsburgh: Carnegie Mellon University, 1999.

[19] Avizienis, A., Laprie, J.-C., Randell, B., and Landwehr, C., "Basic concepts and taxonomy of dependable and secure computing," IEEE Transactions on Dependable and Secure Computing, Volume 1, Issue 1, pp. 11-33, Jan. 2004.

[20] Federal Communications Commission, "About the FCC", Retrieved February 2013, from Federal Communications Commission: http://transition.fcc.gov/aboutus.html

[21] Federal Communications Commission, Automated Reporting Management Information System (ARMIS), Retrieved October 2012, from Federal Communications Commission: http://transition.fcc.gov/wcb/armis/

[22] Timmins, P., Telecommunications Database. Retrieved from TelcoData.us: http://www.telcodata.us/data-downloads 2012.

[23] Switch Search. (n.d.). Retrieved October 2012, from Local Calling Guide: http://www.localcallingguide.com/

[24] Johnson, B., and Christensen, L. Educational Research: Quantitative, Qualitative, and Mixed Approaches, SAGE Publications, 2012.

[25] National Oceanagraphic and Atmospheric Administration, NOAA Reviews Record-Setting 2005 Atlantic Hurricane Season: Active Hurricane Era Likely To Continue.2006, April 13. Retrieved March 24, 2013, from NOAA News: http://www.noaanews.noaa.gov/stories2005/s2540.htm.

Applying the Accumulation of Cross-Power Spectrum Technique for Traditional Generalized Cross-Correlation Time Delay Estimation

Radu-Sebastian Marinescu[*,#], Andi Buzo[*], Horia Cucu[*], Corneliu Burileanu[*]
[#]Research & Development, Rohde & Schwarz Topex
[*]SpeeD Laboratory, University Politehnica of Bucharest
Bucharest, Romania
radu-sebastian.marinescu@rohde-schwarz.com, {andi.buzo, horia.cucu, corneliu.burileanu}@upb.ro

***Abstract*—In many real time applications, time delay estimation requires a special solution. Despite the various approaches, which were proposed over the years, the topic remains hot for digital signal processing because of its large field of applications and implementation forms. Among different classes of methods for this issue, general cross-correlation method is wildly used. It offers good results and does not need an adaptation time, like those based on adaptive filtering. In this paper, we make a survey and compare the most popular generalized cross-correlation methods. We extend the analysis, by applying the accumulation of cross-power spectrum technique, for all well known generalized cross-correlation methods. The comparisons are provided by detailed numerical and simulation analysis, using several metrics. Based on the accuracy rate, error rate, standard deviation of relative error and computing time we provide new considerations for traditional generalized cross-correlation methods.**

Keywords - Time Delay Estimation, General Cross-Correlation, Accumulated Cross-Power Spectrum.

I. INTRODUCTION

In this paper, we continue to evaluate the performances of our recently proposed time delay estimation methods. This work pushes further the analysis done in [1]-[3] for time delay estimation (TDE). Despite the various techniques developed over the years, the topic continues to be interesting. As technology evolved, more and more applications demanded a real time solution for time delay estimation. For echo canceling, acoustics, radar and sonar localization, seismic and medical processing, pattern detection and speech enhancement, scientists are still looking to improve the existent solutions. However, the variety of TDE applications, implementation aspects and proper constraints, inhibit the design of a unique solution. Instead, various approaches have been developed based on application specific aspects [1].

The various approaches for TDE can be grouped into three categories: a) *generalized cross-correlation* (GCC), b) least-mean squares (LMS) adaptive filtering [4]-[10], and c) adaptive eigenvalue value decomposition (AEVD). Based on the specific aspects required by an application, an optimal solution has to be chosen. As showed in [11] by Benetsy, AEVD technique offers an efficient solution for audio applications from reverberant environment. The adaptive filtering methods have a different approach. This leads to a high accuracy results, which need an adaption time. This solution can be very effective for some applications, but for real time systems the adaptation time makes them unusable. For the last ones, an optimal solution is represented by the generalized cross-correlation methods. They provide fast results, keeping also an acceptable accuracy level.

The main contributions of this paper are multiple. We provide an in-depth analysis of the previous and proposed methods by comparing them from the accuracy and processing speed points of view. We perform a new evaluation for the most used GCC methods and extend the accumulating cross-power spectrum scheme to all well known GCC methods, for a deeper evaluation. Finally, we show that our recently proposed methods, for multiple frames TDE, outperform the previous GCC approaches, offering a lower computation time and a higher accuracy rate even at low signal-to-noise ratios.

The rest of the paper is organized as follows. In the next section, we review the related work over the years. Section II also contains the description of our recently proposed methods. Section III is reserved for experimental results and analysis discussions, grouped in different parts: A) experimental setup, B) calibration of the proposed methods, and C) extended evaluation for accuracy and processing time of all presented methods. Finally, the main conclusions and further work are addressed in Section IV.

II. RELATED WORK FOR GENERALIZED CROSS-CORRELATION TIME DELAY ESTIMATION

For two signals $y_1(t)$ and $y_2(t)$, which are two noisy and delayed versions of the same transmitted signal $x(t)$, time delay estimation aims at finding the relative delay between them. Among the various developed approaches to TDE, the most popular and time-efficient method remains the one based on the cross-correlation of the two signals. In 1976, Knapp and Carter proposed in [12] the generalized cross-correlation methods. They pointed out that a common

method of determining the time delay is to compute the cross-correlation function:

$$R^{g}_{y_1y_2}(\tau) = E\left[y_1(t) \cdot y_2(t-\tau)\right] \quad (1)$$

where E denotes expectation. The argument τ that maximizes (1) provides an estimation of delay.

The cross-correlation between $y_1(t)$ and $y_2(t)$ is related to the cross-power spectral density function by the well known Fourier transform relationship:

$$R^{g}_{y_1y_2}(t) = \int_{-\infty}^{\infty} G_{y_1y_2}(f) \cdot e^{j2\pi ft} df \quad (2)$$

To improve the accuracy of delay estimation, a pre-filtering of the inputs is necessary before calculating the cross correlation. When signals $y_1(t)$ and $y_2(t)$ have been filtered with filters having transfer functions $H_1(f)$ and $H_2(f)$ the cross power spectrum between the filter outputs is given by:

$$G^{g}_{y_1y_2}(f) = H_1(f) \cdot H_2^{*}(f) \cdot G_{y_1y_2}(f). \quad (3)$$

Therefore, the generalized cross-correlation between $y_1(t)$ and $y_2(t)$ is:

$$R^{g}_{y_1y_2}(t) = \int_{-\infty}^{\infty} \Psi(f) \cdot G_{y_1y_2}(f) \cdot e^{j2\pi ft} df \quad (4)$$

where:

$$\Psi(f) = H_1(f) \cdot H_2^{*}(f) \quad (5)$$

and denotes the general frequency weighting [12].

Over the years, different weighting functions were proposed to improve the estimation process of the basic cross-correlation. In Table I, we present the various well known weighting functions, used in this work for a detailed analysis, where G_{y1y1} and G_{y2y2} are auto power spectrum of the noisy signals and $\gamma^2_{y_1y_2}(f)$ is the signal's coherence function.

$$G_{y_1y_1}(f) = \int_{-\infty}^{\infty} R_{y_1y_1}(t) \cdot e^{j2\pi ft} df \quad (6)$$

$$G_{y_2y_2}(f) = \int_{-\infty}^{\infty} R_{y_2y_2}(t) \cdot e^{j2\pi ft} df \quad (7)$$

$$\gamma^2_{y_1y_2}(f) = \frac{\left|G_{y_1y_2}(f)\right|^2}{G_{y_1y_1}(f) \cdot G_{y_2y_2}(f)} \quad (8)$$

For the normal Cross-Correlation (CC) the weighting function $\Psi(f)$ is 1. This is the basic and the fastest computing GCC, because it has no weighting operations.

The Eckart filter derives its name from work done in this area in [13], published in 1951. It maximizes the deflection criterion, i.e., the ratio of the change in mean correlator output due to signal present to the standard deviation of the correlator output due to noise alone [12].

TABLE I. GCC WEIGHTING FUNCTIONS

GCC name	Weighting function
CC	1
Eckart	$\frac{\left\|G_{y_1y_2}(f)\right\|}{\left[G_{y_1y_1}(f)-\left\|G_{y_1y_2}(f)\right\|\right]\cdot\left[G_{y_2y_2}(f)-\left\|G_{y_1y_2}(f)\right\|\right]}$
ROTH	$1/G_{y_1y_1}(f)$
HT (ML)	$\frac{\left\|\gamma_{y_1y_2}(f)\right\|^2}{\left\|G_{y_1y_2}(f)\right\|\cdot\left[1-\left\|\gamma_{y_1y_2}(f)\right\|^2\right]}$
SCOT	$1/\sqrt{G_{y_1y_1}(f)\cdot G_{y_2y_2}(f)}$
CSP (PHAT)	$1/\left\|G_{y_1y_2}(f)\right\|$
CSP-m	$1/\sqrt[m]{G_{y_1y_1}(f)\cdot G_{y_2y_2}(f)}$
HB	$\left\|G_{y_1y_2}(f)\right\|/G_{y_1y_1}(f)\cdot G_{y_2y_2}(f)$
Wiener	$\left\|\gamma_{y_1y_2}(f)\right\|^2$
ρ-CSP	$1/\left\|G_{y_1y_2}(f)\right\|^{\rho}$
ρ-CSPC	$\frac{1}{\left\|G_{y_1y_2}(f)\right\|^{\rho}+\min\left[\gamma_{y_1y_2}(f)^2\right]}$

Twenty years later, in 1971, Roth proposed a new processor in [14]. It has desirable effect of suppressing those frequency regions where G_{y1y1} is large and the estimate of G_{y1y2} is more likely to be in error [12].

In the same year it was proposed another weighting function, the HT processor, by Hannan and Thomson. This assigns greater weight in regions of frequency domain where the coherence is large [15]. In [12], it was shown that HT processor is a maximum likelihood (ML) estimator for time delay under usual conditions. Under a low signal-to-noise ratio restriction, the HT processor is equivalent to Eckart prefiltering and cross-correlation.

The SCOT (Smoothed Coherence Transform) was introduced by Carter, Nuttall and Cable in 1973, to reduce the influence of a strong tonal [16]. However, for smoothed

signal and noise spectra, Hassab and Boucher [17][18] have noted that the additional SCOT weighting function has weakened the performance of the basic cross correlator, while other functions have improve it.

Phase Transform (PHAT) or Cross-power Spectrum Phase (CSP) was developed purely as an ad-hoc technique to avoid spreading of the above two presented operators. Ideally, PHAT does not suffer the spreading that other processors do. Also, because it weights G_{y1y2} as the inverse of $|G_{y1y2}|$, the errors are accentuated where signal power is smallest [12].

In 1979, the HB processor was presented by Hassab and Boucher. It is similar to SCOT in that, for highly dynamic spectra, in addition to suppressing the cross-spectral estimate in frequency regions of low signal-to-noise ratio, high signal-to-noise ratio regions are also suppressed in attempt to reject strong tonals in the observations [19].

The Wiener processor was proposed in 1985 by Hero and Schwartz. Based on channel's linearity it tries to estimate the original signal from the observation $y_1(t)$ and channel output signal from $y_2(t)$, by minimizing the mean-square errors. In this way, given the channel characteristics, the solution results in Wiener filters, which yield the Wiener weighting function [20].

In 1996 it was presented a new weighting function, for acoustic localization, by Rabinkin et al., the ρ-Cross-power Spectrum Phase (ρ-CSP). It adds to the normal CSP the tuning parameter ρ (with values between 0 and 1) as a whitening parameter, which discards the non-speech portion (below 200Hz) of the CSP [21].

Relatively recently, in addition to the above work, in 2009 was proposed ρ-Cross-power Spectrum Phase with Coherence (ρ-CSPC), by Shean and Liu. The presence of the minimum of the coherence function in the weighting function helps to reduce errors for relatively small energy signals [22].

For the above presented GCC methods, an implementation block diagram is presented in Fig. 1. First, the analysis frames of input signals are converted into frequency domain using the Fast Fourier Transform (FFT) block. Then, the cross-power spectrum is computed by multiplications of resulted spectra and weighting function. Going further, the generalized cross-correlation is obtained through an Inverse Fast Fourier Transform (IFFT). The final step consists in finding the argument, which maximizes GCC and estimating the delay. This is the basic way to obtain an estimation of delay.

For a large window with L samples, FFT's complexity order is $O(L \cdot logL)$, with L a power of 2. Because these consume important processing time, it is natural to search for solutions, which increase the computing speed. A way to achieve this is to divide the larger analysis window into smaller frames, as it is shown in Fig. 2. Thus, the larger analysis window of L samples is divided in K smaller frames, of n samples each. If the length of the frame l is also a power of 2, then the new complexity order is $O(K \cdot l \cdot logl) = O(L \cdot log(L/K))$, which needs a smaller processing time. For each smaller analysis frame, the partial estimated delay is obtained similarly as in Fig. 1. Then, the final estimated delay yields as the average of all partial estimated delays. In this way, it is also easier to estimate a variable delay. This approach is recommended especially when the estimated delay is expected to be considerable less than the length of the larger window.

An alternative way for the above multi-frame approach is accumulated Cross-power Spectrum Phase (acc-CSP), proposed in 2006, by Matassoni and Svaizer [23]. It accumulates the cross-power spectrum over multiple frames in frequency domain, as showed in Fig. 3. This scheme leads to a new computing time decrease, because the number of IFFT and peak detector is reduced to 1. In frequency domain it can be expressed as follows:

$$G_{acc-CSP}(f) = \sum_{k=1}^{K} \frac{G_{y_1 y_2,k}(f)}{\left|G_{y_1 y_2,k}(f)\right|}, \qquad (9)$$

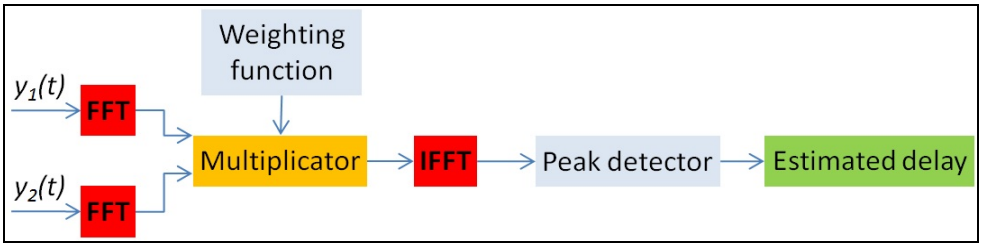


Figure 1. Block diagram for a single frame GCC implementation

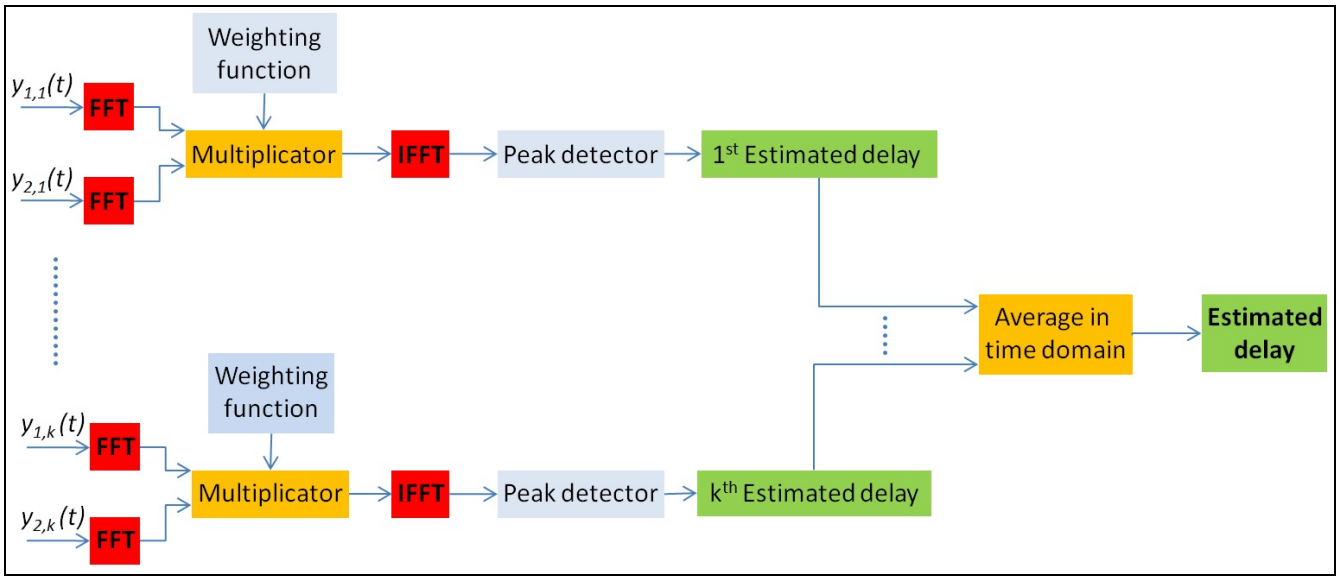


Figure 2. Block diagram for multiple frames GCC with time domain average estimation

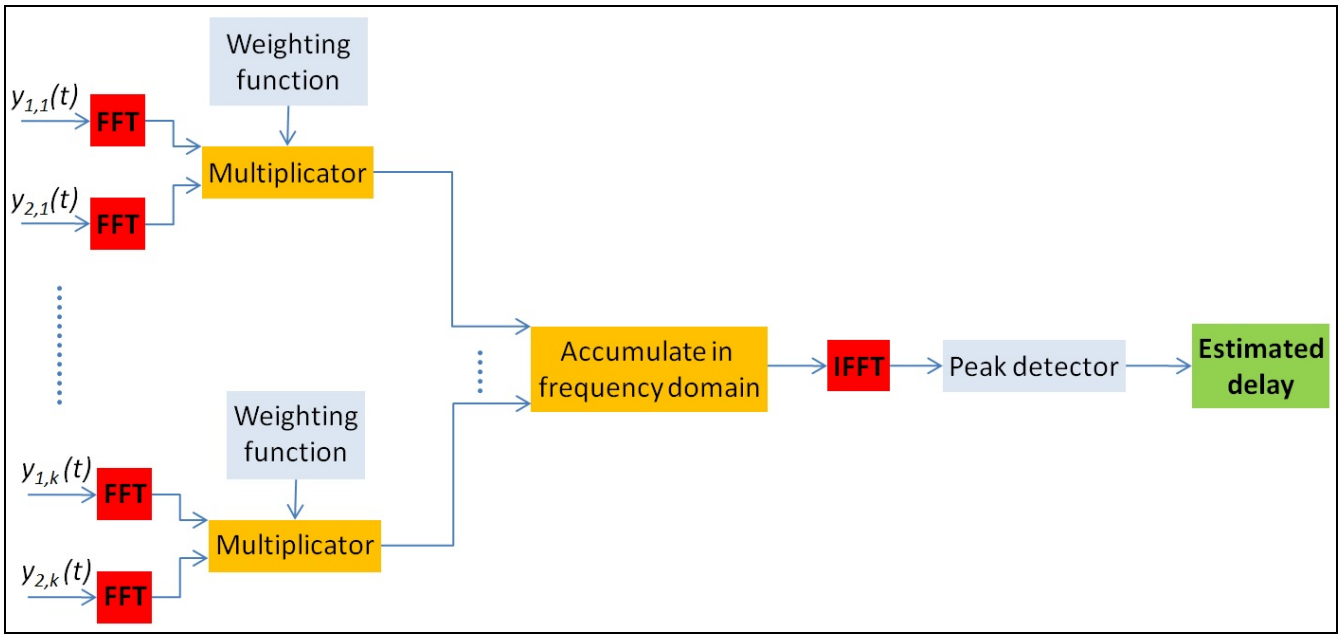


Figure 3. Block diagram for accumulating multiple frames GCC in frequency domain

where K represents the number of accumulated frames. Beside the reduced computational complexity, the *acc-CSP* method enhances the estimation by intrinsic integration for fixed delay during the analysis window [23].

The acc-CSP method proposes the accumulation scheme of cross power spectrum in frequency domain, increasing the computation speed. Methods based on the approach presented in Fig. 2 computes the TDE as the average of all partial estimated delays of each frame from the analysis window. In this way, for K frames, the number of total FFT operations is equal to $3xK$, because two FFT are used to transform the signals from time to frequency domain, and then one IFFT is used on the cross power spectrum to return in the time domain, for each frame. Instead, the accumulation scheme from Fig. 3 is faster because it does not calculate any partial TDEs. Because the cross-power spectrum averaging is computed in frequency domain, only one estimate will result, for any number of K frames. Thus, only one IFFT is needed for the final estimation and $2xK$ FFTs for time to frequency transformations. This leads to a total number of $2xK + 1$ FFT for the accumulating scheme, which is less than the $3xK$ FFT needed by previous methods [3]. Also, a small increase to the computation speed is due to the reduction to only one peak detector call.

Based on ρ-CSPC and ρ-CSP, in combination with the accumulated cross-power spectrum scheme, we recently proposed two new methods in [2] the new *accumulated ρ-Cross Power Spectrum Phase with Coherence* (*acc-ρCSPC*) and *accumulated ρ-Cross Power Spectrum Phase* (*acc-ρCSP*). In frequency domain they are expressed as

$$G_{acc-\rho CSPC}(f)=\sum_{k=1}^{K}\frac{G_{y_1y_2,k}(f)}{\left|G_{y_1y_2,k}(f)\right|^{\rho}+\min\left[\gamma^2_{y_1y_2,k}(f)\right]}. \quad (10)$$

and

$$G_{acc-\rho CSP}(f)=\sum_{k=1}^{K}\frac{G_{y_1y_2,k}(f)}{\left|G_{y_1y_2,k}(f)\right|^{\rho}} \quad (11)$$

In this way, for (10) it is possible to take advantage of both ideas (ρ-CSPC and accumulation scheme). Its effectiveness was proven by experimental results from [2], which showed a better accuracy even for low signal-to-noise ratios (SNR).

The new approach, summarized by (10), leads to faster computations compared to its previous methods, because it uses the accumulating scheme, presented in Fig. 3. It can also provide better results in unfavorable conditions for smaller frame sizes. Beside this, emphasis of speech regions from the spectrum is achieved by the whitening parameter (ρ), which reduces, at the same time, the impact of noise outside the speech region. For parts of the signal with small energy, the addition of the minimum coherence function limits the effect of a very small denominator [1][21][22].

The approach from (11) appeared as a faster variant of (10) for applications where relatively small energy signals are not encountered. In these conditions, the minimum coherence function can be omitted from (10). Thus, there is no need to compute the coherence function and to find its minimum, resulting a substantial computing time decrease.

As shown in [2], a high accuracy rate of TDE with *acc-ρCSPC* and *acc-ρCSP* is achieved if a calibration step is performed first. This procedure will be detailed and commented in the next section.

Over the years, several other studies discussed the details about time delay estimation based on generalized cross correlations, like in [24]-[31]. In this paper, we extend the analysis with the accumulating cross-power spectrum scheme, not only to our recently proposed methods, but also to the others well known GCC. We will apply the accumulation of cross-power spectrum technique (Fig. 3) to the traditional GCC functions from Table I, which are implemented as in Fig. 2 in all current applications.

III. Experimental Results and Discussions

In several previous papers [1]-[3] we proposed and evaluated acc-ρCSPC and acc-ρCSP. These methods derived from their primitive forms, ρCSPC [22] and ρCSP [21], at which we applied accumulation of cross-power spectrum in frequency domain [23]. To extend our research on this topic, in this work we apply accumulating scheme to all well known GCC functions. To the best of our knowledge, this technique was not presented in any other previous study.

A. Experimental Setup

Evaluation tests were performed in Matlab and C language. The input signals were taken from Noizeus data base corpus [32]. It contains 30 sentences (produced by three male and female speakers at a sampling rate of 8 kHz) corrupted by 8 different real-world noises (suburban train, babble, car, exhibition hall, restaurant, street, airport and train station noise), from the AURORA database [33] at 4 different SNRs (0, 5, 10 and 15 dB).

We used four metrics in our experiments: accuracy and error rate, standard deviation of the relative error and computing time. We define the accuracy rate as the ratio between the number of correctly estimated delays and total number of estimations performed (we imply that a correct estimation as one where the estimated delay is equal to the reference delay in terms of samples). Complementary to this we define the error rate as the ratio between the number of incorrectly estimated delays and total number of estimation performed.

For the first three metrics we used Matlab implementations. The forth metric, is the processing time, for which we evaluated the C implementations, compiled with gcc-4.7.3, on a machine with an Intel "Core i5" processor.

B. Calibration stage

It is easy to observe that for ρCSP and ρCSPC, the whitening factor ρ is not defined yet. In [22] it is used with values between 0.78 and 0.9. Also, ρ parameter requires particular attention because it characterizes our recently proposed methods. In our approach, we need to maximize the accuracy rate for accumulated cross-power spectrum versions, so we have to find the optimum value for ρ. Thus, we divided the Noizeus database in two parts, like in [2]. 50% of the sentences were used for *development* and the other 50% were used for the *evaluation.* The signal pairs chosen for alignment cover all the combinations of noise types: $C_2^8 = 28$. By using 4 different SNR levels and 5 artificially introduced delay values (5, 10, 25, 50 and 100 ms), the total number of test pairs becomes 28x15x4x5=8400.

In order to obtain efficient results, it is also important to set adequately the methods parameters (number of frames, frame size, overlap factor and ρ), which influence the accuracy and error rate of the *acc-ρCSPC* and *acc-ρCSP*. The first three parameters have to be chosen based on the nature of the application, making a trade-off between computing time, accuracy and fast response of the system. For this operation, here we used 4 averaging frames of 1024 samples each, and an overlap factor of 25%.

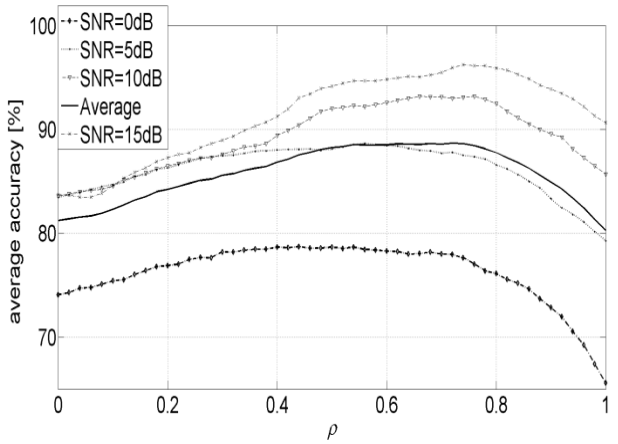


Figure 4. The influence of SNR and ρ over the *acc-ρCSPC* accuracy

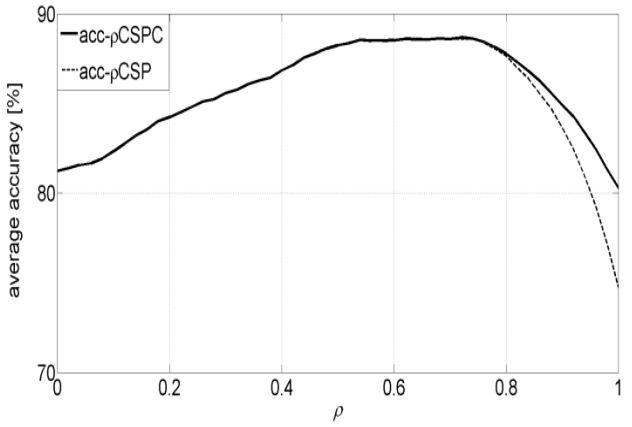


Figure 5. The dependence of *acc-ρCSP* and *acc-ρCSPC* on ρ

Fig. 4 confirms that there is an optimal value for ρ, which depends on SNR. For higher SNRs, the optimal ρ has a greater value. For the comparison with other methods we have chosen $\rho = 0.73$ as this is the value that maximizes the average accuracy in the SNR range 0-15 dB. If the development database is not available or limited, Fig. 4 can be used to choose the optimum ρ, for a general or a specific narrow SNR domain.

The accuracy effect of the omitted coherence term from formula (11) of the acc-ρCSP is visible in Fig. 5. This term makes the difference between acc-ρCSPC and acc-ρCSP. It can be observed for $\rho \in [0, 0.77]$ the accuracy of the two methods is equal. For $\rho > 0.77$ acc-ρCSPC outperforms acc-CSP. However, these are not usual values for ρ (which has the optimum value at 0.73). Hence, in order to improve the computation speed, the acc-ρCSP method can be chosen instead of acc-ρCSPC (i.e., omitting the coherence term). But, more precise computing time results will be shown in next part of this section.

C. Extended evaluations for generalized cross-correlation time delays estimation methods

After the calibration stage for *acc-ρCSPC* and *acc-ρCSP* we continue to evaluate the traditional generalized cross-correlation methods for time delay estimation. In this scenario we tested all GCC approaches, which were presented in Table I. We implemented all three block diagrams, shown in Fig. 1, 2 and 3. Then, we used the *evaluation* part of Noizeus database, for all four metrics. As a special notice, we used m=4 for CSP-m method.

The large analysis window was set to 2048 samples. This corresponds to a single large frame analysis for first scheme represented in Fig. 1, and K=4 smaller frames analysis (of 512 samples each) for the next two schemes represented in Fig. 2 and 3. As a naming convention, extending the accumulation cross-power spectrum technique to any method from Table I, the name of an approach is changed to *acc-* approach. In this way, our recently proposed methods were named *acc-ρCSPC* and *acc-ρCSP,* after applying the accumulation scheme to ρCSPC and ρCSP.

The length of the smaller frame size, of 512 samples, is 64ms on a sampling frequency of 8 kHz. Thus, we automatically vary the inserted delay from 5 to 50 ms, with a step of 5ms. In this way, the number of estimated perform for each delay-SNR configuration is $C_2^8 \cdot 15 = 420$ from a total of $C_2^8 \cdot 15 \cdot 4 \cdot 10 = 16800$ performed estimations. Next, the results were divided in two parts: one for the delays smaller than half of the frame size (from 0 to 30ms), and the other for delays larger than half of the frame size (from 35 to 50ms). Notice that this is meaningful only for schemes 2 and 3, because in scheme 1 we use a large frame of 2048 samples, equivalent with 256ms. In this case, all variable delays (from 5 to 50ms) are less than 128 ms, which correspond to half of the largest frame.

In the next four tables we present the error rate and standard deviation of the relative error for all 3 schemes in the above presented configuration. These metrics were computed for different SNR and delays combinations. In Table II we present results for low SNR and small delays, in Table III the results for low SNR and large delays, in Table IV the results for high SNR and small delays, and finally, in Table V the results for high SNR and large delays.

For scheme 2, in this configuration, it is important to notice that the error rate is not a relevant metric. This is due to the fact that for scheme 2 increasing the number of frames leads to a higher probability of wrong partial estimations. But, for this scheme the standard deviation of relative error is a confident metric.

Results from Table II-V confirm that if SNR increases the GCC methods perform better, yielding smaller error rate and standard deviation of relative error. Also, it is confirmed that a smaller delay has more chances to be estimated correctly than a larger one.

We notice also that not all the proposed methods perform better than the basic cross-correlation, which was already shown in [17][18][31]. This could be explained by the fact that the papers where some methods were proposed contain only mathematical presentation and no simulated results.

As expected, scheme 1 offers the smallest error rate, in all four combinations (low/high SNR and small/large delays). This is thanks to the larger analysis frame. On the other hand, scheme 1 is the slowest one. Scheme 3 is the second best scheme regarding error rate and for some weighting functions it has acceptable performances. It is faster than scheme 1 and represents a tradeoff between the computing time and the error rate. For this scheme, results confirm that *acc-ρCSPC* and *acc-ρCSP*, outperforms other methods regarding error rate results. Yet, a low error rate is observed for *acc-ρCSP*. *Acc*-CC, *acc*-CSP and *acc*-HB have reasonable error rate results. Also, the results pointed out that *acc-CPS-m* (with m=4) is an intermediate solution, between *acc-CC* and *acc-ρCSP*.

We expected that the higher the SNR, the lower the estimated delays are, better results have to be achieved. However, Tables II-V show some exceptions for standard deviation of relative error. At a first look, the results of this metric seem incomprehensible for scheme 2, because they are decreasing at larger estimated delays. To answer to this remark, we have to corroborate them with the error rate.

TABLE II. GCC METHODS EVALUATION, WITH LOW SNR (0 DB), AND SMALL DELAYS (5..30MS)

GCC name	Error rate (%)			Standard Deviation of Relative Error		
	Scheme 1	Scheme 2	Scheme 3	Scheme 1	Scheme 2	Scheme 3
CC	0.53	94.52	8.42	5.13	35.69	**13.4**
Eckart	40.46	99.72	82.71	564.68	109.59	196.95
ROTH	34.99	98.97	65.05	758.9	117.82	220.88
SCOT	51.75	99.87	64.65	201.27	78.29	78.94
CSP	4.46	95.74	19.54	228.85	99.69	94.07
CSP-m	0.48	95.29	6.93	**1.29**	33.6	17.9
HT (ML)	57.82	99.81	92.81	786.93	122.57	272.72
HB	4.46	95.74	19.54	230.8	102.75	101.65
Wiener	31.8	99.01	55.68	116.64	44.2	66.52
ρCSP	**0.44**	93.27	**4.69**	6.35	34.71	22.09
ρCSPC	0.56	93.94	5.53	3.48	**33.53**	20.37

TABLE III. GCC METHODS EVALUATION, WITH LOW SNR (0DB), AND LARGE DELAYS (35..50MS)

GCC name	Error rate (%)			Standard Deviation of Relative Error		
	Scheme 1	Scheme 2	Scheme 3	Scheme 1	Scheme 2	Scheme 3
CC	3.77	99.91	47.22	7.94	**21.54**	**26.47**
Eckart	50.6	100	97.64	223.31	40.5	79.11
ROTH	45.14	100	93.44	272.55	39.37	83.94
SCOT	55.6	100	83.16	113.09	36.73	70.90
CSP	8.07	99.96	62.42	106.65	44.31	78.89
CSP-m	1.81	99.94	31.55	**5.23**	22.34	29.99
HT (ML)	66.46	100	99.15	261.3	40.44	86.63
HB	8.07	99.96	62.42	111.73	44.84	83.10
Wiener	45.32	100	86.99	45.17	21.64	42.92
ρCSP	**1.25**	99.52	**27.52**	7.12	22.8	39.05
ρCSPC	1.35	99.57	30.87	10.31	22.91	41.78

TABLE IV. GCC METHODS EVALUATION, WITH HIGH SNR (15DB), AND SMALL DELAYS (5..30MS)

GCC name	Error rate (%)			Standard Deviation of Relative Error		
	Scheme 1	Scheme 2	Scheme 3	Scheme 1	Scheme 2	Scheme 3
CC	0	86.93	4.79	0	**27.41**	7.37
Eckart	13.63	99.69	66.95	355.88	89.89	161.42
ROTH	15.83	95.84	29.69	461.29	93.21	135.05
SCOT	49.6	99.75	56.68	0.51	55.79	39.92
CSP	0	88.82	5.86	0	61.69	38.25
CSP-m	0	84.98	1.72	0	29.89	5.76
HT (ML)	35.24	98.88	80.36	626.69	112.74	246.01
HB	0	88.82	5.86	0	67.21	45.55
Wiener	7.44	97.83	46.21	20.27	31.99	35.49
ρCSP	0	81.88	0.6	0	32.55	**3.86**
ρCSPC	0	84.13	0.6	0	35.11	4.37

TABLE V. GCC METHODS EVALUATION, WITH HIGH SNR (15DB), AND LARGE DELAYS (35..50MS)

GCC name	Error rate (%)			Standard Deviation of Relative Error		
	Scheme 1	Scheme 2	Scheme 3	Scheme 1	Scheme 2	Scheme 3
CC	0.08	100	39.62	0.51	**22.21**	17.93
Eckart	16.56	100	92.19	152.04	44.35	88.87
ROTH	34.79	100	77.77	233.78	40.71	82.75
SCOT	52.53	100	65.52	2.97	38.41	75.74
CSP	1.58	99.3	28.83	8.66	43.58	68.82
CSP-m	0	100	12.8	0	22.73	**8.33**
HT (ML)	41.83	100	96.21	213.36	42.89	86.43
HB	1.58	99.3	28.83	8.66	47.53	75.62
Wiener	28.89	100	85.85	14.04	20.58	37.10
ρCSP	0	99.4	**4.29**	0	26.99	18.18
ρCSPC	0	99.43	4.83	0	31.17	20.05

Notice that, for larger delays the error rate is 100% for almost all methods, because of incorrect estimations. In this way, it is clear that the decrease of standard deviation of relative error is due to the fact that almost all delays were much frequently estimated incorrectly, with a smaller variation.

Another important remark regards the standard deviation of relative error. In spite of a smaller error rate for GCC implemented with scheme 1, the standard deviation of the relative error is higher when comparing with implementation schemes 2 and 3. This is because scheme 1 uses a four times larger frame size, and any incorrect estimated delay varies in a larger domain. Thus, the variations of relative error are larger, leading to higher standard deviation values.

The forth evaluation metric is the processing time. Table VI provides detailed data for presented GCC methods, in all three implementation schemes. As we expected, the results confirm that the implementation scheme 1, which uses a large analysis frame, is slower than those which divided the large analysis frame in several smaller frames, like scheme 2 and 3. Moreover, the computing time for the accumulating methods is reduced even more, thanks to the benefit offered by scheme 3 (which reduces the total FFT number, form $3K$ to $2K+1$, as we presented in Section II).

TABLE VI. COMPUTING TIME EVALUATION

GCC name	Computing time (μs)		
	Scheme 1	Scheme 2	Scheme 3
CC	92	78	61
Eckart	227	237	220
ROTH	114	100	84
SCOT	147	134	117
CSP	139	127	111
CSP-m	323	307	288
HT (ML)	268	226	209
HB	175	163	146
Wiener	199	188	171
ρCSP	296	285	268
ρCSPC	406	399	382

In all schemes, the normal cross-correlation has the fastest processing time. This is because this is the basic GCC form, which does not compute any weighting. On the

opposite side, we find ρCSPC as the slowest method. Besides the ordinary FFTs, it has to spend expensive time in computing its weighting function.

Based on the presented results we conclude that, after the calibration step, ρCSP and ρCSPC provide the lowest error rate. Considering the demand of a real time application, between them, ρCSP is an obvious solution because of its lower computing time. For this kind of applications we have to take into account CC also. In any implementation scheme, CC is much faster than ρCSP, yielding acceptable error rate.

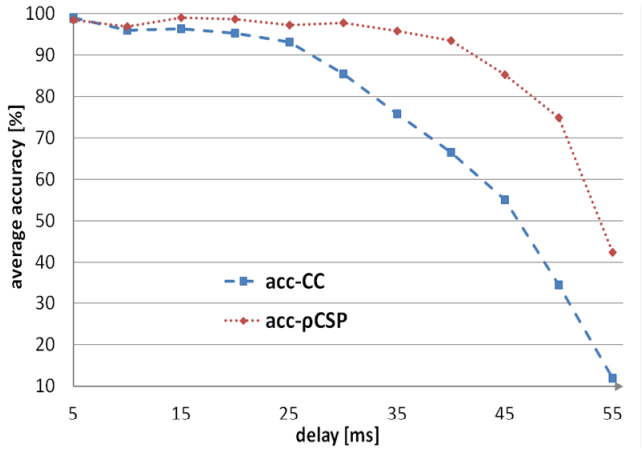


Figure 6. Comparison between acc-CC and acc-ρCSP

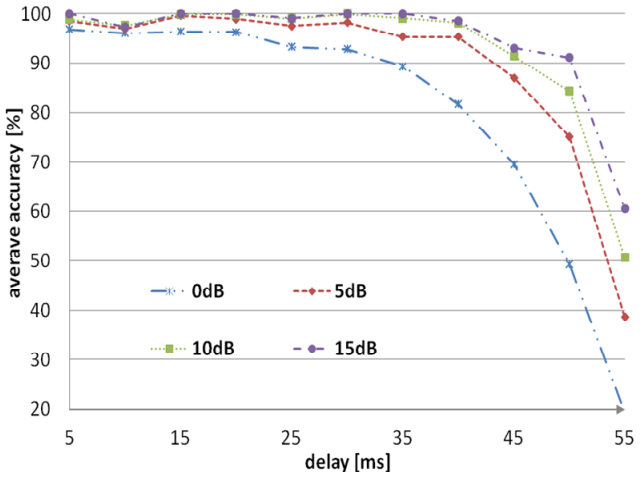


Figure 7. The influence of SNR and delay over the *acc-ρCSP* accuracy

For these reasons, we investigated the dependence of accuracy rate on the estimated delay, for *acc-ρCSP* and *acc*-CC. In Fig. 6 we used $K = 4$ frames of 512 samples each, computing the average accuracy from 0, 5, 10 and 15 dB SNR. For 50 ms delays, which represent 78% of the frame size length (64 ms), the accuracy rate for *acc-ρCSP* is around 75%, while for *acc*-CC is less than 35%. This confirms that while most of GCC methods are able to estimate accurate delays smaller than half of the frame size, *acc-ρCSP* outperform them and continues to provide reasonable accuracy for larger delays.

In Fig. 7, for *acc-ρCSP,* we present the variation of the accuracy rate by the SNR and delay. It is shown that the higher the SNR, the higher the accuracy is. For delays up to 50% of the frame size, the difference between accuracies on various levels of SNR remains almost the same. Once the delay increases over 50% of the frame size, the accuracy decreases much faster for lower SNR.

The influence of the number of frames over the GCC error rate, for scheme 3, is highlighted in Table VII. The GCC accuracies were computed for 1, 4 and 8 frames, with frame sizes of 512 samples, at 15 dB SNR. It is shown that the error rate decreases when it is used a higher number of frames. This is due to the fact that the accumulated cross power spectrum domain keeps the spectral information over multiple frames. In this way, the correlation between the frames is maintained.

TABLE VII. ERROR RATE DEPENDENCE ON THE NUMBER OF FRAMES

	Number of frames		
GCC names	**1**	**4**	**8**
CC	37.17	4.79	2.21
Eckart	63.67	66.95	59.69
ROTH	53.8	29.69	21.22
SCOT	64.92	56.68	45.96
CSP	26.35	5.86	1.66
CSP-m	24.94	1.72	0.05
HT (ML)	76.97	80.36	80.03
HB	26.35	5.86	1.66
Wiener	68.42	46.22	26.56
ρCSP	18.2	0.6	0.56
ρCSPC	19.3	0.6	0.56

The numerical and simulated results confirm that our recently proposed methods *acc-ρCSP* and *acc-ρCSPC* in [2] achieve the highest accuracy rate for multi frame processing.

IV. CONCLUSION AND FUTURE WORK

In this paper, we evaluated traditional generalized cross-correlation time delay estimation methods, applying them the accumulated cross-power spectrum technique. The experiments were performed using the standard Noizeus database. The obtained results showed that, a single large frame yields the smallest error rate when comparing with different multi frame implementation. On the other hand, the *accumulation scheme* over smaller multiple frames is faster than the above approach, providing acceptable error rates for a part of GCC methods. For this scheme also, the increasing number of frames leads to a smaller error rate.

Our analysis showed that ρCSP and ρCSPC provide the lowest error rate, with a small benefit for the first one. CPS-m, CC, CSP and HB have reasonable error rate results. In the same conditions, the others methods like Eckart, ROTH, SCOT, HT (ML) and Wiener do not offer acceptable error rate results.

Regarding the processing time, the normal cross correlation is the faster method because it does not compute any weighting. On the opposite side, ρCSPC has to perform many time consuming operations to calculate weighting function, so it is the slowest one. Between the three presented schemes, the first, which analyzes the signals using one large frame is the slowest, but offers the highest accuracy. The second scheme, which works on multiple frames of fewer samples averaging the final estimate in time domain, is a little faster, but does not provide any usable accuracy results. The third scheme works on smaller frames, like scheme 2, but it accumulates the cross-power spectrum in frequency domain. This leads to a good accuracy and is also the fastest scheme.

The results from this work could be used for a better decision regarding the implementation of GCC methods, based on applications' demands. For expected delays that are comparable with the available analysis window, it is recommended to use a single large frame implementation. But, if the expected delays are much smaller than the available analysis window a faster solution is represented by the accumulation scheme of the cross-power spectrum in frequency domain. Each of these schemes can be efficiently implemented to provide solution for realigning noisy signals in applications such as speech enhancement, echo canceling, seismic and medical processing, radar and sonar localization, and pattern detection.

Future work will involve analysis *for acc-CSP-m,* from which we expect better accuracy results after a proper calibration for *m*. We will continue to focus on these methods and their applications in the VoIP environment and multi-channel speech enhancement.

REFERENCES

[1] R.S. Marinescu, A. Buzo, H. Cucu, and C. Burileanu, "New Considerations for Accumulated ρ-Cross Power Spectrum Phase with Coherence Time Delay Estimation", In Procedings of ICDT 2013, The Eight International Conference of Digital Telecommunications, Apr 2013, pp. 55-59.

[2] R.S. Marinescu, A. Buzo, H. Cucu, and C. Burileanu, "Fast Accurate Time Delay Estimation Based on Enhanced Accumulated Cross-Power Spectrum Phase", European Signal Processing Conference (EUSIPCO 2013), Morocco, 2013, "in press"

[3] R.S. Marinescu, A. Buzo, H. Cucu, and C. Burileanu, "Extensive Evaluation Experiments for the Recently Proposed Accumulated Cross-Power Spectrum Methods for Time Delay Estimation", SpeD 2013, "in press"

[4] B. Widrow and S.D. Strearns, "Adaptive signal processing", Penitence-Hall, ISBN 0130040290, USA, 1985.

[5] S.N. Lin and S.J. Chern, "A new adaptive constrained LMS time delay estimation algorithm", Signal Processing, Volume 71, Issue 1, Nov. 1998, pp. 29-44.

[6] A.W.H. Khong and P.A. Naylor, "Efficient Use Of Sparse Adaptive Filters", In Proceedings of ACSSC '06, Fortieth Asilomar Conference on Signals, Systems and Computers, article ID 10.1109/ACSSC.2006.354982, Nov. 2006, pp. 1375-1379.

[7] R.A. Dyba, "Parallel Structures for Fast Estimation of Echo Path Pure Delay and Their Applications to Sparse Echo Cancellers", In Proceedings of CISS 2008, 42nd Annual Conference on Information Sciences and Systems, article ID 10.1109/CISS.2008.4558529, Mar. 2008, pp. 241-245.

[8] D. Hongyang and R.A. Dyba, "Efficient Partial Update Algorithm Based on Coefficient Block for Sparse Impulse Response Identification", In Proceedings of CISS 2008, 42nd Annual Conference on Information Sciences and Systems, article ID 10.1109/CISS.2008.4558527, Mar. 2008, pp. 233-236.

[9] D. Hongyang and R.A. Dyba, "Partial Update PNLMS Algorithm for Network Echo Cancellation", In Proceedings of ICASSP 2009, IEEE International Conference on Acoustics, Speech and Signal Processing, article ID 10.1109/ICASSP.2009.4959837, Apr. 2009, pp. 1329-1332.

[10] K. Sakhnov, E. Verteletskaya, and B. Simak, "Partial Update Algorithms and Echo Delay Estimation," Communications – Scientific Journal of the University of Zilina, Zilina – Slovakia, vol. 13, no. 2, Apr. 2011, pp. 14-19.

[11] J. Benesty, "Adaptive eigenvalue decomposition algorithm for passive acoustic source localization", J. Acoust. Soc. Am. Volume 107, Issue 1, Jan. 2000, pp. 384-391.

[12] C. Knapp and G.C. Carter, "The Generalized Correlation Method for Estimation of Time Delay", IEEE Transactions on Acoustics, Speech and Signal Processing, vol. 24, issue 4, Aug. 1976, pp. 320-327.

[13] C. Eckart, "Optimal Rectifier System for Detection of Steady Signals", Univ. California, Scripps Inst. Oceanography, Marine Physical Lab. Rep SIO 12692, SIO Ref 52-11, 1952.

[14] P.R. Roth, "Effective measurements using digital signal analysis", IEEE Spectrum, vol 8, Apr 1971, pp. 62-70.

[15] E.J. Hannan and P.J. Thomson, "The Estimation of Coherence and Group Delay", Biometrika, vol. 58, Dec 1971, pp. 469-481.

[16] G.C. Carter, A.H. Nutall, and P.G. Cable, "The Smoothed Coherence Transform", Proc. IEEE (Lett.), vol. 61, Oct 1973, pp 1497-1498.

[17] J.C. Hassab, R.E. Boucher, "A Cuantitative Study of Optimum and Suboptimum filters in Generalized Correlator", IEEE 1979 Int. Conf. Acoust., Speech, Signal Processing Conf. Rec. 79CH 1379-7 ASSP, 1980, pp. 124-127.

[18] J.C. Hassab, R.E. Boucher, "Performance of the Generalized Cross Correlator in the Presence of a Strong Spectral Peak in the Signal", IEEE Transaction on Acoustics, Speech, and Signal Processing, vol. assp-29, no. 3, Jun. 1981, pp. 549-555.

[19] J.C. Hassab, R.E. Boucher, "Optimum Estimation of Time Delay by Generalized Correlator", IEEE Transaction on Acoustics, Speech, and Signal Processing, vol. assp-27, no. 3, Aug. 1979, pp. 373-380.

[20] A.O. Hero and S.C. Schwartz, "A New Generalized Cross Correlator", IEEE Transactions, Acoustics, Speech and Signal Processing, vol. 33, Feb 1985, pp. 38-45.

[21] D.V. Rabinkin, R.J. Renomeron, A. Dahl, J.C. French, J.L. Flanagan, and M.H. Bianchi, "A DSP Implementation of Source Location Using Microphone Arrays", The Journal of the Acoustical Society of America, Volume 99, Issue 4, Apr 1996, pp. 2510-2527.

[22] M. Shean and H. Liu, "A Modified Cross Power-Spectrum Phase Method Based on Microphone Array for Acoustic Source Localization," IEEE International Conference on System, Man and Cybernetics, San Antonio, TX, USA, Oct. 2009, pp. 1286 – 1291.

[23] M. Matassoni and P. Svaizer, "Efficient Time Delay Estimation Based on Cross-Power Spectrum Phase", European Signal Processing Conference (EUSIPCO), Florence - Italy, Sep 2006.

[24] D.H. Youn, N. Ahmed, and G.C. Carter, "On the Roth and SCOTH Algorithms: Time-Domain Implementations", In Proceedings of the IEEE, vol. 71, issue 4, 1983, pp. 536-538.

[25] Q. Tianshuang and W. Hongyu, "An Eckart-weighted adaptive time delay estimation method", IEEE Transactions on Signal Processing, vol. 44, issue 9, Sep. 1996, pp. 2332-2335.

[26] M. Omologo and P. Svaizer, "Acoustic event localization using a crosspower-spectrum phase based technique", Proceedings of ICASSP, Australia, Apr. 1994, pp. 273-276.

[27] M. Omologo and P.Svaizer, "Use of the crosspower-spectrum phase in acoustic event location", IEEE Transactions on Speech Audio Process, May 1997, pp. 288-292.

[28] V. Zetterberg, M.I. Pettersson, and I. Claesson, "Comparison Between Whitened Generalized Crosscorrelation and Adaptive Filter for Time Delay Estimation", In Proceedings of TS/IEEE, OCEANS, vol. 3, article ID 10.1109/OCEANS.2005.1640117, Sept. 2005, pp. 2356 – 2361.

[29] K.W. Wilson and T. Darrell, "Learning a Precedence Effect-Like Weighting Function for the Generalized Cross-Correlation Framework", IEEE Transactions on Audio, Speech, and Language Processing, vol. 14, issue 6, Nov. 2006, pp. 2156-2164.

[30] Y. Sun and T. Qiu, "The SCOT Weighted Adaptive Time Delay Estimation Algorithm Based on Minimum Dispersion Criterion", In Proceedings of the ICICIP Conference on Intelligent Control and Information Processing, Aug. 2010, pp. 35-38.

[31] K. Sakhnov, E. Verteletskaya, and B. Simak, "Echo Delay Estimation Using Algorithms Based on Cross-correlation," Journal of Convergence Information Technology, Volume 6, Number 4, Apr. 2011, pp. 1 – 11.

[32] NOIZEUS: A noisy speech corpus http://www.utdallas.edu/~loizou/speech/noizeus/ [retrieved: Aug, 2013]

[33] AURORA database, http://www.elda.org/article52.html [retrieved: Aug, 2013]

Spectrum Sensing in the TV White Spaces

Daniel Riviello, Roberto Garello
Dipartimento di Elettronica e Telecomunicazioni
Politecnico di Torino
Torino, Italy
Email: {daniel.riviello, roberto.garello}@polito.it

Sergio Benco, Floriana Crespi, Alberto Perotti
Networks and Wireless Communications
CSP-ICT Innovation
Torino, Italy
Email: {sergio.benco, floriana.crespi, alberto.perotti}@csp.it

***Abstract*—Cognitive radio networks operating in the digital television white spaces are of particular interest for their practical applications. In this paper, we review several parametric and non-parametric test statistics commonly used in spectrum sensing. Both single-antenna as well as multiple antenna techniques are considered. For a selected subset of these techniques, an accurate performance assessment is carried out in the presence of a DVB-T primary signal generated using a software-defined real-time transmitter. Sensing performance is assessed both through Monte Carlo simulation and using a real *software-defined radio* implementation. Different channel profiles are considered. The obtained results show the performance of each algorithm in terms of detection probability under fixed false alarm rate and of *Receiver Operating Curve* (ROC). Moreover, these results permit to establish clear relationships between the considered algorithms in case of DVB-T primary signals.**

Keywords-Cognitive radio networks; spectrum sensing; white spaces; software-defined radio; DVB-T; eigenbased detectors.

I. Introduction

The following paper is an extended version of a previous work that the same authors presented at COCORA 2013 [1]. Here, a wider analysis is performed and a broader set of results is shown.

The ever increasing demand for higher data rates in wireless communications is strongly driving the search for new communication technologies able to exploit transmission opportunities wherever available. Unused channels in licensed radio-frequency (RF) bands, like those assigned to TV broadcasters, are one of the most attractive opportunities, since these enjoy favorable propagation conditions and feature fairly large bandwidths. As a matter of fact, some of the most relevant developments in this context are aimed at exploiting the so-called *TV white spaces* to provide broadband Internet access or other kind of wireless services through high-speed wireless communications carried out in the unused TV channels.

The cognitive radio concept [2] has been readily applied to TV white spaces networks and today white-space cognitive systems are one of the most active research fields in cognitive radio networks.

Some of the fundamental features and requirements of white-space cognitive systems include their ability to avoid interference to those who hold the right to use a certain band (the Primary Users, PU), while simultaneously performing high-speed communications. Such systems must as well be able to harmoniously coexist with other cognitive white-space systems (also called Secondary Users, SU) who operate on an non-interference basis with respect to PUs and may be operating in the same area. To these purposes, cognitive radio networks and the systems of which they are composed need to gain awareness of the surrounding electromagnetic environment. This is typically performed by using dedicated subsystems or relying on data provided by external entities.

Remarkable techniques used to gain said awareness include the access to geographically-referenced data bases containing spectral occupation information, *spectrum sensing* techniques, or a combination thereof. In this paper, we are mainly concerned with the investigation, analysis and performance evaluation of spectrum sensing techniques for white-space cognitive radio systems.

According to the cognitive radio paradigm, each node in a CR network must be equipped with an efficient *spectrum sensor* [3] a unit that makes the node able to gain awareness of the available transmission opportunities through the observation of the surrounding electromagnetic environment in a given range of frequencies. The sensor's ultimate goal consists in providing an indication to the node regarding whether a primary transmission is taking place in the observed channel. Such indication is determined as the result of a binary hypothesis testing experiment wherein hypothesis $\mathcal{H}_0$ ($\mathcal{H}_1$) corresponds to the absence (presence) of the primary signal. Thus, the sensing unit collects $1 \leq n \leq N$ samples of kind

$$y(n)|_{\mathcal{H}_0} = w(n) \quad (1)$$

$$y(n)|_{\mathcal{H}_1} = x(n) + w(n), \quad (2)$$

where $x(n)$ are samples of the primary signal as it is received by the spectrum sensor and $w(n)$ are noise samples.

Given the vector $y = (y(1), ..., y(n), ..., y(N)$ of acquired samples, the sensing algorithm computes a test statistics $T(y)$ and compares it against a predefined threshold θ. The performance of each detector is usually assessed in terms of *probability of detection* P_d and *probability of false alarm* P_{fa}

$$P_d = \mathbb{P}(T(y) > \theta|\mathcal{H}_1) \quad (3)$$

$$P_{fa} = \mathbb{P}(T(y) > \theta|\mathcal{H}_0), \quad (4)$$

as a function of the signal-to-noise ratio (SNR) ρ, which is defined as

$$\rho = \frac{\mathbb{E}\|x(n)\|^2}{\mathbb{E}\|w(n)\|^2}. \quad (5)$$

Several methods have been proposed for the computation of the test statistics: a comprehensive description can be found in [4] and references therein. In this paper, we consider a subset that includes most of the widely used algorithms: energy detection, multi-antenna eigenvalue based techniques, either under known or unknown noise variance hypothesis, and techniques exploiting specific signal characteristics, like the lagged-autocorrelation technique used to detect the cyclic prefix of OFDM signals [5][6][7].

Non-parametric methods are generally applicable to any kind of primary signal, since these methods do not make any assumption on the characteristics of such signal. Parametric methods, instead, assume a partial knowledge of some of the primary signal characteristics. The parametric method herein considered (lagged autocorrelation) has been applied to OFDM signals with cyclic prefix, but it can be straightforwardly applied to any other kind of cyclostationary signal, i.e., to almost all digitally modulated signals.

The main contributions of this paper are (i) an introductory review of the most relevant spectrum sensing algorithms, (ii) the application of such algorithms to real DVB-T signals generated using a software-defined real-time transmitter [8], (iii) the application of different realistic channel profiles and (iv) the complexity assessment performed through the software-radio implementation of the considered sensing algorithms. The algorithms' performance and complexity are evaluated and compared in realistic conditions, providing useful results for practical realizations.

II. Primary signal

The DVB-T standard [9] specifies a set of coded OFDM transmission schemes to be used for broadcasting of multiplexed digital television programs.

The transmitted signal consists of a sequence of fixed-duration OFDM symbols. A cyclic prefix (CP) is prepended to each symbol in order to avoid inter-symbol interference over frequency-selective fading channels. The most relevant parameters of DVB-T signals are shown in Table I.

The signal bandwidth is approximately 7.61 MHz, with an intercarrier frequency spacing of 8MHz. A subset of the available 2048 subcarriers (in 2k mode) or 8192 subcarriers (in 8k mode) are used to carry higher layer data and PHY-layer signalling information. The latter consists of pilot sequences, either allocated to fixed subcarriers (continual pilots) or scattered throughout OFDM symbols according to a periodic pattern, which are used for channel estimation at the receiver side, and Transmission Parameter Signaling (TPS) information, wherein encoded information about the current transmission parameters used on data subcarriers is delivered.

OFDM symbols with CP are grouped into frames and superframes: each frame consists of 68 symbols and each superframe consists of 4 frames.

In our study, we used a real encoded and modulated MPEG transport stream (TS) with code rate 5/6, 64-QAM constellation and CP ratio 1/4. The resulting bit rate is approx. 24.88 Mbits/s. At the sensing unit, the DVB-T signal was sampled at the nominal rate of 64/7 Msamples/s.

TABLE I
Main parameters of DVB-T.

	2k mode	8k mode
Symbol duration (T_U)	$224\mu s$	$896\mu s$
Guard interval duration (Δ)	$7-56\mu s$	$28-224\mu s$
Number of active subcarriers	1705	6817
Subcarrier spacing (approx.)	4464Hz	1116Hz
CP duration ratio (Δ/T_U)	1/4, 1/8, 1/16, 1/32	
Constellations	QPSK, 16-QAM, 64-QAM	
Code rate	1/2, 2/3, 3/4, 5/6, 7/8	

A. DVB-T signal characteristics

As a common assumption in the literature on spectrum sensing, the primary signal is modeled as a Gaussian process. Fig. 1 shows that, in the case of DVB-T signals, this assumption is well motivated. In fact, Fig. 1(a) shows the *pdf* of the real and imaginary parts of the DVB-T signal's complex envelope. Clearly, the Gaussian distribution is very well approximated. A more accurate evaluation is provided in Fig. 1(b), where the *quantile-quantile* plot of the DVB-T distribution vs. a zero-mean Gaussian distribution with same variance is shown.

B. Channel characteristics

Let us assume that the primary signal is detected through K sensors (receivers or antennas). Let $\boldsymbol{y}(n) = [y_1(n) \dots y_k(n) \dots y_K(n)]^T$ be the $K \times 1$ received vector at time n, where the element $y_k(n)$ is the n-th discrete baseband complex sample at receiver k.

Typically, a flat Rayleigh fading channel is considered in the literature. In such case, the received signal can be modeled as a linear mixture model of kind

$$\boldsymbol{y}(n) = \boldsymbol{h}x(n) + \boldsymbol{v}(n), \tag{6}$$

where $\boldsymbol{h}$ is the K-element channel vector of size $K \times 1$ whose elements $h_n \sim N_{\mathbb{C}}(0, \sigma_h^2)$ are mutually independent. Moreover we apply the following normalization:

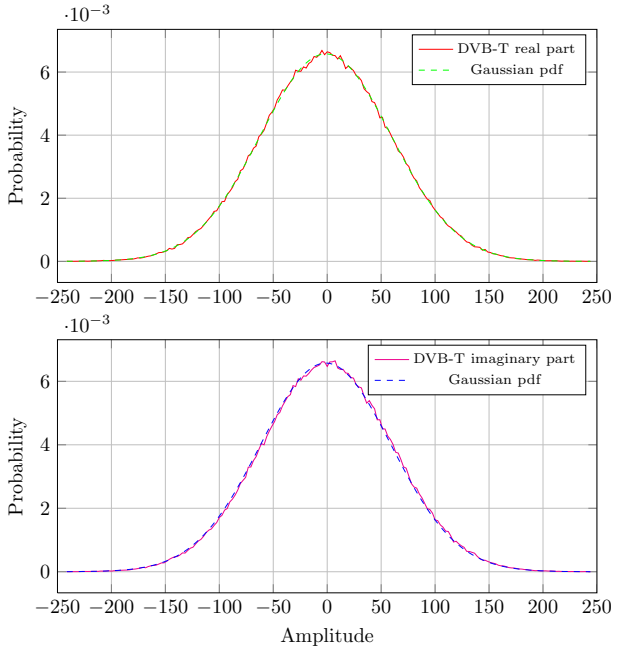
$$\sum_{n=1}^{K} h_n h_n^* = K. \tag{7}$$

Moreover, $\boldsymbol{v}(n)$ is the additive white Gaussian noise distributed as $N_{\mathbb{C}}(\mathbf{0}_{K\times 1}, \sigma_v^2 \mathbb{I}_{K\times K})$.

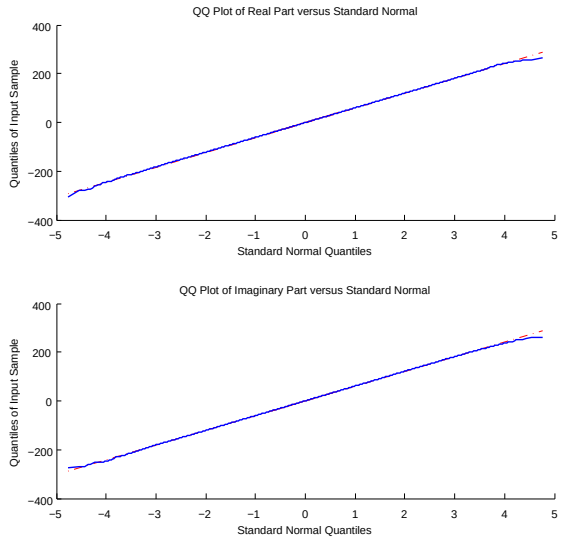
In order to assess the performance of the considered algorithms in a more realistic case, we used a frequency- and time-selective channel model, the 6-path Typical Urban (TU6) mobile radio propagation model developed by the COST 207 European project [10]. This profile is a frequency- and time-selective Rayleigh fading channel model. Given $x(t)$ and $y(t)$ the input and output signal respectively, it can be expressed as follows:

$$y(t) = \sum_{i=1}^{M} \gamma_i e^{-j\theta_i} x(t - \tau_i), \tag{8}$$

where:



(a) Estimated probability density function.



(b) Quantile-quantile plot.

Fig. 1. Statistics of the DVB-T signal.

- M is the number of paths equal to 6;
- γ_i is the average path gain of the ith path (listed in Tab. II-B);
- θ_i is the phase shift from scattering of the i'th path, modeled as a uniformly distributed random variable in $[-\pi, \pi]$;
- τ_i is the relative delay of the ith path (listed in Tab. II-B);

where the classical doppler spectrum is defined as:

$$G(f; f_D) = \frac{1}{\sqrt{1-(f/f_D)^2}}. \quad (9)$$

In our simulation the Doppler spread f_D has been set to 10 Hz, corresponding to a pedestrian mobile profile.

TABLE II
TYPICAL URBAN PROFILE (TU6).

Tap number	Delay τ_i (μs)	Average gain γ_i (dB)
1	0.0	-3
2	0.2	0
3	0.5	-2
4	1.6	-6
5	2.3	-8
6	5.0	-10

III. TEST STATISTICS

As explained in the previous section, the detector builds its test statistic from K sensors (receivers or antennas) and N time samples. The symbol $\boldsymbol{y}(n) = [y_1(n) \dots y_k(n) \dots y_K(n)]^T$ denotes the $K \times 1$ received vector at time n, where the element $y_k(n)$ is the n-th discrete baseband complex sample at receiver k.

The noise is modeled as an additive white Gaussian noise process with zero mean and variance $\sigma_v^2 = N_0/2$, N_0 being the two-sided power spectral density of noise.

The received samples are stored in a $K \times N$ matrix:

$$\boldsymbol{Y} \triangleq [\boldsymbol{y}(1) \dots \boldsymbol{y}(N)]. \quad (10)$$

The sample covariance matrix $\boldsymbol{R}$ is:

$$\boldsymbol{R} \triangleq \frac{1}{N} \boldsymbol{Y}\boldsymbol{Y}^H. \quad (11)$$

We will denote by $\lambda_1 \geq \dots \geq \lambda_K$ the eigenvalues of $\boldsymbol{R}$, sorted in decreasing order.

The usual criterion for comparing two tests is to fix the false alarm rate P_{fa} and look for the test achieving the higher P_d . The Neyman Pearson (NP) lemma [11] is known to provide the Uniformly Most Powerful (UMP) test, achieving the maximum possible P_d for any given value of P_{fa} . The NP criterion is applicable only when both $\mathcal{H}_0$ and $\mathcal{H}_1$ are simple hypotheses. In our setting this is the case when both the noise level σ_v^2 and the channel vector $\mathbf{h}$ are a priori known. The NP test is given by the following likelihood ratio:

$$T_{NP} = \frac{p_1(\mathbf{Y}; \mathbf{h}, \sigma_s^2, \sigma_v^2)}{p_0(\mathbf{Y}; \sigma_v^2)}. \quad (12)$$

For the considered model, the expressions of p_0 and p_1 can be found in [12]. The NP test provides the best possible performance, but requires exact knowledge of both h and σ^2. For most practical applications, the knowledge of h is questionable. The noise variance is somewhat easier to know: since we only consider thermal noise, if the temperature is constant some applications may possess an accurate estimation of it.

Many spectrum sensing algorithms have been proposed in the literature. Reviews and comparisons can be found, for example, in [4][13][14]. In this paper, we consider some of the most popular tests, with the aim of comparing them against true DVB-T signals. The considered tests are divided in three classes.

A. Non-parametric tests, known noise variance

These tests are non-parametric, i.e., do not exploit the knowledge of the signal characteristics. An excellent estimation of the noise variance σ_v^2 is supposed (obtained, for example, during a long training phase).

Roy's Largest Root Test (RLRT): this method tests the largest eigenvalue of the sample covariance matrix against the noise variance. The test statistic is

$$T_{RLRT} = \frac{\lambda_1}{\sigma_v^2}. \quad (13)$$

The RLRT was originally developed in [15]. Performance analysis can be found, for example, in [16][17][14].

Under $\mathcal{H}_1$, the asymptotic behavior of λ_1/σ_v^2 has a *phase transition phenomenon* [18], which depends on the SNR. In case of a single signal the critical value can be expressed as [19]:

$$\rho_{\text{crit}} = \frac{1}{\sqrt{KN}}. \quad (14)$$

If the SNR ρ is lower than the critical value, the limiting distribution of λ_1/σ_v^2 is the same as that of the largest noise eigenvalue, thus nullifying the statistical power of a largest eigenvalue test. If the SNR is higher, the largest eigenvalue λ_1 depends on the signal plus noise power, while the smallest $K1$ eigenvalues depends on the noise only. The optimality of RLRT in the class of "semi-blind" algorithms was pointed out in [17]. For a single emitting source, if the SNR is above the identifiability threshold given by (14), the signal is detectable by the largest eigenvalue λ_1 value. Starting from the NP test and using the asymptotic expansion of the hypergeometric function, it was shown in [17] that, under known noise variance, distinguishing between $\mathcal{H}_0$ and $\mathcal{H}_1$ in the asymptotic regime ($N \to \infty$ with K fixed) depends to leading order only on λ_1. For Gaussian signals and not too low signal-to-noise ratio, the RLRT is the best test statistics in this class.

Energy Detection (ED): the test statistic is the average energy of the received samples, normalized by the noise variance:

$$T_{ED} = \frac{1}{KN\sigma_v^2} \sum_{k=1}^{K} \sum_{n=1}^{N} |y_k(n)|^2 . \quad (15)$$

The energy detection method is probably the most popular technique for spectrum sensing, also thanks to its simplicity. Analytical performance expressions for this detector are well-known in the literature (e.g., [20]).

Likelihood Ratio Tests (LRT): different LRT-based detectors were given in [13]. The complete, noise-dependent, log-likelihood ratio test statistic is given by

$$T_{LRT} = 2(N-1)\left[\log\left(\frac{\sigma_v^{2K}}{\det \boldsymbol{R}}\right) + \left(\frac{\text{tr}\boldsymbol{R}}{\sigma_v^2} - K\right)\right]. \quad (16)$$

Performance analysis for this test can be found, for example, in [13].

B. Non-parametric tests, unknown noise variance

These tests are again non-parametric, but the noise variance is supposed unknown.

Generalized Likelihood Ratio Test (GLRT): this method uses as test statistic the ratio

$$T_{GLRT} = \frac{\lambda_1}{\frac{1}{K}\text{tr}(\boldsymbol{R})}. \quad (17)$$

Performance analysis can be found for example in [21].

The class of "blind" detection tests includes all algorithms where the noise variance σ_v^2 is unknown and not used in the test statistic. It has been recently proved in [22] that the GLRT procedure is optimal, asymptotically and even for finite sample size, in a combined Neyman-Pearson/Bayesian sense, such that it minimizes the average mean-square parameter estimation error subject to an upper bound constraint on the false-alarm probability. Therefore, the GLRT detector [12] is indeed the best performing method in the blind eigenvalue-based class.

By a simple transformation, it is possible to relate GLRT to RLRT. We first note that, since

$$\frac{1}{T_{GLRT}} = \frac{\sum_{i=1}^{K}\lambda_i}{\lambda_1} = 1 + \frac{\sum_{i=2}^{K}}{\lambda_1}, \quad (18)$$

the GLRT is equivalent to

$$T_{GLRT}^{'} = \frac{\lambda_1}{\frac{1}{K-1}\sum_{i=2}^{K}\lambda_i}. \quad (19)$$

In case of a single signal, the new denominator is the average of the noisy eigenvalues and can be interpreted as the Maximum-Likelihood (ML) estimate of the noise variance:

$$\hat{\sigma}_v^2 = \frac{1}{K-1}\sum_{i=2}^{K}\lambda_i. \quad (20)$$

As a consequence, we have obtained this equivalent formulation of the GLRT:

$$T_{GLRT}^{'} = \frac{\lambda_1}{\hat{\sigma}_v^2}. \quad (21)$$

We can observe that the structure of RLRT (13) and GLRT (17) is very similar: RLRT uses the true noise variance, while GLRT uses its ML estimation computed within the sensing slot by using the noise eigenvalues.

Eigenvalue Ratio Detector (ERD): the test statistic (also called maximum-minimum eigenvalue, or condition number test) is the ratio between the largest and the smallest eigenvalue of $\boldsymbol{R}$

$$T_{ERD} = \frac{\lambda_1}{\lambda_K}. \quad (22)$$

Performance analysis can be found, for example, in [23][19].

Noise-independent LRT (LRT-): an alternative log-likelihood ratio was derived in [13], under the assumption of unknown noise variance:

$$T_{LRT-} = 2(N-1)\left[\frac{\frac{1}{K}\sum_{i=1}^{K}\lambda_i}{\left(\prod_{i=1}^{K}\lambda_i\right)^{1/K}}\right]^K. \quad (23)$$

In statistics, this method has been known for many years as the *sphericity test* [24][25]. Performance analysis for cognitive radio applications can be found, for example, in [13].

C. Detectors based on cylic prefix autocorrelation

Primary signal detectors that exploit the presence of the cyclic prefix (CP) in OFDM transmissions have been proposed in the literature. In [5], the detectors based on CP correlation described in [6] have been improved, applied to a real scenario and implemented using a software-defined radio (SDR) platform.

As previously stated, a DVB-T signal consists of OFDM-modulated symbols of which a-priori knowledge of transmission parameters is assumed: the number of subcarriers, cyclic prefix length, constellation type and the code rate. The aim of parametric test statistics is to exploit the knowledge of signal parameters (i.e., signal features) in order to improve primary signals detection with high sensitivity.

The algorithm implemented in [5] using SDR is the well known CP-based spectrum sensing. Assuming that each OFDM symbol consists of $N_s = N_c + N_d$ samples (where N_c is the number of samples in the cyclic prefix and N_d is the number of data samples), the CP correlation function (Eq. 3 in [5]), becomes:

$$R_{xx}^{(\mathrm{CP})}[\tau] = \frac{1}{KN_s}\left|\sum_{k=0}^{K-1}\sum_{n=\tau-kN_s}^{\tau-kN_s-N_c+1} x^*[n]x[n-N_\mathrm{d}]\right|. \quad (24)$$

where K is the number of OFDM symbols on which we compute coherent averaging and τ represents the synchronization mismatch between our acquisition and the symbol start.

The coherent averaging taking the absolute value allows us to improve sensitivity in presence of AWGN noise at the cost of a larger observation interval.

Moreover, to enable the practical implementation of these algorithms, it is necessary to define a noise estimation algorithm to set a threshold that guarantees certain detection performance in terms of probability of false alarm (P_{fa}) and probability of detection (P_d). A slight improvement in terms of noise estimation accuracy (i.e., correlation noise) with respect to [5] is obtained by devising a suitable noise level estimator: without any a-priori knowledge, the correlation noise estimation should be performed by analyzing the received samples when the H_0 hypothesis is true. Hence, training periods with only noise samples must be performed periodically (e.g., to track system temperature changes). To avoid dedicated training, we observed that noise samples can be gathered in a suitable interval between two consecutive correlation maxima. The correlation function $\tilde{R}$ used to estimate the average correlation noise level correspond to the function R, excluding $2N_c$ samples around the detected maxima. In these intervals, $\tilde{R}(\tau) = 0$.

Our optimized CP-based algorithm can thus be summarized as follows:

1) Receive K(N_c+N_d)+N_d samples
2) Perform (24) over acquired samples
3) Record the correlation maximum and its index i
4) Copy only correlation noise values from function in 2. excluding values that have index in the range $i - N_c < i < i + N_c$
5) Decide if the channel is occupied by evaluating the following metric:

$$\frac{\max R_{xx}^{(CP)}[\tau]}{\mathrm{avg}\tilde{R}_{xx}^{(CP)}[\tau]} \gtrless \gamma. \quad (25)$$

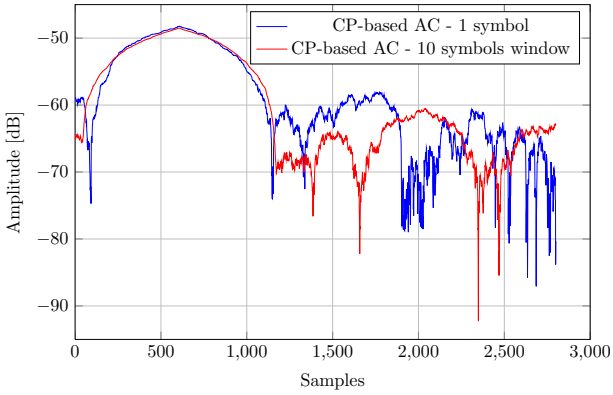


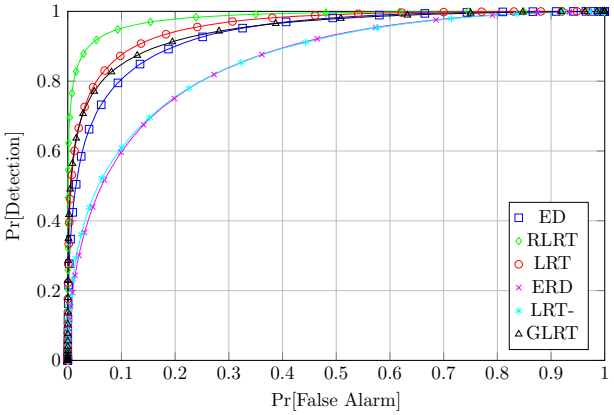
Fig. 2. Amplitude of the CP-based auto-correlation - 1 symbol vs. 10 symbols observation.

Fig. 2 shows the amplitude of the CP-based autocorrelation function with an observation window of respectively $K = 1$ and $K = 10$ symbols in an ideal channel scenario (infinite SNR, no thermal noise). It can be observed that, for $K = 10$, the ratio of the peak value over the maximum value observed outside the cyclic prefix window is improved with respect to $K = 1$.

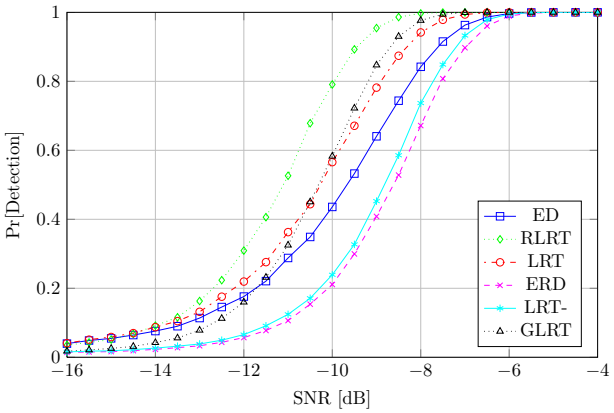
IV. GNU Radio Implementation

The SDR concept has been introduced almost twenty years ago in [26] and it is innovative even nowadays. The main scope of SDR is to improve the flexibility of radio communication devices by implementing the digital sections entirely in software using dedicated or general-purpose processors.

Different SDR platforms have been developed in recent and past years. Among these, an open platform called GNU Radio has prevailed firstly in the academic community and it is becoming frequently adopted even in industrial projects. GNU Radio is a free and open-source software development toolkit for the SDR implementation of transceivers [27]. GNU Radio is not primarily a simulation tool, although it can be used for this purpose. When paired with suitable RF front-ends, a real radio transceiver can be realized. Typically, GNU Radio uses, as RF front-end, the Universal Software Radio Peripheral (USRP) devices by Ettus Research. The GNU Radio environment is structured as a layer of C++ classes organized in components: the *gnuradio-core* component includes a scheduler able to orchestrate the execution of signal processing blocks. Transceiver systems can be modeled through a flow graph, i.e., a set of interconnected signal processing blocks that

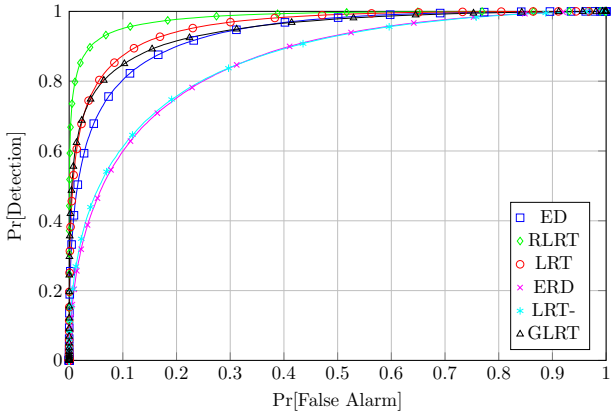


(a) Receiver operating curves (SNR = -10dB).

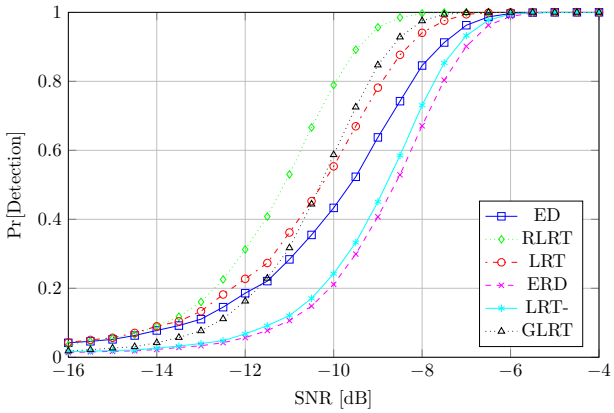


(b) Detection probability.

Fig. 3. DVB-T signal through flat-fading channel.



(a) Receiver operating curves (SNR = -10dB).



(b) Detection probability.

Fig. 4. Gaussian signal through flat-fading channel.

the scheduler executes. Concerning the software structure, it is possible to define models using a high-level object-oriented language such as Python, thanks to an interfacing layer, the Simple Wrapper and Interface Generator (SWIG), that permits to call methods of classes written in C++ from Python. The GNU Radio blocks can be connected to each other to form a flow graph either by using Python or C++, which would result in a higher efficiency. Finally, a graphical tool called GNU Radio Companion (GRC) can be used to compose flow graphs in a visual environment. GRC makes use of the eXtensible Markup Language (XML) to store the flow graph designs. The GNU Radio core components and libraries make available several basic and advanced signal processing and math blocks that make the implementation of transceivers much faster. Users can create new blocks to implement new algorithms that have not been developed yet.

A. *Signal processing blocks*

A module identifies a structure, which is possible to integrate in the overall GNU Radio framework. It consists of one or more signal processing blocks implemented as C++ classes with a well-defined interface.

The main section of a signal processing block is its C++ core engine. The design of a block starts with the definition of the input and output ports, their data types and multiplicities. These characteristics are defined in the block “signature”. A constructor (the “make” method) contains all the initialization code. The sample rates of input and output ports are, in general, different, although some constraints are imposed by some framework design choices. A difference in sample rates between input and output must be notified to the runtime component by calling the “forecast” method, which tells the scheduler how many input items are required to produce a given number of output items. The main signal processing algorithm is defined in the “general_work” method. It receives the data stream provided by the runtime through its input ports, processes it generating output items that are forwarded to blocks connected to its output ports.

Special types of blocks are the sink and source. These are characterized by the lack of either input or output ports. Each flow graph must contain at least one source block and one sink block. As an example, a special sink block is used to connect the flow graph to a USRP device, which will be used to transmit the radio signal.

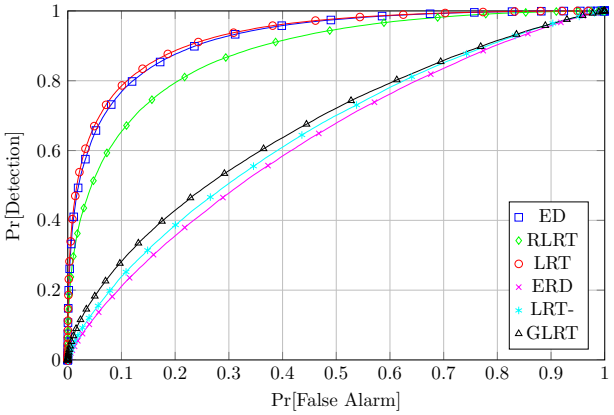
V. EIGENBASED DETECTOR BLOCK

The two optimal non-parametric tests, RLRT and GLRT, respectively for the class of “semi-blind” and “blind” detec-

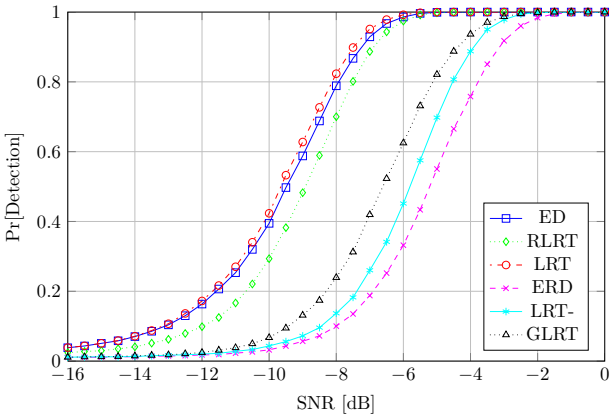
tors, have been implemented and tested in the GNU Radio platform.

A. Eigenvalue algorithm

The most significant part of the GNU Radio block consists in the computation of eigenvalues of the covariance matrix $\boldsymbol{R}$. Such computation is performed through 2 stages: the Lanczos algorithm and the bisection method.



(a) Receiver operating curves (SNR = -10dB).



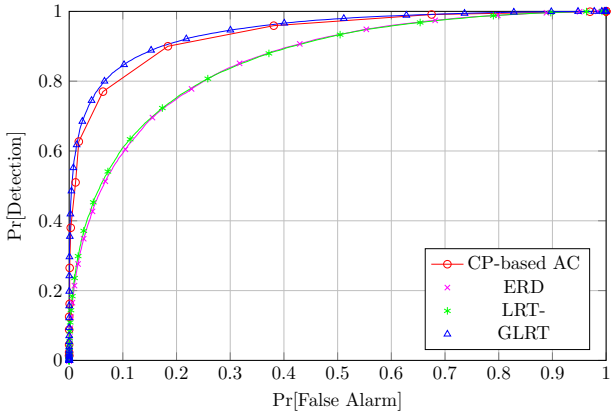
(b) Detection probability.

Fig. 5. DVB-T signal through TU6 channel.

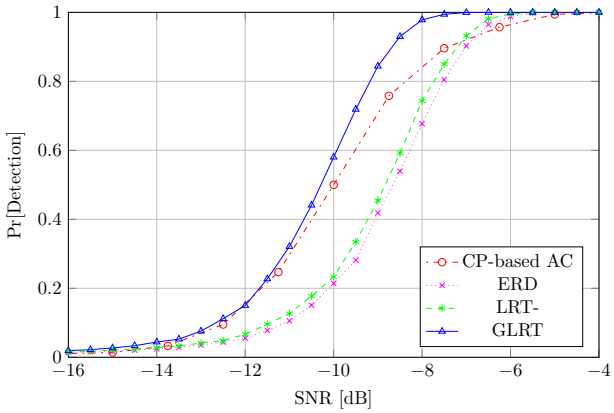
1) Lanczos algorithm: In order to choose the best possible solution we have to take into account that the covariance matrix $\boldsymbol{R}$ is Hermitian, hence the Lanczos method can be applied.

The Lanczos algorithm can be viewed as a simplified Arnoldi's algorithm [28] in that it applies to Hermitian matrices. In the algorithm a series of orthonormal vectors, $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_n$, is generated, which satisfy:

$$\boldsymbol{T} = \boldsymbol{Q}^T \boldsymbol{R}\, \boldsymbol{Q}. \tag{26}$$



(a) Receiver operating curves (SNR = -10dB).



(b) Detection probability.

Fig. 6. Comparison with the CP correlation method.

The matrix $\boldsymbol{T}$ is tridiagonal:

$$\boldsymbol{T}_j = \begin{pmatrix} \alpha_1 & \beta_2 & & & & 0 \\ \beta_2 & \alpha_2 & \beta_3 & & & \\ & \beta_3 & \alpha_3 & \ddots & & \\ & & \ddots & \ddots & \beta_{j-1} & \\ & & & \beta_{j-1} & \alpha_{j-1} & \beta_j \\ 0 & & & & \beta_j & \alpha_j \end{pmatrix},$$

the iterative procedure is based on this three-term recurrence relation:

$$\beta_{j+1}\, \boldsymbol{q}_{j+1} = \boldsymbol{R}\, \boldsymbol{q}_j - \alpha_j\, \boldsymbol{q}_j - \beta_j\, \boldsymbol{q}_{j-1}. \tag{27}$$

More details on the procedure can be found in [29][30]. Here is a short description of the simulation algorithm in pseudocode:

1: **function** LANCZOS($\mathbf{R}, K$)
2: $\mathbf{q_1}$ uniformly distributed random vector
3: $\mathbf{q_1} = \mathbf{q_1}/||\mathbf{q_1}||$
4: $\alpha_1 = \mathbf{q}_1^H\, \mathbf{R}\, \mathbf{q}_1$
5: $\mathbf{w}_1 = \mathbf{R}\, \mathbf{q}_1 - \alpha_1\, \mathbf{q}_1$
6: $\beta_1 = 0$
7: $\beta_2 = ||\mathbf{w}_1||$
8: **for** $i = 2 \rightarrow K$ **do**
9: $\mathbf{q}_i = \mathbf{w}_{i-1}/\beta_i$

```
        α_i = q_i^H R q_i
        w_i = R q_i − α_i q_i − β_i q_{i−1}
        if i < K then
            β_{i+1} = ||w_i||
        end if
    end for
end function
```

The main problem of Lanczos algorithm is the stability, using floating point arithmetic the orthogonality of the vectors $\boldsymbol{q}_j$ is quickly lost. Several stable orthogonalization schemes have been proposed, such as [31], although none of these methods have been applied in our GNU Radio block, since we are basically interested in the computation of the maximum eigenvalue of the covariance matrix. In fact, even with such loss of orthogonality, the algorithm generates very good approximations of the largest eigenvalue. As proven in [32], the Lanczos algorithm produces faster and more accurate results than the power method.

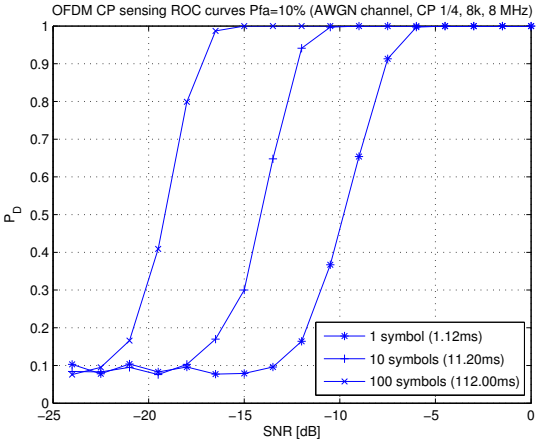


Fig. 7. CP based algorithm probability of detection curves considering a different number of consecutive OFDM symbols.

2) Bisection algorithm: Once, after K iterations, the tridiagonal matrix T has been obtained, the maximum eigenvalue can be simply computed through the bisection algorithm (also called spectral bisection). The algorithm is an iterative procedure based on the computation of the modified Sturm sequence, as explained in [33][34]. The original Sturm sequence, for any number c, is based on the following recursive relation:

$$\begin{aligned} p_0(c) &= 1 \\ p_1(c) &= \alpha_1 - c \\ p_i(c) &= (\alpha_i - c)p_{i-1}(c) - \beta_i^2 p_{i-2}(c), \end{aligned} \tag{28}$$

where $[\alpha_1, \ldots, \alpha_K]$ and $[\beta_2, \ldots, \beta_K]$ are respectively the diagonal and off-diagonal elements of the matrix $\mathbf{T}$. The number $f(c)$ of disagreements in sign between consecutive number of the sequence is equal to the number of eigenvalues smaller than c. However, the original sequence suffers from underflow and overflow problems in floating-point arithmetic, thus, the original sequence $p_i(c)$ is replaced by the sequence $q_i(c)$ defined as

$$q_i(c) = p_i(c)/p_{i-1}(c), \tag{29}$$

now the number of eigenvalues smaller than c, $f(c)$ is given by the number of negative $q_i(c)$. Hence, the new recursive relation is:

$$\begin{aligned} q_1(c) &= \alpha_1 - c \\ q_i(c) &= (\alpha_i - c) - \beta_i^2/q_{i-1}(c). \end{aligned} \tag{30}$$

The algorithm implemented in the GNU Radio block is a slightly modified version of the bisection method described in [33], which computes only the largest eigenvalue. The first part of the algorithm exploits the Gershgorin circle theorem to estimate the upper and lower bounds of the eigenspectrum of the matrix $\mathbf{T}$. In the following pseudocode the bisection function has 6 parameters:

- $\boldsymbol{\alpha}$: diagonal elements of the matrix $\mathbf{T}$;
- $\boldsymbol{\beta}$: off-diagonal elements of $\mathbf{T}$;
- K: order of the tridiagonal matrix;
- m: eigenvalue λ_m is computed, in this algorithm λ_K is the largest eigenvalue (opposite notation w. r. t. III), so $m = K$;
- γ: required precision for the computation of the eigenvalue, which affects the number of iterations;
- ϵ: machine epsilon, the smallest number for which $1+\epsilon > 1$ in the computer.

```
function BISECTION(α, β, K, m, γ, ε)
    xmin = α_K − |β_K|
    xmax = α_K + |β_K|
    for i = K − 1 → 1 do
        h = |β_i| + |β_{i+1}|
        if α_i + h > xmax then
            xmax = α_i + h
        end if
        if α_i − h < xmin then
            xmin = α_i − h
        end if
    end for
    u = xmax
    x = xmax
    v = xmin
    while (u − v) > 2ε(|v| + |u| + γ) do
        p = (u + v)/2
        a = 0
        q = 1
        for i = 1 → K do
            if q ≠ 0 then
                q = α_i − p − β_i^2/q
            else
                q = α_i − p − |β/ε|
            end if
            if q < 0 then
                a = a + 1
            end if
```

```
        end for
        if a < m then
            v = p
        else
            u = p
        end if
    end while
    x = (u + v)/2
    return x
end function
```

B. Description of the "eigenbased" block

In order to work in the GNU Radio framework, the eigenbased detector required the development of a new GNU Radio block. All the details and guidelines on how to write a block can be found, for example, in [35]. Fig. 8 shows the "eigenbased" block.

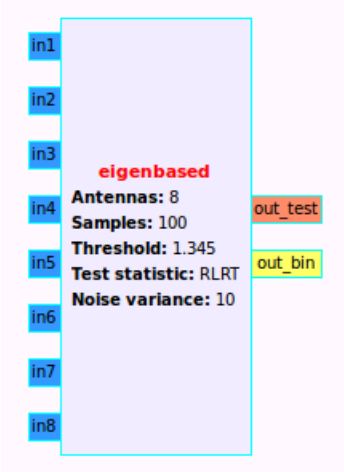


Fig. 8. "Eigenbased" GNU Radio block.

The block has been developed in C++, furthermore a very short part of code has been written in Python and XML in order to make the block available in GNU Radio Companion. The eigenvalue algorithm has been successfully tested in a standalone C++ program up to $K = 100$, while in GNU Radio, for real-time computational effort reasons, we limited K and consequently the block input ports to 32. For simplicity we have shown in Fig. 8 an 8-port block. Since we usually deal with complex baseband signals, all the input must be of type complex.

Five parameters can be set by the user:

- *Antennas*: it corresponds to K, the number of sensors or the oversampling factor in case of single receiving antenna.
- *Samples*: it corresponds to N, the number of samples stored on each row of the Y matrix.
- *Threshold*: the test statistic is compared against this value.
- *Test statistic*: it can be chosen between RLRT or GLRT.
- *Noise variance*: this is required only for RLRT.

The output ports are described as follows:

- *out_test*: type float, it's the floating point value of the computation of the test statistic.
- *out_bin*: type short, after comparing the test statistic against the threshold, it is equal to 1 if the *out_test* is larger than the threshold or equal to 0 otherwise.

VI. CP-BASED DETECTOR BLOCK

A well-known class of techniques that exploits a-priori knowledge of parameters of the received signal is the so-called feature-based detectors. In this subsection, an SDR implementation of an OFDM feature-based detector is discussed.

In order to enable fast symbol synchronization and also avoid inter-symbol interference, the OFDM modulation provides a Cyclic Prefix (CP) insertion between consecutive symbols. Every CP is obtained by copying the tail portion of the OFDM symbol before its head, thus performing a cyclic extension. The result is that the CP length uniquely characterizes each OFDM-based standard. The DVB-T standard, for example, provides four different CP lengths: 1/4, 1/8, 1/16, 1/32 of the OFDM symbol duration. The repetitive pattern produced by the CP can be detected using the CP autocorrelation function (24). Using GNU Radio and the USRP2, the CP-based detection algorithm has been implemented and tested in a real scenario. The resulting spectrum sensing unit is capable of sampling the received signal at 12.5 MHz and calculating (24) to verify the presence of the Cyclic Prefix (CP). In [5], the detectors based on CP autocorrelation described in [6] have been implemented, improved, and applied to a real scenario. The implemented algorithm computes the accumulated autocorrelation function over K OFDM symbols (24) and the test statistics (25).

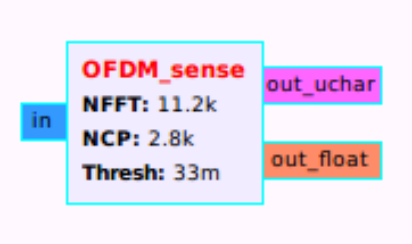


Fig. 11. "OFDM_sense" GNU Radio block.

The developed GNU Radio block, shown in Fig. 11 features the following parameters: number of data samples ($NFFT \equiv N_d$), the number of CP samples ($NCP \equiv N_c$) and the threshold value. The block features one input port through which the received signal is fed and two output ports through which the result of the test statistics is provided in two different numerical formats.

In this work, the USRP2 has been set to process a bandwidth of 12.5 MHz.

Since the DVB-T signal consists of a continuous flow of transmitted symbols, it is possible to perform the sample processing in two steps:

1) buffering;

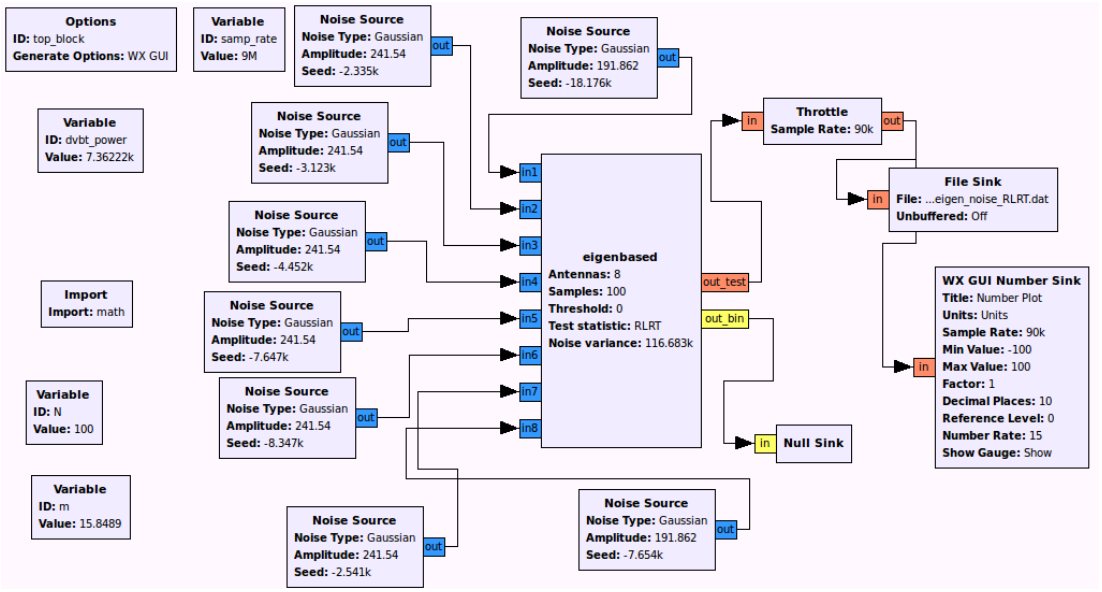


Fig. 9. GNU Radio flow graph corresponding to the $\mathcal{H}_0$ case.

2) compute CP autocorrelation.

This aims at keeping the system responsive and effective. The samples generated by the USRP2 are copied to the input buffer in the general work function that returns no samples until the input buffer is full. Once the buffer is ready, the input from USRP2 is temporarily paused and the processing starts. The samples processed in the general work function are sent at the first output port using a stream of floats. The second output port produces the estimated minimum correlation value, which will be used to set the threshold. A probe block is thus connected to the second output port in order to copy the correlation minimum to the threshold parameter.

VII. Simulation and Results

The performance of the algorithms have been assessed firstly through a numerical computing environment and then using the SDR platform GNU Radio. The performance of each algorithm is shown in terms of:

- the ROC (Receiver Operating Characteristic) curve obtained by plotting the detection probability versus the false alarm probability;
- the detection probability as a function of the signal-to-noise ratio, with fixed false alarm rate (typically, $P_{fa} = 0.01$).

A. Simulation procedure

The Monte Carlo method has been used for our simulation. In order to estimate the values of P_{fa} and P_d, $N_T = 10000$ trials have been performed for each SNR value.

For each considered detection algorithm, we computed the test statistics T_1 corresponding to $\mathcal{H}_1$ (signal plus noise) and T_0 corresponding to $\mathcal{H}_0$ (noise). The estimated P_{fa} (resp. P_d) value corresponding to a given threshold θ is then computed as the empirical complementary cumulative distribution of T_0 (resp. T_1) at θ.

Here is a short description of the simulation algorithm in pseudocode:

$N_B = 1000$ ▷ N_B = number of threshold values
for all *SNR* values **do**
  for $i = 1 \rightarrow N_T$ **do**
    for all detectors **do**
      compute signal+noise test statistic $T_1(i)$
      compute only-noise test statistic $T_0(i)$
    end for
  end for
  for all detectors **do**
    create vector of thresholds $\boldsymbol{\theta}$ of N_B equally spaced values from min($\boldsymbol{T_0}$) to max($\boldsymbol{T_1}$)
    create vector of N_B elements $\boldsymbol{P_{fa}}$
    create vector of N_B elements $\boldsymbol{P_d}$
    for $i = 1 \rightarrow N_B$ **do**
      for $j = 1 \rightarrow N_T$ **do**
        if $T_0(j) > \theta(i)$ **then**
          $P_{fa}(i) \leftarrow P_{fa}(i) + 1$
        end if
        if $T_1(j) > \theta(i)$ **then**
          $P_d(i) \leftarrow P_d(i) + 1$
        end if
      end for
      $P_{fa}(i) \leftarrow P_{fa}(i)/N_T$
      $P_d(i) \leftarrow P_d(i)/N_T$
    end for
  end for
end for

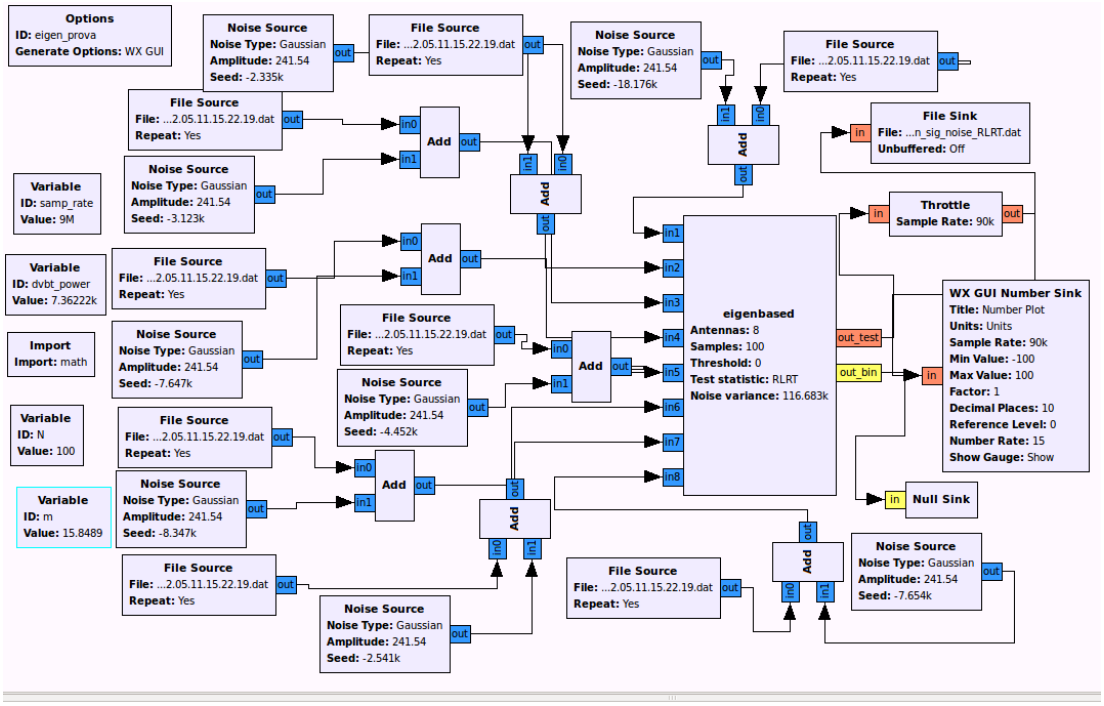


Fig. 10. GNU Radio flow graph corresponding to the $\mathcal{H}_1$ case.

B. GNU Radio flow graphs

At a later stage, we assessed the performance of the algorithms using SDR tools. We defined two different flow graphs containing the "eigenbased" block shown in Fig. 8. In the first flow graph (Fig. 9), corresponding to $\mathcal{H}_0$, only noise sources have been connected to the "eigenbased" block inputs. This configuration has been used to set the threshold θ according to a fixed value of P_{fa}.

In the second flow graph, (corresponding instead to $\mathcal{H}_1$), both noise and signal sources have been connected to the "eigenbased" block inputs (Fig. 10). The primary signal is read from a file containing a segment of DVB-T signal with 8k subcarriers, code rate 5/6, 64-QAM constellation and CP ratio 1/4. The signal is sampled at the nominal rate of 64/7 Msamples/s.

C. Results

The results obtained for the DVB-T signal under linear mixture models provided by flat fading Rayleigh channel are reported in Fig. 3. First, we report the ROC curve then, by setting the false alarm rate to 0.01, we plot the detection probability as a function of the signal-to-noise ratio. By fixing the detection probability, this allows to evaluate the differences in terms of SNR between the algorithms, at the parity of detection and false alarm probability.

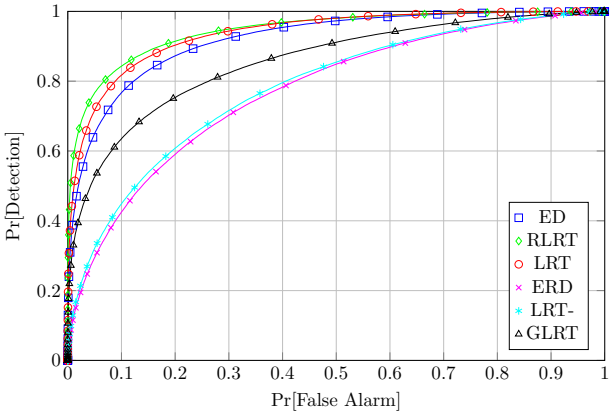
Simulations have been performed assuming $K = 10$ antennas and and observation interval corresponding to $N = 50$ samples.

Looking at Fig. 3, we observe that the best algorithm for known noise variance is the RLRT, while GLRT is the best under unknown variance. We note that these results are in agreement with the results provided in the literature for Gaussian signals. As a reference, results for the same algorithms obtained by simulating Gaussian signal samples are reported in Fig. 4 and are essentially identical to the previous ones (as expected, since the Gaussian properties of the DVB-T signal have been verified).

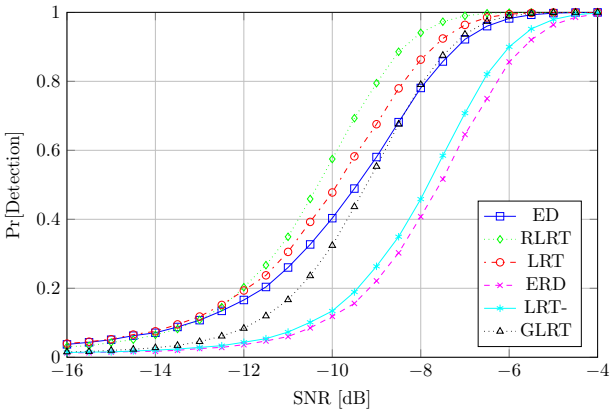
Under a more realistic model, the TU6 channel, the performance of the algorithms are different, as can be observed in Fig. 5. We see how both GLRT and RLRT lose their first place when the received signal model is different from the linear mixture one: simple energy detection becomes highly competitive in this case. The difference between algorithms with known and unknown noise variance is larger, too.

It is important to note that, in this work, we have assumes a perfect knowledge of the noise variance for RLRT, LRT and energy detection. Further analysis will be applied to study their performance under imperfect noise variance knowledge, and address its impact for real DVB-T signals (analysis for Gaussian signals can be found, for example, in [36] for energy detection and [14] for RLRT).

Furthermore, we compare the algorithms for unknown noise variance against the technique exploiting the cyclic prefix autocorrelation of the received signal [5][6] described before. For such method, an interval corresponding to one OFDM symbol has been considered. Moreover, in such case, the signal was sampled at 12.5 Msamples/s. Here, the AWGN channel model is adopted. In this case, we can observe that the performance of this algorithm is similar to that of the GLRT. This single-antenna algorithm does not require the computation of the

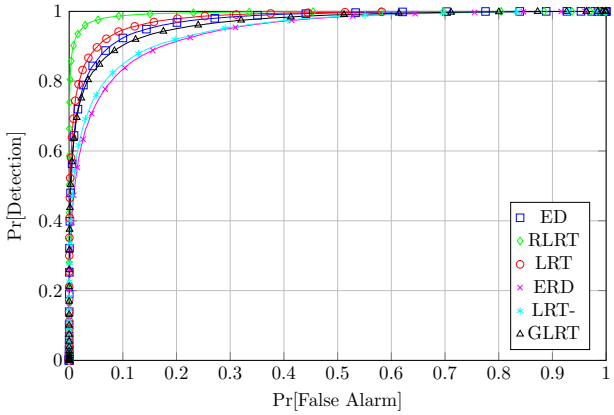


(a) Receiver operating curves (SNR = -10dB).

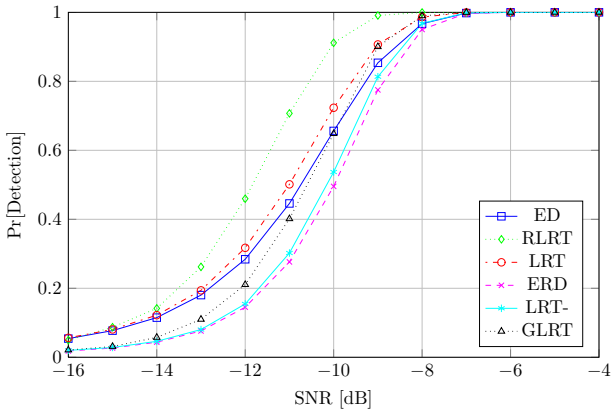


(b) Detection probability.

Fig. 12. DVB-T signal through flat-fading channel with emulated sensors.



(a) Receiver operating curves (SNR = -10dB).



(b) Detection probability.

Fig. 13. DVB-T signal through flat-fading channel, $K = 4$, $N = 200$.

eigenvalues of the sample covariance matrix, but resorts to a precise knowledge of the signal characteristics.

The performance of the proposed feature-based spectrum sensing algorithm has been further evaluated sampling a different number of consecutive OFDM symbols (1, 10, 100 symbols). Considering a constant false alarm probability of 0.1, the curves of Fig. 7 show how the detector sensitivity improves with the number of OFDM symbols. Both in Fig. 6 and Fig. 7 the performance of the CP based method has been estimated using GNU Radio through the procedure described in Sec. VI.

WE also devised a solution for using the eigenbased detectors in a single-antenna scenario. In Fig. 12, corresponding to a flat-fading scenario, $K = 10$ sensors have been emulated through oversampling of a factor K and demultiplexing the received signal stream into K parallel streams. The K demultiplexed and decimated outputs are written in matrix $\boldsymbol{Y}$ so that the generic sample i of the original oversampled signal is stored in row $mod(i, K)$. As it can be noticed, the performance is slightly worse w. r. t. Fig. 3, where the same parameters $N = 50$ and $K = 10$ have been used.

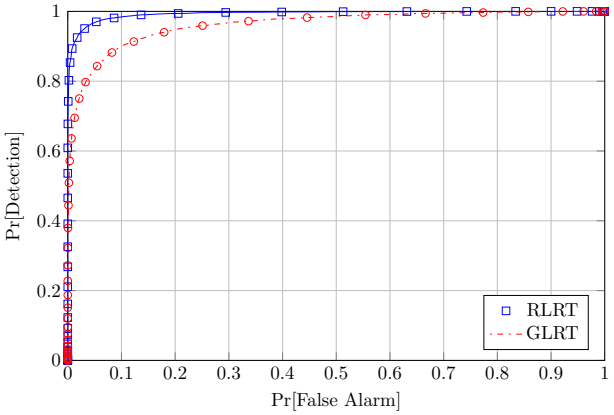
Figs. 13 and 14 show the ROC plot and P_d vs. SNR in a flat fading channel with a different set of paramenters: $K = 4$ and $N = 200$, respectively simulated in a numerical computing environment and in GNU Radio through the flow graphs shown in Figs. 9 and 10. It can be noticed that the results are almost identical, thus validating the correctness of the "eigenbased" block.

Finally, in Fig. 15 we plot the detection probability of GLRT as a function of the observation interval (expressed both in time units and number of received samples per sensor) and the number of sensors for a specific SNR value of -10dB and -15dB, while the false alarm probability remains fixed to 10^{-2}. The channel is Rayleigh flat-fading. As shown, it is possible to obtain the same performance achieved in Fig. 3 using $K = 10$ sensors even with a lower and hence more realistic number of antennas.

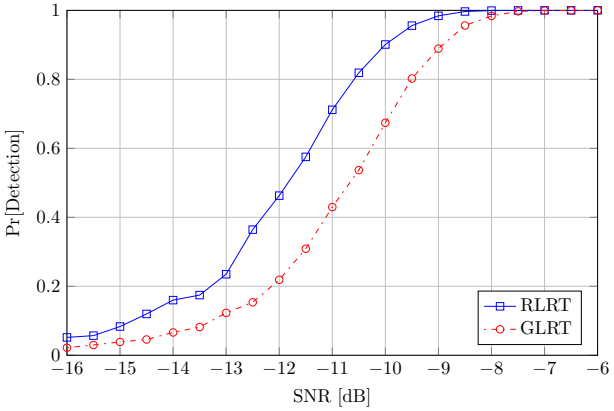
VIII. Conclusions

Unused channels in licensed RF bands are one of the most attractive opportunities for secondary cognitive radio systems as long as they enjoy favorable propagation conditions and feature fairly large bandwidths.

In order to avoid interference to PUs while simultaneously performing high-speed communications, TV white-space systems must be equipped with efficient spectrum-sensing procedures.

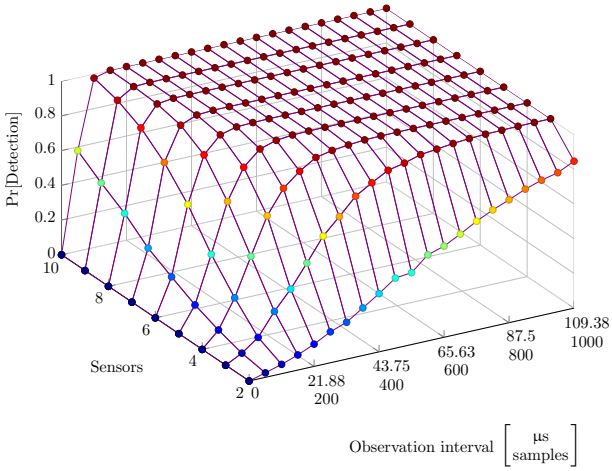


(a) Receiver operating curves (SNR = -10dB).

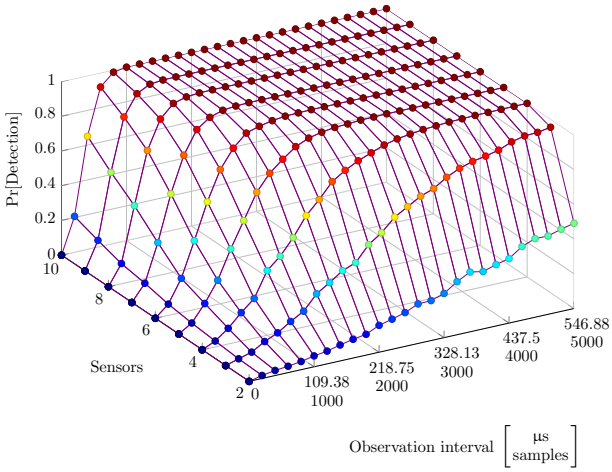


(b) Detection probability.

Fig. 14. DVB-T signal thorugh flat-fading channel, GNU Radio simulation, $K = 4$, $N = 200$.



(a) SNR = -10dB, P_{fa} = 0.01.



(b) SNR = -15dB, P_{fa} = 0.01.

Fig. 15. GLRT detection probability as a function of time (samples) and sensors through flat-fading channel.

For a large selection of sensing algorithms among the most relevant ones, performance has been assessed applying real DVB-T under different channel profiles: he flat fading channel analysis confirms the results previously obtained by simulation using linear mixture models of Gaussian signals while, under a more realistic multi-path channel model, the performance relationships between the same algorithms is completely different. Simpler algorithms like energy detection becomes highly competitive in this case.

A SDR implementation of an *eigenbased* spectrum sensing detector has been carried out to prove that the complexity of these algorithms is manageable in real-time systems.

Overall, these results show that spectrum sensing algorithms applied to OFDM signals can reach a high accuracy in terms of primary signal detection and false-alarm rate, hence they are well suited for TV white-space secondary networks.

In our research, further effort will be devoted to investigating broader sets of sensing techniques, both with a single antenna and with multiple antennas. Distributed sensing algorithms will be considered as well in order to achieve further improved performance.

REFERENCES

[1] D. Riviello, S. Benco, F. Crespi, A. Perotti, and R. Garello, "Sensing of DVB-T signals for white space cognitive radio systems", *COCORA 2013, The Third International Conference on Advances in Cognitive Radio*, Apr. 2013, pp. 12-17.

[2] J. Mitola, III, "Cognitive radio - an integrated agent architecture for software defined radio," Ph.D. dissertation, Royal Institute of Technology (KTH), May 2000.

[3] S. Haykin, D. J. Thomson, and J. H. Reed, "Spectrum sensing for cognitive radio," *Proceedings of the IEEE*, , May 2009, Vol. 97, No. 5, pp. 849-877.

[4] Y. Zeng, Y.-C. Liang, A. T. Hoang, and R. Zhang, "A review on spectrum sensing for cognitive radio: challenges and solutions," *EURASIP Journal on Advances in Signal Processing*, Jan. 2010, pp. 1-15.

[5] S. Benco, F. Crespi, A. Ghittino, and A. Perotti, "Software-defined white-space cognitive systems: implementation of the spectrum sensing unit", in *The 2nd Workshop of COST Action IC0902*, Oct. 2011, pp. 1-3.

[6] D. Danev, E. Axell, and E. G. Larsson, "Spectrum sensing methods for detection of DVB-T signals in AWGN and fading channels," in *Proc. IEEE 21st International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, Sep. 2010, pp. 2721-2726.

[7] K. Manolakis, D. M. Gutiérrez Estévez, V. Jungnickel, W. Xu, and C. Drewes, "A closed concept for synchronization and cell search in 3GPP LTE systems," *IEEE Wireless Communications and Networking Conference (WCNC 2009)*, Apr. 2009, pp. 1-6.

[8] F. Crespi, M. Maglioli, A. Perotti, and S. Benco, "A real-time video broadcasting system based on the GNU Radio-USRP2 platform", *Proc. Karlsruhe Workshop on Software Radios (WSR)*, Mar. 2012.

[9] European Telecommunications Standards Institute, "ETSI EN 300 744, digital video broadcasting (DVB); framing structure, channel coding and modulation for digital terrestrial television," Jan. 2009.

[10] COST207, "Digital land mobile radio communications (final report)," Commission of the European Communities, Directorate General Telecommunications, Information Industries and Innovation, Tech. Rep., 1989.

[11] J. Neyman and E. Pearson, "On the problem of the most efficient tests of statistical hypotheses", *Philosophical Transactions of the Royal Society of London*, 1933, Series A, vol. 231, pp. 289-337.

[12] P. Bianchi, M. Debbah, M. Maida, and J. Najim, "Performance of statistical tests for single-source detection using random matrix theory", *IEEE Transactions on Information Theory*, 2011, pp. 2400-2419.

[13] Q. Zhang, "Advanced detection techniques for cognitive radio", in *Proc. of International Conference on Communications (ICC 2009)*, Jun. 2009, pp. 1-5.

[14] B. Nadler, F. Penna, and R. Garello, "Performance of eigenvalue-based signal detectors with known and unknown noise level", in *Proc. of International Conference on Communications (ICC 2011)*, Jun. 2011, pp. 1-5.

[15] S. N. Roy, "On a heuristic method of test construction and its use in multivariate analysis", *Annals of Mathematical Statistics*, Jun. 1953, vol. 24, no. 2, pp. 220-238.

[16] L. Wei and O. Tirkkonen, "Cooperative spectrum sensing of OFDM signals using largest eigenvalue distributions", in *IEEE International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC 2009)*, Sep. 2009, pp. 2295-2299.

[17] S. Kritchman and B. Nadler, "Non-parametric detections of the number of signals: hypothesis testing and random matrix theory," *IEEE Transactions on Signal Processing*, Oct. 2009, Vol. 57, No. 10, pp. 3930 -3941.

[18] J. Baik and J. W. Silverstein, "Eigenvalues of large sample covariance matrices of spiked population models",*J. Mult. Anal.*, 2006, vol. 97, no. 6, pp. 1382-1408.

[19] F. Penna, R. Garello, and M. A. Spirito, "Probability of missed detection in eigenvalue ratio spectrum sensing", in *5th IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, Oct. 2009, pp. 117-122.

[20] H. Urkowitz, "Energy detection of unknown deterministic signals", *Proceedings of the IEEE*, Apr. 1967, vol. 55, no. 4, pp. 523-531.

[21] P. Bianchi, J. Najim, G. Alfano, and M. Debbah, "Asymptotics of eigenbased collaborative sensing", in *Proc. IEEE Information Theory Workshop (ITW 2009)*, Oct. 2009, pp. 515-519.

[22] G. V. Moustakides, "Finite sample size optimality of GLR tests", *IEEE Transactions on Information Theory*, Nov. 2009, pp. 1-20.

[23] Y. H. Zeng and Y.-C. Liang, "Eigenvalue based spectrum sensing algorithms for cognitive radio", *IEEE Transactions on Communications*, Jun. 2009, Vol. 57, No. 6, pp. 1784-1793.

[24] W. J. Krzanowski, *Principles of multivariate analysis: a user's perspective*. Oxford University Press, 2000.

[25] J. W. Mauchley, "Significance test for sphericity of a normal n-variate distribution", *Annals of Mathematical Statistics*, Jun. 1940, vol. 11, no. 2, pp. 204-209.

[26] J. Mitola, III, "Software radios-survey, critical evaluation and future directions",*IEEE Aerospace and Electronic Systems Magazine*, 1993,vol. 8, no. 4, pp. 25-36.

[27] GNU Radio. [Online]. Available:http://gnuradio.org 06.12.2013

[28] W. E. Arnoldi, "The principle of minimized iterations in the solution of the matrix eigenvalue problem", *Quarterly of Applied Mathematics*, 1951, vol. 9, pp. 1729.

[29] M. T. Jones and M. L. Patrick, *The use of Lanczos's method to solve the large generalized symmetric definite eigenvalue problem*, Defense Technical Information Center, 1989.

[30] J. K. Cullum and R. A. Willoughby, *Lanczos algorithms for large symmetric eigenvalue computations: Vol. I: Theory*, SIAM, 2002.

[31] C. Guo and S. Qiao, "A stable Lanczos tridiagonalization of complex symmetric matrices", *Advanced Signal Processing Algorithms, Architectures, and Implementations XV, Proc. of SPIE*, 2005, pp. 1-12.

[32] J. Kuczynski and H. Wozniakowski, "Estimating the largest eigenvalue by the power and Lanczos algorithms with a random start", *SIAM Journal on Matrix Analysis and Applications*, Oct. 1992, vol. 13, no. 4, pp. 1094-1122.

[33] W. Barth, R. S. Martin, and J. H. Wilkinson, "Calculation of the eigenvalues of a symmetric tridiagonal matrix by the method of bisection", *Numerische Mathematik*, 1967, Vol. 9, pp. 386-393.

[34] D. J. Evans, J. Shanehchi, and C. C. Rick, "A modified bisection algorithm for the determination of the eigenvalues of a symmetric tridiagonal matrix", *Numerische Mathematik*, 1982, Vol. 38, pp. 417-419.

[35] G. Verma and P. Yu, "Developing signal processing blocks for software-defined radios", *Adelphi, MD : Army Research Laboratory*, Jan. 2012.

[36] R. Tandra and A. Sahai, "SNR walls for signal detection", *IEEE Journal of Selected Topics in Signal Processing*, Feb. 2008, vol. 2, no. 1, pp. 4-17.

Improving Retransmission Performance of IP-Based Transport Protocols

Stan McClellan
Ingram School of Engineering
Texas State University
San Marcos, TX USA
stan.mcclellan@txstate.edu

Wuxu Peng
Department of Computer Science
Texas State University
San Marcos, TX USA
wuxu@txstate.edu

***Abstract*—This paper analyzes the algorithm used for estimating retransmission timeouts in connection-oriented IP-based transport protocols, such as the Transmission Control Protocol (TCP) and the Stream Control Transmission Protocol (SCTP). The estimation algorithm uses historical values of the round-trip time to estimate future round-trip delays, and so creates a maximum waiting time before triggering retransmission attempts. The purpose of the analysis is to question / validate some of the fundamental assumptions used in the estimation algorithm. The conclusion of the analysis is that the algorithm is somewhat mismatched to the dynamics of the current Internet. Alternative algorithms are discussed, and potential modifications are presented. Impact of the suggested alternative algorithm on the well-known selective acknowledgement and fast retransmit mechanisms is discussed.**

***Keywords*-SCTP; retransmission timeout; round-trip time; RTT; RTO; selective acknowledgment; fast retransmits; Jacobson algorithm; Chebyshev approximation; parameter estimation; upper bound.**

I. Introduction

This paper looks at the timing aspect of IP-based transport protocols (TCP and SCTP) of the Internet. Although this topic has been studied extensively in the literature in the past, we believe that our paper has made valuable new contribution on it. The theme of this paper is that the timeout aspects of current TCP and SCTP protocols can no longer reflect the current infrastructure and traffic dynamics of today's Internet. Current and future research should work on new and practical timeout mechanisms that meet the need of changed/changing landscapes of Internet. The paper is an extended and enhanced version of our conference paper [1].

IP-based transport protocols such as the Transmission Control Protocol (TCP) [2] and the Stream Control Transmission Protocol (SCTP) [3], [4] estimate maximum round-trip times using data from prior successful transmissions. The purpose of this estimation process is to create a triggering mechanism for retransmission procedures when transmissions are lost or seriously delayed. Estimation of the maximum round-trip time is performed via the Jacobson algorithm [5], which is codified in several IETF RFC's, including RFC 6298 [6]. The Jacobson algorithm has an interesting basis in fundamental theory, but suffers from some performance issues due to a mismatch between the theory and the application area. Performance issues related to the Jacobson algorithm and other retransmission procedures have been noted and addressed in several alternative approaches, including [7]–[12]. The paper discusses the Jacobson algorithm, the theory which motivates it, and several alternative algorithms including a new approach which is a modified form of Jacobson. Section II describes the estimation process and its use in establishing timeouts for retransmission procedures. Related work in retransmission optimization and timeout boundaries is also summarized. In Section III, the parameters of the existing algorithm are analyzed, and in Section IV an alternative approach is presented based on similar theoretical concepts, and achieving improved results. Performance results based on implementation and simulation are summarized in Section V. In Section VI, the impact on SCTP selective acknowledgments (SACK) and fast retransmits (FRT) are discussed. Our simulation results showed that proposed approach has either minimal or manageable impact on both SACK and FRT. Section VII concludes the paper. Note that much of this work is presented in the context of SCTP, but is also applicable to TCP since the timeout estimation processes are identical.

II. Retransmission mechanisms

When an SCTP sender transmits a unit of data, called a *chunk*, it also initializes a *retransmission timer* with an estimated value of the round-trip time (RTT). The value of this timer is the *retransmission time-out* (RTO). When an acknowledgment arrives, the timer is cancelled. If the timer expires before an acknowledgment arrives, the chunk may be retransmitted. The value of RTO is calculated from observed / actual values of RTT using the *Jacobson Algorithm*, which is detailed in Section III. A too optimistic retransmission timer may expire prematurely, producing *spurious timeouts* and *spurious retransmissions*, reducing a connection's effective throughput. On the contrary, a retransmission timer that is too conservative may cause long idle times before lost packets are detected and retransmitted. This can also degrade performance [7]. So, the difficulty lies in finding an algorithm which has a solid theoretical basis, is not computationally expensive, and can predict RTT with enough

bias to minimize retransmission events and waiting time *simultaneously*.

Performance issues related to retransmission procedures, including alternatives to the Jacobson Algorithm, have been noted and addressed several times in the literature. Much work has been focused on late retransmission and other optimizations of the overall retransmission scenarios [13], [14]. Many authors approach this problem with a "holistic" or overall perspective on the retransmission procedures where RTT estimation contributes to triggering these procedures. Other authors specifically address the estimation of RTT, and propose completely new algorithms. However, the Jacobson Algorithm is deeply rooted in the fabric of connection-oriented IP transport protocols, and its basis in fundamental theory is well-established. In following subsections, we briefly summarize several important approaches to retransmission and RTO estimation.

A. Holistic approaches

Holistic or overall approaches to improving retransmission performance typically address the state machines surrounding the retransmission process, including the estimation algorithm which may be used to trigger these procedures. Examples of holistic approaches include *Early Fast Retransmit* (EFR) [11], *Early Retransmit* (ER) [12], and *Window-Based Retransmission* (WB-RTO) [15] as well as protocol-specific techniques such as *Thin Streams* [10], [11], [16].

Early Fast Retransmit (EFR) is an optional mechanism in FreeBSD which is active whenever the congestion window is larger than the number of unacknowledged packets, and packets remain to be sent. When the RTO timer expires and the entire congestion window is not used, EFR retransmits all packets that could have been acknowledged [11]. The *Early Retransmit* (ER) algorithm [12] suggests that a mechanism should be in place to recover lost segments when there are too few unacknowledged packets to trigger Fast Retransmit. The Early Retransmit algorithm reduces waiting time in four specific situations [11]. The *Window-Based Retransmission Timeout* (WB-RTO) [15] asserts that timeout mechanisms based solely on RTT estimates lead to unnecessary retransmissions and unfair resource allocation, and proposes to schedule flows on the basis of their contribution to congestion. *Thin Streams* [10], [11], [16] optimizes throughput for "thin streams" which are often used in control applications, and often depend on SCTP for transport. When stream characterization is accurate, throughput can be improved by adapting specific sections of the retransmission procedures to match flow characteristics.

B. Alternative estimation algorithms

Alternative estimation algorithms address specific performance issues which have been noted in the Jacobson Algorithm. These issues may be related to the over-estimated value, spurious behaviors for certain traffic characteristics, or inefficient bounding computations. In some cases, heuristic state-machine approaches are also included because of complexities associated with the retransmission process. Examples of alternative or modified RTT estimation algorithms include *Peak-Hopper* [8], *Eifel* [7], and *Weighted Recursive Median* (WRM) [9].

The *Eifel* approach [7] notes a particular style of erroneous performance in the Jacobson algorithm, and adapts the algorithm in several ways to compensate for this performance oddity. As a result, Eifel eliminates unnecessary retransmissions which can result from spurious RTO violations. Similar to Eifel, the *Peak-Hopper* algorithm [8] also observes that the Jacobson algorithm responds inappropriately to certain fluctuations in RTT. This behavior produces "spikes" in RTO values because the algorithm does not distinguish between positive and negative variations. The modification proposed in [8] reduces this effect for a wide range of cases, and the findings in [10] concur. However, this solution results in higher average RTO values than the RFC approach [6], which can be a problem [11]. The *Weighted Recursive Median* (WRM) algorithm [9] redefines RTT estimation from a signal processing standpoint. WRM is effective, but tends to be computationally expensive even in a recursive form, which is a problem for per-packet operations.

C. Other considerations

The remainder of this paper addresses the estimation process for the maximum RTT, or the value which establishes the RTO timer. When the RTO timer has expired, retransmission procedures commence, and may include various conditionally executed processes. We do not address those processes, or the overall retransmission procedure. In most cases, we use the Jacobson and Eifel estimation algorithms for comparison because they are widely accepted or implemented.

For reference, the performance of the Jacobson and Eifel algorithms are shown in Figure 1 along with the modified version of the Jacobson Algorithm which is discussed in detail in Section IV. The traces in the figure are all driven by a common RTT sequence created using our testbed of systems with modified networking stacks, as described in more detail in Section V. In the figure, the quiescent sections of the RTT sequence ($t < 170$ms, $260\text{ms} < t < 430$ms) have a fairly low mean, with similarly low standard deviation. This is typical of modern, high-speed networks. Also note that the RTT sequence has abrupt increases ($t \approx 170$ms, $t \approx 430$ms) followed by a stable period ($170\text{ms} < t < 260$ms), and then an abrupt decrease ($t \approx 260$ms). Note that Jacobson and Eifel both "overshoot" after the abrupt positive discontinuity in RTT. However, at the second discontinuity which is abrupt but negative, Eifel corrects downward, whereas Jacobson again corrects upward. This is a primary beneficial feature of the Eifel approach. Unfortunately, the tuning of the

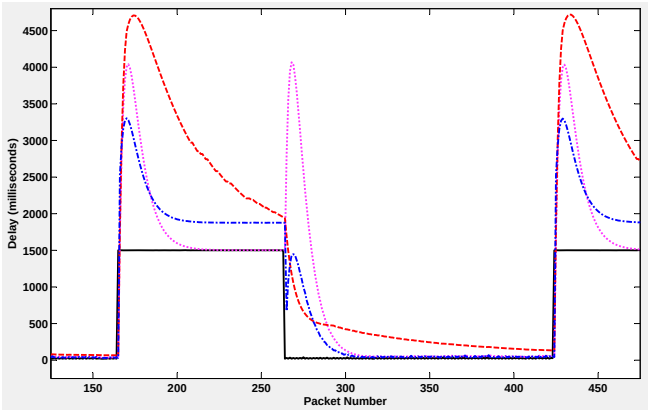


Figure 1. Relative performance of RTO estimators to large discontinuities in RTT. Estimators shown are the Jacobson Algorithm (dotted line), the Eifel Algorithm (dashed line), and the Modified Jacobson Algorithm (dash-dot line). The common RTT sequence driving all algorithms is shown as a solid line.

algorithm creates a larger deviation or "upper bound" for the RTT sequence than is available through the Jacobson Algorithm.

Abrupt changes in RTT (such as those shown in Figure 1) cause the RTO timer to expire, resulting in binary backoff of the timer (BEBO) and retransmission procedures. The RTO values associated with BEBO and retransmissions are not shown, because they do not affect future estimated values of RTO. Post-processing of the RTT estimate to enforce minimum values for RTO is also not considered here. In the RFCs for TCP and SCTP, a hard-coded minimum value for RTO is specified as 1000 milliseconds [2]–[4], [6], whereas the minimum value of the Eifel algorithm is defined as "RTT + 2 $\times$ ticks" (or 200 msec) [7]. The definition of the minimum RTO in TCP-Lite is typically 2 times the clock granularity, which is often taken as 500 msec [4], [7]. This minimum procedure is secondary to the estimation algorithm, and often completely replaces the RTT estimate. For example, in a broadband network with large minimum RTO, the RTT estimate can be orders of magnitude below the minimum, resulting in long wait-times for triggering retransmission. So, in all comparisons here we disable the enforcement of a minimum RTO and focus on the performance of the estimation algorithms.

In [17], a new algorithm to improve the SCTP's retransmission mechanism is proposed. The algorithm uses several fixed values of minimum RTO values, instead of a single minimum RTO value of 1000 msec. It is shown that the algorithm can improve SCTP's performance by as much as five percent.

III. The Jacobson Algorithm

The Jacobson Algorithm, originally proposed in [5], uses the Chebyshev Bound [18] to produce a reasonable value of RTO, or the maximum time the sender will wait for an acknowledgment. After exceeding RTO, a transmission is declared lost and retransmission procedures commence. Interestingly, Jacobson noted the poor performance of the algorithm, since loads higher than 30% resulted in retransmission of packets that were only delayed in transit (i.e., not lost) [5]. This behavior is also noted in later literature, including [7], [8], [16].

The specific computations for the bounding procedure are driven by estimates of the mean (μ) and the standard deviation (σ) of an assumed RTT distribution. The estimates of μ and σ are not conventional parameter estimates, as independent and identically distributed samples from a population. Rather, $\hat{\mu}$ and $\hat{\sigma}$ are *predicted* using prior RTT values. Jacobson estimates $\hat{\mu}$ and $\hat{\sigma}$ from observed values of RTT, and then computes an "overbound" for RTO using Chebyshev. This is the same as saying "we waited a reasonable time ($\hat{\mu}$) and *then* some ($K\hat{\sigma}$), but the ACK didn't come back, so the packet must have gotten lost". This RTO calculation is invoked by the sender for each unique transmission. As such, it is optimized for integer arithmetic and all coefficients are diadic.

Regardless of the specific assumptions or optimizations, Jacobson computes the RTO threshold as

$$x_{thr} = \hat{\mu}_{n+1|n} + K \cdot \hat{\sigma}_{n+1|n} \tag{1}$$

where x_{thr} is the RTO threshold or "overbound" for RTT, K is the number of "standard deviations past the mean," and $\hat{\mu}_{n+1|n}$ and $\hat{\sigma}_{n+1|n}$ are the estimates (predictions) of the mean and standard deviation of the RTT distribution for the next iteration (subscript $n+1$) given some data up to the current time (subscript n).

A. Jacobson and Chebyshev

The Chebyshev bound (2) is a universal bound applicable even for unknown distributions. Chebyshev shows that the probability of the random variable occurring outside a range around the mean (μ) depends on the standard deviation (σ).

$$\Pr\left[|X - \mu| \geq \epsilon\right] \leq \frac{\sigma^2}{\epsilon^2} \tag{2}$$

For RTT estimation, the Jacobson Algorithm uses a fixed offset from the mean ($K\sigma$) [3], [4]. This simplifies the bound and allows a direct computation of the "violation probability". With $\epsilon = K\sigma$ and $K = 4$ in (2),

$$\Pr\left[|X - \mu| \geq K\sigma\right] \leq \frac{1}{K^2} \tag{3}$$

and the fixed, double-sided "violation probability" is $\frac{1}{16} = 0.0625$. However, the RTO timeout is single-sided because the timeout algorithm is only concerned with the case where $X > \mu$. So for a symmetric distribution and $K = 4$, the RTO timeout will be exceeded roughly 3% of the time, with the assumption that the RTT values are reasonably independent, identically distributed.

B. Predicting μ and σ

Rather than using conventional parameter estimation algorithms requiring storage of historical values and more complex calculation, Jacobson estimates μ and σ using simple *prediction algorithms* as in (4). These algorithms rely on current values of the quantities (subscript n) as well as the measured RTT value (x_n). Interestingly, the use of *filtered* (predicted) values for μ and σ implicitly contradicts the assumption of a valid distribution in (2). However, this approach gains computational efficiency and reflects the nonstationary nature of RTT values.

$$\begin{aligned} \hat{\mu}_{n+1|n} &= \hat{\mu}_n + \alpha(x_n - \hat{\mu}_n) \text{ and} \\ \hat{\sigma}_{n+1|n} &= \hat{\sigma}_n + \beta\left(|\hat{\mu}_n - x_n| - \hat{\sigma}_n\right). \end{aligned} \tag{4}$$

When viewed as a time-series prediction or filtering algorithm, it is clear that the relations in (4) are single-pole, lowpass IIR (infinite impulse response) filters (predictors).

C. Selective Acknowledgment

Selective acknowledgment (SACK) was initially introduced in RFC 1072 [19], [20] to handle multiple dropped packets within a window [21]. This feature was omitted in RFC 1323 [22] as the authors felt the details of SACK was still to be worked out. SACK was reintroduced in RFC 2018 [23] and is now an important performance enhancement mechanism in TCP initially and now SCTP as well.

The usefulness and effectiveness of SACK mechanism have been extensively studied. In [24], it was found that "SACK improves TCP throughput significantly in moderate congestion (with a packet loss rate between 2 and 4%), and that the negative impact of SACK on competing non-SACK TCP connections is small".

SACK was proposed to enhance TCP performance over *long delay paths* [19], [23]. We observe that the RTT over the Internet has been steadly decreasing over the last thirty years. This decrease is primarily due to the improvement of Internet infrastructure/software that shortens the store/forward delays and transmission delays.

D. Fast Retransmit

TCP fast retransmit (FRT) and fast recovery algorithm, originally formally defined in [25], and subsequently refined in [26]–[28], is another timing related performance enhancement mechanism. The algorithm is based on a simple heuristic.

In Berkeley-derived kernel implementation FRT works by retransmitting a specific un-acknowledged TCP segment B that was sent after segment A when the third duplicate acknowledgment for segment A is received. The logic reasoning is that by the time the third duplicate acknowledgment for segment A is received, segment B must have lost and hence should be retransmitted. If segment B was not lost, then it should have been acknowledged because four acknowledgments for segment A (one original plus three duplicate acknowledgments) have been received.

It is observed here that the FRT action for segment B occurs before the timer associated with B expires. Otherwise, the retransmission action would be a normal retransmission, not a fast retransmission.

IV. The Modified Jacobson Algorithm

The use of Chebyshev to bound the retransmission timeout is reasonable, since it provides a "target probability" for the timeout calculation. However, the Markov bound [18] is also applicable for RTT estimation since it explicitly uses knowledge of the positivity of the random variable, as in (5) which is valid when $f_X(x) = 0$ for $x < 0$ and $\alpha > 0$.

$$\Pr[X \geq \alpha] \leq \frac{\mu}{\alpha}. \tag{5}$$

Using $\alpha = \mu + K\sigma$ in (5) produces an expression similar to (2). However, the Markov formulation results in a *variable* probability for RTO violation. This concept is particularly unappealing for small, relatively stable values of RTT, since the overbound RTO might have a high probability of violation, which would create spurious retransmissions and a large amount of unwanted network traffic. Recall that in the Chebyshev case, the choice of $\epsilon = K\sigma$ produced a fixed violation probability around 3%. The violation probability of the Markov bound can also be fixed as in (3) if $\alpha = 32\mu$:

$$\Pr[X \geq 32\mu] \leq \frac{\mu}{32\mu} = \frac{1}{32} = \frac{1}{2K^2}. \tag{6}$$

Unfortunately, a bias of 32 times the mean would not produce a viable estimate of RTT, and the overbound RTO would be extremely loose. Therefore, the Markov bound *alone* is not be a reasonable choice to estimate the RTO timeout. However, a *combination* of Markov and Chebyshev approaches seems to produce an effective estimator.

Combining an estimate of σ as in the Jacobson Algorithm with a *biased* estimate of μ as in the Markov bound results in a formulation that retains the Chebyshev structure but improves certain performance aspects. So, we use a slightly revised version of (1) and call this approach the *Modified Jacobson Algorithm*,

$$x_{thr} = \mathcal{A} \cdot \hat{\mu}_{n+1|n} + \mathcal{B} \cdot \hat{\sigma}_{n+1|n}. \tag{7}$$

In (7), the standard deviation estimator $\hat{\sigma}_{n+1|n}$ is identical to the estimator (4) used in the Jacobson Algorithm. However, the multiplier for $\hat{\sigma}_{n+1|n}$ is *reduced* (i.e., $\mathcal{B} = 2$ whereas $K = 4$). Further, the mean estimator $\hat{\mu}_{n+1|n}$ is replaced with the current value of RTT (x_n), and is *biased* in the spirit of a Markov estimator. In this case, we choose $\mathcal{A} = 1.25$ as a reasonable bias term, whereas Jacobson uses $\mathcal{A} = 1.0$.

The values of $\mathcal{A} = 1.25$ and $\mathcal{B} = 2.0$ which are used in (7) were determined experimentally, using the Jacobson values ($\mathcal{A} = 1.0$ and $\mathcal{B} = 4.0$) as a starting point.

Jacobson's mean estimator μ is a filtered version of the time series x_n. The filter structure (in (4)) is a single-pole predictor, or infinite impulse response (IIR) filter. Since the structure is IIR, it embodies some instabilities. As a result, under certain circumstances in the input sequence (such as abrupt changes in x_n) the output value of the IIR filter ca become very large. This is problematic for the combined estimation problem which involves estimators of both μ and σ, particularly when σ is also estimated using an IIR structure. Furthermore, the filtered estimate of μ contains at least a one-sample (one-packet) delay. This can be seen from (4), where μ for time $(n+1)$ is composed using values of μ and x from time (n). This delay in estimation also produces a delay in reaction to instantaneous changes in x, and it is precisely these changes in x that we are attempting to model more accurately. Thus, we replace the IIR-filtered estimate of μ with the instantaneous value x_n which is a viable estimator of the instantaneous mean for this sequence. After replacing μ with x_n, we adjust the bias value of the coefficient according to Markov's relation so that in cases where σ goes to zero, the combined estimation of RTT will not "collapse" exactly onto x_n, which would trigger many more timeouts. In other words, in (7), as σ goes to zero, the value of x_{thr} remains biased above x due to the use of $\mathcal{A} = 1.25$ rather than $\mathcal{A} = 1.0$, as in Jacobson, and this reduces timeouts in quiescent channels. The use of $\mathcal{A} = 1.0$ in Jacobson can be problematic in this regard, but introducing a constant bias (e.g., $\mathcal{A} = 1.25$) in Jacobson can be counterproductive because of the IIR nature of the prediction of μ and the IIR nature of the prediction of σ.

The use of $\mathcal{B} = 2.0$ in the Modified Algorithm follows similar logic. Because the predictor of σ in (4) is an IIR structure, the possibility of overshoot can be dramatic in cases where the RTT sequence x_n has sudden, sharp variations. A multiplier of $\mathcal{B} = 4.0$ or $K = 4.0$ (as in Jacobson) accentuates this overshoot, and is largely responsible for the massive over-estimation of RTT in cases where the underlying network exhibits certain types of instabilities. The specific choice of $\mathcal{B} = 2.0$ was driven by experimentation and is heuristically motivated. The conceptual explanation in terms of the Chebyshev relation is that we "tighten" the timeout boundary by using a smaller multiplier for σ.

Jacobson essentially computes the RTO boundary as "a few sigmas past the mean" where both μ and σ are estimated using IIR filters. We essentially compute a similar RTO boundary as "a few sigmas past the mean" where the estimator of the mean is the actual sequence value plus a small bias (this is a Markov-like formulation), and the estimator of σ is the same as in Jacobson. However, because of the relatively "more accurate" estimate of the mean in the Modified Algorithm (it has no delay), and because the mean estimator in the Modified Case is already biased ($\mathcal{A} = 1.25$ rather than 1.0), we essentially "back off" the multiplier of σ to achieve a similar overall formulation.

Note that the bias term $\mathcal{A}$ for $\hat{\mu}_{n+1|n}$ in (7) is *not equivalent* to the use of gain in the prediction of $\hat{\mu}_{n+1|n}$ in (4). The structure of the prediction filter for $\hat{\mu}_{n+1|n}$ causes delay in the formulation of the overbound, which is problematic. There are no coefficients for the prediction filter which will simultaneously improve delay and maintain stability in the estimation of μ, and incorporating gain in the prediction does not improve the estimate. These undesirable effects are completely eliminated in the Modified approach with the use of x_n as the estimator of μ. This adjustment allows for the use of a Markov-like bias term $\mathcal{A}$ and significantly enhances the performance of the Modified Algorithm.

Several factors must be specifically noted for the Modified Algorithm. First, dependence on the variance of the RTT sequence is preserved via $\hat{\sigma}_{n+1|n}$ and the Chebyshev-like formulation. Some dependence on σ must be maintained in the estimation procedure for cases where the RTT values exhibit significant variability. However, the multiplier $\mathcal{B}$ can be different (smaller) than in Jacobson. This reduced dependence on σ mediates undesirable "overshoot" which is problematic in Jacobson, and has been addressed heuristically in Eifel. Refer to abrupt changes in RTT as shown in Figure 1 for examples.

Secondly, dependence on the mean of the RTT sequence is preserved via the use of x_n for $\hat{\mu}_{n+1|n}$, and a bias is incorporated via the Markov-like formulation for cases where $\sigma \rightarrow 0$. Some dependence on μ is important, since this allows isolation of the variability. However, in cases where $\sigma \rightarrow 0$, Jacobson tends to "settle" directly onto RTT, leading to heuristic modifications including static minimum values which override the Jacobson estimates. This undesirable behavior of Jacobson is clearly evident in Figure 1.

Thirdly, explicit dependence on both μ and σ is retained via the hybrid Markov/Chebyshev formulation which *biases* the estimate higher and reduces the need for secondary minimum computations. Also, the prediction structure for $\hat{\mu}$ and $\hat{\sigma}$ is preserved, which is an important consideration.

Finally, the computational complexity of the Modified Algorithm is essentially the same as the original Jacobson Algorithm. Elimination of the prediction structure for $\hat{\mu}$ and the use of a bias term along with a simplified multiplier for $\hat{\sigma}$ results in an algorithm with the same operational complexity and no heuristic conditional logic steps.

Besides, the Modified Algorithm is designed so that two important performance enhancement mechanisms of current TCP/SCTP implementation will not be adversely affected or impacted.

V. Performance results

To validate algorithm performance, we constructed a "real-world" test environment which pairs client and server computers with modified network stacks via a controllable network infrastructure. In the network testbed, a client

system transmits data to a server system using a specially-constructed user-space application which can vary overall payload length and SCTP chunk size, as well as other parameters such as the number of test iterations. The network stack of the client systems also implement user-selectable timeout estimation algorithms and record important parameters for each transmission. Parameters recorded by the client's network stack include the timestamped, per-chunk values of the actual (measured) RTT, estimated RTO, $\hat{\mu}$, $\hat{\sigma}$, and so on. As described in Section II-C, post-processing of the RTO estimate to enforce minimum values for RTO is disabled since we are investigating the effect of estimation algorithms.

Additionally, the network devices and SCTP server are modified to introduce algorithmically controllable delays in acknowledgments of chunks and delays in delivery of various classes of network traffic. This feature results in an ability to introduce specific "delay profiles" which duplicate other known results (as in Figure 2 [7]) or randomize the round-trip time of the network. Using trace data collected directly from the network stacks of the client & server computers, we were also able to create simulations of system performance which have been cross-checked for accuracy against the delay and estimation performance of the actual systems. All performance data described in this paper was produced using our "real-world" testbed, and has been incorporated into accurate simulations of the estimation algorithms.

The performance of the Modified Jacobson Algorithm is shown in the context of various RTT sequence characteristics or "delay profiles" in Figure 2, Figure 3, and Figure 4. Figure 2 reproduces an important delay profile from literature describing the Eifel algorithm (Fig. 6 of [7]). Figure 3 contains the delay profile for a real, quiescent network with 100 msec average delay and short, artificially induced delay spikes. Figure 4 contains the delay profile for a long-term delay burst on an otherwise quiescent network with 80 msec average delay.

Quantitative assessment of algorithm performance gathered from a large number of packet transmissions is shown in Table I and Table II for each of the delay profiles described by Figure 2, Figure 3, and Figure 4. The data in Table I is presented in terms of Mean Absolute Error (MAE) in milliseconds between the RTO estimate and the actual RTT value for the same packet transmission, according to (8). Table II shows the number of induced timeout events (TO) for each estimation algorithm, and is measured in number of events per 10,000 packets transmitted.

$$\text{MAE} = \frac{1}{N} \sum_{N} |x_{thr} - x_n| \qquad (8)$$

Using our network testbed, we duplicated the RTT sequence in Figure 2 from [7]. This "delay profile" clearly shows the improvement of Eifel over Jacobson, particularly at the termination of the "ramp" sequences. Note

Table I
ESTIMATION PERFORMANCE GATHERED USING NETWORK TESTBED AND SIMULATION. PERFORMANCE IS MEASURED AS MEAN ABSOLUTE ERROR (MAE) IN MILLISECONDS.

Delay profile	Figure	Jacobson	Eifel	Modified
Eifel ramp	Figure 2	1731	2091	1577
Quiet/Spikes	Figure 3	39.81	140.7	43.29
Delay burst	Figure 4	11.31	41.79	2.00

Table II
ESTIMATION PERFORMANCE GATHERED USING NETWORK TESTBED AND SIMULATION. PERFORMANCE IS MEASURED IN TERMS OF TIMEOUT EVENTS (TO) PER 10,000 PACKETS TRANSMITTED.

Delay profile	Figure	Jacobson	Eifel	Modified
Eifel ramp	Figure 2	1	1	1
Quiet/Spikes	Figure 3	99	51	79
Delay burst	Figure 4	378	82	12

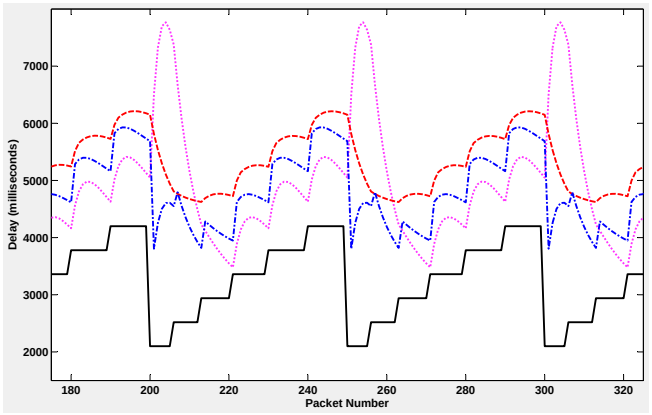


Figure 2. Performance of RTO estimators for the "ramp" RTT sequence described in [7]. Jacobson (dotted), Eifel (dashed), and Modified (dash-dot). The common RTT sequence driving all algorithms (solid).

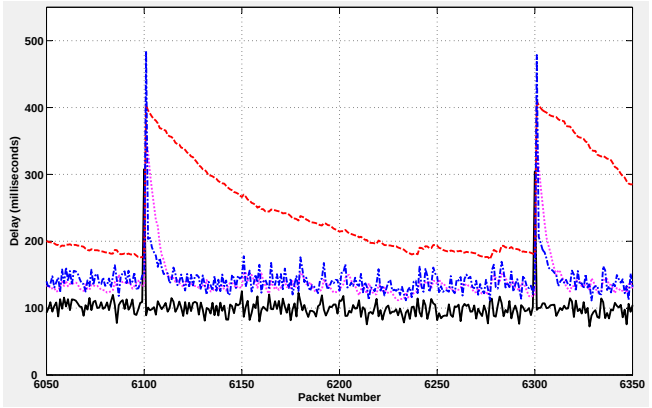


Figure 3. Performance of RTO estimators for a quiescent network with short, artificially induced delay spikes. Jacobson (dotted), Eifel (dashed), and Modified (dash-dot). The common RTT sequence driving all estimation algorithms (solid).

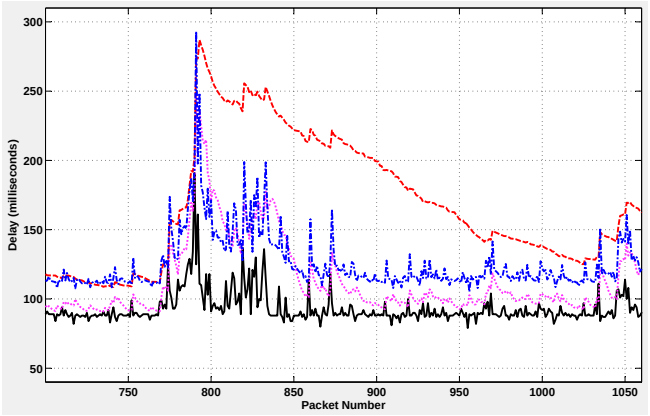


Figure 4. Performance of RTO estimators for a quiescent network with a large, naturally-occurring multi-packet delay burst. Jacobson (dotted), Eifel (dashed), and Modified (dash-dot). The common RTT sequence driving all algorithms (solid).

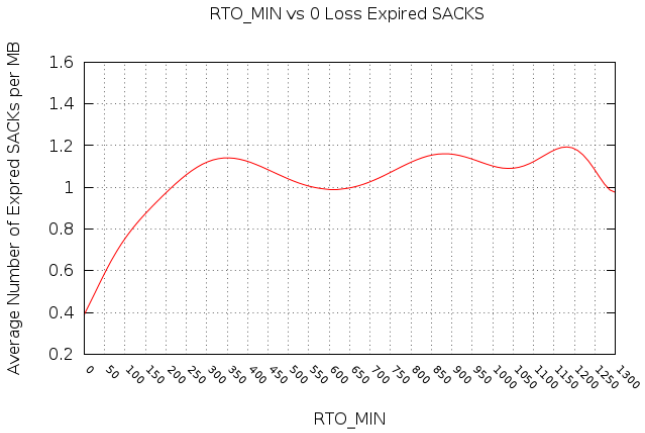


Figure 5. Relationship between RTO-min and expired SACKs

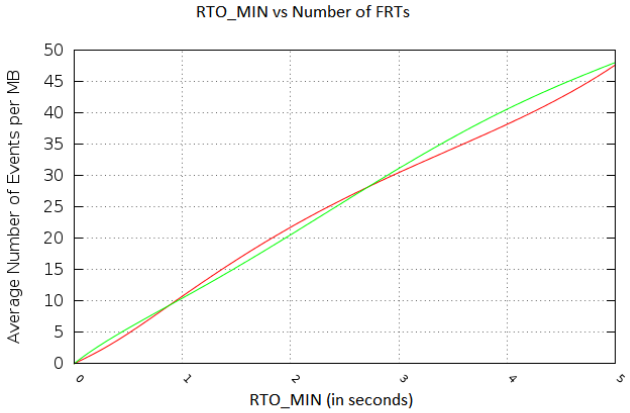


Figure 6. Relationship between RTO-min and FRTs

that Jacobson overshoots *upward* even though the RTT sequence has rapidly declining values. Eifel compensates for overshoot, but at the expense of slightly higher bias from the RTT sequence. Note that the Modified approach mimics Jacobson during ramp ascension, and also compensates for the overshoot at ramp termination. Modified has a smaller bias than Eifel. Interestingly, each algorithm induces a single timeout event, but the Modified Jacobson algorithm does so with an MAE 154 msec smaller than Jacobson, and 514 msec smaller than Eifel. Such a significant difference in MAE can translate to a smaller task-completion time in cases where a small differential in timeouts is encountered.

A delay profile from a relatively quiescent network with artificially induced delay spikes is shown in Figure 3. The network exhibits an average delay of around 100msec, and has artificially induced delay spikes which are typical of an unstable link. Note that with this network profile, delay spikes occur every 200 packets, forcing a timeout. After each delay spike, the estimation algorithms recover in very different manners: Modified and Jacobson fall quickly toward the quiescent RTT sequence, while Eifel decays very slowly, creating a relatively large wait-time for over 100 subsequent packets. As a result, the MAE for the Eifel algorithm with this delay profile is more than 3 times larger than the other estimation algorithms while still creating only 35% fewer timeout events. Regardless, each estimation algorithm creates a very small proportion of timeout events relative to the number of packets transmitted.

A delay profile from another network test is shown in Figure 4. In this figure, the network exhibits an average delay during quiescent periods of around 80 msec. However, between packets 750 and 850 a large, naturally occurring, correlated delay burst is observed, which disrupts the estimation algorithms. Note that Modified and Jacobson both fall quickly after individual, large delay spikes. However, Modified maintains a larger offset during the quiescent period between packets 900 and 1000 due to the specific bias for $\hat{\mu}$. The failure of Jacobson to maintain a bias during periods of low RTT variance ($\sigma \rightarrow 0$) is responsible for many timeout events, with Jacobson inducing *30 times more timeouts* than Modified, and *almost 5 times more* than Eifel. Eifel again exhibits a bias which is significantly larger than Jacobson or Modified, with an MAE *20 times larger* than Modified, and *4 times larger* than Jacobson. Also note that Jacobson tracks RTT fairly well, but has large, positive overshoot when RTT drops suddenly. Modified compensates for the Jacobson "overshoot" problem in cases where the RTT sequence drops suddenly (cf. packet 800 & 840).

VI. Impacts on SACK and FRT

As discussed in Section III-C and Section III-D, SACK and FRT are two important mechanisms to enhance the performance of TCP [21]. Naturally, an important question

of any revision to the current SCTP timeout mechanism is how the revision will impact SACK and FRT. The merits of any proposed revision would be questionable if the revision significantly decreases SACK or incurs unnecessary FRT.

Figure 5 shows the relationship between the number of bytes transmitted and the minimum RTO values (in millisecond). It is clear to see that the number of expired SACKs is only significantly impacted when the minimum RTO value is smaller than 200 msec. Otherwise, the minimum RTO has no direct impact on number of SACKs.

As discussed in Section III-C, SACK is proposed to enhance TCP performance over *long delay paths* [19], [23]. Our algorithm will set mininum RTO to values smaller than 200 msec when it is confident that the values of the current and most recent RTT are very small (than 200 msec). Naturally in that case the TCP communication sessions are not on *long delay paths*. Hence, it is understandable that the number of SACKs decreases.

The impact of our algorithm on FRT is less intuitive. Figure 6 shows that the number of fast retransmits is almost unaffected by the RTO algorithms. In the diagram, the X-axis is the percentage of packet losses, which were created in our simulation in order to force timeout and fast retransmits, and the chunk size is 50 bytes. The Y-axis is the number FRT events per MByte of data transmitted. The green line is the FRT for our algorithm and the red line is for the current RTO algorithm.

Recall that FRT is based on the heuristic that a TCP/SCTP segment must be lost and hence should be retransmitted because another segment sent before that segment has been acknowledged multiple times. This fast retransmission occurs before the timer for that segment expires. Smaller RTT values means that the network is more reliable and it will in turn set smaller minimum RTO values. Similarly larger RTT values implies a less reliable network, which in turn will set larger minimum RTO values. Therefore, regardless of algorithms used, the number of FRTs is directly proportional to the number of timer expirations, which is illustrated in Figure 6.

The current minimum RTO algorithm (Jacbson's Algorithm) starts by setting a 1000 msec minimum RTO value. On the other hand, our algorithm sets minimum RTO values dynamically according its own heuristics. If it sets minimum RTO values smaller than the mininum RTO value of current algorithm, then more timeouts will occur. But it will not incur more FRTs.

More details about the relationships of proposed algorithm and SACKS/FRTs will be presented in a forthcoming writing.

VII. Conclusion and Future Work

The theme of this paper is that the timeout aspects of current TCP and SCTP protocols can no longer reflect the current infrastructure and traffic dynamics of today's Internet. Current and future research should work on new and practical timeout mechanisms that meet the need of changed/changing landscapes of Internet.

This paper analyzes the methods for computing and using RTT and RTO estimates in IP-based transport protocols such as SCTP and TCP. The theoretical basis of the Jacobson Algorithm is discussed and an alternative approach is presented, which retains the fundamentally sound theoretical basis and operational structure of the algorithm, but improves the performance over other well-known techniques without introducing heuristic modifications. It is shown that our alternative approach will not adversely affect the SACK and FRT mechanisms.

Currently, we are implementing an RTO-min optimization algorithm based on the limited past history of RTO values. The implementation is done by modifying the SCTP modules of Linux kernels. The new algorithm is being tested over the Internet under various traffic dynamics.

Future work involves the continued optimization of the modified algorithm as well as investigation into the effects of variable minimum bound for the RTO timer. We believe that the work that we have done is innovative and solid. However, an implementation and testing of our algorithm over a sizable large real network will be helpful to pinpoint the strength and weakness of the algorithm.

Acknowledgement

The authors would like to express appreciation to Mr.Eduardo Gonzalez for providing data used in Section VI.

References

[1] S. McClellan and W. Peng, "Estimating retransmission timeouts in IP-based transport protocols," in *Proc. of ICDT 2013*, Apr. 2013, pp. 26–31.

[2] J. Postel, "Transmission control protocol," RFC 793, Sep. 1981.

[3] R. Stewart *et al.*, "Stream control transmission protocol," RFC 2960, Oct. 2000.

[4] R. Stewart, "Stream control transmission protocol," RFC 4960, Sep. 2007.

[5] V. Jacobson, "Congestion avoidance and control," *in Proc. ACM Comput. Commun. Review* (SIGCOMM'88),vol. 18, Aug. 1988, pp. 314–329.

[6] V. Paxson, M. Allman, J. Chu, and M. Sargent, "Computing TCP's retransmission timer," RFC 6298, June 2011.

[7] R. Ludwig and K. Sklower, "The Eifel retransmission timer," *in Proc. ACM Comput. Commun. Review* (SIGCOMM'00), vol. 30, July 2000, pp. 17–27.

[8] H. Ekstrom and R. Ludwig, "The Peak-Hopper: A new end-to-end retransmission timer for reliable unicast transport," in *Proc. 23rd Joint Conf. of the IEEE Computer & Communications Societies (INFOCOM 2004)*, vol. 4, Nov. 2004, pp. 2502–2513.

[9] L. Ma, G. Arce, and K. Barner, "TCP retransmission timeout algorithm using weighted medians," *IEEE Sig. Proc. Letters*, vol. 11, pp. 569–572, June 2004.

[10] J. Pedersen, C. Griwodz, and P. Halvorsen, "Considerations of SCTP retransmission delays for thin streams," in *Proc. 31st IEEE Conf. on Local Computer Networks*, Tampa, FL, Nov. 2006, pp. 135–142.

[11] A. Petlund, P. Beskow, J. Pedersen, E. S. Paaby, C. Griwodz, and P. Halvorsen, "Improving SCTP retransmission delays for time-dependent thin streams," *Multimedia Tools and Applications*, vol. 45, pp. 33–60, Oct. 2009.

[12] M. Allman, K. Avrachenkov, U. Ayesta, J. Blanton, and P. Hurtig, "Early retransmit for TCP and stream control transmission protocol (SCTP)," RFC 5827, Apr. 2010.

[13] M. Allman and V. Paxson, "On estimating end-to-end network path properties," in *Proc. Conf. Appl., technol., arch., and protocols for comput. commun.*, ser. SIGCOMM '99. New York, NY, USA: ACM, 1999, pp. 263–274.

[14] L. Coene and J. Pastor-Balbas, "Telephony signaling transport over stream control transmission protocol (SCTP) applicability statement," RFC 4166, Feb. 2006.

[15] I. Psaras and V. Tsaoussidis, "Why TCP timers (still) don't work well," *Computer Networks*, vol. 51, pp. 2033–2048, Nov. 2007.

[16] J. Pedersen, "Evaluation of SCTP retransmission delays," Master's thesis, University of Oslo Department of Informatics, May 2006.

[17] S. Khatri, "SCTP performance improvement based on adaptive retransmission time-out adjustment," Master's thesis, Texas State University Department of Computer Science, Aug. 2011.

[18] A. Papoulis, *Probability, Random Variables, and Stochastic Processes*, 3rd ed. New York, NY: McGraw-Hill, 1991.

[19] V. Jacobson and R. Braden, "TCP extensions for long-delay paths," RFC 1072, Oct. 1988.

[20] L. Eggert, "Moving the undeployed TCP extensions RFC 1072, RFC 1106, RFC 1110, RFC 1145, RFC 1146, RFC 1379, RFC 1644, and RFC 1693 to historic status," RFC 6247, May 2011.

[21] W. R. Stevens, *TCP/IP Illustrated, Volume 1: The Protocols*, 1st ed. Boston, MA: Addison-Wesley, 1994.

[22] V. Jacobson, R. Braden, and D. Borman, "TCP extensions for high performance," RFC 1323, May 1992.

[23] M. Mathis, J. Mahdavi, S. Floyd, and A. Romanow, "TCP selective acknowledgement options," RFC 2018, Oct. 1996.

[24] R. Bruyeron, B. Hemon, and L. Zhang, "Experimentations with TCP selective acknowledgment," in *Proc. Comput. Commun. Review*, vol. 28. ACM, Apr. 1998, pp. 54–77.

[25] W. Stevens, "TCP slow start, congestion avoidance, fast retransmit, and fast recovery algorithms," RFC 2001, Jan. 1997.

[26] M. Allman, V. Paxson, and W. Stevens, "TCP congestion control," RFC 2581, Apr. 1999.

[27] M. Allman, S. Floyd, and C. Partridge, "Increasing TCP's initial window," RFC 3390, Oct. 2002.

[28] M. Allman, V. Paxson, and E. Blanton, "TCP congestion control," RFC 5681, Sep. 2009.

Depth-Included Curvature Inpainting for Disocclusion Filling in View Synthesis

Suryanarayana M. Muddala, Roger Olsson, and Mårten Sjöström

Dept. of Information and Communication Systems
Mid Sweden University
Sundsvall, Sweden 85170
Email: marten.sjostrom@miun.se

***Abstract*—Depth-image-based-rendering (DIBR) is the commonly used for generating additional views for 3DTV and FTV using 3D video formats such as video plus depth (V+D) and multiview-video-plus-depth (MVD). The synthesized views suffer from artifacts mainly with disocclusions when DIBR is used. Depth-based inpainting methods can solve these problems plausibly. In this paper, we analyze the influence of the depth information at various steps of the depth-included curvature inpainting method. The depth-based inpainting method relies on the depth information at every step of the inpainting process: boundary extraction for missing areas, data term computation for structure propagation and in the patch matching to find best data. The importance of depth at each step is evaluated using objective metrics and visual comparison. Our evaluation demonstrates that depth information in each step plays a key role. Moreover, to what degree depth can be used in each step of the inpainting process depends on the depth distribution.**

***Keywords*—3D; video plus depth; multiview video plus depth; 3D warping; depth-image-based rendering; image inpainting; disocclusion filling.**

I. INTRODUCTION

In recent years, Three Dimensional Television (3DTV) and Free Viewpoint Television (FTV) have become attractive in the 3D research area. The 3D video formats video-plus-depth (V+D) and multiview video-plus-depth (MVD) are efficient ways to send the 3D content to the end user. These data types provide backward compatibility with 2D displays, and at the same time enable rendering of virtual views corresponding to the requirement of current and future stereo and multiview 3D displays. Depth-image-based rendering (DIBR) is a fundamental method for producing these virtual views using the video or texture component and the depth per pixel information. A major concern when using DIBR on V+D content is the artifacts caused by disocclusions, i.e., areas that are occluded in the original view become visible in the rendered view when DIBR is used. When using MVD, DIBR methods can access multiple V+D components and thereby fill disocclusions in virtual views from several known views, which reduce the disocclusion problem yet does not solve it. Hence, the disocclusion problem inherent in the above 3D video formats is still an open question that needs to be addressed in order for applications such as 3DTV and FTV to be successful. The present work considered horizontal disparity for creating images on stereo or multiview displays. The depth-based inpainting [1] method does not explain the importance of depth at every step of the inpainting process, which raises the question how the depth information influence the total inpainting process in addressing the disocclusions.

There are several methods that aim to reduce the disocclusions by filling the holes they constitute, including preprocessing of depth map [2], linear interpolation, and inpainting. Preprocessing of the depth map before warping reduces the disocclusion but causes geometrical distortions in the resultant view. Using linear interpolation is a simple approach that gives acceptable results when the area of the disocclusion is smaller and less noticeable. However, when the disocclusions grow bigger, linear interpolation results a stark contrast to the structure of the texture outside the holes borders due to stretching of the border pixels into the holes. In general, inpainting methods aim to solve the missing areas by filling the unknown regions using neighborhood information. The disocclusion problems can be considered as missing texture information alone and it can be recovered by using texture synthesis methods [3]. Due to inpainting being an ill-posed problem the target of inpainting methods is to produce filled disocclusions that are as perceptually plausible as possible. In the following we will classify inpainting methods into the following categories: textural inpainting methods and structural inpaiting methods.

In texture inpainting methods, the missing regions are filled by replicating the repetitive patterns in the image, which surrounds the disocclusion. Many of these methods rely on Markov Random Field (MRF) to model the local patterns, and using small amount of known textures to generate new textures to fill the missing areas [4]. Texture inpainting methods perform well for uniform texture images but faces problems with real world images as they consist of a mixture of different textures and linear structures.

Structural inpainting methods, imitate the idea of manual inpainting that is propagating the linear structures present in the neighborhood of the missing regions using diffusion process [5]. Structural inpainting methods based on partial differential equations (PDE) fill the missing regions iteratively. These methods preserve and propagate linear structures in to missing regions. The disadvantages associated with these methods are blurring and structure discontinuity for large missing areas due to the diffusion process and not knowing the edge information. The total variation approach and curvature driven diffusion (CDD) model follow the anisotropic

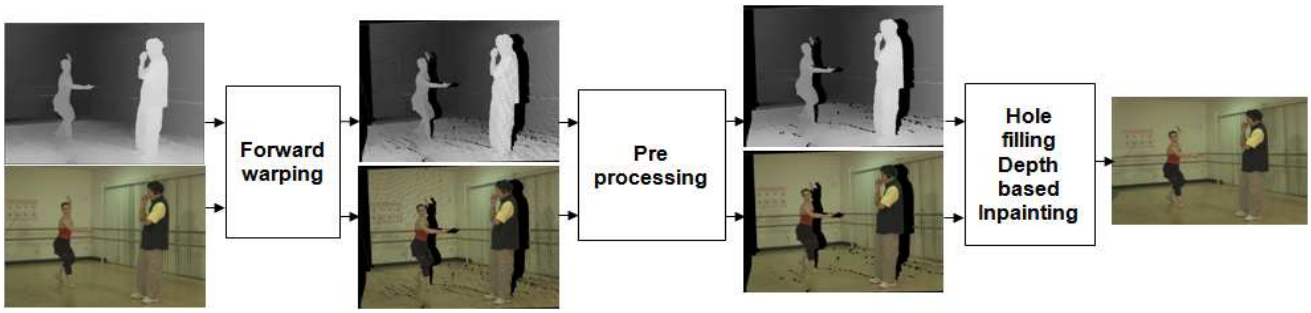


Figure 1. Block diagram of the rendering method using inpainting.

diffusion, which means diffusion strength reduces with respect to the structure variations [6] [7]. Total variation approach depends only on isotropic strength, whereas CDD depends on both isotropic strength and geometries. Although incorporating many constraints and higher order PDEs in the inpainting process, the resulted inpainted views still suffer from blurring artifacts for larger missing regions.

Criminisi et al. [8] proposed an efficient exemplar-based image inpainting technique, a kind of hybrid inpainting that combines the benefits of structural and textural inpainting methods to fill the missing regions. However, this method introduces inconsistent structures in to disocclusions areas due to not considering what is foreground (objects closer to the camera) and background (objects away from the camera) in the rendered view. Daribo et al. [9] extended the exemplar-based inpainting to address this limitation by introducing the depth constraint. However, hole filling with this method causes noticeable artifacts around foreground objects in the virtual view. Gautier et al. [10] extended the Criminisi et al. method by taking into account the structure tensor as a data term that identifies the strongest structure in the neighborhood, and the depth information to fill the holes. Moreover, filling with these methods require true depth information. Usually, the depth at the required camera view position needs to be estimated. By having access to the true depth at the virtual camera view position disocclusions can be filled according to depth level.

We proposed depth-included curvature inpainting method that consider the scene depth in all parts of the inpainting process [1]. This paper presents the method in more detail and also performs an analysis of how the depth information affects the quality of the inpainting result. The proposed method also relies on the fundamental method introduced in [8], but extends and improves upon it by using the available depth information at different stages of the inpainting process. In contrast to [9], [10], the proposed depth-included curvature inpainting method does not rely on having access to a true depth map but instead considers a more general and realistic case of having access to a warped depth when performing the inpainting.

The outline of the paper is as follows: A more general review of depth-image-based rendering is given in Section II. A selected summary of related work within the field of inpainting is presented in Section III and the proposed depth-included curvature inpainting method is described in Section IV. The methodology used in this work, including test arrangement and evaluation criteria, are described in Section V, followed by results and in Section VI. Finally, Section VII concludes the work.

II. Depth-image based rendering

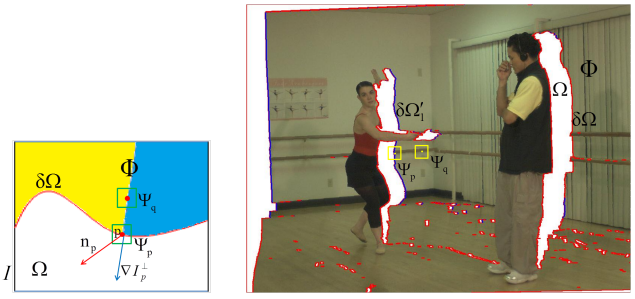
With the development of many rendering algorithms, the rendering methods are categorized into image based rendering and model based rendering methods. Image based rendering methods, use the available image data to render new views whereas model based methods require information about 3D models to compute synthesized views. Image based rendering methods can be further classified depending on the amount of geometry used. DIBR is an approximation of rendering with partial geometry, i.e., depth data. The depth data can be obtained from multiview images or range sensors [11], [12]. Using this depth data and texture, new views can be synthesized using basic principle 3D warping [13]. The warping can be done in two ways either forward warping or backward warping. In forward warping both depth and texture are warped to the virtual view position, whereas in backward warping first the depth map is warped and then pixel locations of the virtual view are located in the original image. The inherent problems with the DIBR are cracks, translucent cracks, missing areas and corona like artifacts when views are synthesized from multi view format. The causes of artifacts and different solutions are presented in [14].

In this work, we present the total DIBR using image inpainting technique to fill the disocclusion problems. Initially, the texture and depth map are warped to the virtual view positions and then the crack filling is applied by using averages of neighboring pixels. In the next step, the ghosting-like artifacts are removed on the borders. They appear as a thin layer of foreground which is projected on the background at the boundary of the disocclusion, due to depth and texture misalignments. These ghosting problems can be removed by simply extending the hole area on the background side. Otherwise, the filling will be affected due to the mixed colour pixels presented at the boundary area and the filling depends on the boundary data. Thus these problems need to be addressed before starting the inpainting process. The total DIBR method is illustrated in the Fig. 1.

III. Related work

The exemplar-based texture synthesis introduced by Criminisi et al. effectively replicates both structure and texture by using the advantages of both partial differential equations (PDE) based inpainting method and non-parametric texture

synthesis. The quality of the inpainted image highly depends on the order of the filling direction.



(a) Notation diagram (b) Warped image with notations

Figure 2. Schematic illustration.

Consider an input image I with an empty region Ω, also known as hole, the source region Φ (the remaining part of the image except the empty region) is defined as $\Phi = I - \Omega$. The boundary between Φ and Ω is denoted as $\delta\Omega$ (see Fig. 2). The basic steps of Criminisi's algorithm are as follows: (i) Identify the boundary and compute the priorities on the boundary region and (ii) Find the patch with the maximum priority and find the best patch that matches the selected patch using patch matching and filling (iii) Update the confidence values. Suppose a patch $\Psi_{\mathbf{p}}$ centered at a pixel $\mathbf{p}$ for some $\mathbf{p} \in \delta\Omega$ and the priority is computed as the product of two terms:

$$P(\mathbf{p}) = C(\mathbf{p}) \cdot D(\mathbf{p}), \tag{1}$$

$$C(\mathbf{p}) = \frac{1}{|\Psi_{\mathbf{p}}|} \sum_{\mathbf{q} \in \Psi_{\mathbf{p}} \cap \Phi} C(\mathbf{q}), \tag{2}$$

$$D(\mathbf{p}) = \frac{\left\langle \nabla I_{\mathbf{p}}^{\perp}, \mathbf{n}_{\mathbf{p}} \right\rangle}{\alpha}, \tag{3}$$

where $C(\mathbf{p})$ is the confidence term indicating the amount of non-missing pixels in a patch and the data term $D(\mathbf{p})$ gives importance to the isophote direction. $|\Psi_{\mathbf{p}}|$ is number of pixels in $\Psi_{\mathbf{p}}$, α is normalization factor (e.g., 255 for gray scale image), $\mathbf{n}_{\mathbf{p}}$ is a unit vector orthogonal to $\delta\Omega$ at a point $\mathbf{p}$, and $\nabla I_{\mathbf{p}}^{\perp}$ is the direction and of the isophote.

After the priorities on boundary $\delta\Omega$ are computed, the highest priority patch $\Psi_{\hat{\mathbf{p}}}$ centered at $\hat{\mathbf{p}}$ is selected to be filled first. Next, a block matching algorithm is used to find the best similar source patch $\Psi_{\hat{\mathbf{q}}}$ in order to fill-in the missing pixels in the target patch:

$$\Psi_{\hat{\mathbf{q}}} = \arg \min_{\Psi_{\mathbf{q}} \in \Phi} \left\{ d(\Psi_{\hat{\mathbf{p}}}, \Psi_{\mathbf{q}}) \right\}, \tag{4}$$

where d is the distance between two patches defined as sum of squared difference (SSD). After the most similar source patch $\Psi_{\hat{\mathbf{q}}}$ is found, the values of the missing pixels in the target patch $\acute{\mathbf{p}}|\acute{\mathbf{p}} \in \Psi_{\hat{\mathbf{p}}} \cap \Omega$ are copied from their corresponding pixels inside source patch $\Psi_{\hat{\mathbf{q}}}$. Once the target patch $\Psi_{\hat{\mathbf{p}}}$ is filled, the update of the confidence term $C(\mathbf{p})$ is as follows:

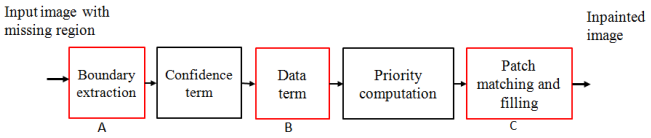


Figure 3. Block diagram of the exemplar-inpainting method; The highlighted blocks indicates the depth-included curvature inpainting method enhancements.

$$C(\mathbf{q}) = C(\hat{\mathbf{p}}), \forall \mathbf{q} \in \Psi_{\hat{\mathbf{p}}} \cap \Omega. \tag{5}$$

However, this method is not aimed at 3DV formats to handle disocclusions and thereby could not recognize the differences between foreground and background in a virtual view. As a result, missing areas are sometimes filled with foreground information instead of background.

Daribo et al. extended the Criminisi et al. method by using depth information, first by introducing a depth regularity term in the priority term calculation $P(\mathbf{p}) = C(\mathbf{p}) \cdot D(\mathbf{p}) \cdot L(\mathbf{p})$ and is given as:

$$L(\mathbf{p}) = \frac{|Z_{\mathbf{p}}|}{|Z_{\mathbf{p}}| + \sum_{\mathbf{q} \in, Z_{\mathbf{p}} \cap \Phi} \left(Z_{\mathbf{p}}(\mathbf{q}) - \overline{Z_{\mathbf{p}}} \right)^2}, \tag{6}$$

The depth regularity term $L(\mathbf{p})$ is approximated as the inverse variance of the depth patch $Z_{\mathbf{p}}$ centered at $\mathbf{p}$. The depth regularity term presented in their inpainting method controls the inpainting process by favoring the filling order that comes from the background. In addition, the patch matching step is modified by adding the depth into the search process to find a best patch in both the texture and the depth domain. Although this method uses the depth in the inpainting process reduces the problem to a degree as it still partly fills the disocclusion regions with the foreground information and wrong textures.

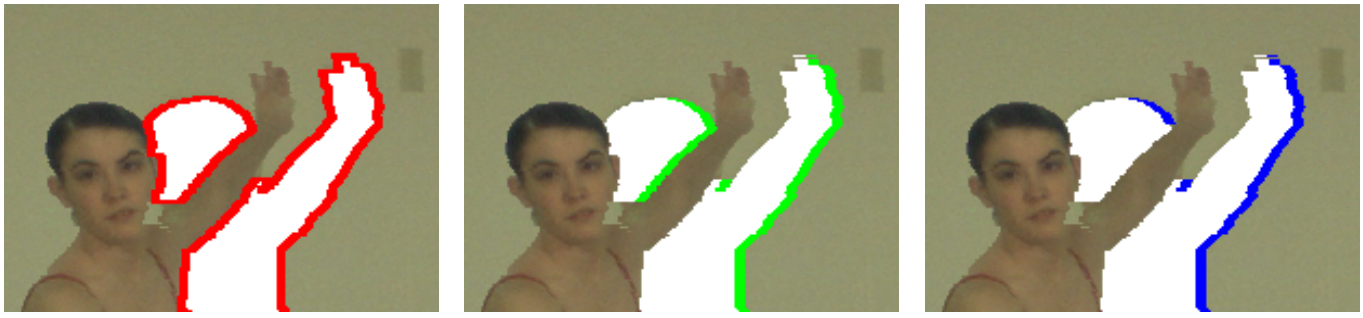
Gautier et al. followed the [9] method in considering depth map and extending the exemplar approach to help the inpainting process. They introduced a 3D tensor to calculate the data term in the priority calculation of (1) and a one-sided priority to restrict the filling direction. In the patch matching step, they also used a weighted combination of the best patches as the final selected patch.

$$J = \sum_{l=R,G,B,Z} \nabla I_l \nabla I_l^T, \tag{7}$$

$$D(\mathbf{p}) = a + (1-a) \exp\left(\frac{-C_1}{(\lambda_1 - \lambda_2)^2} \right), \tag{8}$$

where ∇I_l is local spatial gradient over a 3x3 window. J is the 3D structure tensor and λ_1,λ_2 are the eigen values of J, which gives amount of structure variation, C_1 is a constant positive value and $a \in [0, 1]$.

Moreover, these previous work both Daribo et al. and Gautier et al. rely on having true depth information available at the rendered view position. In general, this assumption is not feasible or realistic since the depth information of the virtual view also must be estimated.



(a) Total boundary (b) One sided boundary (c) Depth-guided one sided boundary

Figure 4. Boundary extraction.

IV. Proposed Depth-included curvature inpainting method

The depth-included curvature inpainting method also followed the Criminisi method by introducing the new data term and depth based constraints in all stages of the inpainting process. The depth information is added in the following steps of the depth-included curvature inpainting process:

A. Depth guided directional priority
B. Depth included curvature data term
C. Depth-based source region selection

Fig. 3 illustrates how these steps relate to the general inpainting process. Step A consists of defining a depth guided directional priority, which helps the selection of background patches to be filled first. In Step B, we used the Curvature Driven Diffusion (CDD) model [6] similarly to [15] as the data term $D(\mathbf{p})$ in the priority computation and extend the data term by incorporating depth to give the importance to the isophote curvature and strength. Finally, Step C prevents the foreground data from the source region, using depth constraints derived from the warped depth to favor the background filling. In addition, a weighted combination of - N best patches is used to define the target patch in the patch matching.

A. Depth guided direction priority

The boundary extraction block of Fig. 3 is improved by using depth information to guide the filling such that it starts from the background. This is because disocclusions result from depth discontinuities between foreground and background, and the disocclusion regions belong to the background which makes filling the disocclusion from the horizontal background side reasonable. Moreover, when the disocclusion appears between two foreground objects, selection of one sided boundary results foreground propagation into the holes. Hence, the one-sided boundary is constrained by using the depth, which controls the foreground scattering problem. The background side of the disocclusion is obtained as follows. First, a one sided boundary $\delta\Omega_1$ of the disocclusion area is obtained by applying the convolution operation $(*)$ on a disocclusion map (DM) as given $\delta\Omega_1 = DM * H$. Disocclusion map is a mask, which represents the hole regions with 1s and remaining regions with 0s so the convolution gives one sided edges. Second, the directional priority selection is constrained by using a depth on $\delta\Omega_1$, such that pixels whose depth values are less than M percent of the maximum depth value in the warped depth map are selected. The one-sided boundary and depth guided boundaries are shown in the Fig. 4.

$$\delta\Omega_1^{'} = \delta\Omega_1(\mathbf{q})|_{\mathbf{q}\in\delta\Omega_1\cap(Z(\mathbf{q})<M\cdot max(Z))}, \tag{9}$$

where $\delta\Omega_1^{'}$ is the depth guided boundary, Z is the depth map and $Z(\mathbf{q})$ is the depth value at pixel location $\mathbf{q}$. Convolution kernel H is defined as follows according to the warped view:

$$H = \begin{cases} [1 \quad -1 \quad 0] & \text{if left warped view ;} \\ [0 \quad -1 \quad 1] & \text{if right warped view.} \end{cases} \tag{10}$$

Once the hole boundary is obtained, using (9), priorities are calculated according to (1) utilizing the data term (14). Then the holes in the background regions are filled using the selected depth guided direction priority. However, the filling with the depth-guided directions handles holes to a certain depth level, the remaining holes are filled with one sided boundary priority. Moreover, when the virtual view camera is not horizontal, the holes do not appear according to the assumption that a right virtual view has holes on right side. In that case, hole filling is processed with the total boundary extraction.

B. Depth included curvature data term

As the data term in the general inpainting process we adopt, and add depth to, the CDD model in order to consider the depth curvature along with the texture. The CDD model is a variational approach, which solves the PDE to fill the missing regions using diffusion [6]. It mainly aims at satisfying the connectivity principle, i.e., broken lines should connect over missing areas. The CDD introduces the curvature along with the gradient strength to achieve that goal by using diffusion process.

Although the present approach is not a diffusion-based inpainting, the CDD model computes the structures information using the curvature as it gives the geometry of the isophote. Hence, we adopt and add the depth while calculating the data term to propagate the structure details into the missing areas.

$$g(s) = s^{\alpha}, s > 0, \alpha \geq 1 \tag{11}$$

(a) "Ballet" depth map (b) "Lovebird" depth map

Figure 5. Different depth distributions in the background.

$$k_{\mathbf{p}} = \nabla \cdot \left(\frac{\nabla I_{\mathbf{p}}}{|\nabla I_{\mathbf{p}}|} \right) \tag{12}$$

$$\frac{\partial I_{\mathbf{p}}}{\partial t} = \nabla \cdot \left(\frac{g(|k_{\mathbf{p}}|)}{|\nabla I_{\mathbf{p}}|} \nabla I_{\mathbf{p}} \right), \tag{13}$$

where $k_{\mathbf{p}}$ is the curvature of the isophote through some pixel $\mathbf{p}$, $\nabla\cdot$ is the divergence at $\mathbf{p}$, and g is the control function to adjust the curvature. By incorporating the CDD model as a data term and setting $\alpha = 1$ in (11), the data term becomes:

$$D(\mathbf{p}) = \left| \nabla \cdot \left(\frac{k_{\mathbf{p}}}{|\nabla I_{\mathbf{p}}|} \nabla I_{\mathbf{p}} \right) \right|. \tag{14}$$

The data term is calculated using the depth gradient. Therefore, less variation in depth reduces the influence of the data term. Fig. 5 shows two examples of depth variations, where Fig. 5(a) exhibits more changes in depth than Fig. 5(b).

C. Depth-based source region selection

The patch-matching step is an improvement to the method of [9] and [10], where they add depth information in the patch matching step to find the best texture according to the depth range, and moreover, they assume the original depth map is available. If the original depth information is available, hole filling will be straight forward to find the source patch according to the right depth level. The improvement to the reference methods consists of classifying the source region using warped depth information, in order to select similar patches from the nearest depth range. The source region is divided into background region and foreground regions using depth information in the target patch. By considering Φ to be the known source region, which contains both foreground and background regions, the best source patch selection from foreground region is avoided by sub-dividing Φ using depth threshold Z_c according to:

$$\Phi_b = \Phi - \Phi_f, \tag{15}$$

where Φ_f is the source region whose depth values are higher than the depth threshold Z_c.

The depth threshold has two different values selected adaptively from the variance of the known pixel values of the target depth patch. When the depth patch lies near to foreground (See Fig. 6(a)), the variance of the target depth patch is greater than a threshold γ, and the patch might contain unwanted foreground values (See Fig. 6(b)). The average value of the depth patch is then chosen instead as the depth threshold in order to deduct the foreground parts. Otherwise, the patch contains the uniform or continuous depth values, so the maximum value in the depth patch is used as the depth threshold in order to get the best patch according to the depth level. The depth threshold Z_c is defined as follows:

$$Z_c = \begin{cases} \overline{Z_{\hat{\mathbf{p}}}} & \text{if } var(Z_{\hat{\mathbf{p}}}(\mathbf{q})|_{\mathbf{q} \in \Psi_{\hat{\mathbf{p}}} \cap \Phi}) > \gamma; \\ \max(Z_{\hat{\mathbf{p}}}) & \text{otherwise.} \end{cases} \tag{16}$$

$\Psi_{\hat{\mathbf{p}}}$ is the highest priority patch, $Z_{\hat{\mathbf{p}}}$ is the depth patch centered at $\hat{\mathbf{p}}$; and $\overline{Z_{\hat{\mathbf{p}}}}$ is the average value of the depth patch. $Z_{\hat{\mathbf{p}}}(\mathbf{q})$ is the depth value at pixel $\mathbf{q}$ and γ is the depth variance threshold.

Once the highest priority patch $\Psi_{\hat{\mathbf{p}}}$ from the priority term and depth-based source region Φ_b defined in (15) are computed, the target patch is filled with the best N number of patches within the source region.

$$\Psi_{\hat{\mathbf{q}}} = \arg \min_{\Psi_{\mathbf{q}} \in \Phi_b} \left\{ d(\Psi_{\hat{\mathbf{p}}}, \Psi_{\mathbf{q}}) + \beta \cdot d(Z_{\hat{\mathbf{p}}}, Z_{\mathbf{q}}) \right\}, \tag{17}$$

where d is SSD, and β is a value to give weight to the depth as important as the texture. In contrast to the reference methods, the depth-included curvature method used the warped depth information in the inpainting process. In order to help the inpainting process using the depth information, the holes in the depth image should be filled simultaneously, or the hole free depth image should be available for depth-guided inpainting process. In depth-included curvature method, we considered that the depth map should be filled simultaneously along with the texture.

From the idea of [16], we used a weighted average of N patches from the patch matching to fill the missing information in the disocclusion. Weighted average minimizes the noise in the selected patch and helps the smooth continuation filling process. The depth information helps the inpainting process to fill with the background textures otherwise when the depth information in the patch-matching is not considered, we put the filling in risk, i.e., the holes being filled with foreground textures.

(a) Warped depth-information (b) Source region before separation (c) Depth-based background source region

Figure 6. Depth-based source region selection.

In [1], the data term is calculated iteratively after every target patch is filled because the new copied texture to fill the hole region is a combination of the N best patches. In addition, the source region is updated such that the filled area is also available as source region for the next target patch.

V. Test arrangement and evaluation criteria

A. Compared to related work

Initially, the depth-included curvature inpainting method results are evaluated by using objective measurements as well as using visual comparison. A set of 10 frames are selected from the three MVD sequences "Ballet", "Break dancers" and "Lovebird1" for objective evaluation. All three sequences have a spatial resolution of 1024x768 pixels. The former two sequences are captured with 8 cameras and a baseline of 300 mm and 400 mm respectively [12]. The latter sequence is captured with 12 cameras and a baseline of 35 mm [17]. The test sequences have different depth and texture characteristics that make them suitable for testing different disocclusion filling attributes of inpainting methods as the method relies on depth. The "Ballet" sequence has large depth discontinuities at different depth ranges, which results in large disocclusions at different depth levels. The "Break dancers" sequence has a large number of objects located in almost the same depth level and smaller holes due to gradual depth discontinuities. The "Lovebird1" sequence has complicated scene with complex texture, structured background and large depth discontinuities.

B. Sensitivity analysis

The sensitivity of the depth at various steps of the proposed inpainting process is analyzed by restricting the depth at those stages. The depth information is used in three steps of the proposed inpainting process. In the first step, sensitivity of the depth in boundary extraction (SZB) is analyzed by performing the inpainting method without considering the depth information, i.e., using one sided boundary. In the next step, sensitivity of depth in data term (SZD) is analyzed by computing the data term using R, G, and B channels without incorporating the depth information as an additional channel. In the last step, the depth is used in source region selection and in block matching. Sensitivity of depth in patch matching (SZP) is analyzed separately without using depth. The sensitivity of depth in source region selection (SZSR) and sensitivity of depth in block matching (SZBM) are analyzed by using the Φ as a source region and finding SSD for only R, G, and B channels. While measuring the sensitivity at one step, the depth is not changed in other steps.

All test sequences are used in a DIBR of V+D scenario with possibility of full reference evaluation, i.e., access to ground truth texture and available depth at the required camera view positions. For the first two sequences, camera view 4 is rendered from camera view 5 and compared with ground truth image at camera view 4. In the third "Lovebird1" sequence, camera view 4 is rendered from camera view 6 and compared with ground truth image at camera view 4. The virtual views are first rendered and small holes and ghosting artifacts are removed using preprocessing step (see Fig. 1). Thereafter, the processed warped views are used as inputs to the inpainting method presented in Section IV. Important parameter and values of the proposed inpainting method is given in [1]. The best exemplars are searched in the warped depth and texture images, where as in [9] and [10] methods exemplars are searched in the warped texture and original depth at the virtual camera position. The implementation of inpainting method [8] is available on [18]. We have also implemented the method [9] for the comparison. The following two objective evaluation metrics are considered for assessing the results: peak signal to noise ratio of the luminance component (Y-PSNR) and mean structural similarity index (MSSIM). Both of these metrics measure the image quality where as MSSIM mostly corresponds to the perceptual visual quality [19]. Both metrics are applied to the full image, although the disocclusion areas only corresponds to $3-15\%$ of the total number of pixels. This results in the quality change shown in the sensitivity analysis to only manifest in the fractional part of the metric values. Another approach that would emphasize the changes would be to evaluate disocclusion pixels only. However, that would produce results that are not comparable with previous work.

VI. Results and analysis

A. Compared to related work

The objective evaluation results from the related work and proposed methods are shown in Fig. 7. The PSNR and MSSIM graphs consistently demonstrate that the depth-included curvature inpainting method performs better than the

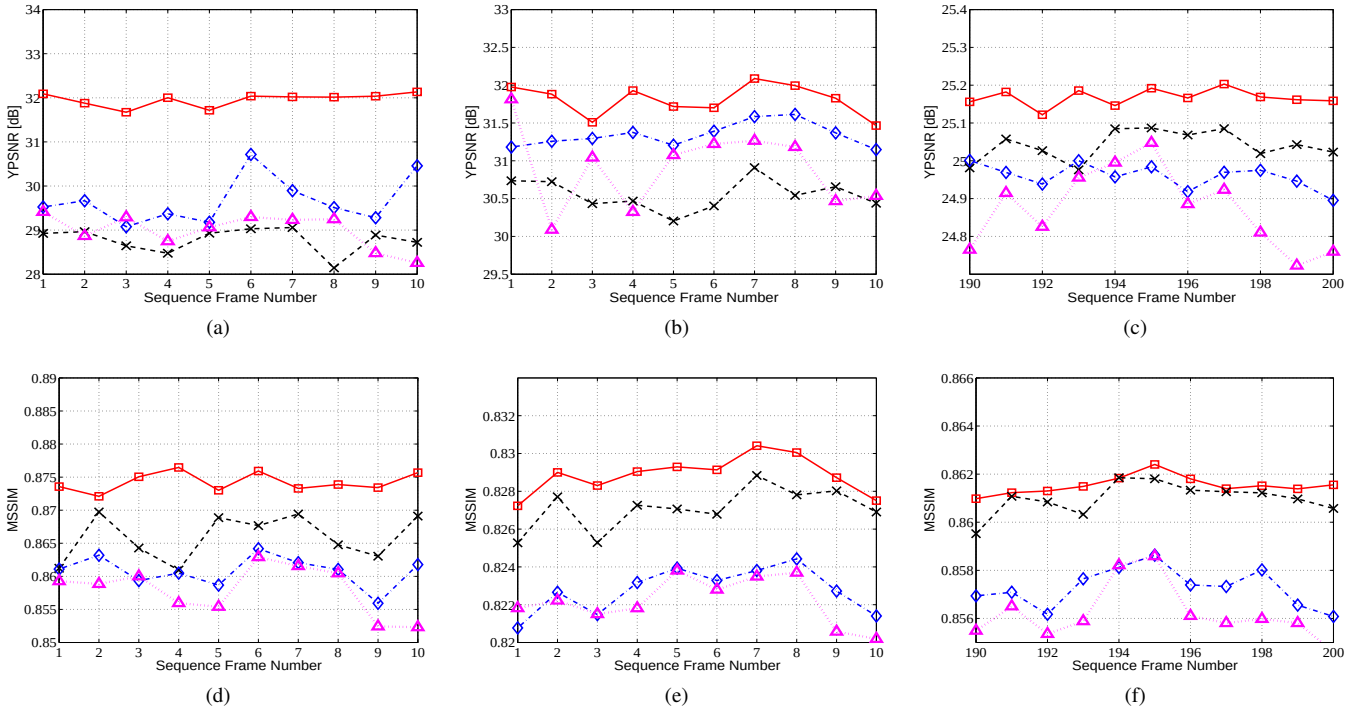


Figure 7. Compared to related work: Objective metrics PSNR and MSSIM of the investigated sequences; Proposed method (−□−), Gautier et al. method (− × −), Daribo et al. method (− · ◊ · −) and Criminisi method (· · △ · ·), PSNR for each rendered frame at view position 4 of "Ballet" (a), at view position 4 of "Break dancers" (b), at view position 4 of "Lovebird1" (c), MSSIM for each rendered frame at view position 4 of "Ballet" (d), at view position 4 of "Break dancers" (e) and at view position 4 of "Lovebird1" (f).

reference Criminisi, Daribo and Gautier methods. In addition to the objective results, Fig. 8 shows the synthesized views of the "Ballet" and "Lovebird1" images with the missing areas and inpainted images from different inpainting methods for visual assessment. Missing regions in Fig. 8(b) are filled with foreground information since no information about the depth is used to assist the filling process. Although the Daribo and Gautier methods are aided with true depth information, the missing areas are filled with the unwanted textures due to the lack of depth constraints and filling order.

The depth-included curvature inpainting method operates in a more realistic setting that only depends on warped depth and it shows visual improvements compared to the reference methods. The results from Fig. 8(e) show that the depth-included curvature inpainting method propagates the necessary neighboring information into the missing areas, by retaining both smooth areas (at the left side of the "Ballet" image) and propagating neighborhood structure (on the curtain in the "Ballet" image and at the head of the women in the "Lovebird1" image). The inpainting method might not reproduce the exact structure as in ground truth images due to the lack of knowledge about the scene contents, but it almost replicates the main structure.

B. Sensitivity analysis

The sensitivity of the depth in various stages of the depth-included curvature inpainting method is analyzed by using both the objective metrics and visual comparison. The results for the objective measurements are presented in Table I and Table II. Fig. 9 and Fig. 10 show the synthesized views of the "Ballet" and "Lovebird1" images with the missing areas and sensitivity of the depth in depth-included curvature inpainting method for visual assessment. The average PSNR and MSSIM values and visual comparison consistently demonstrate that the influence of the depth depends on the scene content and available depth information.

1) Sensitivity of depth in boundary extraction: The results from the sensitivity of depth in boundary extraction show that the depth information is important to handle the disocclusions, plausibly when they occur between foregrounds (see Fig. 9(b)). Although depth information is used in source region selection and block matching in order to fill holes from the background, still the holes are filled with the foreground due to the lack of knowledge about the depth on boundary, and as a result foreground boundary is selected in the boundary extraction. Thus adding the depth constraint on the selection of the boundary improves the hole filling process.

2) Sensitivity of depth in data term: The results from the sensitivity of depth in data term show that the depth information is less important for filling holes in the inpainted view (see Fig. 9(c) and Fig. 10(c)). However, the depth information gives the priority to structures when the depth contains several layers. For example, Fig. 5(a) contains several depth layers, whereas the other depth image does not contain less depth layers (see Fig. 5(b)). Moreover, the depth characteristics depend on the depth acquisition method. Thus, if there are some layers in the depth map enriching the depth information, it favors the inpainting process.

TABLE I. Sensitivity analysis: Average Y-PSNR.

Test sequence	SZB	SZD	SZSR	SZBM	Proposed
Ballet	31.91	31.88	31.76	**31.99**	31.96
Break Dancer	**31.84**	31.82	31.82	31.80	31.80
Love bird1	**25.16**	25.15	24.91	25.13	**25.16**

TABLE II. Sensitivity analysis: Average MSSIM.

Test sequence	SZB	SZD	SZSR	SZBM	Proposed
Ballet	0.8745	0.8745	0.8741	**0.8748**	0.8742
Break Dancer	0.8289	0.8289	**0.8298**	0.8296	0.8289
Love bird1	0.8614	0.8613	0.8607	**0.8615**	**0.8615**

3) Sensitivity of depth in patch matching: The results from the sensitivity of depth in source region selection (SZSR) in the patch matching step demonstrate that the depth information is necessary in order to avoid the selection of the similar texture regions from the foreground. Without using the depth information, regions between two foreground objects are filled with the foreground texture (see Fig. 9(d) and Fig. 10(d)). Other results from the sensitivity of depth in source region selection (SZBM) in the patch matching step demonstrate that the use of depth in the source region is not so important when the depth data contains no layers (see Fig. 9(e) and Fig. 10(e)). In contrast, the depth data is essential for filling holes when the scene contains multiple depth layers in order to propagate the similar texture according to the depth level.

In summary, the quality of the virtual view is highly dependent on the available depth information. Moreover, the depth-information plays a crucial role in filling the missing regions in the synthesized views by guiding the filling process to proceed from the background direction and copying the best texture from the background data. It is important that the depth map is filled with the background information otherwise errors will propagate because all stages of the proposed inpainting method depend on the depth information.

VII. Conclusion

We have analyzed the importance of the depth information in the proposed depth-based inpainting method to fill disocclusions in a virtual view. The depth information guide the direction from which the filling should proceed, helping the data term calculation and, moreover, in the patch matching to select the best texture from the background. The influence of the depth information at each stage of the inpaiting process is analyzed by using objective measurements and visual inspection and the results are compared with the depth-included curvature inpainting method results. The evaluation demonstrates that the proposed method performs better than related work.

To what degree depth can be used in each step of the inpainting process depends on the depth distribution, which is presented in our visual analysis and objective evaluation. More elaborate knowledge about the depth distribution allows for tradeoffs that may reduce computational requirements without sacrificing quality. One such example is in the case where the scene has no disocclusions between foreground, and depth may be excluded from the boundary extraction step. Moreover, when the depth map contains only foreground and background layers, the depth information in the data term and block matching demonstrate less impact on visual quality.

For the future work, we will focus on reducing the computation time as the current method uses third order PDEs and is iteratively calculating the data term. Another focus in the future work will be on temporal coherency by using information from neighboring frames with valid tests and analysis.

Acknowledgment

This work has been supported by grant 00156702 of the EU European Regional Development Fund, Mellersta Norrland, Sweden, and by grant 00155148 of Länsstyrelsen Västernorrland, Sweden. We would like to acknowledge our colleagues Yun Li and Mitra Damghanian for their help. We would also like to acknowledge reviewers for their valuable suggestions.

References

[1] S. M. Muddala, R. Olsson, and M. Sjöström, "Disocclusion handling using depth-based inpainting," in *The Fifth International Conferences on Advances in Multimedia (MMEDIA)*, April 2013, pp. 136–141.

[2] W. J. Tam, G. Alain, L. Zhang, T. Martin, and R. Renaud, "Smoothing depth maps for improved steroscopic image quality," *Three-Dimensional TV, Video, and Display III*, vol. 5599, pp. 162–172, 2004.

[3] Z. Tauber, Z. N. Li, and M. S. Drew, "Review and preview: Disocclusion by inpainting for image-based rendering," *IEEE Transactions on Systems, Man and Cybernetics, Part C: Applications and Reviews*, vol. 37, no. 4, pp. 527–540, 2007.

[4] A. Efros and T. Leung, "Texture Synthesis by Non-parametric Sampling," in *International Conference on Computer Vision*, 1999, pp. 1033–1038.

[5] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in *Proceedings of ACM Conf. Comp. Graphics (SIGGRAPH)*, 2000, pp. 417–424.

[6] T. F. Chan and J. Shen, "Non-texture inpainting by curvature-driven diffusions (cdd)," *J. Visual Comm. Image Rep*, vol. 12, pp. 436–449, 2001.

[7] T. Chan and J. Shen, "Mathematical Models for Local Nontexture Inpaintings," CAM TR 00-11, March 2000.

[8] A. Criminisi, P. Pérez, and K. Toyama, "Region filling and object removal by exemplar-based image inpainting," *IEEE Transactions on Image Processing*, vol. 13, pp. 1200–1212, 2004.

[9] I. Daribo and B. Pesquet-Popescu, "Depth-aided image inpainting for novel view synthesis," in *IEEE International Workshop on Multimedia Signal Processing*, 2010.

[10] J. Gautier, O. L. Meur, and C. Guillemot, "Depth-based image completion for view synthesis," in *3DTV conference*, 2011, pp. 1–4.

[11] R. Lange and P. Seitz, "Solid-state time-of-flight range camera," *IEEE Journal of Quantum Electronics*, vol. 37, pp. 390–397, March 2001.

[12] C. L. Zitnick, S. B. Kang, M. Uyttendaele, S. Winder, and R. Szeliski, "High-quality video view interpolation using a layered representation," *ACM Trans. Graph.*, vol. 23, no. 3, pp. 600–608, Aug. 2004.

[13] C. Fehn, "Depth-image-based rendering (DIBR), compression, and transmission for a new approach on 3D-TV," *Proc. SPIE Stereoscopic Displays and Virtual Reality Systems XI*, pp. 93–104, Jan. 2004.

[14] S. M. Muddala, M. Sjöström, and R. Olsson, "Edge-preserving depth-image-based rendering method," in *International Conference on 3D Imaging 2012 (IC3D)*, December 2012.

[15] S. Li, R. Wang, J. Xie, and Y. Dong, "Exemplar image inpainting by means of curvature-driven method," in *Computer Science and Electronics Engineering (ICCSEE)*, March 2012, vol. 2, pp. 326–329.

[16] Y. Wexler, E. Shechtman, and M. Irani, "Space-time completion of video," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 3, pp. 463–476, 2007.

[17] G. M. Um, G. Bang, N. Hur, J. Kim, and Y. S. Ho, "3d video test material of outdoor scene," ISO/IEC JTC1/SC29/WG11/M15371, April 2008.

[18] Sooraj Bhat, "Matlab implementaion of inpainting," http://www.cc.gatech.edu/~sooraj/inpainting/.

[19] Zhou Wang, Alan C. Bovik, Hamid R. Sheikh, and Eero P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," *IEEE Transactions on Image Processing*, vol. 13, pp. 600–612, 2004.

(a) Warped images of Ballet and Love bird1

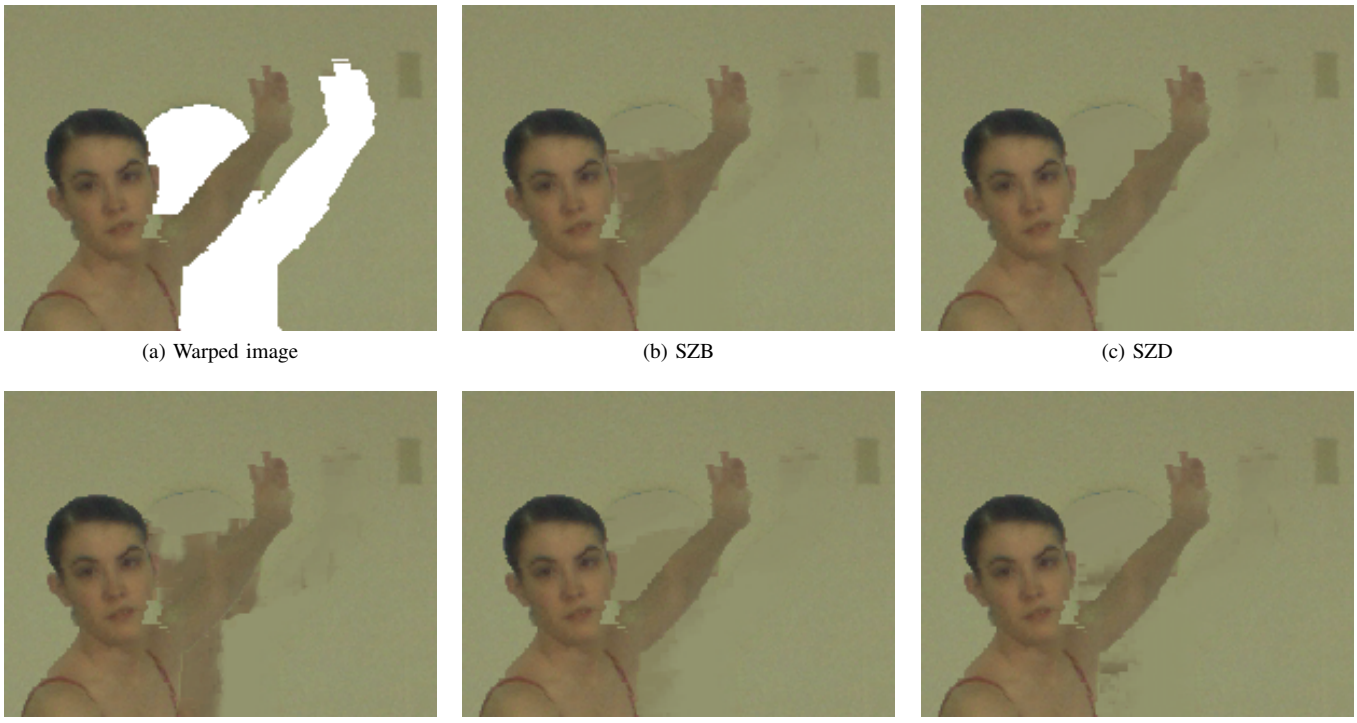
(b) Inpainted with Criminisi et al. method

(c) Inpainted with Daribo et al. method

(d) Inpainted with Gautier et al. method

(e) Inpainted with Proposed method

Figure 8. Compared to related work: inpainting method results for the investigated sequence frames “Ballet” first frame in the coulmn1 and “Lovebird1” 190th frame in column 2.



(a) Warped image (b) SZB (c) SZD

(d) SZSR (e) SZBM (f) Proposed method (depth in all steps)

Figure 9. Sensitivity analysis results for the investigated “Ballet” sequence second frame.

(a) Warped image (b) SZB (c) SZD

(d) SZSR (e) SZBM (f) Proposed method (depth in all steps)

Figure 10. Sensitivity analysis results for the investigated “Lovebird1” sequence frame198.

Performance of LTE Turbo Codes with Joint Source Channel Decoding, Adaptive Scaling and Prioritised QAM Constellation Mapping

Tulsi Pawan Fowdur, Yogesh Beeharry, and K.M. Sunjiv Soyjaudah
Dept of Electrical and Electronic Engineering
University of Mauritius
Réduit, Mauritius
e-mail: p.fowdur@uom.ac.mu, yogesh536@hotmail.com, ssoyjaudah@uom.ac.mu

Abstract— **Turbo coded Quadrature Amplitude Modulated (QAM) systems have been adopted by standards such as Code Division Multiple Access (CDMA) 2000 and Long Term Evolution (LTE) to achieve high data rates. Although several techniques have been developed to improve the performance of Turbo coded QAM systems, combinations of these techniques to produce hybrids with better performances, have not been fully exploited. This paper investigates the performance of LTE Turbo codes with Joint Source Channel Decoding (JSCD), adaptive Sign Division Ratio (SDR) based scaling and prioritised QAM constellation mapping. JSCD exploits a-priori source statistics at the decoder side and SDR based scaling provides a scale factor for the extrinsic information as well as a stopping criterion. Additionally, prioritised constellation mapping exploits the inherent Unequal Error Protection (UEP) characteristic of the QAM constellation and provides greater protection to the systematic bits of the Turbo encoder. Simulation results show that with 16-QAM at Bit Error Rates (BERs) above 10^{-1}, the combination of these three techniques achieves an average gain of 1.7 dB over a conventional LTE Turbo coded 16-QAM system. For BERs below 10^{-1}, the combination of prioritized constellation mapping, JSCD and SDR scaling provides an average gain of 0.6 dB. The proposed scheme with 64-QAM gives a performance gain of 3 dB on average over the conventional LTE Turbo coded 64-QAM system over the whole E_b/N_0 range.**

Keywords- Turbo Code; QAM; JSCD; SDR; Priortised Mapping.

I. INTRODUCTION

This paper extends the work of [1] by considering both 16-QAM and 64-QAM with LTE Turbo codes. Since the inspection of Turbo codes by Berrou *et.al* in 1993 [2], several communication standards have adopted this powerful near Shannon limit error correcting code. For example, Turbo coded QAM systems have been widely exploited to achieve reliable transmission at high data rates in several standards such as Long Term Evolution (LTE) [3] [4], CDMA 2000 [5] and HomePlug Green PHY [6]. These systems have also been reported to be promising for IEEE 802.11a [7]. The major impact of Turbo codes has led to the emergence of several techniques such as Joint Source Channel Decoding (JSCD) [8] [9] [10] [11], extrinsic information scaling and iterative detection [12] [13] [14] [15] to improve its error performance and lower its decoding complexity. Moreover, certain characteristics of the 64-QAM constellation have also been exploited to improve the performance of Turbo coded QAM [16]. An overview of these techniques is given next.

JSCD essentially involves the use of a-priori source statistics and the exploitation of residual redundancy to enhance the channel decoding process. For example, Murad and Fuja [8] proposed a composite trellis, made up of a Markov source, a Variable Length Code (VLC), and a channel decoder's state transitions, to exploit a priori source statistics. A low complexity version of the technique in [8] was developed by Jeanne *et.al* [9] and more recently Xiang and Lu [10] proposed a JSCD scheme for Huffman encoded multiple sources, which could exploit the a-priori bit probabilities in multiple sources. Also, Fowdur and Soyjaudah [11] proposed a JSCD scheme with iterative bit combining, which incorporated two types of a-priori information, leading to significant performance gains. On the other hand, extrinsic information scaling aims at improving the Turbo decoder's performance by scaling its extrinsic information with a scale factor. For example, Vogt and Finger [12] used a fixed scale factor to improve the Max-Log-MAP Turbo decoding algorithm, while Gnanasekaran and Duraiswamy [13] proposed a modified MAP algorithm using a fixed scale factor. Interestingly though Lin *et.al* [14] proposed a scaling scheme that extended the Sign Division Ratio (SDR) technique of Wu *et.al* [15] to adaptively determine a scaling factor for each data block at every iteration. Finally, the Turbo decoding process can be further enhanced by exploiting the UEP characteristic of the QAM constellation to give more protection to the systematic bits of the Turbo encoder. This technique has been applied to LTE Turbo codes by Lüders *et.al* [16].

In contrast with previous works, which have mostly considered the schemes developed to improve the performance of Turbo codes independently, this paper builds upon the work of [1] and analyses the performance of LTE Turbo codes with QAM by integrating three different techniques. Firstly, at the encoder side, prioritized constellation mapping [16] is performed so that the systematic bits output by the Turbo encoder are given the highest protection when they are mapped onto the QAM constellation. The second technique employed is JSCD [8] [11], which exploits a-priori source statistics during Turbo decoding. The final technique used is adaptive extrinsic scaling based on the SDR criterion [14]. Significant performance gains are obtained for both LTE Turbo coded

16-QAM and 64-QAM systems with the combination of these three techniques.

The organization of this paper is as follows. Section II describes the complete system model. Section III presents the simulation results and analysis. Section IV concludes the paper and lists some possible future works.

II. SYSTEM MODEL

The complete transmission system is shown in Fig. 1. A random alphabet source is first generated with a non-uniform probability distribution and then encoded into bits with the Reversible Variable Length Code (RVLC) of [17].

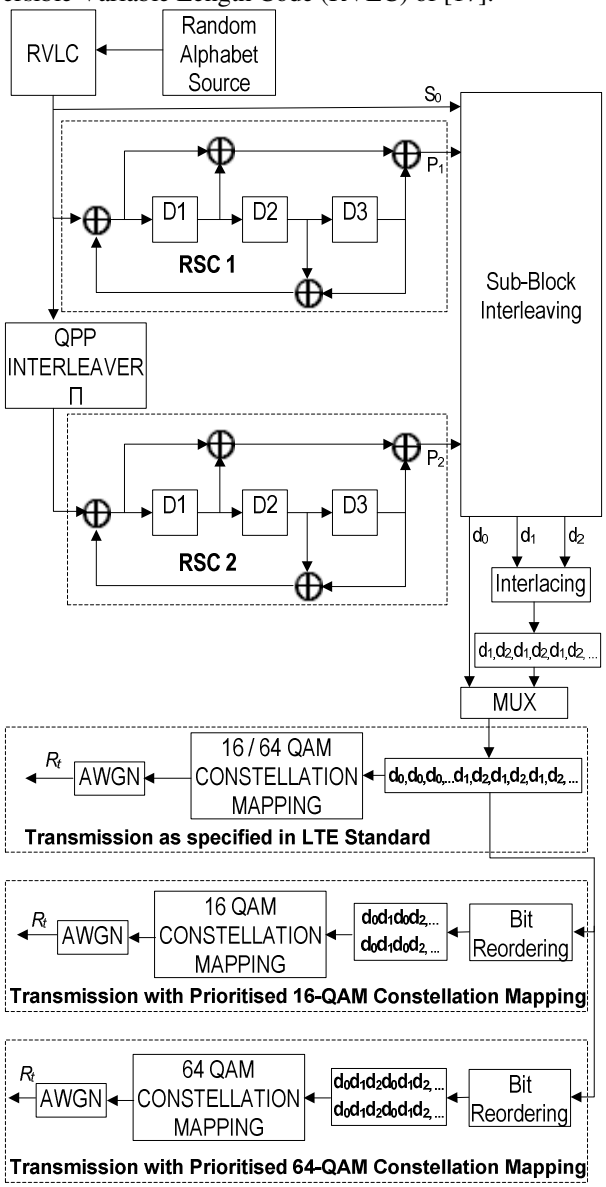


Fig. 1. Complete Transmission system

The coded bits are fed to an LTE Turbo encoder, which consists of a parallel concatenation of two Recursive Systematic Convolutional (RSC) encoders, RSC1 and RSC2, separated by an interleaver (Π). Fig. 2 shows the trellis diagram of the RSC encoders.

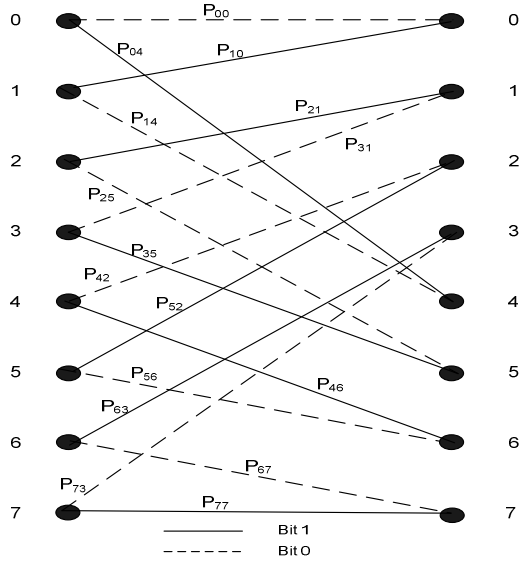


Fig. 2. Trellis diagram for LTE Turbo decoder

The interleaver used in the design of the LTE Turbo encoding system is a Quadratic Polynomial Permutation (QPP) interleaver. As the name suggests, QPP interleaver is based on a quadratic polynomial to obtain the interleaved permutation. The relationship between the output index z and input index $\pi(z)$ is expressed by the following equation [18]:

$$\pi(z) = (f_1.z + f_2.z^2)\ mod\ K \qquad (1)$$

where
f_1 and f_2 are the permutation coefficients,
K is the size of the information block.

After the encoding process, each of the three output streams S_0, P_1, and P_2 is interleaved with its own sub-block interleaver resulting in d_0, d_1, and d_2 respectively. Sub-block interleaving is performed as follows:

a. The input block is written row-wise in a rectangular matrix to form a R_s by C_s matrix.
b. The columns of this matrix are permuted.
c. The output is read column-wise from the permuted matrix.

The number of columns, C_s, is fixed to 32 and the number of rows, R_s, is (K / 32). The inter-column permutation is achieved using the Bit-Reversal-Order (BRO) pattern given below:
Q(j) = [0, 16, 8, 24, 4, 20, 12, 28, 2, 18, 10, 26, 6, 22, 14, 30, 1, 17, 9, 25, 5, 21, 13, 29, 3, 19, 11, 27, 7, 23, 15, 31]

Q(j) is the original column position of the j^{th} permuted column.

The multiplexing in the LTE Turbo coded standard is done such that the rearranged systematic bits, d_0, are placed

in the beginning followed by bit-by-bit interlacing of the two rearranged parity streams, d_1 and d_2, in order to form a single output buffer [18]. The LTE Turbo encoder generates a systematic stream, d_0 and two parity streams d_1 and d_2. To achieve prioritized constellation mapping, such that the systematic bits, d_0, are placed at the most strongly protected points on the QAM constellation, bit reordering [16] must be performed after the multiplexing process. The bit reordering is performed either on a group of four bits at a time since four bits are mapped onto one complex 16-QAM symbol or on a group of six bits at a time since six bits are mapped onto one complex 64-QAM symbol. The modulated symbols are then sent over an AWGN channel and the receiver obtains the corrupted symbols, R_t.

With 16-QAM, it is observed that after bit re-ordering, the systematic bits d_0 occupy the first and third positions of the four bits that are mapped onto one symbol of the 16-QAM constellation shown in Fig. 3. In this constellation, the bits found in the first and third positions are most protected, while the bits found in the second and fourth positions receive the lowest protection.

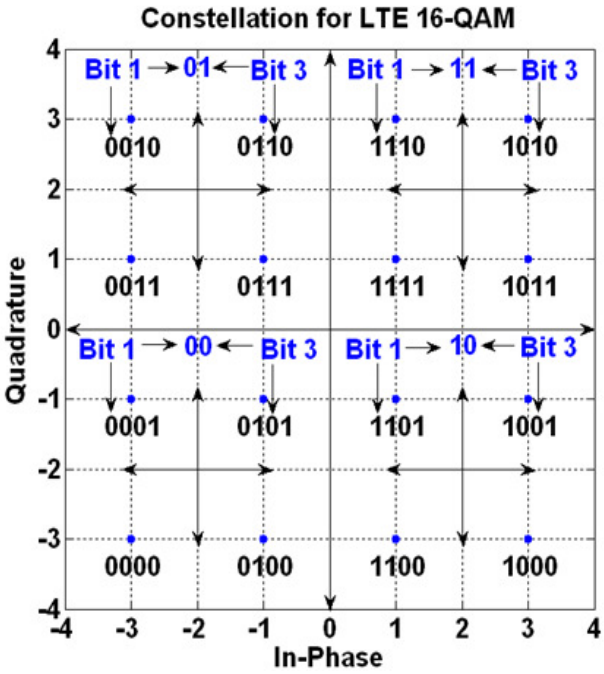


Fig. 3. 16-QAM constellation with major quadrants.

This can be explained by considering the four major quadrants of the 16-QAM constellation. The major quadrants are distinguished by the bits in the first and third positions in the 16-QAM constellation point. For example, in the upper right major quadrant of the 16-QAM constellation, the first and third bits are 11. Hence, if the de-mapper only distinguishes between the four quadrants correctly, the first and third bits of the 16-QAM constellation are correctly de-mapped.

With 64-QAM, after bit re-ordering, the systematic bits d_0 occupy the first and fourth positions of the six bits that are mapped onto one symbol of the 64-QAM constellation shown in Fig. 4. In this constellation, the bits found in the first and fourth positions are most protected, while the bits found in the third and sixth positions receive the lowest protection.

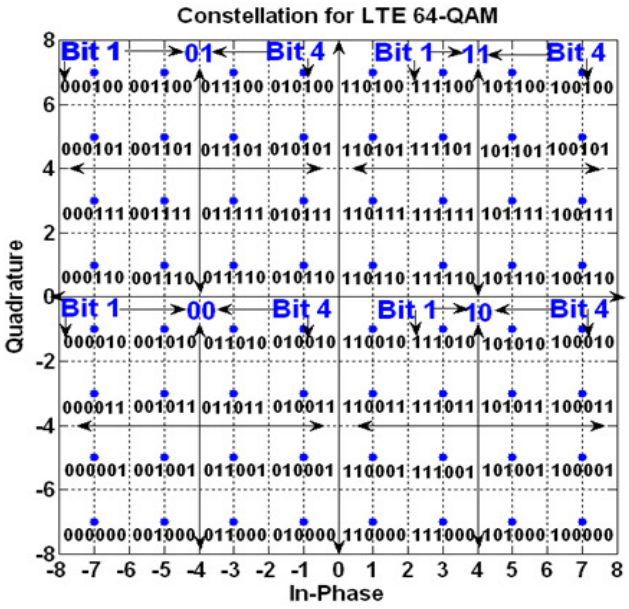


Fig. 4. 64-QAM constellation with major and minor quadrants.

This can be explained by considering both the four major and 16 minor quadrants of the 64-QAM constellation. The major quadrants are distinguished by the bits in the first and fourth positions in the 64-QAM constellation point. The first and fourth bits are 11 in the upper right major quadrant of the 64-QAM constellation. Hence, if the de-mapper only distinguishes between the four quadrants correctly, the first and fourth bits of the 64-QAM constellation are correctly de-mapped. Each major quadrant of the 64-QAM constellation is divided into four minor quadrants, which are distinguished using the 2nd and 5th bits of the constellation points. Therefore, with bit ordering [16], the systematic bits S_0 receive the highest protection while the second parity bits, P_2, receive the lowest. Since the systematic bits of a Turbo encoder have the greatest impact on its performance, the re-ordering scheme improves the performance of the Turbo decoder. The modulated QAM symbols are then transmitted over a complex Additive White Gaussian Noise (AWGN) channel and the corresponding received sequence is denoted by R_t.

The complete system model for the receiver is shown in Fig. 5. The received symbols R_t are fed to a soft-output QAM de-mapper to produce soft bits. These soft bits are then ordered and de-multiplexed. The parity soft bits are de-interlaced. The systematic and parity soft information are sub-block de-interleaved and the soft bits are sent for Turbo decoding. The first Turbo decoder is modified so that it can

incorporate a-priori source statistics by combining the trellis of the Turbo decoder with the trellis of the RVLC decoder as described in [8] and [11]. This results into a composite trellis structure with, which JSCD can be performed. The joint trellis is obtained by merging the trellises of Fig. 2 and Fig. 7. The RVLC codewords are as shown in Table I. The corresponding RVLC decoder state diagram is given in Fig. 6 from, which the bit-level trellis shown in Fig. 7 is obtained. With JSCD the computation of the branch transition probability is modified.

Fig. 5. Turbo decoding system with JSCD and SDR scaling.

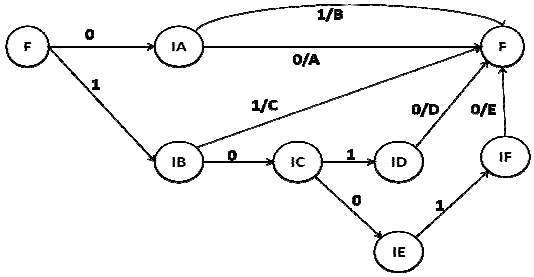


Fig. 6. RVLC decoder state diagram

Assuming that the Max-Log-MAP algorithm [19] is used, the branch metric probability for the joint decoder is computed as follows:

$$\overline{\gamma_t^{1(i)}}(l',l) = \log\left[p_t^1(i).\exp\left(-\frac{[r0_t - x0_t]^2 + [r1_t - x1_t]^2}{2\sigma^2} \right).\Pr\{C_i\} \right]$$

$$= \log\left[p_t^1(i)\right] - \left(\frac{[r0_t - x0_t]^2 + [r1_t - x1_t]^2}{2\sigma^2} \right) + \log[\Pr\{C_i\}] \quad (2)$$

where

$\overline{\gamma_t^{1(i)}}(l',l)$ is the branch transition probability from state l' to l of bit i (i = 0 or 1) at time instant t,

$p_t^1(i)$ is the a-priori probability of bit i derived from the channel extrinsic information and input to the joint (first) decoder,

$\Pr\{C_i\}$ is the a-priori probability of bit i obtained from source statistics,

$r0_t$ and $r1_t$ are the de-mapped soft bits corresponding to the bipolar equivalent of the transmitted systematic bits, $x0_t$ and first parity bits, $x1_t$. σ^2 is the noise variance [11] [19].

With the joint decoder, the a-priori statistics, $\Pr\{C_i\}$ can be incorporated into the Turbo decoding process. The derivation of the a-priori source statistics for the RVLC source given in Table I is now explained. The RVLC decoder's bit-level trellis is shown in Fig. 7 [11] [19].

TABLE I. RVLC CODEWORDS

Symbol	Probability	RVLC [17]
A	0.33 (PA)	00
B	0.30 (PB)	01
C	0.18 (PC)	11
D	0.10 (PD)	1010
E	0.09 (PE)	10010

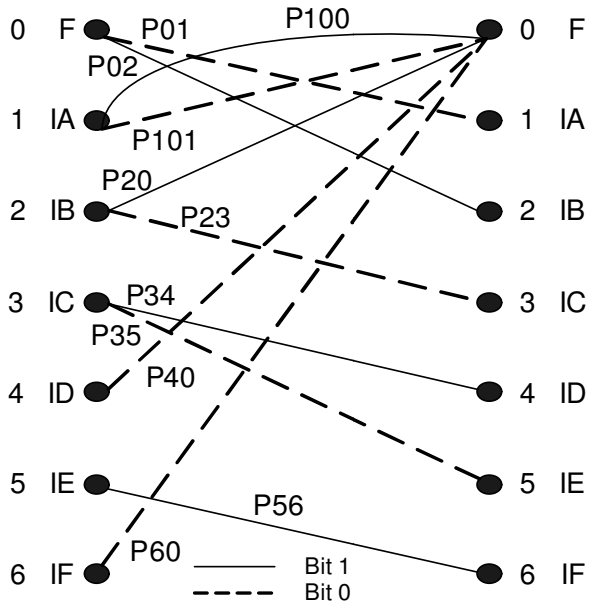


Fig. 7. Bit level trellis of RVLC decoder [11].

From the bit level trellis, the probability of the transition from state $M_{t-1} = l'$ to $M_t = l$, where l', $l \in (F,IA,IB,IC,ID,IE,IF)$, given an input bit i at time instant t, can be derived for all possible state transitions. For example, the probability of the transition from the final state F to the intermediate state IA, is given by [11]:

$$\Pr\left(M_t = IA, i = 0 \mid M_{t-1} = F\right) = PA + PB = P01 \quad (3)$$

For simplicity, the state transition probability for any state corresponding to bit *i* is denoted as $\Pr\{C_i\}$ and the joint decoder exploits this probability by computing the branch metric probability given by (2) [11]. The forward recursive variable, $\overline{\alpha_t^1}(l)$, at time *t* and state *l* is computed as follows for a joint decoder with M_j states:

$$\overline{\alpha_t^1}(l) = \max\left(\overline{\alpha}_{t-1}^1(l') + \overline{\gamma_t^{1(i)}}(l', l)\right) for \;\; 0 \le l' \le M_j - 1 \quad (4)$$

The number of states of the joint decoder, M_j is greater than the number of states, M_s of the second decoder (DEC2), because the joint decoder is obtained by merging the states of the RVLC decoder with the states of the Turbo decoder as described in [11]. A joint trellis is obtained by merging the trellis of the turbo decoder and the bit-level trellis of the RVLC decoder. The backward recursive variable, $\overline{\beta_t^1}(l)$, is computed as follows:

$$\overline{\beta_t^1}(l) = \max\left(\overline{\beta}_{t-1}^1(l') + \overline{\gamma_t^{1(i)}}(l, l')\right) for \;\; 0 \le l' \le M_j - 1 \quad (5)$$

The Log-Likelihood Ratio (LLR), $\Lambda_1^{(r)}(t)$ at iteration *r* and time *t* for the joint decoder is computed as follows:

$$\Lambda_1^{(r)}(t) = \max\left(\overline{\alpha}_{t-1}^1(l') + \overline{\gamma_t^{1(1)}}(l', l) + \overline{\beta}_t^1(l)\right) \quad (6)$$
$$- \max\left(\overline{\alpha}_{t-1}^1(l') + \overline{\gamma_t^{1(0)}}(l', l) + \overline{\beta}_t^1(l)\right) for \;\; 0 \le l' \le M_j - 1$$

The extrinsic information $\Lambda_{1e}^{(r)}(t)$ at iteration *r* and time *t* for the joint decoder is computed as follows:

$$\Lambda_{1e}^{(r)}(t) = \Lambda_1^{(r)}(t) - \frac{2}{\sigma^2} r0_t - \overline{\Lambda}_{2e}^{(r-1)}(t) \quad (7)$$

where $\overline{\Lambda}_{2e}^{(r-1)}(t)$ is the de-interleaved extrinsic information obtained from the second decoder at iteration *r*-1.

The extrinsic information, $\Lambda_{1e}^{(r)}(t)$ and the LLR, $\Lambda_1^{(r)}(t)$ are then sent to a SDR scaling mechanism, which computes a scale factor S_{1r} as follows:

$$S_{1r} = \frac{1}{N} \sum_{t=1}^{N} f\left(\Lambda_{1e}^{(r)}(t), \Lambda_1^{(r)}(t)\right) \quad (8)$$

where $f\left(\Lambda_{1e}^{(r)}(t), \Lambda_1^{(r)}(t)\right)$=1 if $\Lambda_{1e}^{(r)}(t)$ and $\Lambda_1^{(r)}(t)$ have the same sign, otherwise $f\left(\Lambda_{1e}^{(r)}(t), \Lambda_1^{(r)}(t)\right)$=0. *N* is the frame size in bits.

When S_{1r} takes its maximum value of 1.0, the switch T1 is opened, the iterative decoding process is stopped and a hard decision is made on $\Lambda_1^{(r)}(t)$. However, when $S1r$ is less than one, T1 remains closed and the extrinsic information $\Lambda_{1e}^{(r)}(t)$ is scaled with S_{1r} and interleaved to obtain $\overline{\Lambda}_{1e}^{(r)}(t)$. Hence, the SDR scaling mechanism acts both as a stopping criterion and a scale factor generator. The mechanism is derived from the one proposed in [14], but, in this work $\Lambda_{1e}^{(r)}(t)$ and $\Lambda_1^{(r)}(t)$ are used to compute the scale factor and not $\overline{\Lambda}_{2e}^{(r-1)}(t)$ and $\Lambda_{1e}^{(r)}(t)$. Another difference is that in this work only the extrinsic information has been scaled and not the soft channel inputs, as was the case in [14]. The a-priori probability, $p_t^2(i)$, is computed as follows and sent to decoder 2:

$$p_t^2(i) = \left\{ \begin{array}{l} \dfrac{\exp(\overline{\Lambda}_{1e}^{(r)}(t))}{1 + \exp(\overline{\Lambda}_{1e}^{(r)}(t))} \; for \;\; i = 1 \\ \dfrac{1}{1 + \exp(\overline{\Lambda}_{1e}^{(r)}(t))} \; for \;\; i = 0 \end{array} \right\} \quad (9)$$

The branch metric probability for the second decoder is computed as follows:

$$\overline{\gamma_t^{2(i)}}(l', l) = \log\left[p_t^2(i)\right] - \left(\frac{[\overline{r0}_t - x0_t]^2 + [r2_t - x2_t]^2}{2\sigma^2} \right) \quad (10)$$

where $r2_t$ is the de-mapped soft bits corresponding to the bipolar version of the transmitted second parity bits $x2_t$ and $\overline{r0}_t$ is the interleaved counterpart of $r0_t$.

The forward and backward recursive variable, $\overline{\alpha_t^2}(l)$ and $\overline{\beta_t^2}(l)$ at time *t* and state *l* are computed as follows:

$$\overline{\alpha_t^2}(l) = \max\left(\overline{\alpha}_{t-1}^2(l') + \overline{\gamma_t^{2(i)}}(l', l)\right) for \;\; 0 \le l' \le M_s - 1 \quad (11)$$

$$\overline{\beta_t^2}(l) = \max\left(\overline{\beta}_{t-1}^2(l') + \overline{\gamma_t^{2(i)}}(l, l')\right) for \;\; 0 \le l' \le M_s - 1 \quad (12)$$

The LLR, $\Lambda_2^{(r)}(t)$ and extrinsic information, $\Lambda_{2e}^{(r)}(t)$ at iteration *r* and time *t* are computed as follows:

$$\Lambda_2^{(r)}(t) = \max\left(\overline{\alpha}_{t-1}^2(l') + \overline{\gamma_t^{2(1)}}(l', l) + \overline{\beta}_t^2(l)\right) - \max\left(\overline{\alpha}_{t-1}^2(l') + \overline{\gamma_t^{2(0)}}(l', l) + \overline{\beta}_t^2(l)\right) for \;\; 0 \le l' \le M_j - 1 \quad (13)$$

$$\Lambda_{2e}^{(r)}(t) = \Lambda_2^{(r)}(t) - \frac{2}{\sigma^2}\overline{r0}_t - \overline{\Lambda}_{1e}^{(r)}(t) \quad (14)$$

The scale factor S_{2r} is computed as follows:

$$S_{2r} = \frac{1}{N}\sum_{t=1}^{N} f\left(\Lambda_{2e}^{(r)}(t), \Lambda_2^{(r)}(t)\right) \quad (15)$$

where $f\left(\Lambda_{2e}^{(r)}(t), \Lambda_2^{(r)}(t)\right)$=1 if $\Lambda_{2e}^{(r)}(t)$ and $\Lambda_2^{(r)}(t)$ have the same sign. Finally, the a-priori probability, $p_t^1(i)$, is computed as given by (9) but using $\overline{\Lambda}_{2e}^{(r)}(t)$. If S_{2r} = 1.0, T2 is opened to stop the iterative decoding process and a hard decision, (HD) is made on $\overline{\Lambda}_2^{(r)}(t)$.

The combination of prioritized constellation mapping, JSCD and adaptive scaling certainly lead to an enhanced LTE Turbo coded QAM system, but at the cost of greater computational complexity and delay. The complexity increase due to the bit re-ordering scheme is negligible and may even be integrated with the multiplexer. JSCD on the other hand leads to the greatest increase in complexity and delay because as mentioned previously the joint decoder is obtained by merging the states of the RVLC decoder with the states of the Turbo decoder. The number of computations involved in computing S_{1r} and S_{2r} to perform adaptive scaling also increase the delay. However, this is compensated by the faster convergence achieved with the use of the scale factor and the possibility of stopping the iterative decoding process once convergence is achieved. This prevents the decoder from performing unnecessary iterations.

III. Simulation Results and Analysis

The performances of the following four LTE Turbo coded QAM schemes are compared:

Scheme 1 – The LTE Turbo coded QAM system as specified in the standard with conventional decoding. No adaptive scaling and prioritised constellation mapping are used. The encoding framework is as given in Fig. 1. The decoding is as shown in Fig. 5 but without JSCD and adaptive scaling.

Scheme 2 - This scheme uses prioritised constellation mapping with highest priority given to the systematic bits. The encoding is as shown in Fig. 1. The decoding is similar to that of Scheme 1.

Scheme 3 – This scheme uses JSCD and adaptive scaling with the LTE Turbo coded QAM as specified by the standard. The decoding framework is given in Fig. 5.

Scheme 4 – This scheme is an LTE Turbo coded QAM system with prioritised constellation mapping, JSCD and SDR scaling.

In all simulations, a random alphabet source with the probability distribution given in Table I has been used. After generating the alphabets, they are grouped into packets of size P = 64 symbols. The packets are then Reversible Variable Length Coded to obtain an RVLC bit-stream as shown in Fig. 1. Normally, the length in bits, L, of each packet is transmitted as side-information because L is different for each packet. The packetization is important to prevent error propagation [11]. The RVLC bit-streams of all packets are grouped into blocks of 4096 bits since an interleaver size of 4096 bits has been used in the simulations. The parameters for the LTE Turbo code used are as follows [20]:

Generator: G = [1, g1/g2], where g1 = 15 and g2 = 13 in Octal.
Interleaver size, N = 4096 bits.
QPP Interleaver parameters: f_1= 31 and f_2 = 64
Maximum number of iterations, I = 12.
Code-rate = 1/3 and channel model: Complex AWGN.

The graphs of Bit Error Rate (BER) and Levenshtein Error Rate (LER) as a function of E_b/N_0 have been plotted over an E_b/N_0 range: 0 dB ≤ E_b/N_0 ≤ 11 dB in steps of 0.5 dB for 16-QAM and 64-QAM. The Levenshtein distance denoted by $d_L(u_1, u_2)$ between two symbol sequences u_1 and u_2, which are not necessarily of equal lengths, is defined as the number of insertions, deletions, and substitutions necessary to transform one sequence into the other [21] [22] [23]. The LER is a more appropriate measure for Variable Length Coded (VLC) symbol sources than the Symbol Error Rate (SER) because the length of the transmitted and received symbols may not necessarily be the same. Hence, LER provides a fairer measure than the SER for VLC coded symbols. E_b/N_0 is the ratio of the bit energy, E_b, to the noise power spectral density, N_0. The performance analysis has been made for both iterative and non-iterative decoding. Non-iterative curves were given to analyse the baseline turbo coding performance as was done in [16].

Fig. 8 shows the graph of BER against E_b/N_0 for iterative decoding with 16-QAM. In the low E_b/N_0 range (0 dB ≤ E_b/N_0 ≤ 4 dB) it is observed that the LTE Turbo coded 16-QAM system with JSCD, adaptive scaling, and prioritised constellation mapping (Scheme 4) provides the best performance with an average gain of 1.7 dB for BER > 10^{-1} over the conventional LTE Turbo coded system (Scheme 1). At an E_b/N_0 of 1 dB, Scheme 2, which uses only prioritised constellation mapping, outperforms Scheme 1 by 1dB. It is to be noted from a theoretical point of view the performance of the system for a BER > 10^{-1} is important only because it is revealing a new characteristic of the system whereby it is

seen that significant gains can be obtained in this BER region using the proposed technique. However, from a practical point of view the performance of the system for BER > 10^{-1} is not really relevant. In the high E_b/N_0 range (4.5 dB $0 \leq E_b/N_0 \leq 11$ dB), it is observed that prioritised constellation mapping is not really beneficial. For instance, Scheme 2 provides the worst performance. A possible explanation is that with iterative decoding, in the high E_b/N_0 range (4.5 dB $0 \leq E_b/N_0 \leq 11$ dB), convergence takes place. As such, giving more protection to the systematic bits does not provide further gains. Also, since lower protection has been given to the parity bits, this can lead to performance degradation. Over this E_b/N_0 range, Scheme 4 that uses prioritised constellation mapping with JSCD and adaptive scaling provide the best performance with an average gain of 0.6 dB in E_b/N_0 over Scheme 1.

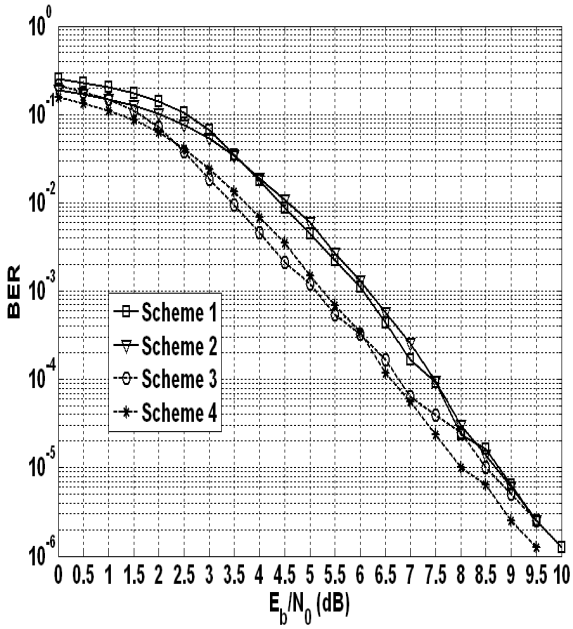


Fig. 8. Iterative BER performance with N = 4096 and 16-QAM.

Fig. 9 shows the graph of LER against E_b/N_0 for iterative decoding with 16-QAM. In the low E_b/N_0 range (0 dB $\leq E_b/N_0 \leq 4$ dB) it is observed that the LTE Turbo coded 16-QAM system with JSCD, adaptive scaling, and prioritised constellation mapping (Scheme 4) provides the best performance with an average gain of 1.5 dB for LER > 10^{-1} over the conventional LTE Turbo coded system (Scheme 1). At an E_b/N_0 of 1 dB, Scheme 2, which uses only prioritised constellation mapping outperforms Scheme 1 by 0.6 dB. In the high E_b/N_0 range (4.5 dB $0 \leq E_b/N_0 \leq 11$ dB), Scheme 2 provides the worst performance and Scheme 4, which uses prioritised constellation mapping with JSCD and adaptive scaling provides the best performance with an average gain of 0.8 dB in E_b/N_0 over Scheme 1.

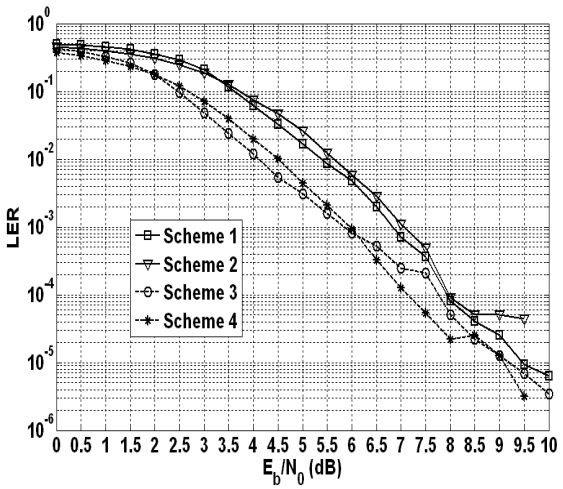


Fig. 9. Iterative LER performance with N = 4096 and 16-QAM.

Fig. 10 shows the graph of average number of iterations versus E_b/N_0 over the range 4.5 dB $\leq E_b/N_0 \leq$ 12 dB, when 16-QAM is used. Schemes 3 and 4, which employ SDR based scaling with a stopping criterion, show a progressive decrease in the number of iterations required as the E_b/N_0 increases. For example at an E_b/N_0 of 8.5 dB, Scheme 4 consumes seven iterations less than Schemes 1 and 2. However, at the same E_b/N_0 of 8.5 dB, Scheme 4 consumes 1.5 iterations more than Scheme 3 due to performance loss as a result of using prioritised constellation mapping after convergence. The number of iterations required by Schemes 1 and 2 remains fixed at 12 throughout the E_b/N_0 range.

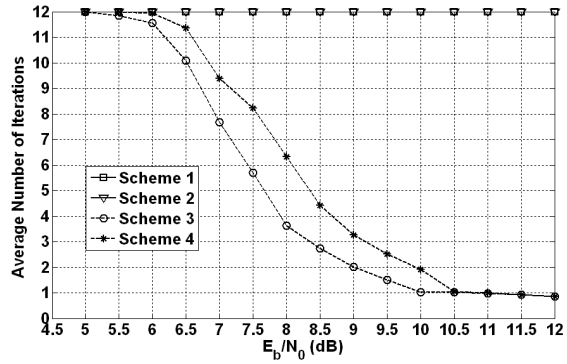


Fig. 10. Average number of iterations vs E_b/N_0 for N = 4096 and 16-QAM.

Fig. 11 shows the graph of BER against E_b/N_0 for non-iterative decoding with 16-QAM. In the low E_b/N_0 range (0 dB $\leq E_b/N_0 \leq 4$ dB), it is observed that Scheme 4 outperforms all the other schemes with an average gain of 1.5 dB over Scheme 1 and 0.8 dB over Scheme 2. However, with non-iterative decoding, convergence does not take place and the use of prioritized constellation mapping does not lead to degradation at BERs below 10^{-1}. This is observed in the high E_b/N_0 range (4.5 dB $0 \leq E_b/N_0 \leq 11$ dB) whereby Scheme 4 outperforms Scheme 1 by 0.8 dB on average.

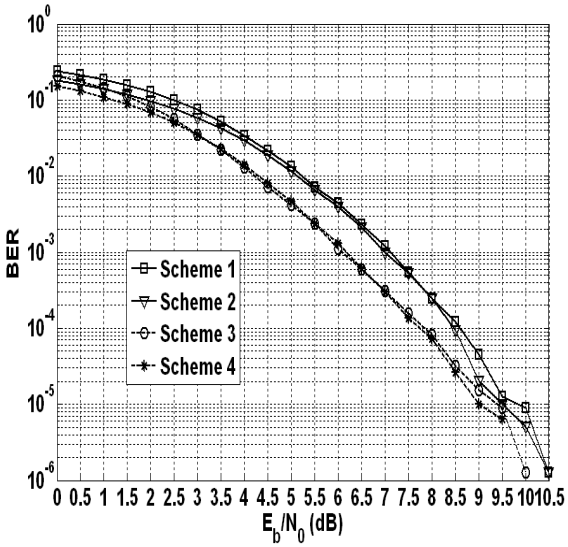


Fig. 11. Non-Iterative BER performance with N = 4096 and 16-QAM.

Fig. 12 shows the graph of LER against E_b/N_0 for non-iterative decoding with 16-QAM. In the low E_b/N_0 range (0 dB $\leq E_b/N_0 \leq$ 4 dB), it is observed that Scheme 4 outperforms all the other schemes with an average gain of 1.5 dB over Scheme 1 and 1 dB over Scheme 2. However, the use of prioritized constellation mapping does not lead to degradation at BERs below 10^{-1} since with non-iterative decoding, convergence does not take place. This is observed in the high E_b/N_0 range (4.5 dB $0 \leq E_b/N_0 \leq 11$ dB) whereby Scheme 4 outperforms Scheme 1 by 1 dB on average.

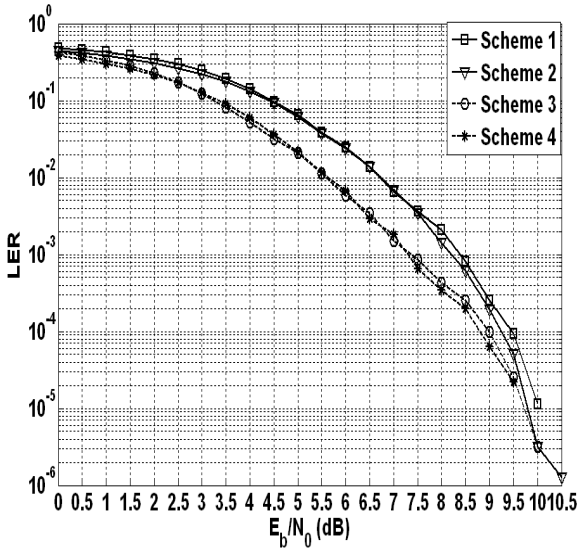


Fig. 12. Non-Iterative LER performance with N = 4096 and 16-QAM.

Fig. 13 shows the graph of BER against E_b/N_0 for iterative decoding with 64-QAM. In the low E_b/N_0 range (0 dB $\leq E_b/N_0 \leq$ 4 dB), it is observed that Scheme 4 outperforms all the other schemes with an average gain of 4.5 dB over Scheme 1 and 1 dB over Scheme 2 at BERs above 10^{-1}. However, with 64-QAM, convergence does not take place and the use of prioritized constellation mapping does not lead to degradation at BERs below 10^{-1}. This is observed in the high E_b/N_0 range (4.5 dB $0 \leq E_b/N_0 \leq 11$ dB) whereby Scheme 2 outperforms Scheme 1 by 2 dB on average and Scheme 4 outperforms Scheme 1 by 3 dB on average. The combination of LTE-Turbo codes with 16-QAM did not lead to performance improvements in the high E_b/N_0 range when prioritized constellation mapping was used. However, with 64-QAM, prioritized constellation mapping provides an improvement over the whole E_b/N_0 range. In [1], when 64-QAM was combined with a 4-state turbo code with g1 = 5 and g2 = 7 in octal, a degradation was obtained due to convergence. However, with LTE Turbo code, this was observed with only 16-QAM. With 64-QAM, prioritised constellation mapping still provided the gain especially when combined with JSCD and SDR scaling. One possible explanation is that convergence depends on both the free distance of the turbo code and the modulation scheme. Since LTE Turbo codes have a higher free distance than the turbo code in [1], a gain was observed for 64-QAM throughout the E_b/N_0 range.

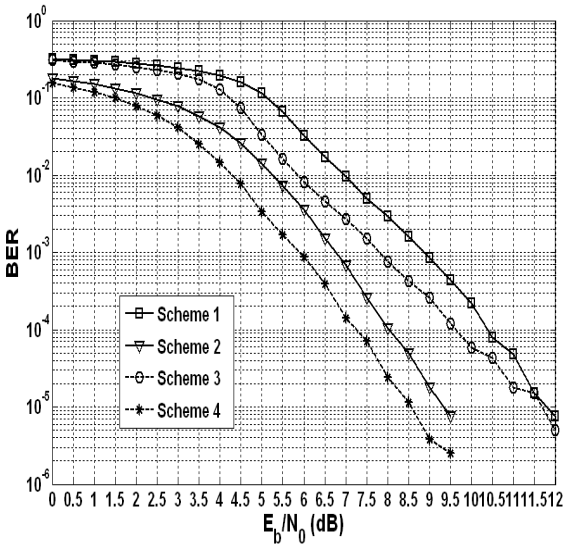


Fig. 13. Iterative BER performance with N = 4096 and 64-QAM.

Fig. 14 shows the graph of LER against E_b/N_0 for iterative decoding with 64-QAM. In the low E_b/N_0 range (0 dB $\leq E_b/N_0 \leq$ 4 dB), it is observed that Scheme 4 outperforms all the other schemes with an average gain of 4.5 dB over Scheme 1 and 1.5 dB over Scheme 2. However, with 64-QAM, convergence does not take place and the use of prioritized constellation mapping does not lead to degradation at LERs below 10^{-1}. This is observed in the high E_b/N_0 range (4.5 dB $0 \leq E_b/N_0 \leq 11$ dB) whereby scheme 2 outperforms Scheme 1 by 1 dB on average and Scheme 4 outperforms Scheme 1 by 2 dB on average over this range of E_b/N_0. Also, it can be noted that Scheme 4 outperforms Scheme 3 by 2.7 dB on average in the low E_b/N_0 range and by almost 2.2 dB in the high E_b/N_0 range.

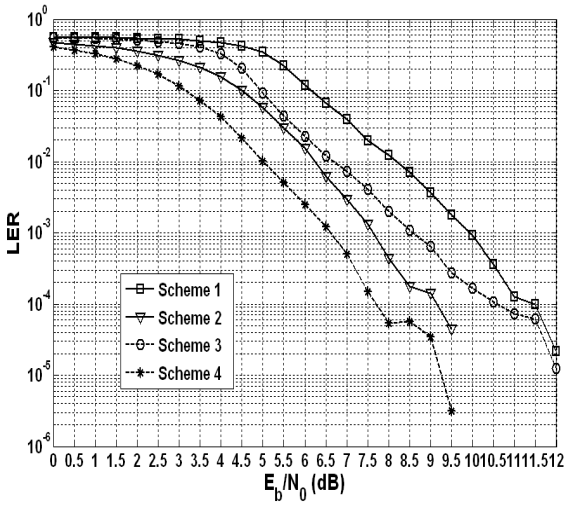


Fig. 14. Iterative LER performance with N = 4096 and 64-QAM.

Fig. 15 shows the graph of average number of iterations versus E_b/N_0 over the range 4.5 dB $\leq E_b/N_0 \leq$ 12 dB, when 64-QAM is used. Schemes 3 and 4, which employ SDR based scaling with a stopping criterion, show a progressive decrease in the number of iterations required as the E_b/N_0 increases. For example at an E_b/N_0 of 10 dB, Scheme 4 consumes nine iterations less than Schemes 1 and 2. However, it is very interesting to note that at the same E_b/N_0 of 10 dB, Scheme 4 also consumes 8 iterations less than Scheme 3, which is not the case when 16-QAM is used. The number of iterations required by Schemes 1 and 2 remains fixed at 12.

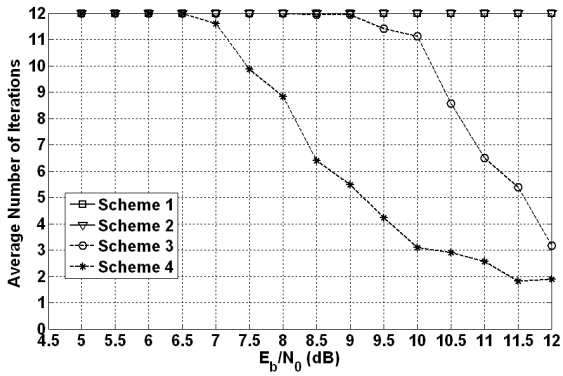


Fig. 15. Average number of iterations vs E_b/N_0 for N = 4096 and 64-QAM.

Fig. 16 shows the graph of BER against E_b/N_0 for non-iterative decoding with 64-QAM. In the low E_b/N_0 range (0 dB $\leq E_b/N_0 \leq$ 4 dB), it is observed that Scheme 4 outperforms all the other schemes with an average gain of 4.5 dB over Scheme 1 and 0.5 dB over Scheme 2. However, with non-iterative decoding, convergence does not take place and the use of prioritized constellation mapping does not lead to degradation at BERs below 10^{-1}. This is observed in the high E_b/N_0 range (4.5 dB $0 \leq E_b/N_0 \leq 11$ dB) whereby Scheme 4 outperforms Scheme 1 by 4.5 dB on average and Scheme 2 by 0.7 dB on average over this range of E_b/N_0. Also, it can be noted from Fig. 16 that Scheme 4 outperforms Scheme 3 by 3.5 dB on average in both the low E_b/N_0 range and the high E_b/N_0 range.

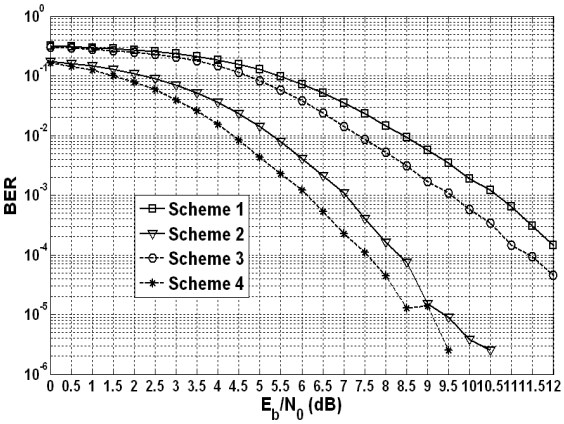


Fig. 16. Non-Iterative BER performance with N = 4096 and 64-QAM.

Fig. 17 shows the graph of LER against E_b/N_0 for non-iterative decoding with 64-QAM. In the low E_b/N_0 range (0 dB $\leq E_b/N_0 \leq$ 4 dB), it is observed that Scheme 4 outperforms all the other schemes with an average gain of 4.5 dB over Scheme 1 and 1 dB over Scheme 2. However, the use of prioritized constellation mapping does not lead to degradation at LERs below 10^{-1}. This is observed in the high E_b/N_0 range (4.5 dB $0 \leq E_b/N_0 \leq 11$ dB) whereby Scheme 4 outperforms Scheme 1 by 4.5 dB on average and Scheme 2 by 1 dB on average over this range of E_b/N_0. Also, it can be noted from Fig. 17 that Scheme 4 outperforms Scheme 3 by 3.5 dB on average both in the low E_b/N_0 range and the high E_b/N_0 range.

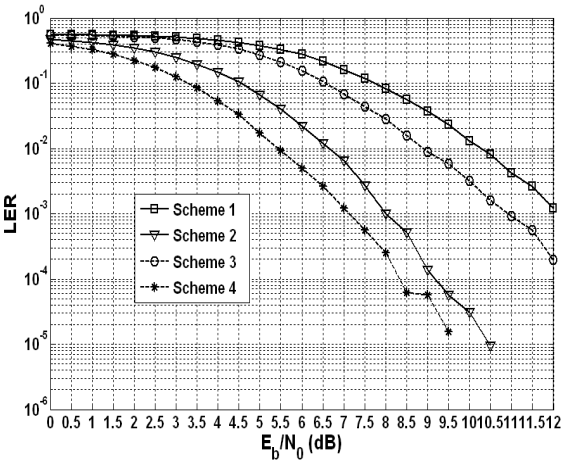


Fig. 17. Non-Iterative LER performance with N = 4096 and 64-QAM.

IV. CONCLUSION AND FUTURE WORKS

This paper proposed an efficient LTE Turbo coded QAM scheme with JSCD, adaptive scaling, and prioritised constellation mapping. At the encoder side a re-ordering mechanism is used to map the systematic bits of the Turbo

encoder on the most strongly protected points of the QAM constellation. To enhance the decoding process, JSCD is used to incorporate a-priori source statistics and adaptive SDR based scaling is also performed. When 16-QAM is used, at BERs above 10^{-1}, the proposed scheme provides a significant performance gain of 1.7 dB over a conventional LTE Turbo coded scheme. For BERs below 10^{-1}, the use of prioritised constellation mapping with conventional decoding degrades performance as a result of convergence. Hence, for BERs below 10^{-1}, it is preferable to use JSCD and SDR scaling with prioritised constellation mapping with 16-QAM, which achieves a gain of 0.6 dB on average over a conventional LTE Turbo coded scheme. However, with 64-QAM, the proposed scheme outperforms a conventional LTE Turbo coded scheme at all BERs because there is no performance degradation due to prioritised constellation mapping. The proposed scheme provides a significant average performance gain of 3 dB over a conventional LTE Turbo coded scheme. Overall, the combination of prioritised constellation mapping with JSCD and SDR based scaling appears promising for LTE Turbo coded QAM systems.

Several interesting future works can be envisaged from the scheme proposed in this work. For, e.g., the prioritised constellation mapping scheme could be optimised so that performance gains could be obtained in the high E_b/N_0 range (4.5 dB $0 \leq E_b/N_0 \leq 11$ dB) also when 16-QAM is used with conventional decoding. A more complex channel, such as selective fading, could be considered for the performance analysis of Scheme 4. Puncturing could be used to obtain higher coding rates, e.g., ½ to extend the analysis of the four schemes presented.

ACKNOWLEDGMENT

The authors would like to thank the University of Mauritius for providing the necessary facilities for conducting this research as well as the Tertiary Education Commission of Mauritius.

REFERENCES

[1] Tulsi Pawan Fowdur, Yogesh Beeharry, and K. M. Sunjiv Soyjaudah, "Performance of Turbo Coded 64-QAM with Joint Source Channel Decoding, Adaptive Scaling and Prioritised Constellation Mapping," CTRQ, 6th International Conference on Communciation Theory, Reliability and Quality of Service, Venice, Italy, April 2013, pp. 35-41.

[2] C. Berrou, A. Glavieux, and P. Thitimajshima, "Near Shannon limit error-correcting coding and decoding: Turbo-codes," IEEE International Conference on Communications, (ICC), Geneva, May 1993, vol. 2, pp. 1064-1070.

[3] S. Sesia, I. Toufik, and M. Baker, LTE – The UMTS Long Term Evolution: From Theory to Practice, John Wiley & Sons, Ltd, 2009, ISBN: 978-0-470-69716-0.

[4] 3GPP, "Technical Specifications Rel.8.", 2009.

[5] 3GPP2 C.S0024-B, "CDMA 2000 High Rate Packet Data Air Interface Specification" Version 1.0, May 2006. Available Online: http://www.3gpp2.org/Public_html/specs/C.S0024-B_v1.0_060522.pdf [Accessed: November 2012].

[6] J. Zyren, "Home Plug Green PHY Overview," Technical Paper, Atheros Communications, 2010.

[7] I. Lee, C.E.W. Sundberg, S. Choi, and W. Lee, "A modified medium access control algorithm for systems with iterative decoding," IEEE Transactions on Wireless Communications, vol. 5, no. 2, 2006, pp. 270-273.

[8] A.H. Murad and T.E. Fuja, "Joint source‑channel decoding of variable length encoded sources," Proceedings of the Information Theory Workshop (ITW). Killarney, Ireland, June. 1998, pp. 94‑95.

[9] M. Jeanne, J.C. Carlach, and P. Siohan, "Joint source‑channel decoding of variable length codes for convolutional codes and turbo codes," IEEE Trans Commun, vol. 53, no. 1, 2005, pp. 10‑15.

[10] W. Xiang and P. Lu, "Bit-Based Joint Source-Channel Decoding of Huffman Encoded Markov Multiple Sources," Journal of Networks, vol. 5, no. 4, 2010, pp. 443-450.

[11] T.P. Fowdur and K.M.S. Soyjaudah "Performance of joint source–channel decoding with iterative bit combining and detection," Ann. Telecommun. vol. 63, 2008, pp. 409-423.

[12] J. Vogt and A. Finger, "Improving the MAX-Log-MAP Turbo decoder," Electr. Lett., vol. 36, no. 23, Nov. 2000, pp. 1937-1939.

[13] T. Gnanasekaran and K. Duraiswamy, "Performance of Unequal Error Protection Using MAP Algorithm and Modified MAP in AWGN and Fading Channel," Journal of Computer Science, vol. 4, no. 7, 2008, pp. 585-590.

[14] Y. Lin, W. Hung, W. Lin, T. Chen, and E. Lu, "An Efficient Soft-Input Scaling Scheme for Turbo Decoding," IEEE International Conference on Sensor Networks, Ubiquitous, and Trustworthy Computing Workshops, vol. 2, 2006, pp. 252-255.

[15] Y. Wu, B. Woerner, and J. Ebel, "A Simple Stopping Criterion for Turbo Decoding," IEEE Commun. Lett., vol. 4, no. 8, Aug. 2000, pp. 258-260.

[16] H. Lüders, A. Minwegen, and P. Vary, "Improving UMTS LTE Performance by UEP in High Order Modulation," 7th International Workshop on Multi-Carrier Systems & Solutions (MC-SS 2009), Herrsching, Germany, 2009, pp. 185-194.

[17] Y. Takishima, M. Wada, and H. Murakami, "Reversible variable length codes," IEEE Trans. on Commun., vol. 43, 1995, pp. 158–162.

[18] Farooq Khan, "LTE for 4G Mobile Broadband Air Interface Technologies and Performance," Cambridge University Press, 2009.

[19] B. Vucetic and J. Yuan, Turbo Codes: Principles and Applications, Kluwer Academic Publishers, 2000.

[20] 3GPP TS 36.212 version 9.2.0: "Evolved Universal Terrestrial Radio Access (E-UTRA); Multiplexing and Channel Coding," May 2010.

[21] V. I. Levenshtein, "Binary codes capable of correcting spurious insertions and deletions of ones," Probl. Perdachi Inform., vol. 1, no. 1, pp. 12-25, 1965.

[22] V. I. Levenshtein, "Binary codes with correction of deletions, insertions and substitutions of symbols," Dokl. Akad. Nank. SSSR, vol. 163, pp. 845-848, 1965.

[23] T. Okuda, E. Tanaka, and T. Kasai, "A method for the correction of the garbled words based on the Levenshtein metric," IEEE Trans. Comput., vol. C-25, no. 2, pp. 172-176, Feb. 1976.

High Quality Video Conferencing: Region of Interest Encoding and Joint Video/Audio Analysis

Christopher Bulla*, Christian Feldmann*, Magnus Schäfer†, Florian Heese†, Thomas Schlien†, and Martin Schink‡
*Institut für Nachrichtentechnik, RWTH Aachen University, Aachen, Germany,
Email: {bulla,feldmann}@ient.rwth-aachen.de
†Institute of Communication Systems and Data Processing, RWTH Aachen University, Aachen, Germany,
Email: {schaefer,heese,schlien}@ind.rwth-aachen.de
‡MainConcept GmbH, Aachen, Germany, Email: Martin.Schink@rovicorp.com

***Abstract*— In this paper, we present a high quality video conferencing system, that has been developed in the collaborative project "Connected Visual Reality (CoVR)[1] – High Quality Visual Communication in Heterogeneous Networks" and was designed to reduce bitrate while preserving a constant visual quality. We utilize the fact that the main focus in a typical video conference lies upon the participating persons to save bitrate in less interesting parts of the video and introduce a scene composition concept that is merely based on the detected regions of interest. The region of interest encoding and the scene composition will be supported by a joint video and audio analysis. On the video analysis side we use a Viola-Jones face detector to detect, and a MeanShift tracker to track the regions of interest. The audio analysis exploits the information from the video analysis about the detected participants by a beamforming algorithm and creates an activity index for each participant. To represent the detected region of interests for the encoder we use a quality map on the level of macro-blocks, which allows the encoder to choose its quantization parameter individually for each macro-block. Finally, the proposed scene composition omits the background and shows only the most active participants of the conference, thus visual quantization artifacts introduced by the encoder get irrelevant. Experiments on recorded conference sequences demonstrate bitrate savings up to 50% that can be achieved with the proposed system.**

Keywords-object detection; object tracking; region of interest coding; beamforming; scene composition; video-conferencing

I. Introduction

Video conferencing greatly enhances traditional telephone conferencing, with applications ranging from every day calls of friends and family to cutting management expenses by replacing business trips with video conferences. The solutions for such systems range from Telepresence systems, e.g., by Tandberg or Polycom, specially designed rooms for video conferences with high-performance hardware which create the impression to sit at the same table with the far end participants, to mobile clients or classical dial-in telephones. Till now high acquisition costs, limited quality or bad usability often prevent the acceptance of this systems. Another drawback, which has already been addressed in our previous work [1], is the very high operation costs caused by the high bandwidth requirements.

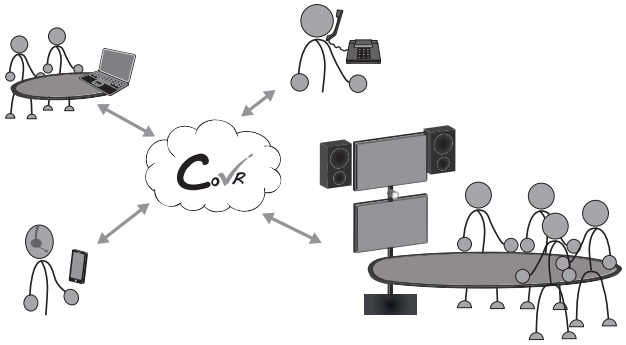


Fig. 1. CoVR - Video Conference.

The collaborative project "Connected Visual Reality (COVR) – High Quality Visual Communication in Heterogeneous Networks" [2] was working on the goal of improving the audio and video quality and interoperability of video conference systems like shown in Fig. 1. The participants of such a conference can differ significantly from each other in terms of their location, their network connectivity and in terms of the hardware. For a high heterogeneity of systems the project aims the best possible quality for every participant.

For this purpose, new methods of audio and video signal processing, coding and transmission have been developed and investigated. In the field of video processing, the project partners worked on the detection of persons, the development and standardization of HEVC (ITU-T/MPEG High Efficiency Video Coding) and Region of Interest (ROI) coding, among others [3] [4]. In the area of audio signal enhancement, algorithms for echo cancellation, noise reduction, dereverberation for multi-channel communication and artificial bandwidth extension have been considered [5].

This paper focuses on combining different parts of the work of the CoVR project. Most of the operation costs of yearly expenses for a video conference system has to be spend on the provision of the necessary bandwidth. So the saving of bandwidth by developing more efficient video codecs, like the HEVC, is a big advantage. Another approach to

[1]CoVR was funded by the NRW Ziel 2-Programm "Regionale Wettbewerbsfähigkeit und Beschäftigung" 2007-2013 and the ERDF 'European Regional Development Fund'. Participating project partners are Ericsson GmbH, MainConcept GmbH, part of Rovi, as well as two institutes of the RWTH Aachen, Institut für Nachrichtentechnik and Institute of Communication Systems and Data Processing.

EUROPEAN UNION
EUROPEAN REGIONAL DEVELOPMENT FUND
INVESTING IN YOUR FUTURE

save bandwidth is the ROI encoding, which can save up to 50% of the necessary bandwidth by coding the regions of interest (participants) with high quality and the background with low quality. This is particularly interesting for the scene composition developed in CoVR. The combination of face detection, tracking, region of interest encoding and activity index calculation allows for a scene composition which only shows the most active participants of the conference and omits the background. This way the bitrate can be reduced while preserving a constant visual quality.

The paper is organized as follows. In Section II, we will explain our video analysis, region of interest encoding concept, multi channel audio analysis and interaction between the video and audio analysis for the scene composition in detail. The evaluation of the achieved bitrate savings and audio analysis will be presented in Section III. Final conclusions as well as an outlook for future work will be given in Section IV.

II. High Quality Video Conferencing

In classical multi party video conferencing approaches each connected endpoint has one camera. The captured video is encoded into two video streams: One with a high resolution (e.g., 720p) and a second one with a lower resolution which is used as a thumbnail. Both streams are transmitted to a server, which decides upon the most active party and forwards the high resolution video stream of this party to all the other parties. The thumbnail views are always routed to all endpoints. Of course, the active party does not receive the high resolution video of itself but the high resolution video of the last active party. Hereby, each party can see a high resolution video of the active speaker and thumbnails of the other parties. But, especially in the uplink, this classical approach wastes a lot of bandwidth, and is not capable of handling simultaneous activity of multiple participants at different terminals.

The CoVR high quality video conferencing system overcomes the above mentioned limitations. Taking the requirements w.r.t. bandwidth and audio-visual quality into account our video conferencing system aims to reduce the required bandwidth of the system while the audio-visual quality remains constant. To achieve this, we use on the one hand the concept of region of interest encoding and on the other hand adaptive scene composition. A region of interest encoded video stream saves bitrate by encoding interesting parts of a video in a better quality then less interesting parts. This may result in visual coding artifacts in the less interesting areas, which in turn is controversial to our aim of high visual quality. We therefore introduce a scene composition concept, where we only display the last n active speakers, which could be located at different terminals. The transmission of a thumbnail to the server becomes thus superfluous and further bandwidth can be saved. Note that as long as our proposed scene composition is used, the complete background could be omitted leading to an even lower data rate. However, to ensure interoperability with other video conferencing systems that do not include a scene composition, at least a low quality background has to be transmitted.

Therefore, in contrast to a classical video conferencing system, we need active video and audio analysis to detect participants at each terminal and separate individual speakers. Fig. 2 gives an overview of the system and shows the interaction of video analysis, audio analysis, video encoding and scene composition in the sending and receiving client. The video analysis consists of a face detection and an object tracking part. It steers the region of interest encoder with information about the position of all conferees and provides the audio analysis with information about their angle in the room. The audio analysis can then separate each speaker and create an activity index for each individual speaker. This information together with the ROI encoded video will be transmitted to a media mixer that decides which individual person is visible at which client. Finally, the receiving client decodes the video streams and displays the last n active speakers on the screen.

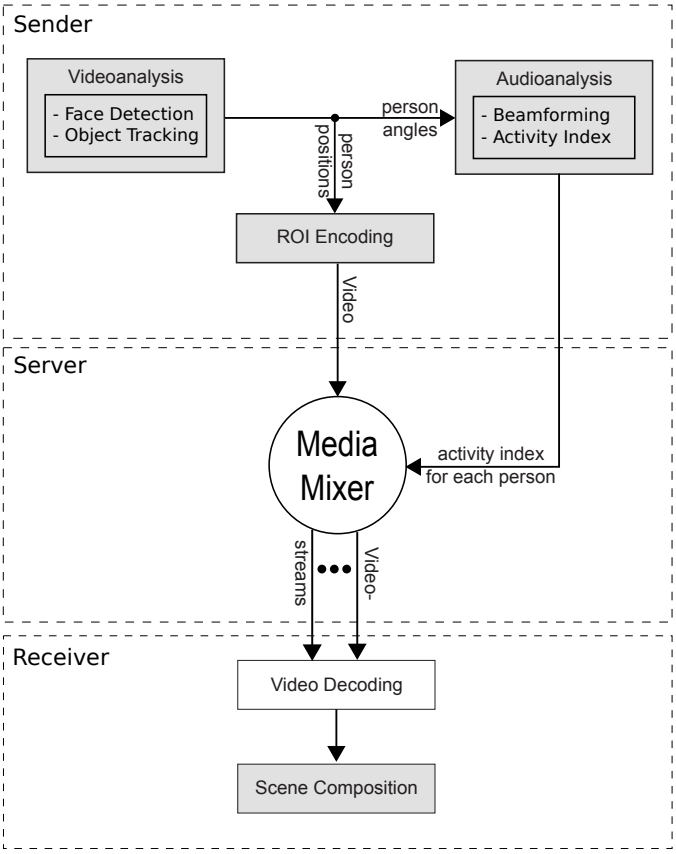


Fig. 2. System overview. Interaction of videoanalysis, audioanalysis, video encoding and scene composition in sending and receiving client.

The following subsections will explain each component in detail. We start with an explanation of the video analysis part in Section II-A. The modification on the video encoder will be presented in Section II-B and Section II-C describes the audio analysis part. Finally, Section II-D explains the CoVR scene composition.

A. Video Analysis

In order to feed the encoder with information about the position of all conferees and to provide information about their angle in the room for the audio analysis, an analysis of the video data is performed. The conferees will be detected with a

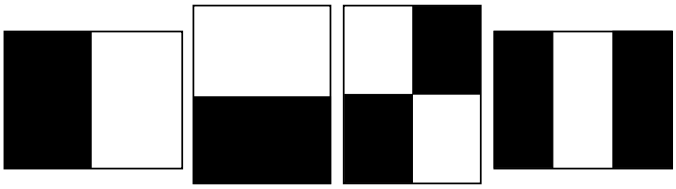


Fig. 3. Rectangle features used for face detection. Left to right: horizontal and vertical two-rectangle features, diagonal four-rectangle feature and horizontal three-rectangle feature.

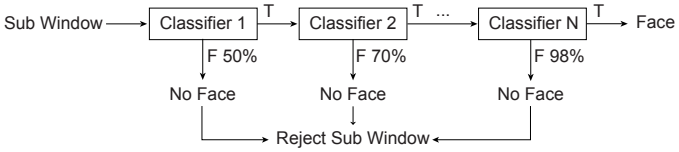


Fig. 4. Cascaded classifier structure. Simple classifier reject many negative sub-windows while complex classifiers reduce the false positive rate.

Viola Jones object detector [6], that has been trained to detect frontal views of faces. Once a face of a conferee has been detected a Mean Shift [7] tracker will be initialized to track it. The tracker is necessary for two reasons: The face detection algorithm may not provide a result in every frame, however, the encoder expects a result for each frame. Tracking of the detected persons across consecutive frames will provide the encoder with the necessary information in those frames. A second motivation for the use of a tracker is given by the fact that persons may not look at the camera all the time. In this case, the face detector would not be able to detect these persons which might finally lead to a classification of these areas as not being of interest.

In the following subsections, we will explain the used face detection and tracking algorithms in detail.

1) Face detection: Our face detection algorithm is based on the Viola-Jones object detection framework [6]. It has three key components. In a first step, a learning algorithm selects significant features in order to build efficient classifiers. The features used in this classifiers are Haar like and can be computed efficiently using an integral image representation. To speed up the classification process the single classifiers will be combined in a cascade.

Fig. 3 depicts exemplarily the features that were used in the object detection system. The response of each feature is the sum of all pixels inside the black area subtracted form the sum of all pixels inside the white area. Using an alternative image representation, the integral image $II(x, y)$, these features can be computed very efficiently:

$$II(x,y) = \sum_{x' \leq x, y' \leq y} I(x', y') \ , \tag{1}$$

with $I(x', y')$ denoting the original image.

The integral image allows for the computation of the sum of all pixels inside a rectangle with only four memory access operations. The response of each feature can thus be computed very efficiently. The features are so called weak features, that means, that a classifier based on each single feature is only able to distinguish between a face and something else in a limited extend. However, a combination of these weak classifiers can yield a strong classifier.

For a detection window of 24x24 pixel the entire set of possible rectangle features is about 45000. Since not all of them are necessary to detect faces in an image, a set of significant features has to be selected from all possible features which is done by AdaBoost [8].

Given a set of positive and negative training examples, the rectangle features that best separate the positive and negative examples need to be selected. The learning algorithm therefore determines the optimal threshold for a classification function such that the minimum number of examples are misclassified. The weak classifier $h_j(\mathbf{x})$ is then given by the function:

$$h_j(\mathbf{x}) = \begin{cases} 1, & \text{if } p_j f_j(\mathbf{x}) \leq p_j \theta_j \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

with f_j denoting the feature, θ_j a threshold, p_j a parity for the direction of the inequality and $\mathbf{x}$ a patch of the image.

The final classifier $h(\mathbf{x})$ is then a linear combination of the selected weak classifiers:

$$h(\mathbf{x}) = \begin{cases} 1, & \text{if } \sum_{j=1}^{J} w_j h_j(\mathbf{x}) \leq \frac{1}{2} \sum_{j=1}^{J} w_j \\ 0, & \text{otherwise} \end{cases} \tag{3}$$

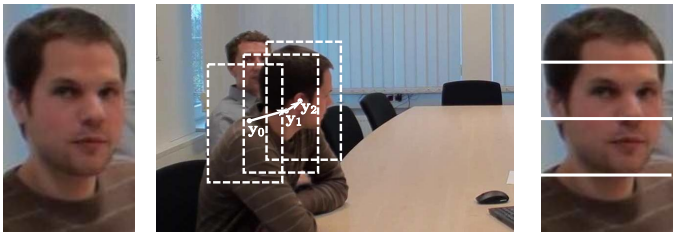
with J denoting the total number of weak classifier and w_j a specific weight for each weak classifier. More information on the determination of the weights can be found in [6].

In order to reduce computation time and increase the detection performance the classifiers are arranged in a cascaded structure. An example of such a structure is depicted in Fig. 4. Classifiers with relatively large false positive rates at the beginning of the cascade can be used to reject many negative sub-windows. Computationally more complex classifiers are used at the remaining sub-windows to reduce the false positive rate. The idea is motivated by the fact that many sub-windows within an image won't contain a face.

For example, a single rectangle feature classifier at the beginning of the cascade can be adjusted to detect 100% of the faces while rejecting 50% of all negative sub-windows. This simple classifier can thus significantly reduce the number of sub-windows for subsequent classification stages if subsequent stages are evaluated just in case of a positive result from the previous stage. A negative result in any of the classification stages causes the sub-window to be rejected.

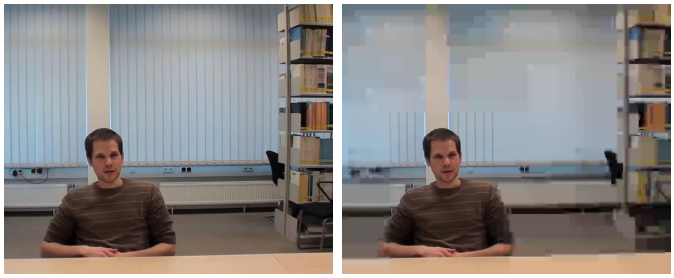
2) Mean Shift Tracking: Since the face detection does not provide a detection result for each frame, a tracking of the face positions across consecutive frames is necessary. In the general case, given the object location and its representation in frame t we want to estimate the object location in frame $t+1$. We will use a Mean Shift based tracking algorithm in order to fulfill this task. Mean Shift is an iterative technique for locating the mode of a density estimation based on sample observations $\{\mathbf{x}_n\}$ [7]. In the context of video object tracking, the samples $\{\mathbf{x}_n\}$ represent the pixel positions within the object region. In the following, we will refer to the object that will be tracked as target, while possible locations of that object will be denoted as target candidates.

Let a kernel function G be given, the Mean Shift procedure estimates the new position of the target candidate $\mathbf{y}_j$ based on



(a) holistic target representation (b) target candidates and their position estimations (c) multi-part target representation

Fig. 5. Target representation and new location estimation by iterative mean shift updates.



(a) QP 26 in ROI and background (b) QP 26 in ROI and 51 in background

Fig. 6. Comparison of image qualities within and outside of region of interest.

a previous estimate of the target candidate position $\mathbf{y}_{j-1}$ as follows:

$$\mathbf{y}_j = \frac{\sum_{n=1}^{N} w_n \mathbf{x}_n G\left(\frac{\mathbf{y}_{j-1}-\mathbf{x}_n}{h}\right)}{\sum_{n=1}^{N} w_n G\left(\frac{\mathbf{y}_{j-1}-\mathbf{x}_n}{h}\right)} \tag{4}$$

Here, N denotes the number of pixels within the object region, h the width of the kernel and w_n the weight at pixel position $\mathbf{x}_n$. The actual weight is given by:

$$w_n = \sum_{u=1}^{M} \sqrt{\frac{q_u}{p_u(\mathbf{y}_0)} \delta(b(\mathbf{x}_n) - u)} \ , \tag{5}$$

with the normalized kernel-weighted M-bin target and candidate histograms $\mathbf{q} = \{q_u\}_{u=1,\ldots,M}$ and $\mathbf{p}(\mathbf{y}) = \{p_u(\mathbf{y})\}_{u=1,\ldots,M}$:

$$q_u = C \cdot \sum_{n=1}^{N} K(\mathbf{y}_0 - \mathbf{x}_n)\delta(b(\mathbf{x_n} - u)) \tag{6}$$

$$p_u(\mathbf{y}) = C_h \cdot \sum_{n=1}^{N} K\left(\frac{\mathbf{y} - \mathbf{x}_n}{h}\right) \delta(b(\mathbf{x_n} - u)) \ . \tag{7}$$

Here, u denotes an index of a histogram bin, $b(\cdot)$ yields the bin index of the color at pixel location $\mathbf{x}_n$, $\delta(\cdot)$ is the Kronecker delta function and C and C_h are normalization constants.
The kernel functions $K(\mathbf{x})$ and $G(\mathbf{x})$ are connected through their individual profiles $k(x)$ and $g(x)$ for which $g(x) = -k'(x)$ holds [7].

Because the appearance of the target may change over time (e.g., due to a change in the lighting or a change of the 3D object pose), we will update the target representation in each frame:

$$\mathbf{q}_t = \alpha \mathbf{q}_{t-1} + (1-\alpha)\mathbf{p}(\mathbf{y}_{final})_t \ , \ 0 \leq \alpha \leq 1 \ . \tag{8}$$

Fig. 5 shows an example of the iterative Mean Shift procedure in a possible conference scenario. The target is depicted in Fig. 5a, the target candidates and the estimated locations as well as the final object location in Fig. 5b.

In order to get a more distinct object representation and thus an improved and robust tracking result, we divide our object representations according to [9] into parts which will be tracked separately. Fig. 5c shows an example of such a multi-part object representation. In contrast to the holistic representation illustrated in Fig. 5a, a multi-part representation provides information about the distribution of features for each subregion of the object. Further information especially about the influence of the object representation on the performance of the stability of the tracker is given in [10].

B. ROI Video encoding

Implementing a region of interest algorithm strongly alters the behavior of encoders and creates vastly different visual results. A traditional H.264/AVC encoder compresses a video stream, composed by a sequence of frames, by representing the content of these frames in a more efficient way. Although this compression is lossy, resulting in non-recoverable loss of image content, the effects are usually barely noticeable to the viewer. Rate distortion optimization makes sure that content with high importance to the viewer's perception of the videos quality, e.g., high frequency parts like the contours of a face or the pattern on a plant, is compressed less aggressively than content that contributes little to the viewer's perception of the videos quality. Fig. 6a shows a scene with a person at a desk, and a bookshelf in the background; the scene is compressed with a standard H.264/AVC that uses the same quantization parameter (QP) for the person and the bookshelves, thus showing both in about the same visual quality - the contours of both the person and the bookshelf are clearly identifiable, because both contribute equally to the overall visual quality. While this approach is very natural and pleasing to the human eye, it does not take the viewers attention into account: in a video-conference setting we are more interested in the person talking than in the books on the shelves. Taking the viewers attention into account means that the encoder should increase the quality of objects that are currently capturing the viewers attention, while paying for this increase in quality with lower quality on anything that is not important to the viewer; consequently, the goal of region of interest encoding is to redistribute bits for image compression from areas with little interest to areas with high interest. Fig. 6b shows a very extreme case of ROI encoding, where the bookshelf and the background outside the ROI is now encoded in a much lower quality (higher QP) than the face of the person.

A region of interest in its simplest form is a rectangle containing the object of highest interest. In the case of video conferencing this is the face of the person currently speaking and the immediate area around it. However, the shape of the ROI is not limited to a rectangle but is flexible in shape as

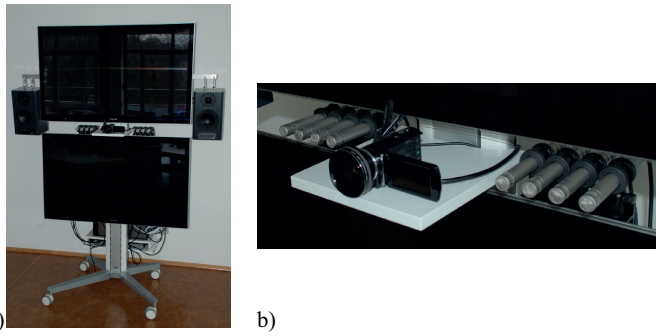
well as in the distribution of weights within the region.

A final thought should be given to H.264/AVC standard compliance. While it is possible to implement proprietary solutions that require an encoder and decoder pair capable of understanding the implemented region of interest algorithm, it is much preferred to stick to the current H.264/AVC video coding standard. Video-conferencing, just like telephone-conferencing, first and foremost requires interoperability. Consequently, a region of interest implementation may only modify the encoder, but must leave the decoder untouched, resulting in decodable content by every standard compliant decoder.

Taking all these conditions into account, we chose the modification of the quantization parameters for each individual macro-block (MB), similar to the approach by Ferreira et al. [11]. In H.264/AVC each frame is divided into MBs, each with a dimension of 16x16 pixels. These MBs are then transformed into the frequency domain using the discrete cosine transform (DCT), and are then quantized before entropy encoding [12]; the decoder performs the inverse steps to recover the final frame. Quantization is used to increase compression efficiency by mapping a large set of data to a smaller set of data. This operation is lossy and introduces a quantization error into the reconstructed values. By applying this technique to the transform coefficients the amount of coded data as well as the quality of the reconstructed picture can be controlled. In H.264/AVC, the quantization can be controlled by a quantization parameter ranging from 0 to 51, 0 being the finest quantization and 51 the coarsest.

We implemented ROI encoding in the MainConcept H.264/AVC encoder by quantizing the MBs within areas of low interest very coarsely, e.g., with QPs in the range from 40 to 51, while quantizing MBs of interesting parts more finely to preserve as much of the original values as possible. Our approach generalizes the approach by Ferreira et al. [11] by allowing arbitrary values for the region of interest. As an example region of interest may include fading, e.g., values of 22 on the MBs covering the face of the active speaker, values of 28 in the MBs adjacent to the face and then QPs of 51 for the remaining background regions. Another reason for allowing a more flexible quantization of the MBs describing a region of interest are our two main use cases for video-conferencing: Without scene composition one will always view the entire frame in contrast to scene composition where parts of the frame are cropped, typically only showing the person and immediately adjacent content; since large parts of the frame are not even seen during scene composition the quantization can easily be set to 51 for the background region that will be discarded during scene composition; likewise, without scene composition the less interesting MBs would probably not be quantized so harshly because they are clearly seen and are, while arguably less interesting, still negatively impacting the perception of quality due to the blocky nature of coarsely quantized MBs.

The quantization parameters for each MB are stored in an array which is the output of the face tracking algorithm. For convenience and to give extra room to rate distortion optimization and rate-control, we changed the values from 0 to 51 to 100 to 0, indicating the percentage of interest the viewer has in a MB - with a value of 0 resulting in the coarsest quantization and a value of 100 resulting in the finest quantization available. We chose to receive a QP array every frame, to allow for maximum flexibility for a region of interest, even though the region typically does not change rapidly due to the fact that people are rarely moving dramatically to warrant constant changes in the ROI.



a) b)

Fig. 7. Hardware setup of the video conferencing system:
a) complete system
b) detail view of the microphone array and the camera.

The benefit of this approach is a flexible region of interest, implemented into the H.264/AVC encoder without breaking standard compliance. This way any client with a compliant H.264/AVC decoder can decode the video. The downside of this approach is the MB based structure which can create blocky artifacts particularly with a very coarse quantization. Furthermore, a region of interest that resembles the exact contours of a face is also not possible due to the block based approach.

C. Audio Analysis

The information from the video analysis stage about the detected participants of the video conference as described in Section II-A is not only the basis for the ROI encoding but also for the audio analysis. The information about the position is exploited by a beamforming algorithm that allows to flexibly target different areas in front of the video conferencing system. With this system, the signals of the participants can be separated, which allows to quantify the current activity of each participant.

This section begins with a short introduction of the hardware setup that forms the basis of the audio analysis which is then described in two parts. First, the beamforming algorithm is presented followed by the determination of the speaker activity.

1) Microphone Array Design: The acoustic analysis of speaker activity requires the use of multiple microphones. A microphone array was specifically designed for a near field beamforming algorithm in a video conference szenario. The algorithm is utilized to extract separate acoustic signals for all speakers that were detected by the video analysis stage as described in Section II-A.

The design of a microphone array for video conferences in general has to consider the constraints that are given by the application. In the CoVR project, the main constraint is the

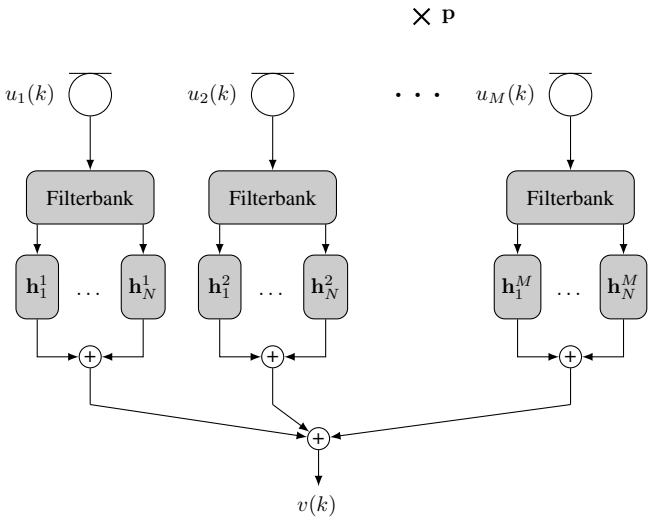


Fig. 8. Filter-and-sum beamformer with M Microphones and N non-uniform sub-bands.

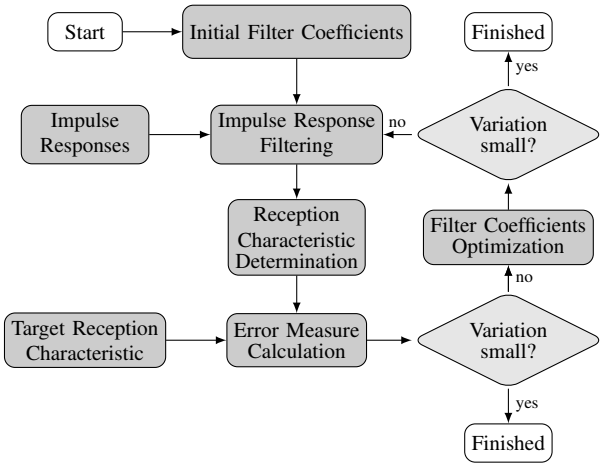


Fig. 9. Block diagram of the optimization process for each frequency sub-band n.

integration of the different components that are necessary for the visual aspects of a business video conference: two video screens and a video camera. The chosen solution is depicted in Fig. 7-a). The most convenient place to integrate a microphone array that allows to extract spatial properties of the acoustic environment is inbetween the displays. Due to the position of the video camera, two groups of four microphones each on both sides of the camera proved to be the best solution. The final setup is depicted in Fig. 7-b).

2) Beamforming Algorithm:

a) System Overview: The developed beamforming algorithm belongs to the class of filter-and-sum beamformers. The most important aspect when designing such a beamformer algorithm is the optimization of the filter coefficients. The target within the video conferencing system is to quantify the speaker's activities. Since the participants are usually located fairly close to the conferencing system, a numerical optimization procedure for the filter coefficients was developed [13], [14], [15] which exploits the acoustic properties of the room including the near field.

During the optimization procedure a predefined reception characteristic is approximated which can be chosen according to the application, e.g., extracting a specific speaker. A simplified block diagram of the proposed microphone array system is depicted in Fig. 8. It consists of a filterbank with N sub-bands followed by different filter-and-sum units represented by the impulse responses $\mathbf{h}_n^m$ $m \in \{1, \ldots, M\}$, $n \in \{1, \ldots, N\}$ with n denoting the sub-band index and m the microphone index at all M microphones. The samples $u_m(k)$ are obtained by analog-digital conversion with a sampling frequency of $f_s = 48\,\mathrm{kHz}$, where k is the discrete time index.

To obtain an almost uniform broadband reception characteristic independently of the operating frequency, a non uniform filterbank [16] is applied to subdivide each microphone signal into N frequency sub-bands. The optimization of the filter-and-sum units is carried out in these frequency bands. Thereby the degrees of freedom for the filter coefficients determination are increased.

A point source $s(k)$ is assumed to be at position $\mathbf{p}$ on an appropriately chosen spatial grid (e.g., in a two-dimensional Cartesian coordinate system: $\mathbf{p} = (x\ y)^T$). With the impulse responses $h_{\mathbf{p}m}(k)$ from the point source to each microphone the microphone signal $u_m(k)$ can be expressed as:

$$u_m(k) = h_{\mathbf{p}m}(k) * s(k). \tag{9}$$

The output $v(k)$ depends on the source location $\mathbf{p}$ and can be calculated according to:

$$v(k) = \sum_{m=1}^{M} \sum_{n=1}^{N} h_n^m(k) * \left(h_n^{\mathrm{FB}}(k) * u_m(k)\right), \tag{10}$$

where $h_n^{\mathrm{FB}}(k)$ represents the filterbank and $h_n^m(k)$ the FIR sub-band filters of length L_h.

For the determination of appropriate impulse responses $\mathbf{h}_n^m$ of the filter-and-sum units an iterative numerical optimization process was developed. The basic concept is depicted in Fig. 9. The optimization method is based on measured or simulated impulse responses $h_{\mathbf{p}m}$. By exciting the system with dirac impulses, the reception characteristic, i.e., the spatial distribution of damped and amplified areas, is determined. The procedure consists of an iterative minimization of an error measure, which is the summed level difference between the predefined target reception characteristic and the calculated one, based on the current state of the filter coefficients $\mathbf{h}_n^m$.

The current reception characteristic can be computed by the following three steps:

- simulating or measuring impulse responses between points on an appropriately chosen spatial grid in the near field and all microphones,
- processing these impulse responses with the sub-band filter-and-sum beamformer (see Fig. 8) to get an overall filter for every point in the near field, and
- determining the amplification and damping for every point from these overall filters.

Therefor the output signal $v(k)$ has to be calculated for each source location $\mathbf{p}$ on the spatial grid which can be expressed

as a filtered version of the source signal:

$$v(k) = \sum_{m=1}^{M} \sum_{n=1}^{N} h_n^m(k) * h_n^{\mathrm{FB}}(k) * h_{\mathbf{p}m}(k) * s(k), \quad (11)$$

the overall filter $g_{\mathbf{p}}(k)$ is thus obtained as:

$$g_{\mathbf{p}}(k) = \sum_{m=1}^{M} \sum_{n=1}^{N} h_n^m(k) * h_n^{\mathrm{FB}}(k) * h_{\mathbf{p}m}(k). \quad (12)$$

The frequency transform of the overall filter $g_{\mathbf{p}}(k)$ results in:

$$G_{\mathbf{p}}(f) = \mathcal{F}\left\{g_{\mathbf{p}}(k)\right\} . \quad (13)$$

Finally the reception characteristic $S_{\mathbf{p}}(f)$ in dB can be obtained at frequency f for every point $\mathbf{p}$ in the vicinity of the microphone array by:

$$S_{\mathbf{p}}(f) = 20 \cdot \log_{10} \left|G_{\mathbf{p}}(f)\right| . \quad (14)$$

This calculated reception characteristic is compared with a predefined target reception characteristic $\hat{S}_{\mathbf{p}}(f)$. The target reception characteristic is specified as a spatial distribution of areas with defined amplification $\mathbb{P}_{\mathrm{high}}$ (target level S_{high}) or damping $\mathbb{P}_{\mathrm{low}}$ (target level S_{low}) in front of the microphone array. It can be defined individually for all frequencies but a frequency-independent target is suitable for many applications:

$$\hat{S}_{\mathbf{p}}(f) = \hat{S}_{\mathbf{p}} = \begin{cases} S_{\mathrm{high}} & \text{for } \mathbf{p} \in \mathbb{P}_{\mathrm{high}} \\ S_{\mathrm{low}} & \text{for } \mathbf{p} \in \mathbb{P}_{\mathrm{low}}. \end{cases} \quad (15)$$

The precise choice of the areas and levels depends on *a priori* knowledge from the application, e.g., in a conferencing scenario, where the target speaker activity should be calculated, the reception characteristic is defined by the number and position of possible speakers. This information is provided by the video analysis.

A quadratic error measure Δ_S between the two reception characteristics is determined as the summed level difference for all points where $\hat{S}_{\mathbf{p}}(f)$ is set according to (15) and over all frequencies f_i ($i \in \{i_{\min}, \ldots, i_{\max}\}$):

$$\Delta_S(n) = \sum_{i=i_{\min}}^{i_{\max}} \sum_{\mathbf{p} \in (\mathbb{P}_{\mathrm{high}} \cup \mathbb{P}_{\mathrm{low}})} \hat{S}_{\mathbf{p}}(f_i) - S_{\mathbf{p}}(f_i), \quad (16)$$

where $f_{i_{min}}$ and $f_{i_{max}}$ denote the lower and upper edge frequencies of sub-band n.

Based on the error measure $\Delta_S(n)$ the optimum filter coefficients for each sub-band n are determined in a minimum mean square error (MMSE) sense by:

$$\left[\mathbf{h}_n^1, \ldots, \mathbf{h}_n^M\right]_{\mathrm{opt}} = \arg\min_{\mathbf{h}} \Delta_S(n)^2 . \quad (17)$$

The optimization is carried out by an iterative interior-point algorithm [17]. The algorithm checks whether the filter coefficients change between each iteration. If the change is sufficiently small, the algorithm is terminated. The optimization of the filter coefficients can take place on the basis of generated and measured impulse responses.

3) Determination of Speaker Activity: The beamforming stage of the audio processing system allows to extract one audio signal $v_s(k)$ per participant s of the video conference. The determination of the activity $a(\lambda)$ of this participant in a 20 ms signal frame λ ($L = 960\ samples$) is done based on the corresponding extracted audio signal. In a first step, a short term energy of the signal of participant s is calculated by

$$V_s(\lambda) = \sum_{i=0}^{L-1} v_s^2(\lambda \cdot L + i) . \quad (18)$$

This energy fluctuates quite strongly on such a short timeframe so a smoothing of the energy is introduced according to

$$\bar{V}_s(\lambda) = \alpha \cdot \bar{V}_s(\lambda - 1) + (1 - \alpha) \cdot V_s(\lambda) . \quad (19)$$

The smoothing factor α is chosen as 0.9 to be able to adapt quickly to changes while the larger fluctuations are leveled out.

This smoothed energy could directly be used as the indicator of activity for the scene composition. However, it can be observed that the smoothed energy values increase steeply between situations with no activity and those with a lot of activity. Hence, as additional step, a mapping to a target scale range from 0 to 100 with 0 indicating no activity and 100 indicating a lot of activity is suitable. The activity index $a_s(\lambda)$ which is finally used for the scene composition is calculated by applying a sigmoid function

$$a_s(\lambda) = \frac{100}{1 + \mathrm{e}^{-\beta \cdot (\bar{V}_s(\lambda) - \gamma)}} . \quad (20)$$

The parameters of the sigmoid function are set as $\beta = 110$ and $\gamma = 0.05$. The resulting activity index is then rounded to the nearest integer and sent to the video encoding and scene composition modules of the video conferencing system.

D. Scene Composition

The proposed region of interest concept combined with the joint audio and video analysis offers the possibility to compose a video based on the detected persons at the receiving client. Inspired by the idea of telepresence video conference systems, which create the impression that all conference participants are sitting on the same table, and the fact that the focus of interest in a typical conference scenario is on the participating persons, an alternative video composition could be achieved by showing only the detected persons. Each person is scaled and placed side by side at the receiving client. This concept can be extended in that way, that only the last n most active speakers will be displayed at the receiving client. Determining the active speaker has been discussed in Section II-C3 and can be achieved through a combined audio and video analysis. The decision which person gets rendered at which client will be made by a central media mixing component that compares the activity indices of all participants.

Fig. 10 shows an example of the scene composition with three parties and three active participants rendered at each party. Of course, due to the fact that no client will render conferees from its own party, different clients may have different scene compositions.

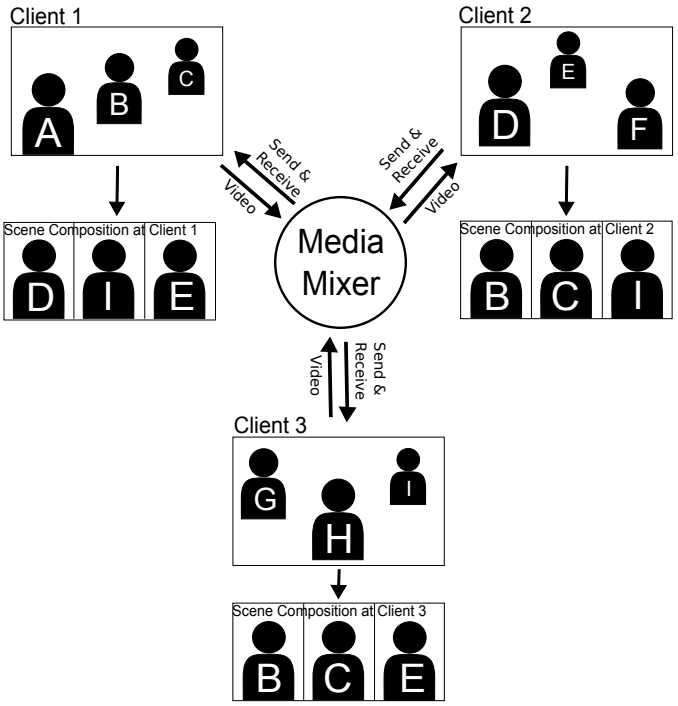


Fig. 10. Exemplary scene composition of three clients with three persons at each client. Different scenes are rendered at each client based on the decisions in the media mixer.

In addition to the advantage that our proposed scene composition depicts only relevant and active conference participants, the roughly quantized background gets discarded and the visual quantization artifacts depicted in Fig. 6 can be neglected. This kind of scene composition thus allows a very coarse quantization of the background what in turn yields large bandwidth savings (cf. Section III-A2).

III. Evaluation

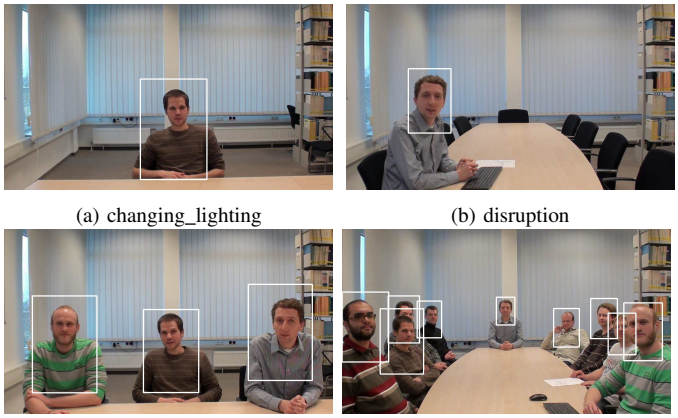
The system that is proposed in Section II has been implemented in a real time video conferencing prototype, which has also been demonstrated publicly.[2] However, due to the complexity of the system a detailed evaluation of the entire system is not feasible. Therefore, we focus our evaluation to some specific aspects of the system, such as the bitrate savings that can be achieved with region of interest encoding and the directivity improvement of the beamforming algorithm.

Due to the complexity of the task and the unavailability of reliable ground truth data, a detailed evaluation of the tracking algorithm and of the activity index is not included in this paper.

A. Region of Interest encoding

Our investigations focus on the bitrate savings achievable through region of interest (ROI) encoding in a video-conference. We thereby assume that the result of the detection and tracking algorithm is reliable. Our goals for visual quality differ when scene composition is turned on or off: in case of scene composition, most of the video is cropped, so the ROI encoding should achieve high bitrate reduction without regard to visual quality outside the ROI; without scene composition the effects of ROI encoding are directly visible to the viewer so our goal here was to find a sweet spot where bitrate savings and visual quality outside the ROI are in balance.

[2]Demonstration of video conferencing prototype at:
International Workshop on Acoustic Signal Enhancement (IWAENC'12), 2012, Aachen, Germany
Centrum für Büroautomation, Informationstechnologie und Telekommunikation (CeBIT'13), 2013, Hannover, Germany



(a) changing_lighting (b) disruption

(c) next_to_each_other (d) individual_spakers

Fig. 11. Sample Images of our test sequences with detected ROIs.

1) Test environment: In order to show the efficiency of our region of interest encoding approach we captured several videos with typical video-conferencing conditions. All of these videos have been recorded with a high-end consumer grade camera at a resolution of 720p and 50fps. All of the videos are 900 frames long. The videos *changing_lighting*, *disruption*, *next_to_each_other*, and *individual_spakers* show scenes with one, one, three and nine tracked people in them. Fig. 11 shows a typical frame from each of these videos. In addition to changing the number of tracked faces, we also included a change of light in the video *changing_lighting*: mid way through the video the light is turned off suddenly and gradually faded back in. Additionally, we included movement of a person in the video *next_to_each_other*. The area covered by the ROI box is 6% for *disruption*, 13% for *changing_lighting*, 23% for *next_to_each_other*, and 26% for the nine people video *individual_spakers*. For the quantization parameters of the ROI only two values have been chosen: all MBs inside the ROI have the same quantization value, just like anything outside has the same values.

The face tracker generates a box shaped region of interest sized with respect to the individual faces, showing head and shoulders. The region of interest encoding has been implemented in MainConcept's H.264/AVC encoder, based on MainConcept Codec SDK 9.5. The encoder itself has been configured to a low-delay setting suitable for video-conferences: bi-prediction disabled, base profile, one intra frame every 300 frames, constant quantization instead of rate-control, and deblocking turned on. The periodic intra frame allows for a re-entry into the decoding process, but does not allow frequent joining of a conference; whenever a new user joins the video-conference, a new intra frame is requested. Deblocking helps improve the visual quality for

highly compressed areas so it has been turned on for the whole test set.

The quantization parameters inside the ROI ranged from 18 to 34; values below 18 no longer provide improved visual quality for the viewer, values above 34 produce artefacts that make reading facial expressions difficult. The outside of the ROI is quantized with a step size which is a multiple of six; The quantization parameters outside the ROI range from +0, to create a non-ROI reference, until they reach +18 for a very coarse quantization.

2) Results: In Fig. 12 the encoder performance for different quantization values for the ROI and the non ROI region are shown for all four sequences in the test set. Each graph represents a constant QP difference between the ROI and the non ROI area. For QP Difference 0 the ROI and the non ROI regions use the same quantization so this is the reference for encoding not using ROI information. With higher QP difference values the quality of the non ROI region decreases. The peak signal-to-noise ratio (PSNR) measure only takes the area inside of the ROI into account.

We can see that especially at high bitrates the bandwidth savings using a coarser quantization for the background are enormous. For example, for the highest data point (ROI QP 22) in the sequence *changing_lighting* we save about 77% using a QP of 28 for the background (QP difference 6) or 86% using a QP of 34 (QP difference 12). However, such high bitrates are unrealistic to be used in video conferencing applications. A more realistic QP range is between QP 26 and 30 where the conventional video coding approach uses a bitrate of about 1-2 Mbit/sec. In this area our ROI based encoding approach yields a coding gain of approximately 50%.

In Table I, the Bjøntegaard delta rate (BD-rate [18]) savings are shown for the test set ar different QP differences between ROI and non ROI regions. Table II shows the average BD-

TABLE I. BD-RATE SAVINGS FOR THE TEST SET AT DIFFERENT QP DIFFERENCES.

QP Difference	Y	U	V
6	-52.75%	-59.45%	-58.86%
12	-57.91%	-64.98%	-64.08%
18	-54.99%	-63.22%	-63.89%

(a) *disruption*

QP Difference	Y	U	V
6	-53.52%	-60.26%	-59.87%
12	-55.49%	-63.94%	-64.00%
18	-57.27%	-61.12%	-64.25%

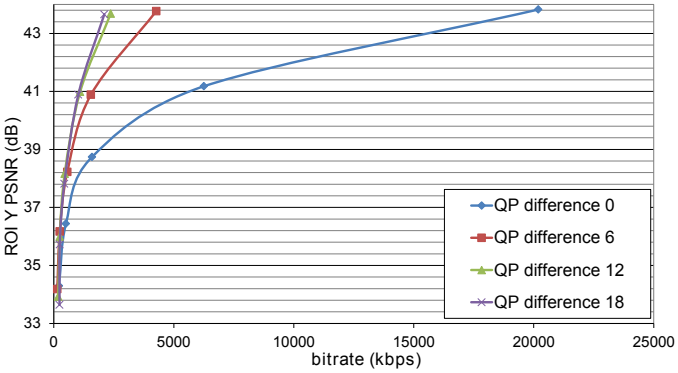
(b) *changing_lighting*

QP Difference	Y	U	V
6	-25.28%	-31.11%	-29.93%
12	-25.48%	-33.81%	-32.33%
18	-20.98%	-33.07%	-32.49%

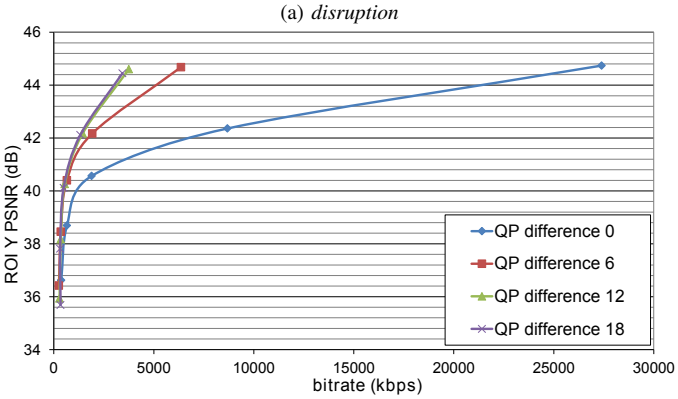
(c) *next_to_each_other*

QP Difference	Y	U	V
6	-42.47%	-44.17%	-45.12%
12	-46.75%	-48.08%	-48.42%
18	-46.35%	-47.59%	-48.50%

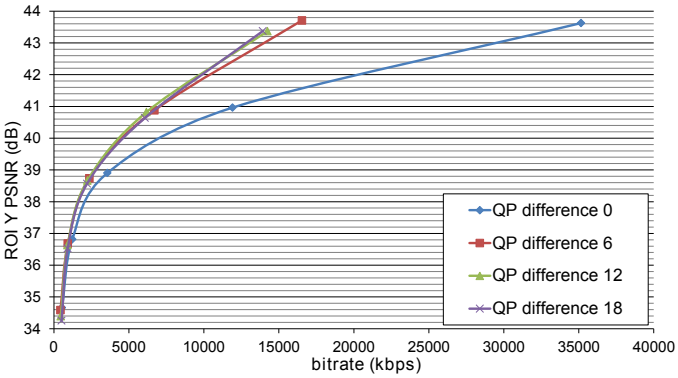
(d) *individual_speakers*



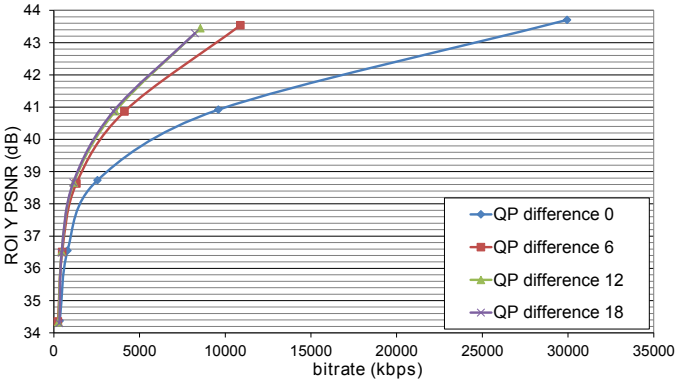
(a) *disruption*



(b) *changing_lighting*



(c) *next_to_each_other*



(d) *individual_speakers*

Fig. 12. ROI Y-PSNR vs Bitrate for each test sequence and different differences relations between the QP inside and outside of the ROI.

rate savings. In the table as well as in Fig. 12 one can see that the rate savings do not grow with the chosen QP difference. While a QP difference of 6 already gives great rate savings a difference of 12 or more does not further decrease the bitrate by the same magnitude. However, the perceived image quality of the not ROI regions suffers badly when the QP difference is increased to 12 or even 18.

TABLE II. AVERAGE BD-RATE SAVINGS FOR THE TEST SET AT DIFFERENT QP DIFFERENCES.

QP Difference	Y	U	V
6	-43.51%	-48.75%	-48.45%
12	-46.41%	-52.70%	-52.21%
18	-44.90%	-51.25%	-52.28%

B. Audio analysis

The performance of the beamformer design will be demonstrated by a comparison with the well known *Minimum Variance Distortionless Response* (MVDR) beamformer [19]. For an identical setup the reception characteristic of both is depicted in Fig. 13 and Fig. 14 at two different frequencies (500Hz and 2000Hz). For better comparison the simulation were carried out under free field conditions for the two algorithms.

1) Test environment: A desired reception characteristic can be defined in front of the microphone array (including the near field) for the numerical optimization of the filter coefficients. Both systems were designed in such way that acoustic sources on the left side ($-0.5\,\mathrm{m} \leq x < 0\,\mathrm{m} \wedge 0.2\,\mathrm{m} < y \leq 0.8\,\mathrm{m}$) are amplified while sources on the right side ($0\,\mathrm{m} < x \leq 0.5\,\mathrm{m} \wedge 0.2\,\mathrm{m} < y \leq 0.8\,\mathrm{m}$) are damped. In Figs. 13 and 14 those areas are marked by white edged boxes. The resolution of the spatial grid for each space dimension (x, y) was set to $1\,\mathrm{cm}$, which results in 3000 points for each area $\mathbb{P}_{\mathrm{high}}$ and $\mathbb{P}_{\mathrm{low}}$. A level difference of 40 dB between the amplified $\mathbb{P}_{\mathrm{high}}$ and damped $\mathbb{P}_{\mathrm{low}}$ areas was chosen.

The microphone array consists according to Fig. 7-b of $M = 8$ sensors with a spacing of [3, 3, 3, 30, 3, 3, 3] cm. Spatial alias can be expected for frequencies greater than approx. 5600 Hz. Thus, the behavior above this frequency can not be clearly controlled. For the developed algorithm a non uniform filterbank [16] is used, which consists of $N = 6$ sub-bands. The frequency range of the sub-bands are given in Table III. For simplicity all sub-band filters have been realized as FIR filters. The length of the impulse responses of the filter-and-sum units $\mathbf{h}_n^m$ was set to $L_h = 8$. The filter length of the MVDR beamformer was set to 96 for the complete frequency range and per microphone with susceptibility K_0 set to 3 [20]. This makes it twice as long as the effective filter length of the new system ($N \cdot L_h = 48$).

2) Results: Fig. 13 depicts the reception characteristics of the MVDR system at $500\,\mathrm{Hz}$ and $2000\,\mathrm{Hz}$. At $500\,\mathrm{Hz}$ the MVDR beamfomer is only able to achieve a low directivity. For the configuration at $2000\,\mathrm{Hz}$ a noticeable level difference between $\mathbb{P}_{\mathrm{high}}$ and $\mathbb{P}_{\mathrm{low}}$ is observable. However, in this case the system has a very inhomogeneous behavior in the stop-band.

TABLE III. FILTERBANK SUB-BANDS

Band	Frequency range [Hz]		Band	Frequency range [Hz]	
1	1	268	4	1549	2614
2	268	839	5	2614	4731
3	839	1549	6	4731	12049

The reception characteristic of the proposed beamformer (see Fig. 14) has a significantly improved directivity characteristic. A significant level difference is recognizable for both operating frequencies between both areas $\mathbb{P}_{\mathrm{high}}$ and $\mathbb{P}_{\mathrm{low}}$. Especially at the edge regions the predefined reception characteristic is well approximated.

In comparison with the widely used MVDR beamformer the new numerical optimization procedure is able to achieve a much better separation of (in this case two) speakers across a wide frequency range.

IV. Conclusion and future work

In this paper, we presented a system that combines video and audio analysis elements to provide information for region of interest based encoding and scene composition to improve the users video conferencing experience. The video analysis stage consists of a face detection with a suitably trained Viola Jones detector and a continuous tracking of found participants by a Mean Shift tracker. The combination of these two stages provides very stable and robust results even in adverse conditions. The information about the position of the participants is utilized as the foundation of both the ROI encoding and the audio analysis stage. The audio analysis is based on a microphone array setup that was specifically designed for the video conferencing system and uses a novel beamforming algorithm which is capable of directly using the information from the video analysis. The output of the beamformer is then used to quantify the activity of the participant. Hence, the combination of the video and the audio analysis allows to simultaneously achieve robust results regarding the position and the activity of every participant of the video conference.

It was shown that the region of interest based encoding allows to either save a significant amount of bandwidth or increase the quality of the video inside the ROI by choosing a coarser quantization for the non ROI regions. When this system is combined with our proposed scene composition, the non ROI regions and their coding artifacts are removed which improves the quality of the video conference. However, also without scene composition the user experience is enhanced by shifting the encoder focus into the regions that are interesting to the user.

The future work will focus on improving the accuracy of the face detection and tracking to provide reliable information also in difficult environments. Additionally, the shape of the ROI region can be better adapted to the speaker (e.g., give a higher priority to the face) then choosing a constant QP in a rectangular region around the face. The output of the beamformer is so far only utilized for the determination of the activity of every participant, the separated signals of the participants could also be used for an enhancement to the scene composition by matching the spatial properties of the audio signals to the spatial position of the participant on the display.

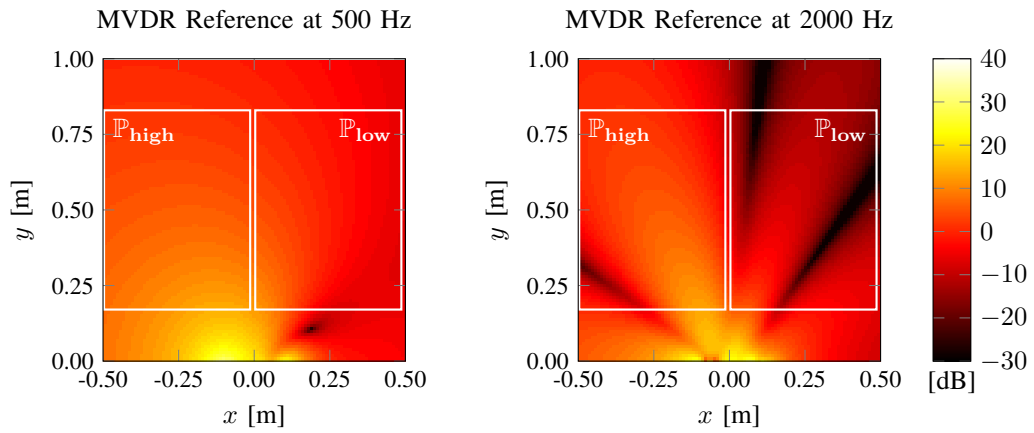


Fig. 13. Reception characteristics of the reference MVDR beamformer.

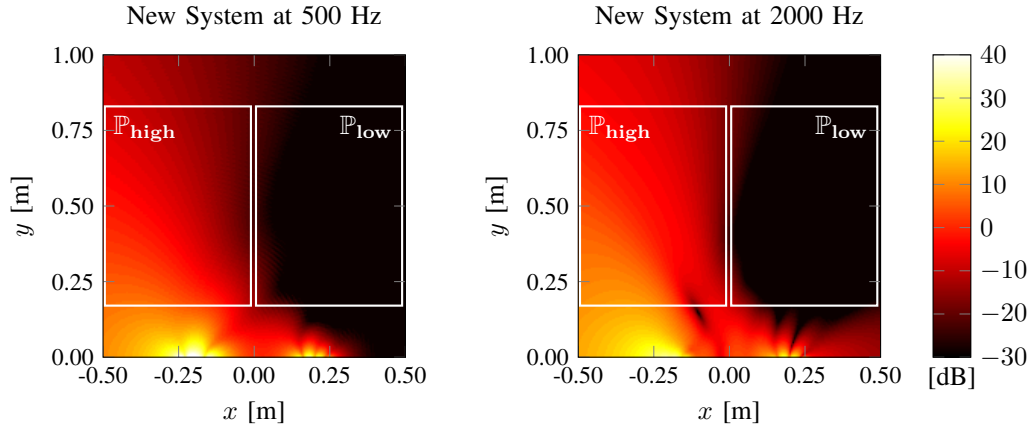


Fig. 14. Reception characteristics of the new beamforming algorithm.

REFERENCES

[1] C. Bulla, C. Feldmann, and M. Schink, "Region of Interest Encoding in Video Conference Systems," in *Proc. of International Conferences on Advances in Multimedia (MMEDIA)*, 2013, pp. 119–124.

[2] "Connected Visual Reality (CoVR) - High Quality Visual Communication in Heterogeneous Networks", 2013, joint project of Ericsson GmbH, MainConcept GmbH, part of Rovi, as well as two institutes of the RWTH Aachen: Institut für Nachrichtentechnik and Institute for Communication Systems and Data Processing, http://www.covr.rwth-aachen.de (last access: 15 Dez. 2013).

[3] C. Feldmann, M. Wien, J. Hsu, and F. Jäger, "Single-loop SNR scalability using Binary Residual Refinement Coding," 12th Meeting, Geneva, Switzerland, Tech. Rep. JCTVC-L0154, Jan. 2013.

[4] C. Feldmann, C. Bulla, and B. Cellarius, "Efficient Stream-Reassembling for Video Conferencing Applications using Tiles in HEVC," in *Proc. of International Conferences on Advances in Multimedia (MMEDIA)*, Venice, Italy, Apr. 2013, pp. 130–135.

[5] T. Schlien, F. Heese, M. Schäfer, C. Antweiler, and P. Vary, "Audiosignalverarbeitung für Videokonferenzsysteme," in *Workshop Audiosignal- und Sprachverarbeitung (WASP)*. Gesellschaft für Informatik, Sep. 2013, workshop im Rahmen der 43. Jahrestagung der Gesellschaft für Informatik.

[6] P. Viola and M. Jones, "Robust Real-Time Face Detection," *International Journal of Computer Vision*, vol. 57, pp. 137 –145, 2004.

[7] Y. Cheng, "Mean Shift, Mode Seeking, and Clustering," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 17, pp. 790–799, 1995.

[8] Y. Freund and R. Schapire, "A Decision-Theoretic Generalization of On-line Learning and an Application to Boosting," *Journal of Computer and System Sciences*, vol. 55, pp. 119–139, 1997.

[9] D. Caulfield and K. Dawson-Howe, "Evaluation of Multi-Part Models for Mean-Shift Tracking," in *Proc. of International Machine Vision and Image Processing Conference*, 2008, pp. 77–82.

[10] P. Hosten, A. Steiger, C. Feldmann, and C. Bulla, "Performance Evaluation of Object Representations in Mean Shift Tracking," in *Proc. of International Conferences on Advances in Multimedia (MMEDIA)*, 2013, pp. 1–6.

[11] L. Ferreira, L. Cruz, and P. Assunção, "H. 264/SVC ROI Encoding with Spatial Scalability," in *Proc. of International Conference on Signal Processing and Multimedia Applications*, 2008, pp. 212–215.

[12] T. Wiegand, G. Sullivan, G. Bjøntegaard, and A. Luthra, "Overview of the H. 264/AVC Video Coding Standard," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 13, pp. 560–576, 2003.

[13] M. Schäfer, F. Heese, J. Wernerus, and P. Vary, "Numerical Near Field Optimization of Weighted Delay-and-Sum Microphone Arrays," in *Proceedings of International Workshop on Acoustic Signal Enhancement (IWAENC)*, Sept. 2012.

[14] F. Heese, M. Schäfer, P. Vary, E. Hadad, S. M. Golan, and S. Gannot, "Comparison of Supervised and Semi-supervised Beamformers Using Real Audio Recordings," in *Proceedings of IEEE 27-th Convention of Electrical and Electronics Engineers in Israel (IEEEI)*, Nov. 2012.

[15] F. Heese, M. Schäfer, J. Wernerus, and P. Vary, "Numerical Near Field Optimization of a Non-Uniform Sub-band Filter-and-Sum Beamformer," in *Proceedings of IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 2013.

[16] H. W. Löllmann, "Allpass-Based Analysis-Synthesis Filter-Banks: Design and Application," Ph.D. dissertation, IND, RWTH Aachen University, Nov. 2011.

[17] R. H. Byrd, J. C. Gilbert, and J. Nocedal, "A Trust Region Method Based on Interior Point Techniques for Nonlinear Programming," *Mathematical Programming*, vol. 89, no. 1, pp. 149–185, 2000.

[18] G. Bjøntegaard, "Calculation of Average PSNR Differences between RD Curves," document VCEG-M33, ITU-T Q6/16, Austin TX, USA, Tech. Rep., Apr. 2001.

[19] P. Vary and R. Martin, *Digital Speech Transmission - Enhancement, Coding & Error Concealment*. John Wiley & Sons, Ltd., Jan. 2006.

[20] M. Dörbecker, "Mehrkanalige Signalverarbeitung zur Verbesserung akustisch gestörter Sprachsignale am Beispiel elektronischer Hörhilfen," Ph.D. dissertation, IND, RWTH Aachen University, Jul. 1998.

www.iariajournals.org

International Journal On Advances in Intelligent Systems

ICAS, ACHI, ICCGI, UBICOMM, ADVCOMP, CENTRIC, GEOProcessing, SEMAPRO, BIOSYSCOM, BIOINFO, BIOTECHNO, FUTURE COMPUTING, SERVICE COMPUTATION, COGNITIVE, ADAPTIVE, CONTENT, PATTERNS, CLOUD COMPUTING, COMPUTATION TOOLS, ENERGY, COLLA, IMMM, INTELLI, SMART, DATA ANALYTICS

issn: 1942-2679

International Journal On Advances in Internet Technology

ICDS, ICIW, CTRQ, UBICOMM, ICSNC, AFIN, INTERNET, AP2PS, EMERGING, MOBILITY, WEB

issn: 1942-2652

International Journal On Advances in Life Sciences

eTELEMED, eKNOW, eL&mL, BIODIV, BIOENVIRONMENT, BIOGREEN, BIOSYSCOM, BIOINFO, BIOTECHNO, SOTICS, GLOBAL HEALTH

issn: 1942-2660

International Journal On Advances in Networks and Services

ICN, ICNS, ICIW, ICWMC, SENSORCOMM, MESH, CENTRIC, MMEDIA, SERVICE COMPUTATION, VEHICULAR, INNOV

issn: 1942-2644

International Journal On Advances in Security

ICQNM, SECURWARE, MESH, DEPEND, INTERNET, CYBERLAWS

issn: 1942-2636

International Journal On Advances in Software

ICSEA, ICCGI, ADVCOMP, GEOProcessing, DBKDA, INTENSIVE, VALID, SIMUL, FUTURE COMPUTING, SERVICE COMPUTATION, COGNITIVE, ADAPTIVE, CONTENT, PATTERNS, CLOUD COMPUTING, COMPUTATION TOOLS, IMMM, MOBILITY, VEHICULAR, DATA ANALYTICS

issn: 1942-2628

International Journal On Advances in Systems and Measurements

ICQNM, ICONS, ICIMP, SENSORCOMM, CENICS, VALID, SIMUL, INFOCOMP

issn: 1942-261x

International Journal On Advances in Telecommunications

AICT, ICDT, ICWMC, ICSNC, CTRQ, SPACOMM, MMEDIA, COCORA, PESARO, INNOV

issn: 1942-2601