



# **ICDT 2011**

The Sixth International Conference on Digital Telecommunications

April 17-22, 2011

Budapest, Hungary

**ICDT 2011 Editors**

Constantin Paleologu, University Politehnica of Bucharest, Romania

# ICDT 2011

## Foreword

The Sixth International Conference on Digital Telecommunications [ICDT 2011], held between April 17 and 22 in Budapest, Hungary, continued a series of special events focusing on telecommunications aspects in multimedia environments. The scope of the conference was to focus on the lower layers of systems interaction and identify the technical challenges and the most recent achievements.

The conference served as a forum for researchers from both the academia and the industry, professionals, and practitioners to present and discuss the current state-of-the art in research and best practices as well as future trends and needs (both in research and practices) in the areas of multimedia telecommunications, signal processing in telecommunications, data processing, audio transmission and reception systems, voice over packet networks, video, conferencing, telephony, as well as image producing, sending, and mining, speech producing and processing, IP/Mobile TV, Multicast/Broadcast Triple-Quadruple-play, content production and distribution, multimedia protocols, H-series towards SIP, and control and management of multimedia telecommunications.

High quality software is not an accident; it is constructed via a systematic plan that demands familiarity with analytical techniques, architectural design methodologies, implementation polices, and testing techniques. Software architecture plays an important role in the development of today's complex software systems. Furthermore, our ability to model and reason about the architectural properties of a system built from existing components is of great concern to modern system developers.

Performance, scalability and suitability to specific domains raise the challenging efforts for gathering special requirements, capture temporal constraints, and implement service-oriented requirements. The complexity of the systems requires an early stage adoption of advanced paradigms for adaptive and self-adaptive features.

On online monitoring applications, in which continuous queries operate in near real-time over rapid and unbounded "streams" of data such as telephone call records, sensor readings, web usage logs, network packet traces, are fundamentally different from traditional data management.

The difference is induced by the fact that in applications such as network monitoring, telecommunications data management, manufacturing, sensor networks, and others, data takes the form of continuous data streams rather than finite stored data sets. As a result, clients require long-running continuous queries as opposed to one-time queries. These requirements lead to reconsider data management and processing of complex and numerous continuous queries over data streams, as current database systems and data processing methods are not suitable.

Event stream processing is a new paradigm of computing that supports the processing of multiple streams of event data with the goal of identifying the meaningful events within those streams.

We take here the opportunity to warmly thank all the members of the ICDT 2011 Technical Program Committee, as well as the numerous reviewers. The creation of such a broad and high quality conference program would not have been possible without their involvement. We also kindly thank all

the authors who dedicated much of their time and efforts to contribute to ICDT 2011. We truly believe that, thanks to all these efforts, the final conference program consisted of top quality contributions.

Also, this event could not have been a reality without the support of many individuals, organizations, and sponsors. We are grateful to the members of the ICDT 2011 organizing committee for their help in handling the logistics and for their work to make this professional meeting a success.

We hope that ICDT 2011 was a successful international forum for the exchange of ideas and results between academia and industry and for the promotion of progress in the field of digital communications.

We are convinced that the participants found the event useful and communications very open. We also hope the attendees enjoyed the historic charm of Budapest, Hungary.

**ICDT 2011 Chairs:**

Saied Abedi, Fujitsu Laboratories of Europe Ltd. (FLE), UK

Bilal Al Momani, Cisco Systems, Inc., Ireland

Gerard Damm, Alcatel-Lucent, France

Javier Del Ser Lorente, TECNALIA RESEARCH & INNOVATION - Zamudio, Spain

Klaus Drechsler, Fraunhofer Institute for Computer Graphics Research IGD - Darmstadt, Germany

Michael Grottke, University of Erlangen-Nuremberg, Germany

Constantin Paleologu, University Politehnica of Bucharest, Romania

Jyrki Penttinen, Nokia Siemens Networks, Spain / Helsinki University of Technology, Finland

Reda Reda, ICT-MC, Austria

Dan Romascanu, Avaya, Israel

Sandra Sendra Compte, Polytechnic University of Valencia, Spain

Weilian Su, Naval Postgraduate School - Monterey, USA

Tomohiko Taniguchi, Fujitsu Laboratories Limited, Japan

Abdulrahman Yarali, Murray State University, USA

# ICDT 2011

## Committee

### ICDT Steering Committee

Reda Reda, ICT-MC, Austria  
Constantin Paleologu, University Politehnica of Bucharest, Romania  
Tomohiko Taniguchi, Fujitsu Laboratories Limited, Japan  
Abdulrahman Yarali, Murray State University, USA

### ICDT Advisory Chairs

Michael Grottke, University of Erlangen-Nuremberg, Germany  
Weilian Su, Naval Postgraduate School - Monterey, USA  
Javier Del Ser Lorente, TECNALIA RESEARCH & INNOVATION - Zamudio, Spain

### ICDT Industry Liaison Chairs

Bilal Al Momani, Cisco Systems, Inc., Ireland  
Saied Abedi, Fujitsu Laboratories of Europe Ltd. (FLE), UK  
Gerard Damm, Alcatel-Lucent, France  
Dan Romascanu, Avaya, Israel

### ICDT Research/Industry Chairs

Jyrki Penttinen, Nokia Siemens Networks, Spain / Helsinki University of Technology, Finland  
Klaus Drechsler, Fraunhofer Institute for Computer Graphics Research IGD - Darmstadt, Germany

### ICDT Publicity Chair

Sandra Sendra Compte, Polytechnic University of Valencia, Spain

### ICDT 2011 Technical Program Committee

Saied Abedi, Fujitsu Laboratories of Europe Ltd. (FLE), UK  
Akbar Sheikh Akbari, Staffordshire University, UK  
Bilal Al Momani, Cisco Systems, Inc., Ireland  
Maria Teresa Andrade, FEUP / INESC Porto, Portugal  
Regina B. Araujo, Federal University of São Carlos, Brazil  
Koichi Asatani, Kogakuin University - Tokyo, Japan  
Ilija Basicevic, University of Novi Sad, Serbia  
Carlos Becker Westphall, Federal University of Santa Catarina, Brazil  
Andi Buzo, University Politehnica of Bucharest, Romania  
Doru Constantin, University of Pitesti, Romania  
Gerard Damm, Alcatel-Lucent, France  
Javier Del Ser Lorente, TECNALIA RESEARCH & INNOVATION, Spain  
Klaus Drechsler, Fraunhofer-Institut für Graphische Datenverarbeitung IGD - Darmstadt, Germany  
António Manuel Duarte Nogueira, University of Aveiro, Portugal  
Andrei Alexandru Enescu, Politehnic University of Bucharest, Romania  
Gerardo Fernández-Escribano, University of Castilla-La Mancha, Spain

Mário Ferreira, University of Aveiro, Portugal  
Alex Galis, University College London, UK  
Andrea Giachetti, Università degli Studi di Verona, Italy  
Stefanos Gritzalis, University of the Aegean, Greece  
Michael Grottke, University of Erlangen-Nuremberg, Germany  
Yang Guo, Research and Innovation, Technicolor, USA  
Georgi Iliev, Technical University of Sofia, Bulgaria  
Theodoros Iliou, Greek Society of Scientists / Greek Computer Society, Greece  
Yuri V. Ivanov, Movidius Ltd., Ireland  
Norbert Jordan, Vienna University of Technology, Austria  
Carlos Juiz, University of the Balearic Islands, Spain  
Ki Hong Kim, The Attached Institute of ETRI, Republic of Korea  
Wen-hsing Lai, National Kaohsiung First University of Science and Technology, Taiwan  
Sathiamoorthy Manoharan, University of Auckland, New Zealand  
Antonio Marcos Alberti, INATEL - Instituto Nacional de Telecomunicações, Brazil  
Jesús Martínez, Universidad Politécnica de Madrid, Spain  
Andreas U. Mauthe, Lancaster University, UK  
Parag S. Mogre, Technische Universität Darmstadt, Germany  
Huan Nguyen, Glasgow Caledonian University, UK  
Patrik Österberg, Mid Sweden University, Sweden  
Constantin Paleologu, University Politehnica of Bucharest, Romania  
Gerard Parr, University of Ulster - Coleraine, UK  
Encarna Pastor, Universidad Politécnica de Madrid, España  
Pubudu Pathirana, Deakin University, Australia  
Jyrki Penttinen, Nokia Siemens Networks, Spain / Helsinki University of Technology, Finland  
Raquel Pérez Leal, Universidad Politécnica de Madrid (UPM), Spain  
Roger Pierre Fabris Hoeffel, Federal University of Rio Grande do Sul, Brazil  
Thilo Pionteck, University of Lübeck, Germany  
Reda Reda, ICT-MC, Austria  
Dan Romascanu, Avaya, Israel  
Tapio Saarelainen, National Defence University, Finland  
Sattar B. Sadkhan, University of Babylon, Iraq  
Nabil J. Sarhan, Wayne State University - Detroit, USA  
Stefan Schmid, Deutsche Telekom Laboratories (T-Labs) / TU Berlin, Germany  
Ahmad M. Shafiee, Islamic Azad University (Najafabad branch), Iran  
Charalampos Skianis, University of the Aegean, Greece  
Himanshu B. Soni, IIT Bombay, India  
Leonel Sousa, INESC-ID/IST, TU-Lisbon, Portugal  
Dirk Staehle, University of Würzburg, Germany  
Weilian Su, Naval Postgraduate School - Monterey, USA  
Nary Subramanian, The University of Texas at Tyler, USA  
Hailong Sun, Beihang University - Beijing, China  
Tomohiko Taniguchi, Fujitsu Laboratories Limited, Japan  
Tony Thomas, Nanyang Technological University, Singapore  
Božo Tomas, University of Mostar, Bosnia and Herzegovina  
Dimitrios D. Vergados, University of Piraeus, Greece  
José Miguel Villalón Millan, Universidad de Castilla-La Mancha, Spain  
Krzysztof Walczak, Poznan University of Economics, Poland  
Qishi Wu, University of Memphis, USA  
Abdulrahman Yarali, Murray State University, USA  
Meng Yu, Virginia Commonwealth University, USA  
Pooneh Bagheri Zadeh, De Montfort University - Leicester, UK  
Yan Zhenyu, Florida International University, USA

## Copyright Information

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

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

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

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

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

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

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

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

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

## Table of Contents

An Analysis of Secure Interoperation of EPC and Mobile Equipments <i>Cristina-Elena Vintila, Victor-Valeriu Patriciu, and Ion Bica</i>	1
Constant-rate Adaptive Space-time Code Selection Technique for Wireless Communications <i>Dimas Mavares Teran, Rafael Pedro Torres Jimenez, and Roberto Antonio Uzcategui Vargas</i>	7
Handover Optimization in WiMAX Vehicular Communications <i>Mihai Constantinescu and Eugen Borcoci</i>	12
Mobile Phone Security Awareness and Practices of Students in Budapest <i>Iosif Androulidakis and Gorazd Kandus</i>	18
Impulse Response and Generating Functions of sincN FIR Filters <i>Jean Le Bihan</i>	25
Multi-Level Log-Likelihood Ratio Clipping in a Soft-Decision Near-Maximum Likelihood Detector <i>Sebastien Aubert, Andrea Ancora, and Fabienne Nouvelle</i>	30
Analysis of the Least-Squares Adaptive Algorithms in Interference Cancellation Configuration <i>Silviu Ciochina, Constantin Paleologu, and Andrei Alexandru Enescu</i>	36
An Evaluation Of Multiwavelet Families For Stereo Correspondence Matching <i>Pooneh Bagheri Zadeh and Cristian Serdean</i>	41
New Adaptation Method Using Two-dimensional PCA for Speaker Verification <i>Chunyan Liang, Xiang Zhang, Lin Yang, Li Lu, and Yonghong Yan</i>	46
Language Recognition With Locality Preserving Projection <i>Jinchao Yang</i>	51
Speech Recognition and Text-to-speech Solution for Vernacular Languages <i>James Tamgno Kouawa, Morgan Richomme, Claude Lishou, Aristide Thomas Mendo'o, Pascal Uriel Elingui, and Seraphin D. Oyono Obono</i>	56
Linguistic Text Compression <i>Ondrej Kazik and Jan Lansky</i>	64
Dynamic Codec Selection Algorithm for VoIP <i>Maja Sulovic, Darijo Raca, Mesud Hadzialic, and Nasuf Hadziahmetovic</i>	74

E-Model based Adaptive Jitter Buffer with Time-Scaling Embedded in AMR decoder <i>Liyun Pang and Laszlo Boszormenyi</i>	80
Service Strategies for Eliminating Digital Divide in Korea <i>Hyongsoon Kim, Sang-Ug Kang, Kwang-Taek Ryu, and Eunyoung Lee</i>	86
UEP for Wireless Video Streaming Using Spatially Multiplexed MIMO Systems With a Suboptimal Joint Detection <i>Sanghyun Bak and Jaekwon Kim</i>	91
Prefilter Bandwidth Effects in Sequential Symbol Synchronizers based on Pulse Comparison by Positive Transitions at Half Rate <i>Antonio Reis, Jose Rocha, Atilio Gameiro, and Jose Carvalho</i>	97
Sequential Symbol Synchronizers based on Pulse Comparison at Quarter Rate <i>Antonio Reis, Jose Rocha, Atilio Gameiro, and Jose Carvalho</i>	101
Prefilter Bandwidth Effects in Sequential Symbol Synchronizers based on Pulse Comparison at Quarter Rate <i>Antonio Reis, Jose Rocha, Atilio Gameiro, and Jose Carvalho</i>	105
Particle Swarm Optimization with Time-Varying Acceleration Coefficients Based on Cellular Neural Network for Color Image Noise Cancellation <i>Te-Jen Su, Jui-Chuan Cheng, and Yang-De Sun</i>	109
Peer Clustering System for Different Start Video Broadcasting <i>Hathairat Ketmaneechairat</i>	116
Mining Web Usage and Content structure Data to Improve Web Cache Performance in Content Aggregation Systems <i>Carlos Guerrero, Carlos Juiz, and Ramon Puigjaner</i>	123
Enhancing Situational Awareness by Means of Combat-ID to Minimize Fratricide and Collateral Damage in the Theater <i>Tapio Saarelainen</i>	131
The MediaSense Framework: Ranking Sensors in a Distributed Architecture <i>Jamie Walters, Patrik Osterberg, and Theo Kanter</i>	140
Evaluation of End-to-End Quality of Service over VPN Networks through Various Priority Mechanisms <i>Nasser-Eddine Rikli and Saad Almogari</i>	145

# An Analysis of Secure Interoperation of EPC and Mobile Equipments

Cristina-Elena Vintilă, Victor-Valeriu Patriciu, Ion Bica  
 Military Technical Academy, Bucharest, Romania  
 cvintila@ixiacom.com, vip@mta.ro, ibica@mta.ro

**Abstract** - 4G architecture is the latest 3GPP development when it comes to mobile networks design and optimization. Designed initially for data, but having a flexible architecture, 4G is capable of integrating IMS, this way bringing in voice and services. 3GPP wanted to facilitate the migration from 3G and non-3GPP solutions to the 4G design, therefore the core network components of 4G are capable of interacting with 3G core network devices, provided these ones have an interoperability feature. This way, even 3G or non-3GPP devices are capable of using the 4G services. Nevertheless, one of the biggest issues when interoperating these solutions and devices is the security aspect. This paper reviews some of the most common access methods and summarizes the security concerns that raise in each case. The 4G security design is a very powerful solution for authenticating the users, even though it has some shortcomings that can be addressed.

**Keywords** – SAE; EPC; AKA; EAP-AKA; HSS; J-PAKE; PKI; key management; security context.

## I. INTRODUCTION

The 4G architecture consists of two main components: the radio access network and the Evolved Packet Core. The radio network is represented by the eNodeB, the antenna and the air medium of transportation. The mobile devices connect to this antenna, which, in turn, has responsibilities in the mobile device authentication to the core network. This core network has several devices that deal with the signaling, traffic routing and prioritization and as well user authentication and charging [1]. The most common core network devices are the following: MME – Mobility Management Entity (that deals with user registration to the network, control-plane or signaling of user's traffic patterns permissions and manages the mobility of the user from one area to the other), SGW – Serving Gateway (this entity does both signaling and user-plane and it is the tracking area entity – where tracking area is a group of cells the user may camp on), PGW – Packet Data Network Gateway (this is the device that connects the 4G network to the intranet or to the Internet, it deals with traffic routing and prioritization based on the PCRF rules for a particular customer; it is also the mobility anchor of the UE – User Equipment, when this user moves around the network), PCRF – Policy Charging and Rules Function (a policy database of a customer's subscription to the operator) and HSS – Home Subscriber Server (a database that contains the mobile device identity and credentials) [9]. Figure 1 represents a simplified 4G architecture that shows the core network devices, as well as the logical interfaces that link their functionality. It also represents an example of 3G and non-3GPP connectivity to the 4G network.

The eNodeB, or the antenna, in the 4G architecture is the user's first point of contact to the network. The 4G mobile device identifies the antenna and tries to connect to it, asking for permission. The antenna then plays the role of an authentication relay agent for the user. The 4G architecture has been designed in such a matter that it can operate with 3G mobile devices, as long as the 3G parts of the network have the capability of communicating with the SGW [3]. The 3G portion of the network has multiple entities: the SGSN – Serving GPRS Support Node (this is the homologous of the MME and part of the SGW in the 4G architecture, with the important difference that it does both signaling and user-plane, unlike MME which is a signaling-only entity), GGSN – Gateway GPRS Support Node (this is the homologous of the PGW in the 4G architecture) and the U-TRAN. U-TRAN stands for UMTS Terrestrial Radio Access Network, and it is composed of multiple antennas (NodeB devices) and a RNC – Radio Network Controller. It is the RNC that actually connects to the SGSN in order to authenticate the user. The procedures for both 4G access and 3G access are similar: UMTS-AKA.

The non-3GPP access may be any other form of access, like WLAN. This time the user authentication can no longer be realized via the classic authentication procedure AKA. Instead, there is a separate architecture of 3GPP-AAA servers that does the authentication of non-3GPP access users using the EAP-AKA procedure [6]. The entities involved in this case are the 3GPP AAA server, a 3GPP AAA Proxy Server, which is used in case where the user is in roaming and an ePDG – Evolved Packet Data Gateway. The AAA acronym stands for Authentication, Authorization and Accounting. The ePDG has an important role in the user authentication; this entity is the peer the UE establishes a security communication with. The ePDG authenticates the user by accessing the AAA servers. The access to the 4G core network can be classified as non-roaming and roaming access. It can also be classified by the type of access network: 4G, 3G, 2.5G, non-3GPP. The roaming scenarios on their own can be further classified as having *home routed traffic* (meaning that the PGW is located in the home network), *local breakout with home operator's application functions only* (the PGW is in the visited network and the user does its signaling and data traffic via the visited PDN – Packet Data Network- – this is the case of a voice mail application) and *local breakout with visited operator's application functions only* (this is the case where the home and visited operators have an agreement to provide services to each other's users; all the user's traffic is served by and routed through the visited network, while the home network only does the authentication and policy verification). It is

not mandatory that the roaming scenarios are of any one type of the three types described; there can be a combination of architectures, where for certain functions the home network offers the services – like the voice mail, while some other services, like access to the Internet can be offered directly by the visited network. Also, the same network operator may have one type of architectural interconnection with one operator, while having a different connection with

another operator. Figure 1 presents a *local breakout scenario*, with both home operator's and visited operator's application functions. This means that some of the services are offered by the home network, while others are offered by the visited operator. In this case there are three users, all connecting from roaming, one is a native 4G device, the other is a 3G device and the third is a non-3GPP device, a laptop that connects via WiFi.

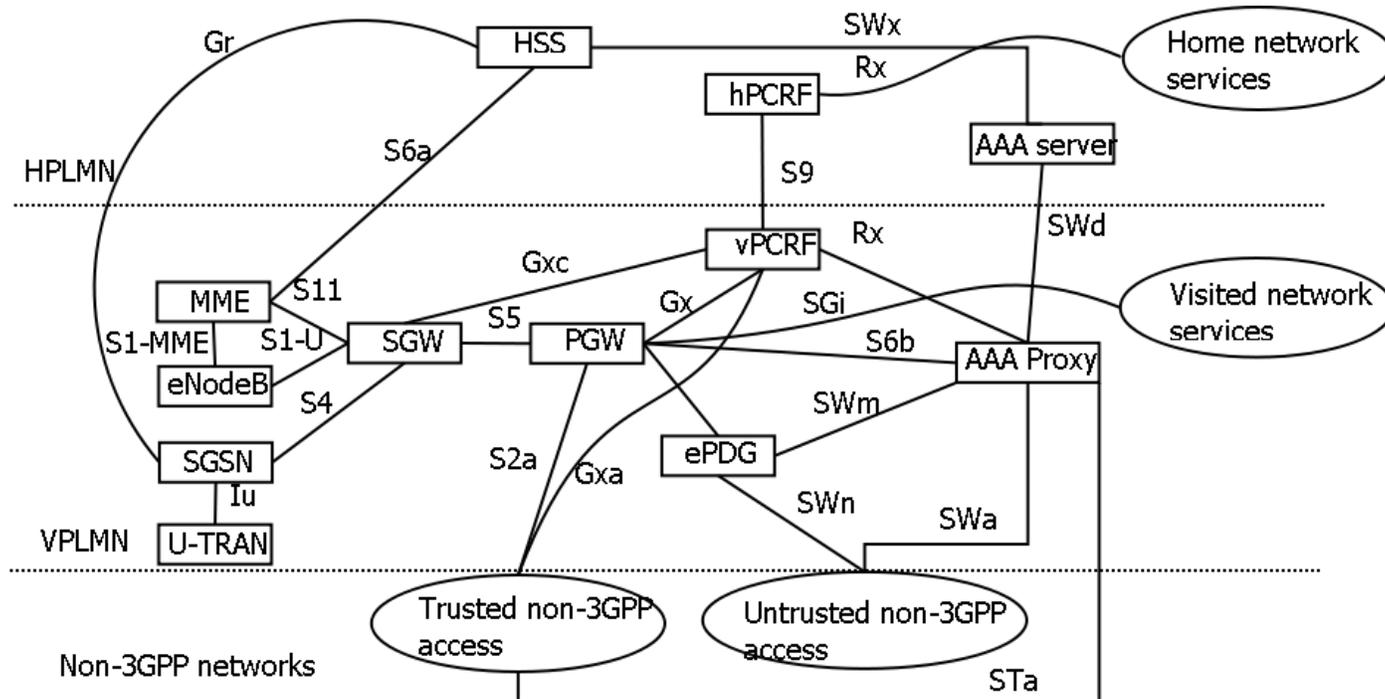


Figure 1. EPS roaming architecture with local breakout

The next two sections describe the security architecture of the 4G network, summarizing the security domains of this architecture, as well as it analyzes the three types of access to this network. The types of access are 4G radio access (via native 4G mobile stations), 3G compatible radio access and non-3gpp access via technologies like WLAN or WiFi.

## II. SECURITY ARCHITECTURE AND REQUIREMENTS

The security architecture involves most of the devices, in less or greater measure. 4G design specifies 5 security areas:

- Network Access Security: this area deals with granting access to the (core) network only to those users that prove their identity, that identity matching a network's registered user, with valid authentication credentials and with a subscription that allows services to be delivered to this user
- Network Domain Security: this area deals with the secure interoperation between the Evolved Packet Core (EPC) network entities; most of these entities of a 4G network are already represented in Figure 1, as well as some of the 3G devices, like SGSN; this security is described by the protocols involved in securing the communications between

EPC nodes: IPsec (recommended by Specs to take place within an operator's premises) and TLS (usually for inter-operator secure communications)

- User Domain Security: this area deals with the secure access to the mobile stations
- Application Domain Security: this area is concerned with how to secure the communication between the applications that reside on the user's mobile device and the core network application servers; as a layer 7 application, this area may implement a large variety of security structures
- Visibility and Configurability of Security: this is an informational area, for the user; the subscriber must have constant access to the information concerning the security features available on his device, whether or not they are functioning properly and whether or not they are required for the secure operation of a certain service

The security requirements for the 4G networks are classified according to the areas above and most of the security requirements are summarized in [6]. The eNodeB, being the access point into the network, has a large variety of security parameters and classes that must be verified and certified in order to assure a secure operation. These classes are the *setup and configuration* (this class deals with the secure communication in terms of confidentiality and

integrity between the eNB and the EPC, over the S1-MME and S1-U interfaces, between eNBs, over the X2 interface, with the secure setup configuration of the eNB and the secure software update), *key management inside the eNB* (as the eNB participates in the user equipment authentication process, it also stores some of the keys derives from the authentication process; these keys should be stored on a secure environment and never leave it, except in the situations specifically mentioned by the Specs), *handling of the user-plane traffic* (this data is transmitted over the S1-U interface to the SGW and via X2-U interface between eNBs; securing this data means assuring its confidentiality, integrity and protection against replay attacks), *handling of control-plane traffic* (this signaling transmissions take place over S1-MME interface towards the MME and via the X2-C interface between eNBs; securing this data means assuring its confidentiality, integrity and protection against replay attacks).

### III. NETWORK ACCESS SECURITY

#### A. 4G mobile device access

As per [6], the preferred access method for the 4G mobile devices is AKA, named EPS-AKA – Evolved Packet System Authentication and Key Agreement, compatible with 3G UMTS – AKA authentication system, but not compatible with the 2G SIM, nor a SIM application on a UICC. The purpose of the AKA is to produce master keying material for protection of 3 classes of traffic: user-plane traffic, RRC – Radio Resource Control and NAS – Non Access Stratum. The exact procedure for deriving and distributing the keys is not important at this point. There are 6 keys that result from this process: K-eNB, K-NASint, K-NASenc, K-UPenc, K-RRCint and K-RRCenc. The AKA procedure is represented in the picture below.

This exchange is triggered by the UE connecting to an antenna. The antenna (called eNodeB) is forwarding to the MME the identity declared by the UE. At the very first attach, this is usually the IMSI – International Mobile Subscriber Identity, afterwards it is a temporary identity called GUTI – Global Unique Temporary Identity. The MME then contacts the HSS, sending the UE’s identity over Diameter. If it finds the identity, the HSS responds with a set of AVs – Authentication Vectors (called generically Authentication Data). An AV contains 4 fields: the RAND – a random challenge string, an AUTN – an authentication token, an XRES – expected authentication response and a session key for the traffic between MME and HSS, named K-ASME – Key for Access Security Management Entity, which in our case is the role assumed by the MME. The MME forwards, via eNB, the RAND and AUTN to the UE. The UE authenticates the network using the AUTN and computes the RES – Response, which is sent back to the MME. The MME compares the RES and XRES and, if they match, the UE is considered authenticated.

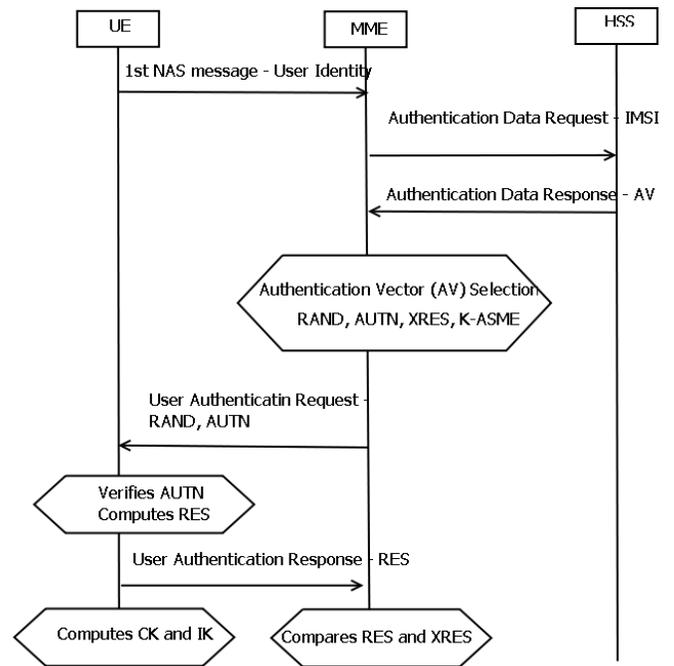


Figure 2. EPS-AKA procedure

There are at least 2 security issues related to the EPS-AKA process. One of them appears at the first Initial Attach of the User Equipment, when the user’s identity IMSI is sent unencrypted over the air and the second one is the lack of PFS – Perfect Forward Secrecy property of the AKA algorithm. The identification of any subsequent requests a particular user may make to the network is done via a temporary identity called GUTI. The new MME reads this identifier from the UE’s message (TAU – Tracking Area Update, for instance), contacts the previous MME in order to obtain the IMSI and then does the actual UE authentication to the HSS. The messages exchanged between the UE and the MME all pass through the eNB. These messages are GTPv2-C packets via the S1-MME interface. The AKA process continues over the S6a interface, where the information is encapsulated in Diameter protocol packets. The actual user-plane, after leaving the radio domain, is forwarded by the eNB directly to the designated SGW, over the S1-U interface, where the encapsulation is GTPv1-U. The keys derived by the AKA process are used over the air interface, then between the eNB and the MME. Traffic protection between the eNB and the SGW is a network domain security field of activity.

There have been multiple research projects done in order to improve the security and speed up the authentication process. Some of these projects are already patented and used in industry: SPEKE algorithm for the authentication between the BlackBerry device and the BES – BlackBerry Enterprise Server. The SPEKE is very similar to Diffie-Hellman, with the exception that the hash of the password is used as the group generator. Still, this method has a lot of vulnerabilities and technical limitations. There is yet another project that was very well received by the

cryptographic community. This is called J-PAKE [19]. It overcomes the vulnerabilities of both Diffie-Hellman and EKE – SPEKE methods. J-PAKE provides *off-line dictionary attacks resistance* (it does not leak any information that allows an attacker to search for the password off-line), *forward secrecy* (the information remains protected even if the original shared secret was disclosed), *known-key security* (even if a session key is disclosed, the information protected with other session keys is not accessible) and *on-line dictionary attacks resistance* (an on-line attacker can only test one password per execution). Even though this protocol requires two computational rounds and 14 exponentiations, it is much stronger and requires a smaller exponent to generate its keys.

This method is not yet used in the mobile industry, even though it is lightweight and applicable to the mobile devices characteristics. This protocol can be used along with AKA to provide efficient end-to-end cryptography for the 4G core network services, like the ones provided by IMS – IP Multimedia Subsystem, as well as instead of AKA in the 4G authentication protocol. Besides the fact that it is strong with regards to the 4 security aspects listed above, this method does not require a PKI implementation, which makes it more flexible and easier to use.

**B. 3G mobile device access**

The 3G security requirements and procedures are described in [8]. The solution used for authentication in the 3G design is the predecessor of the EPS-AKA. It is called UMTS-AKA and it uses the same methods. When authenticating to a 4G network, the user equipment is still connecting through a 3G access network. In the 4G E-UTRAN, the eNB has the entire access control role as an antenna (doing both signaling and data) and it is managed by an MME device, which has only signaling role: authentication, management and mobility management. The 3G U-TRAN design had a pool of antennas managed by a RNC (Radio Network Controller) and it was the SGSN device that did the mobility management, the authentication and also data-plane. The 3G design got simpler in the 4G, creating a smarter antenna and completely segregating the control and data planes in separate entities: MME and SGW. So, in order to effectively connect to a 4G core, the design must keep the SGW in place, as an essential core device, and define the 3G-4G delimitation on the S4 interface, between the SGSN and the SGW. This way, the SGSN still manages the 3G UE, does its authentication and manages also its mobility to the 4G network, but the actual traffic is forwarded to the SGW, and then the PGW in order to be routed to the Internet/Intranet or IMS behind the PDN the user connected to. The authentication of a 3G device is done by the SGSN, which interacts with the HSS over the Gr interface, which is Diameter based.

The interoperability issues between the 3G and 4G devices appear when the UE is moving from 3G cover to 4G cover and viceversa. It is very possible that the HSS already sent multiple AVs to the UE, and this one stores many of

them, so that at a certain moment in time, it authenticates/re-authenticates using one of them, at its choice. The security association that exists between a mobile device and the network is called security context. In EPS, this context is composed of 2 other security contexts: the AS – Access Stratum and NAS – Non-Access Stratum contexts, which are sets of keys between the entities participating in the AKA process, which will provide hop-by-hop security (confidentiality and/or integrity and/or replay protection) for radio bearers, signaling and user-plane traffic. All the entities must be able to do Security Context management, and mostly the UE must be able to store multiple security contexts. It can be a legacy security context (a context created after the UMTS-AKA process from 3G) or a native EPS security context (results from the EPS-AKA procedure) or it may be a mapped security context, where the keys have been generated from the EPS-AKA process, but they are going to be used in a 3G communication (partial native context). As there can be multiple security contexts at one time in the UE and network, only one can be in effect (this one is called current security context); the others are non-current contexts. Table 1 summarizes the states and types of security contexts.

TABLE I. SECURITY CONTEXTS

AGE/EFFECT	CURRENT	NON-CURRENT
FULL	NATIVE / MAPPED	NATIVE
PARTIAL	X	NATIVE

Note that there is possible one single transformation, that is from a partially native security context, there can be generated a fully native context, but not the other way around.

There are multiple scenarios that assume 3G-4G interaction. One case is when the UE moves from the 4G network towards the 3G network, procedure called RAU – Routing Area Update, which can take place when the UE is either in ECM-IDLE mode or in ISR – Idle Signaling Reduction. When the UE is in ECM-CONNECTED mode, the procedure is called handover. As this article describes a 3G mobile connecting to a 4G network, we will detail the procedures required when a device moves from a 3G network towards a 4G network. These are also divided into procedures that apply when the device is in ECM-IDLE (specifically the TAU – Tracking Area Update procedure) and the handover from UTRAN to E-UTRAN. With regards to security, these procedures translate into a mapping of the old/previous security context into a new security context. As the case discussed involves 3G to 4G mobility, the old security context may or may not provide the (new) MME with the UE identity:

a) *It may send a temporary identity that was being used in the 3G context: the UE sends in the TAU Request its former P-TMSI identity used in the 3G context, in the form of an old GUTI IE; this implies that the TAU request is integrity-protected, but not encrypted, and also implies the*

UE sending to the MME more information about its previous security context:

- the KSI, P-TMSI and RAI, so that the (new) MME can find the (old) SGSN
- a P-TMSI signature
- a 32bit Nonce

b) It may not include the previous temporary identity, case where the AKA process takes place again

In case of the handover, the process has 2 steps:

- a) Signaling handover using a mapped EPS security context
- b) Subsequent NAS signaling to determine whether an EPS context can be used – in the cases where the network and UE security properties and requirements don't match – this usually takes place at the first MME change mobility process

Figure 3 describes a basic 3G to 4G handover process. The process is presented in more details in [6].

### C. Non-3GPP mobile device access

A WLAN UE, like a laptop, may also connect to a 4G network, provided it supports EAP-AKA procedures and that the network has an AAA proxy and an 3GPP AAA server. The figure below describes the EAP-AKA process that takes place when a laptop connects via WLAN to a 4G core network.

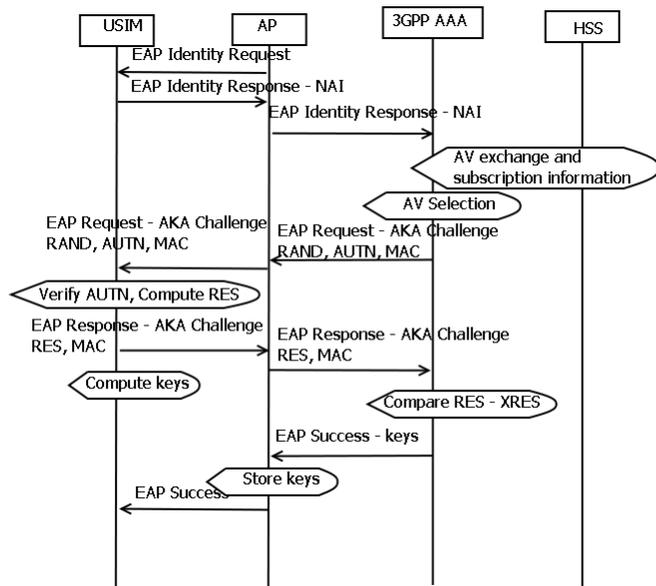


Figure 3. EPS EAP-AKA

The abbreviations and notations have the same meaning as for the classical AKA procedure. The only difference here is that this entire AKA negotiation takes place over the EAP – Extensible Authentication Protocol, an authentication framework used with success in protocols like 802.1x for layer 2 authentication in both wired and wireless

technologies, as well as for upper layer security procedures like EoU – EAP over UDP.

The EPS-AKA system is essentially the same as the UMTS-AKA system. However, there are several distinctive improvements, with regards to both security level and negotiation speed. The EPS-AKA process includes in the authentication the network ID, which means the protection of the mobile station from a fake antenna attack; this solution is not present in the UMTS-AKA system. Both of the 3G and 4G systems provide user ID protection only after the attach process completes (during re-register or during handover process): the IMSI is not protected at Initial Attach. Another improvement of the 4G AKA is the keying hierarchy, which determines the storage of the K-ASME in the MME. This means that the NAS traffic is also protected, between the UE and the MME, which does not happen in the 3G system, where the SGSN stores only the CK and IK for the user-plane traffic. Both the 3G and 4G systems protect the AS level traffic in terms of ciphering and integrity protection. Another key difference between the 3G and 4G is that the former allows the 2G interaction with regards to security. 4G system does no longer allow for handover to 2G systems, considering the security level of this system is not high enough. Also, the 4G system design permits for handover of non-3GPP devices, which was not previously permitted in the 3G system.

### IV. CONCLUSIONS AND FUTURE WORK

Interoperating in the 4G worlds is a complex task, as a large variety of devices, coming from all over the standards and implementations ask for connectivity and for services. The 4G operators must be sure to protect their network and also their customers, with the minimum overload, for both the network and the user equipments. This paper reviewed only 3 basic types of access to the 4G core network. These are the native EPS – 4G mobile access, the traditional UMTS – 3G access, and a generic WLAN device. The very first step when deciding whether to serve or not a potential customer is to make sure this is a valid customer, and not an attacker. Looking from the user perspective, you want to make sure you are not connecting to a rogue network, and that your data remains private. 4G provides and requires mutual authentication. This is done natively via EPS-AKA procedure and can also be mapped from a legacy UMTS-AKA procedure. The WLAN device can enter 4G if it supports EAP-AKA and if the 4G network has 3GPP AAA capabilities.

The university world has come up with newer, faster and more secure procedures for doing mutual authentication. One of these procedures is derived from the PEKE algorithm and it is called J-PAKE. This simple method is very appropriate for applications in a mobile world, so it may be a revolutionary alternative to the way we do mutual authentication, even as an alternative to AKA or as a more

secure proof of knowledge before doing AKA key derivation.

This article is just one of a series of articles that debate the way 4G accomplished its 5 Security Domains duties, discussing the access security. This work continues with the analysis of the security issues that appear at mobility. There are a lot of mobility scenarios, both in Connected and Idle modes, as well as between 3G and 4G, that test the security of a 4G network.

#### REFERENCES

- [1] TS 23.401 – GPRS Enhancements for E-UTRAN access - [http://www.3gpp.org/ftp/Specs/archive/23\\_series/23.401/](http://www.3gpp.org/ftp/Specs/archive/23_series/23.401/) [last access: February 2011]
- [2] TS 23.122 – NAS Functions related to Mobile Stations in idle mode [http://www.3gpp.org/ftp/Specs/archive/23\\_series/23.122/](http://www.3gpp.org/ftp/Specs/archive/23_series/23.122/) [last access: February 2011]
- [3] TS 36.300 – E-UTRAN Overall Description - [http://www.3gpp.org/ftp/Specs/archive/36\\_series/36.300/](http://www.3gpp.org/ftp/Specs/archive/36_series/36.300/) [last access: February 2011]
- [4] TS 43.022 – Functions of the MS in idle mode and group receive mode - [http://www.3gpp.org/ftp/Specs/archive/43\\_series/43.022/](http://www.3gpp.org/ftp/Specs/archive/43_series/43.022/) [last access: February 2011]
- [5] TS 25.304 – UE Procedures in idle mode and procedures for cell reselection in connected mode - [http://www.3gpp.org/ftp/Specs/archive/25\\_series/25.304/](http://www.3gpp.org/ftp/Specs/archive/25_series/25.304/) [last access: February 2011]
- [6] TS 33.401 – SAE – Security Architecture - [http://www.3gpp.org/ftp/Specs/archive/33\\_series/33.401/](http://www.3gpp.org/ftp/Specs/archive/33_series/33.401/) [last access: February 2011]
- [7] TS 33.310 – Network Domain Security; Authentication Framework - [http://www.3gpp.org/ftp/Specs/archive/33\\_series/33.310/](http://www.3gpp.org/ftp/Specs/archive/33_series/33.310/) [last access: February 2011]
- [8] TS 33.102 – 3G Security Architecture- [http://www.3gpp.org/ftp/Specs/archive/33\\_series/33.102/](http://www.3gpp.org/ftp/Specs/archive/33_series/33.102/) [last access: February 2011]
- [9] RFC 5516 - Diameter Command Code Registration for the Third Generation Partnership Project (3GPP) Evolved Packet System (EPS) <http://tools.ietf.org/html/rfc5516> [last access: February 2011]
- [10] TS 29.272 – MME related interfaces based on Diameter - [http://www.3gpp.org/ftp/Specs/archive/29\\_series/29.272/](http://www.3gpp.org/ftp/Specs/archive/29_series/29.272/) [last access: February 2011]
- [11] Tech-Invite - <http://tech-invite.com/> [last access: February 2011]
- [12] TS 29.294 – Tunneling Protocol for Control plane (GTPv2-C) - [http://www.3gpp.org/ftp/Specs/archive/29\\_series/29.274/](http://www.3gpp.org/ftp/Specs/archive/29_series/29.274/) [last access: February 2011]
- [13] TS 33.220 – Generic Authentication Architecture; Generic Bootstrapping Authentication - [http://www.3gpp.org/ftp/Specs/archive/33\\_series/33.220/](http://www.3gpp.org/ftp/Specs/archive/33_series/33.220/) [last access: February 2011]
- [14] TR 33.919 – Generic Authentication Architecture – System Overview - [http://www.3gpp.org/ftp/Specs/archive/33\\_series/33.919/](http://www.3gpp.org/ftp/Specs/archive/33_series/33.919/) [last access: February 2011]
- [15] TS 33.221 – Support for Subscriber Certificates - [http://www.3gpp.org/ftp/Specs/archive/33\\_series/33.221/](http://www.3gpp.org/ftp/Specs/archive/33_series/33.221/) [last access: February 2011]
- [16] “Efficient Remote Mutual Authentication and Key Agreement with Perfect Forward Secrecy” – Han-Cheng Hsiang, Weu-Kuan Shih, Information Technology Journal 8 – 2009, Asian Network for Scientific Information [last access: February 2011]
- [17] RFC 4187 – EAP Method for 3GPP AKA - <http://tools.ietf.org/html/rfc4187> [last access: February 2011]
- [18] RFC 2631 – Diffie-Hellman Key Agreement Method - <http://tools.ietf.org/html/rfc2631> [last access: February 2011]
- [19] “Password Authenticated Key Exchange by Juggling” – J-PAKE – 2008, F. Hao, P.Ryan, Proceedings of the 16th International Workshop on Security Protocols, 2008 - <http://grouper.ieee.org/groups/1363/Research/contributions/hao-ryan-2008.pdf> [last access: February 2011]
- [20] TS 33.234 – WLAN – 3G interworking security - [http://www.3gpp.org/ftp/specs/archive/33\\_series/33.234/](http://www.3gpp.org/ftp/specs/archive/33_series/33.234/) [last access: February 2011]

# Constant-rate Adaptive Space-time Code Selection Technique for Wireless Communications

Dimas Mavares Terán

Electronic engineering department  
Universidad Nacional Politécnica (UNEXPO)  
Barquisimeto, Venezuela  
email:dmavares@unexpo.edu.ve

Rafael P. Torres

Communications department  
Universidad de Cantabria  
Santander, España  
email:torresrp@unican.es

Roberto A. Uzcátegui

Electronic engineering department  
Universidad Nacional Politécnica (UNEXPO)  
Barquisimeto, Venezuela  
email:ruzcategui@unexpo.edu.ve

**Abstract**—A constant-rate space-time code selection technique for transmit antenna diversity systems is proposed. The proposed technique selects both the space-time code and the number of transmitter antennas through a comparison of equivalent SISO channels with a set of predetermined threshold levels, using only four bits for feedback. The constant-rate transmission is based in the inclusion of a complementary transmission mode, which is used whenever no equivalent SISO channel's envelope is found to be over the threshold levels. Simulation results show that the proposed technique outperforms other adaptive transmission techniques, while in comparison with its variable-rate counterpart, there is an increase in spectral efficiency with a slight penalty performance. Additionally to a constant-rate transmission, the new techniques makes BER performance almost independent of the relative velocity between the transmitter and the receiver.<sup>12</sup>

**Index Terms**—Space-time coding, adaptive transmission, antenna selection, wireless communications, baseband signal processing.

## I. INTRODUCTION

Transmit antenna diversity (TAD) is one of the tools to be applied to construct multiple-input multiple-out (MIMO) systems, which are expected to contribute to providing the high data rates needed by fourth generation wireless systems. One way to implement TAD is by using space-time block codes (STBC), obtaining an open loop system whenever channel state information (CSI) is not available at the transmitter end [1], [2]. Closed-loop systems are designed to take advantage of the CSI at the transmission end. However, they are limited by the need to maintain the amount of data in the feedback channel as low as possible. Channel feedback information is one important issue when designing transmit systems.

Several structures that combine antenna selection techniques with space-time coding have been proposed in [3]–[7]. In [7], it was shown that selection of antenna subsets in conjunction with space-time codes produces an increase in the mean signal to noise ratio (SNR), maintaining the diversity order. In this paper, a modification of the space-time code selection technique (STCS) described in [8], [9] is proposed, which

allows to obtain a constant-rate transmission. In the former STCS technique, the STBC and the antennas to be used were selected through a comparison of the equivalent single-input single-output (SISO) channel envelopes (taken as a linear combination of Rayleigh channel envelopes) with a set of predetermined threshold levels. A no-transmission mode was selected whenever no equivalent SISO channel satisfied the predetermined conditions, producing a slight average spectral efficiency penalty, generally no higher than 10%. However, such non-constant rate transmission could prevent its use in constant-rate applications. The proposed modification corrects this drawback by substituting the non-transmission mode by a complementary mode, which consist in transmitting with a predetermined code. In our test, it was found that transmitting with a single antenna was the most convenient choice, thanks to the fastest recovery of the Rayleigh channel compared to other equivalent SISO channels.

Monte Carlo simulations show that, using the complementary mode, almost the same bit-error rate (BER) performance than the variable-rate STCS (VR-STCS) is achieved, but without any spectral efficiency penalty. Additionally, simulations over a range of Doppler frequencies show that the performance of the new technique is almost independent of the relative velocity between the transmitter and the receiver.

The remainder of the paper is organized as follows. In Section II, the system model is introduced. Section III describes the new space-time code algorithm. In Section IV, simulation results are shown and discussed. Finally, Section V ends the paper with the conclusions.

## II. SYSTEM MODEL

In the communication system model, four transmitter and one receiver antennas are employed. Initially, the transmitter sends pilot symbols to perform the vector channel estimation at the receiver. The equivalent SISO channel envelopes are compared with a given set of threshold levels, which are previously chosen according to the mobile speed. Afterwards, the receiver decides on the space-time code and the transmit antennas to use, and sends this decision to the transmitter. Transmission is adapted to that decision, which is maintained until the next decision instant, when the process is repeated.

<sup>1</sup>The authors gratefully acknowledge financial support for this work from the Regional Fund for Digital Innovation in Latin America and The Caribbean (FRIDA).

<sup>2</sup>This work was also partially supported by the Spanish Ministry of Science and Innovation (project number TEC2008-02730).

The transmission is assumed to occur in flat frequency channels. The squared envelopes of the equivalent SISO channels using an orthogonal STBC (OSTBC) are given by

$$\alpha_{n_T}^2 = \sum_{k=1}^{n_T} r_k^2, \quad (1)$$

where  $n_T$  and  $r_k^2$  represent the number of transmitter antennas and the square of Rayleigh distributed channel envelopes, respectively. It is well known that  $\alpha_{n_T}$  follows a Nakagami- $n_T$  distribution [10] with a mean value of  $\Omega_{n_T} = 2n_T$ . Its probability density function is given by

$$p_N(x) = \left( \frac{n_T}{\Omega_{2n_T}} \right)^{n_T} \frac{x^{n_T-1}}{\Gamma(n_T)} e^{-n_T x / \Omega_{2n_T}}. \quad (2)$$

The received signal can be expressed by the following input/output relationship:

$$\mathbf{Y} = \mathbf{S}_{n_T} \mathbf{F} \mathbf{H} + \mathbf{V}, \quad (3)$$

where  $\mathbf{H}$  is the  $4 \times 1$  channel vector,  $\mathbf{S}_{n_T}$  is a  $4 \times 4$  space-time coding matrix,  $\mathbf{F}$  is a  $4 \times 4$  permutation and selection matrix and  $\mathbf{V}$  is the  $1 \times 4$  noise vector. The matrix  $\mathbf{S}_{n_T}$  may take the following structures, based on the space-time codes proposed in [1] and [2], depending on the number of selected antennas [8], [9]:

$$\mathbf{S}_1 = \begin{pmatrix} s_0 & 0 & 0 & 0 \\ s_1 & 0 & 0 & 0 \\ s_2 & 0 & 0 & 0 \\ s_3 & 0 & 0 & 0 \end{pmatrix}, \quad (4)$$

$$\mathbf{S}_2 = \begin{pmatrix} s_0 & s_1 & 0 & 0 \\ -s_1^* & s_0^* & 0 & 0 \\ s_2 & s_3 & 0 & 0 \\ -s_3^* & s_2^* & 0 & 0 \end{pmatrix}, \quad (5)$$

$$\mathbf{S}_3 = \begin{pmatrix} s_0 & s_1 & \frac{s_2}{\sqrt{2}} & 0 \\ -s_1^* & s_0^* & \frac{s_2}{\sqrt{2}} & 0 \\ \frac{s_2^*}{\sqrt{2}} & \frac{s_2^*}{\sqrt{2}} & \frac{(-s_0 - s_0^* + s_1 - s_1^*)}{2} & 0 \\ \frac{s_2^*}{\sqrt{2}} & -\frac{s_2^*}{\sqrt{2}} & \frac{(s_1 + s_1^* + s_0 - s_0^*)}{2} & 0 \end{pmatrix}, \quad (6)$$

$$\mathbf{S}_4 = \begin{pmatrix} s_0 & s_1 & \frac{s_2}{\sqrt{2}} & \frac{s_2}{\sqrt{2}} \\ -s_1^* & s_0^* & \frac{s_2}{\sqrt{2}} & -\frac{s_2}{\sqrt{2}} \\ \frac{s_2^*}{\sqrt{2}} & \frac{s_2^*}{\sqrt{2}} & \frac{(-s_0 - s_0^* + s_1 - s_1^*)}{2} & \frac{(s_0 - s_0^* - s_1 - s_1^*)}{2} \\ \frac{s_2^*}{\sqrt{2}} & -\frac{s_2^*}{\sqrt{2}} & \frac{(s_1 + s_1^* + s_0 - s_0^*)}{2} & -\frac{(s_0 + s_0^* + s_1 - s_1^*)}{2} \end{pmatrix}. \quad (7)$$

### III. SELECTION ALGORITHM

In this section, the selection algorithm, which is carried out at the receiver baseband in the discrete-time domain, is described. The selection algorithm is based on the comparison of possible equivalent SISO channels with a set of four threshold levels, which are selected depending on the maximum Doppler frequency. Since the decision should be maintained until new feedback information is available, it is not sufficient to make it based only on instantaneous CSI; it should also take into account the rate of change of the

equivalent SISO channel envelope. As the rate of change of the fading envelope decreases with  $n_T$ , it could be established that, in order to select an equivalent SISO channel only when its envelope is above a predefined threshold level during a given period of time, the lower the number of antennas that comprise the channel, the higher the requirements for selecting it. For a given instantaneous BER objective, equivalent SISO channels for larger  $n_T$  values will take more time to drop to the envelope level that produces such an instantaneous BER objective. Therefore, the highest threshold level ( $\rho_1$ ) will correspond to the one-antenna equivalent SISO channels, while the lowest threshold level ( $\rho_4$ ) will correspond to the four-antennas equivalent SISO channel. Then, the selection algorithm proceeds as follows:

- *Step 1:* Once the receiver obtains the CSI (which can be estimated based on a pilot sequence sent by the transmitter using all the transmitter antennas), it compares individually the Rayleigh fading envelopes with threshold  $\rho_1$ . If one or more of these envelopes are detected to be over  $\rho_1$ , the corresponding transmitter antennas are marked as selected and the procedure jumps to step 6.
- *Step 2:* The receiver compares all the Nakagami-2 fading envelopes with  $\rho_2$ . If one or more of these envelopes are detected to be over  $\rho_2$ , all the transmitter antennas which compose the detected envelopes are marked as selected and the procedure jumps to step 6.
- *Step 3:* The receiver compares all the Nakagami-3 fading envelopes with  $\rho_3$ . If one or more of these envelopes are detected to be over  $\rho_3$ , all the transmitter antennas which compose the detected envelopes are marked as selected and the procedure jumps to step 6.
- *Step 4:* The receiver compares the Nakagami-4 fading envelope with  $\rho_4$ . If it is detected to be over  $\rho_4$ , all the transmitter antennas are marked as selected and the procedure jumps to step 6.
- *Step 5:* This step is reached only if no Nakagami- $k$  fading envelope is detected to be over  $\rho_k$ . In this case, the highest envelope's Rayleigh channel is selected, that is, a single antenna is selected.
- *Step 6:* The receiver constructs the permutation and selection matrix  $\mathbf{F}$  and sends it to the transmitter.

The only form of feedback is given by the permutation and selection matrix  $\mathbf{F}$ . It is first set as a permutation matrix of order four with the first  $n_T$  columns of the identity matrix corresponding to the position of the selected antennas. Finally, as  $\mathbf{F}$  should specify  $n_T$  (for selecting the corresponding  $\mathbf{S}_{n_T}$ ), its last  $4 - n_T$  rows are set to zero.

For example, if antennas 1 and 4 are selected, matrix  $\mathbf{F}$  should permute columns 2 and 4 of  $\mathbf{S}_2$ . Then,

$$\mathbf{F} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \quad (8)$$

At the transmitter end, the number of ones in  $\mathbf{F}$  specifies  $n_T$ , and the antenna selection into that code is performed

directly through the product  $\mathbf{S}_{n_T}\mathbf{F}$ . Notice that the overhead transmitting matrix  $\mathbf{F}$  is limited to 4 bits.

The main difference between this procedure and that stated in [8] resides in the inclusion of step 5, in which the highest envelope's Rayleigh channel is selected whenever no Nakagami- $k$  fading envelope is detected to be over  $\rho_k$  in previous steps. In this condition, it is likely that all SISO equivalent channels have a low instant envelope value. The selection of a single antenna is based on the fastest recovery of the Rayleigh channel, compared with Nakagami- $k$  channels, for  $k > 1$ .

#### IV. RESULTS

For the simulations, symbol-synchronous receiver sampling and ideal timing have been assumed. Uncorrelated narrow-band Rayleigh channels were used, modeled as a circular complex Gaussian variable with zero mean and unit standard deviation. For simplicity, a zero-delay feedback channel and perfect channel estimation at the receiver have been assumed. Maximum likelihood detection was employed in reception.

The transmitter power was maintained constant independently of the number of transmitter antennas in use, distributing it evenly over them. The symbol time was  $T_s = 3.2\mu\text{s}$  and the carrier frequency  $f_c = 3.5\text{ GHz}$ . The normalized Doppler frequency of the simulated channels was  $1.28 \times 10^{-4}$ ,  $3.84 \times 10^{-4}$ ,  $6.4 \times 10^{-4}$  y  $8.95 \times 10^{-4}$ , corresponding to mobile speeds of approximately  $v = 12\text{ km/h}$ ,  $37\text{ km/h}$ ,  $62\text{ km/h}$  and  $86\text{ km/h}$ , respectively. Pilot symbols used for channel selection (feed-forward and feedback STCS information) were inserted every 140 symbols. Then, for the test channels, a fixed feedback spacing of 140 symbols was used, which corresponds to 10%, 30%, 50% and 70% of the channels' coherence time ( $T_C$ ), when using the relationship between  $T_C$  and  $f_d$  stated by

$$T_C = \frac{9}{16\pi f_d}. \quad (9)$$

In order to obtain the same average spectral efficiency (3 b/s/Hz) when using the different  $\mathbf{S}_{n_T}$  matrices, an 8 phase shift keeing (8PSK) modulation was used when selecting one or two transmitters, while 16 quadrature amplitud modulation (16QAM) was used when selecting three or four transmitters. No channel coding was used. Threshold levels were found for an instantaneous BER objective equal to  $1.2 \times 10^{-3}$  (for details about the how theshold levels were obtained, see [8]). The threshold levels values used in the simulations were  $\rho_1 = \rho_2 = -0.43\text{ dB}$ ,  $\rho_3 = \rho_4 = -1\text{ dB}$ .

To present the simulation results, comparisons should be made with closed loop multiple-input single output (MISO) system techniques that use four radio frequency (RF) chains, i.e., to use four transmitter antennas and a single receiver antenna. A considerable effort has been focused on research regarding MIMO systems, such as in spatial multiplexing and linear precoding [11]–[16]. However, such systems require more than one receiver antenna; in fact, they usually require more receiver than transmitter antennas. The combination of beamforming and orthogonal STBC (OSTBC-BF) proposed in [17] and the combination of orthogonal STBC with adaptive

sub-group antenna encoding (OSTBC-SGE) proposed in [18] are, to the authors' knowledge, two of the most recent closed loop MISO systems with four transmitter antennas found in the literature. As OSTBC-BF requires weighting the transmit signal with a complex channel vector, 512 bits (four complex numbers in standard IEEE 754 double precision) feedback were established for this technique, with the intention of disassociate system performance with feedback information accuracy. On the other hand, OSTBC-SGE requires only four bits feedback.

Figs. 1 to 4 show the BER performance of the constant-rate STCS (CR-STCS), that is, our proposed technique which uses the complementary transmission mode, in comparison with the variable rate STCS (VR-STCS), the OSTBC-SGE and the OSTBC-BF as a function of the SNR for different relative velocities between the transmitter and the receiver. Additionally, the VR-STCS average spectral efficiency is shown. In general, it can be seen that the CR-STCS slightly sacrifices the BER with respect to the VR-STCS, but this sacrifice is compensated with a larger and constant average spectral efficiency. Furthermore, it can be noticed that the BER performance of the CR-STCS hardly depends on the relative velocity of the receiver.

In Fig. 1 the aforementioned systems are compared for a velocity  $v = 12\text{ km/h}$ . It can be seen that the difference in performance between CR-STCS and VR-STCS is around 0.3dB for a BER =  $1 \times 10^{-3}$  and 1dB for a BER =  $1 \times 10^{-5}$ . Compared to the OSTBC-BF and OSTBC-SGE techniques, a significant SNR gain is obtained. For example, for a BER =  $1 \times 10^{-3}$ , 4.8dB and 6.2dB are respectively obtained.

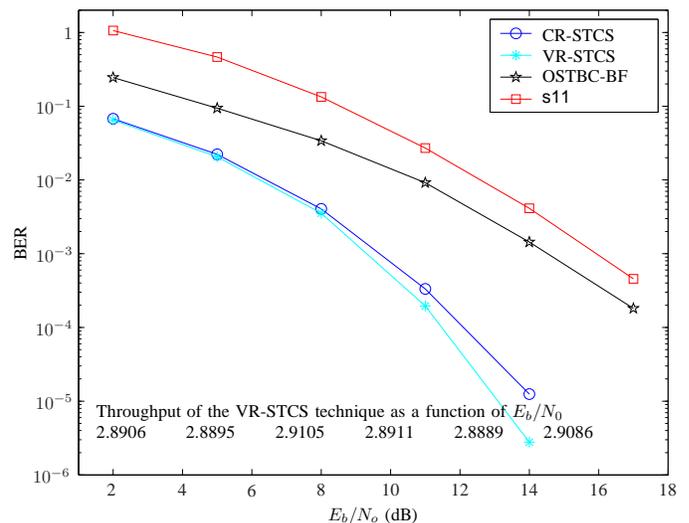


Fig. 1: BER and throughput comparison of the proposed technique (CR-STCS) with other schemes for a speed of 12 km/h. The throughputs of CR-STCS, OSTBC-BF and OSTBC-SGE techniques are constant at 3b/s/Hz.

In Fig. 2 the velocity is  $v = 37\text{ km/h}$ , and when comparing CR-STCS and VR-STCS a difference of around 0.6 and 1.8dB

can be observed for a BER equal to  $1 \times 10^{-3}$  and  $1 \times 10^{-5}$ , respectively. However, the spectral efficiency, besides being constant, is a 6% higher than with VR-STCS. Regarding OSTBC-BF y OSTBC-SGE, again it can be noticed a BER difference of around 5dB and 7dB for a BER=  $1 \times 10^{-3}$ .

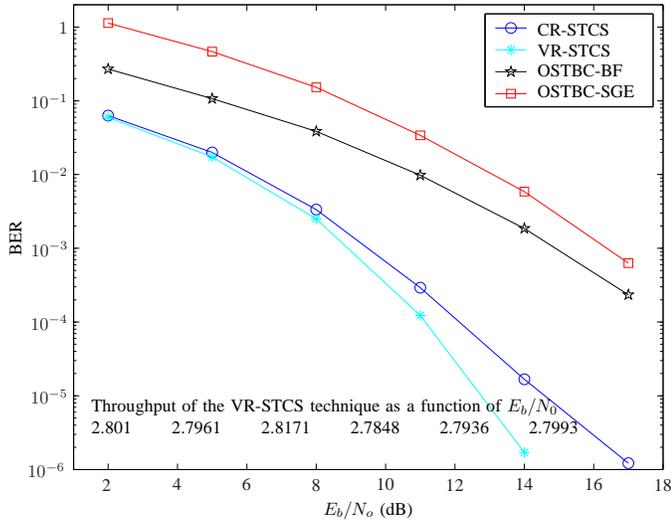


Fig. 2: BER and throughput comparison of the proposed technique (CR-STCS) with other schemes for a speed of 37 km/h. The throughputs of CR-STCS, OSTBC-BF and OSTBC-SGE techniques are constant at 3b/s/Hz.

Figs. 3 and 4, containing results for  $v = 62\text{km/h}$  and  $v = 86\text{km/h}$ , show the same trend. For  $v = 62\text{km/h}$  the CR-STCS suffers a performance degradation with respect to VR-STCS of around 0.7dB and 1.5dB, for a BER equal to  $1 \times 10^{-3}$  and  $1 \times 10^{-5}$ , respectively. For  $v = 86\text{km/h}$ , the degradation is around 1 and 2dB. In comparison with OSTBC-BF and OSTBC-SGE, SNR gains higher than 6dB are obtained for a BER=  $1 \times 10^{-3}$ . The increase of spectral efficiency with respect to VR-STCS is higher than 10%. It can be noticed that VR-STCS achieves a high BER performance through the rate sacrifice. As CR-STCS maintains a constant rate, there is a higher degradation of CR-STCS in terms of SNR compared with VR-STCS. However, results show that BER performance and rate are maintained by the proposed technique.

Finally, Fig. 5 shows the confidence intervals with 95% of confidence for the Monte Carlo simulation for a velocity  $v = 12 \text{ km/h}$  (see Fig. 1). The fluctuation of the results from the different simulations is noticeable only for the lowest BER values. In general, only BER results close to  $1 \times 10^{-6}$  or lower show a significant uncertainty. This is a consequence of using a large amount of samples, which increments with the SNR. As a general rule, for low SNRs half a million samples were used, as no BER lower than  $1 \times 10^{-3}$  was expected. For medium SNRs 1 million samples were used, expecting a BER higher than  $1 \times 10^{-4}$ . Finally, for high SNRs 3 millions samples were used.

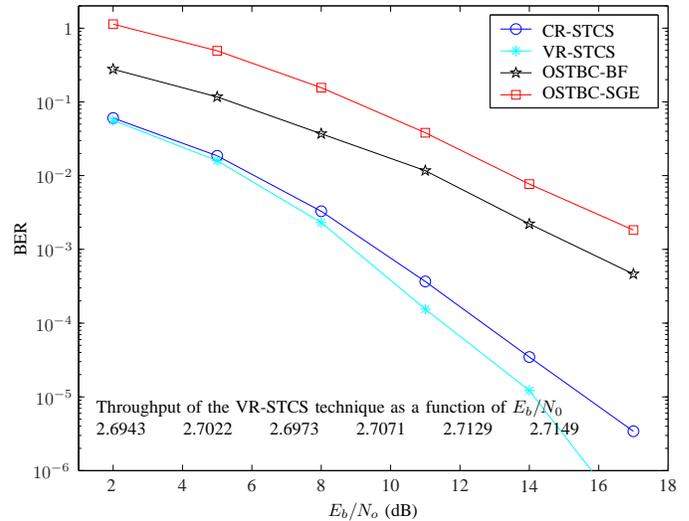


Fig. 3: BER and throughput comparison of the proposed technique (CR-STCS) with other schemes for a speed of 62 km/h. The throughputs of CR-STCS, OSTBC-BF and OSTBC-SGE techniques are constant at 3b/s/Hz.

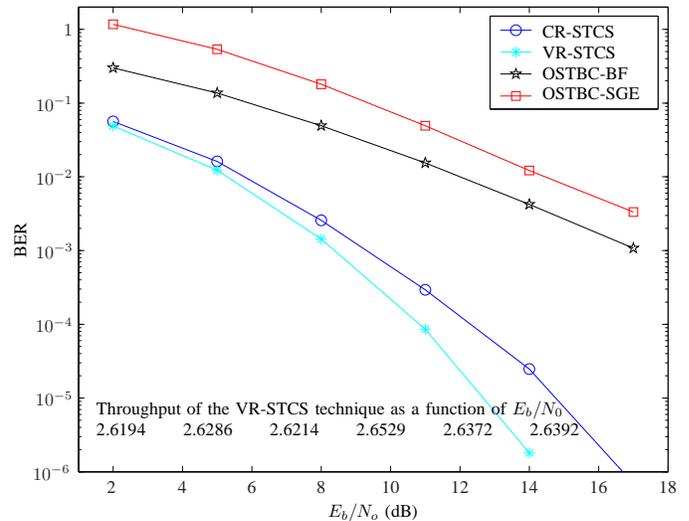


Fig. 4: BER and throughput comparison of the proposed technique (CR-STCS) with other schemes for a speed of 86 km/h. The throughputs of CR-STCS, OSTBC-BF and OSTBC-SGE techniques are constant at 3b/s/Hz.

## V. CONCLUSIONS

An adaptive space-time code selection technique for constant-rate transmit diversity wireless systems has been proposed. The proposed technique selects both the space-time code and the antennas to be energized, using the instantaneous channel state information and comparing it with a set of predefined threshold levels. In case no channel satisfies the established conditions, a complementary mode is chosen, in

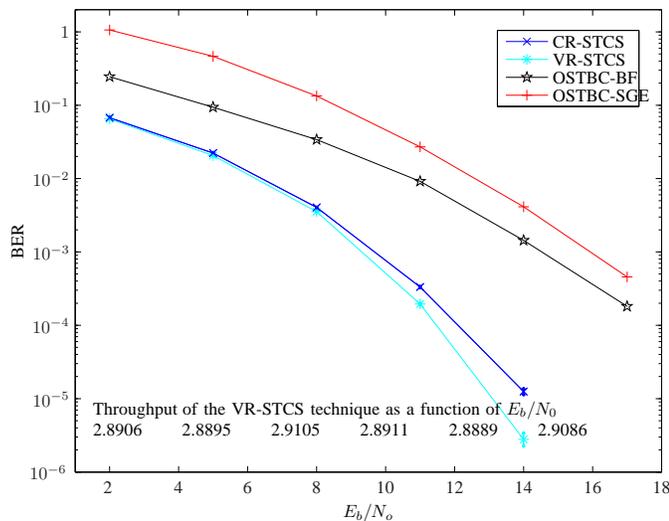


Fig. 5: Confidence intervals of the simulation results for a speed of 12 km/h (see Fig. 1).

which a single antenna is selected. The addition of the complementary mode allows to obtain a constant-rate transmission without sacrificing the spectral efficiency. Simulation results show that, when considering the channel state information obsolescence, the proposed technique widely outperforms other adaptive transmission techniques. Regarding the original space-time code selection technique, which provides a variable rate-transmission, the proposed technique slightly sacrifices the BER performance, in no more than 2dB in the worst case (for high relative velocity and high SNRs), but attains more than a 10% increment in spectral efficiency. Finally, the proposed technique achieves a more stable BER performance with respect to the relative velocity between transmitter and receiver.

## REFERENCES

- [1] S. Alamouti, "A simple transmit diversity technique for wireless communications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 8, pp. 1451–1458, Oct. 1998.
- [2] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block coding for wireless communications: performance results," *IEEE Journal on Selected Areas in Communications*, vol. 17, pp. 451–460, Mar. 1999.
- [3] M. Katz, E. Tirola, and J. Ylitalo, "Combining space-time block coding with diversity antenna selection for improved downlink performance," in *IEEE 54th Vehicular Technology Conference (VTC 2001)*, vol. 1, Atlantic City, NJ, USA, Oct. 2001, pp. 178–182.
- [4] Z. Chen, J. Yuan, B. Vucetic, and Z. Zhou, "Performance of Alamouti scheme with transmit antenna selection," *Electronic Letters*, vol. 39, no. 23, pp. 1666–1668, Nov. 2003.
- [5] I. Bahceci and T. M. Duman, "Antenna selection for multiple-antenna transmission systems: performance analysis and code construction," *IEEE Transactions on Information Theory*, vol. 49, pp. 2669–2681, Oct. 2003.
- [6] W. Wong and E. Larsson, "Orthogonal space-time block coding with antenna selection and power allocation," *Electronic Letters*, vol. 39, no. 4, pp. 379–381, Feb. 2003.

- [7] D. Gore. and A. Paulraj, "MIMO antenna subset selection with space-time coding," *IEEE Transactions on Signal Processing*, vol. 50, no. 10, pp. 2580–2588, Oct. 2002.
- [8] D. Mavares and R. P. Torres, "Space-time code selection for transmit antenna diversity systems," *IEEE Transactions on Vehicular Technology*, vol. 57, no. 1, pp. 620–629, Jan. 2008.
- [9] —, "Space-time code selection for transmit antenna diversity systems," in *Proceedings of the First Mobile Computing and Wireless Communication International Conference, 2006 (MCWC 2006)*, 2006, pp. 83–87.
- [10] H. Shin and J. H. Lee, "Exact symbol error probability of orthogonal space-time block codes," in *GLOBECOM 2002 - IEEE Global Telecommunications Conference*, vol. 21, no. 1, Nov. 2002, pp. 1206–1210.
- [11] D. Lu and D. So, "Performance based receive antenna selection for V-BLAST systems," *IEEE Transactions on Wireless Communications*, vol. 8, no. 1, p. 214, Jan. 2009.
- [12] H. Lee, S. Park, and I. Lee, "Orthogonalized spatial multiplexing for closed-loop MIMO systems," *IEEE Transactions on Communications*, vol. 55, no. 5, p. 1044, May 2007.
- [13] C.-Y. Chen, A. Sezgin, J. Cioffi, and A. Paulraj, "Antenna selection in space-time block coded systems: Performance analysis and low-complexity algorithm," *IEEE Transactions on Signal Processing*, vol. 56, no. 7, p. 3303, Jul. 2008.
- [14] Y. Kim, H. Lee, S. Park, and I. Lee, "Optimal precoding for orthogonalized spatial multiplexing in closed-loop MIMO systems," *IEEE Journal on Selected Areas in Communications*, vol. 26, no. 8, p. 1556, Oct. 2008.
- [15] Y. Fu, C. Tellambura, and W. Krzymien, "Limited-feedback precoding for closed-loop multiuser MIMO OFDM systems with frequency offsets," *IEEE Transactions on Wireless Communications*, vol. 7, no. 11, p. 4155, Nov. 2008.
- [16] Y. Kakishima, T. Kawamura, H. Taoka, and T. Nakamura, "Suboptimal precoding vector selection scheme combined with frequency domain scheduling for closed-loop MIMO transmission in LTE-advanced uplink," in *2010 IEEE 21st International Symposium on Personal Indoor and Mobile Radio Communications (PIMRC)*, 2010, p. 2466.
- [17] G. Jöngren, M. Skoglund, and B. Ottersten, "Combining beamforming and orthogonal space-time block coding," *IEEE Transactions on Information Theory*, vol. 48, no. 3, pp. 611–627, Mar. 2002.
- [18] J. Wu, H. Horng, J. Zhang, J. Olivier, and C. Xiao, "Combining orthogonal space time block codes with adaptive sub-group antenna encoding," *Intern. J. Wireless Info. Networks*, vol. 12, pp. 179–186, Jul. 2005.

# Handover Optimization in WiMAX Vehicular Communications

Mihai Constantinescu, Eugen Borcoci

University POLITEHNICA of Bucharest  
Bucharest, Romania

emails: [mihai.constantinescu@elcom.pub.ro](mailto:mihai.constantinescu@elcom.pub.ro)  
[eugen.borcoci@elcom.pub.ro](mailto:eugen.borcoci@elcom.pub.ro)

**Abstract**—The WiMAX communications for vehicular is a topic of significant interest for research and industry communities, on both V2V (vehicle-to-vehicle) and V2I (vehicle-to-infrastructure) situations. This paper presents results of an experimental study, simulation based, directed to determine multi-dimensional regions where different lower layers parameters have influence on the overall handover performance in mobility scenarios, related to WiMAX V2I communications. The simulations results are consolidated in multi-dimensional graphs, named decision spaces. Based on them, optimal parameter sets can be provided by the network operator to vehicular mobile station, to guide its adaptation of the major WiMAX parameters to its speed and network topology and to help the handover decision.

**Keywords**-WIMAX vehicular communications; antenna gain; scanning; handover; cross layer interaction

## I. INTRODUCTION

The WiMAX communications for vehicular use has gained a continuous attention from research community, on both V2V (vehicle-to-vehicle) and V2I (vehicle-to-infrastructure) situations. This paper presents a detailed experimental study related to WiMAX V2I communications.

This work aimed to determine multi-dimensional regions where different lower layers parameters have influence on the overall handover performance in mobility scenarios related to WiMAX V2I communications.

Given the high number of WiMAX physical and MAC layer parameters influencing in an inter-dependent mode the overall performance in mobility scenarios, experimental simulation studies of complex scenarios are very helpful to determine the combined effect on such parameters. In this work the simulations results are consolidated in multidimensional graphs, named decision spaces. These decision spaces present in aggregated form the performance obtained in a WiMAX V2I mobility scenario, related on a specific trajectory. The results can be used as a method of optimizing the vehicular communications by guiding the handover (HO) decision. Based on decision spaces, optimal parameter sets can be provided by the network operator to vehicular mobile station, in order to adapt the major WiMAX parameters to its speed and network topology.

The effect at different PHY parameters adjustments on the throughput values has been evaluated, to determine some decision regions space usable to optimize working ranges of

parameters in case those policies are applied to govern the handover. That is why we called the aggregated simulation results diagrams as “decision spaces”.

The optimal parameter sets obtained from decision spaces can be provided by a network operator to mobile station (MS), helping it to adapt dynamically its behavior and to obtain the maximum throughput possible from network at different speed, antenna gain, maximum transmission (Tx power), and scanning values.

The paper is organized as follows: the Section II presents some related work. The Section III defines the simulation context. The Section IV describes the simulation results. Conclusion, open issues and future work are shortly outlined in the Section V.

## II. RELATED WORK

In [1], the authors propose study the feasibility of using WiMAX for V2I communication on a static setting in urban environment and perform a comparison with use of WiFi.

Pegasus, a system providing wireless connection roaming at high velocities over multiple interfaces uses network information for user locations and used paths for effective and balanced utilization of the available bandwidth [2].

The reference [3] evaluates an architecture based on IEEE 802.21 framework, integrating both mobility and Quality of Service (QoS) mechanisms, through an advanced mobility scenario using a real WiMAX testbed.

In [4], mobile WiMAX trials are analyzed to investigate the vehicular downlink performance for a number of on-car antenna configurations.

This work presented is a continuation of a set of complex studies on WiMAX mobility. First results have been shown in a study of HO performance for WiMAX mobility [5], continued with an WiMAX HO conditions evaluation towards enhancement through cross-layer interaction proposed in [6], together with a SIP-based cross-layer optimization for WiMAX Hard HO method, described in [7].

## III. SIMULATIONS CONTEXT

There are numerous studies about WiMAX mobility and methods of optimizing the V2I communications. However, an analysis of V2I system behavior on incremental variance of speed, antenna gain, maximum Tx power, and scanning threshold/methods could add new value to existing optimization methods.

All simulations were done in OPNET v.14.5. [8] A typical mobility scenario has been considered (linear trajectory along a road where WiMAX BS stations are located). The BSes use the same set of frequencies and the mobile station (MS) is moving on a linear trajectory along the chain of BSEs. The utility of such a scenario is that it is similar with a road region in which WiMAX station are located along the road and the MS is a vehicle moving on the road.

The parameters taken into considerations (used in pairs in batteries of simulations) have been: (Table 1)

- MS Transmission Power (W)
- MS Antenna Gain (dBi)
- MS Scanning Threshold (dB)
- MS HO Threshold Hysteresis (dB)
- BS Transmission Power (W)
- BS Antenna Gain (dBi)

Table 1 WiMAX Parameters

WiMAX Parameter	1	2	3	4	5	6	7	8	9	10
A MS Maximum Tx Power (W)	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
B MS Antenna Gain (dBi)	-1	0	2	4	6	8	10	12	14	16
C MS Scanning Threshold (dB)	3	6	9	12	15	18	21	24	27	54
D MS HO Threshold Hysteresis (dB)	0.4	2	4	6	8	10	12	14	16	18
E BS Maximum Tx Power (W)	0.5	0.8	1	1.3	1.5	1.8	2	2.3	2.5	2.8
F BS Antenna Gain (dBi)	6	8	8	9	10	12	14	15	16	21

Each set of simulations considered the variance of a pair of parameters, the rest of them remain unchanged. The final value of average application throughput for each simulation is represented as a single point in the related decision space. In that way, each decision space covers 100 instances of a scenario, describing in details the behavior of V2I communication quality under effects of WiMAX parameters pair variance, on that particular trajectory and network topology.

#### IV. SIMULATION RESULTS

Due paper size limitations, only MS Maximum TX Power-MS Antenna Gain, MS Antenna Gain-MS Scanning Threshold, MS Scanning-Method MS Speed, MS HO Threshold Hysteresis-MS Speed, MS Maximum TX Power-BS Antenna Gain, and MS Maximum TX Power - MS Antenna Gain decision spaces will be presented. Each vertical section on decision space provides an analysis of system behavior under influence of a single parameter variance (depending on the selected axis).

##### A. Influence of Maximum transmission power and MS antenna gain

This experiment has been simulated while letting the Maximum TX power and antenna gain to pass through all specified values, for a MS speed of 10m/s and respectively 50 m/s. The scanning method had parameters: N=4 P=240 T=10. N=scanning (frames), P=interleaving (frames), T=iterations. The rest of parameters have been constant.

Simulations set: A\_B\_C01 D01 E01 F01 for low speed (s=10m/s).

The diagrams from Fig. 1 and Fig. 2 show MS Maximum TX Power-MS Antenna Gain decision space for MS speed of 10m/s and respectively 50 m/s, giving an overall idea on the relative performance without presenting details on each HO action.

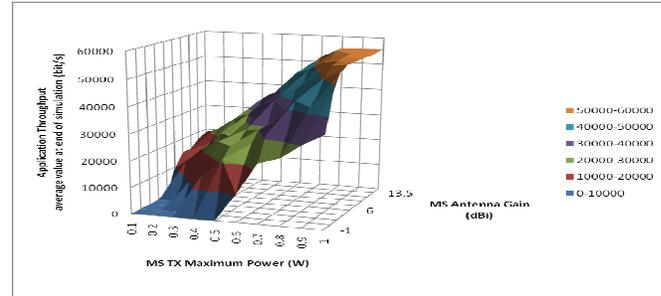


Figure 1. MS Maximum TX Power - MS Antenna Gain -- Decision Space for MS Speed 10m/s

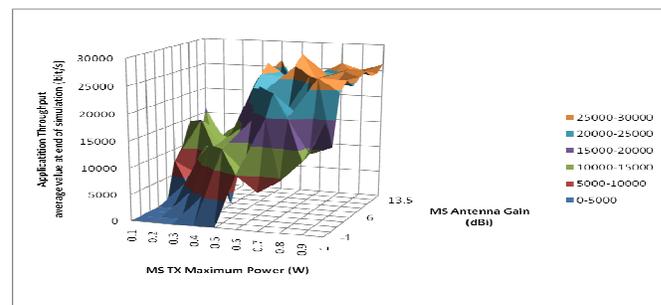


Figure 2. MS TX Maximum Power - MS Antenna Gain -- Decision Space for MS Speed 50m/s

It is seen that speed has a major impact on the throughput. Out of theoretical throughput of 64 kbps (source rate) for high speed  $s=50$  m/s, the results are significantly worse than for low speed as 10 m/s, for the same range of TX power and antenna gain. While for  $s=10$  m/s we get for sufficient TX power ( $>0.5$ W) and antenna gain ( $>8$ Db) a throughput closer to 64 kbps, while in the same conditions we get for the  $s=50$  m/s case, only something close to 30kbps, i.e. half. Therefore, in such cases MS need a higher TX power and higher antenna gain. Also in the case of high speed, the effect of vehicular channel is higher than for low speed (see the non-monotonic behavior of the second diagram).

Fig. 3 shows details of important antenna gain effect in the conditions where only Maximum TX power is varied (vertical section on MS Maximum TX Power-MS Antenna Gain decision space for MS speed 10m/s).

The diagrams show that even if we have an increase of maximum TX power ten times ( i.e. from 0.1W to 1W), the overall throughput achievable is modest one ( i.e. only 22kbps), if  $g=0$ dB, while for a gain of  $g=10$ dB ( right diagram) it is seen that the throughput increases more than 100%. This result clearly shows the benefit of a directional antenna, which may have a high gain versus a conventional omni directional one.

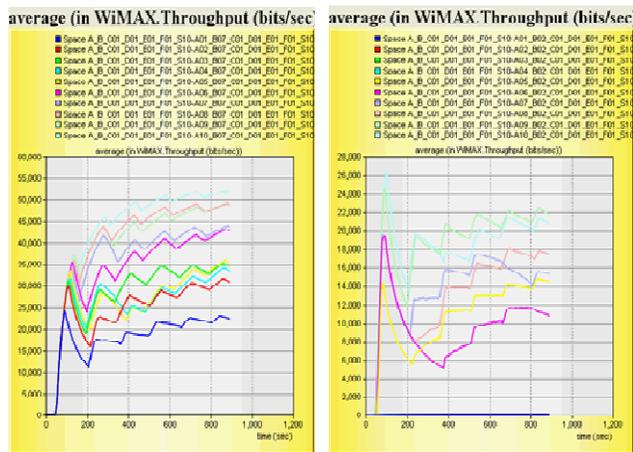


Figure 3. Effect of antenna gain ( $g = 0\text{dBi}$  – left, and  $g = 10\text{dBi}$  – right) on throughput in Maximum TX power variation conditions,  $s = 10\text{m/s}$ .

Fig. 4 shows the same diagrams as in Fig. 5 but for  $s = 50\text{ m/s}$  (vertical section on MS Maximum TX Power-MS Antenna Gain decision space for MS speed 50m/s).

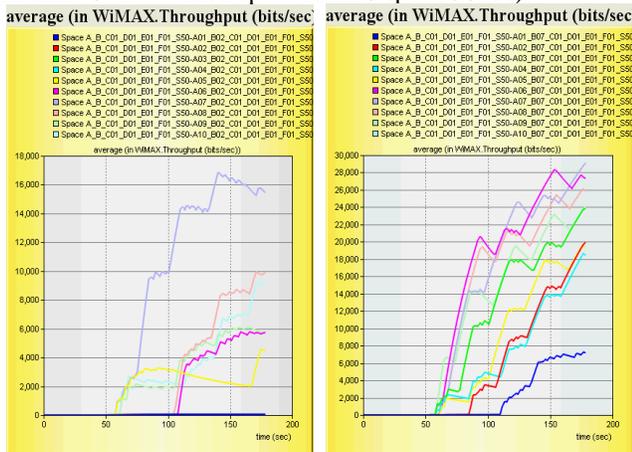


Figure 4. Effect on throughput of Maximum TX power variation conditions, while the antenna gain is fixed ( $g = 0\text{dBi}$  – left, and  $g = 10\text{dBi}$  – right);  $s = 50\text{m/s}$  (180 km/h).

It is seen that high speed would worsen the behavior. For instance, for  $s = 10\text{ m/s}$  we get a throughput of 55 kbps (out of 64 kbps), in the conditions ( $P = 1\text{W}$  and  $g = 10\text{dBi}$ ), while for  $s = 50\text{ m/s}$ , even the gain and maximum TX power is high, the maximum throughput at the end of simulation is only close to 28 kbps. Still the gain of antenna is important (increase from 16kbps to 28 kbps).

Therefore, even in adverse condition (related to Maximum TX power), the gain of the antenna can significantly improve the performance. The left diagram confirms that, even if MS has high Maximum TX Power, a very low antenna gain prevents the system of getting sufficient performance. The HO gaps have still large values in time (tens of seconds) at  $P = 1\text{W}$ . On the other side, if MS has a gain of  $g = 10\text{dBi}$  (right diagram), it is shown that throughput is good in the range  $P > 0.7\text{W}$  and increasing this power does not bring significant additional increase in throughput. These are the confirmation on detailed

experiments of the conclusions drawn for synthesis diagrams.

### B. Influence of MS antenna gain and scanning threshold

This section studies the scanning process influence on the performance in the given linear configuration and together with other parameters among which we are mainly interested in antenna gain effects.

A short summary of scanning process is given here as reminder. When the fading SNR reaches the scanning threshold, the MS begins with scanning process on the announced DL channels by sending the MOB-SCN-REQ message. The BS allows for scanning by replying with the MOB-SCN-RSP message that contains the parameters for scan duration ( $N$ ), interleaving interval ( $P$ ) and the start frame ( $M$ ). After receiving the MOB-SCN-RSP from the target BS, the SS starts the scanning after  $M$  received frames (start frame). The SS changes after  $M$  frames to the next channel and stays there for an  $N$  frames period (scanning interval/duration) to detect a BS and to assess its SNR.

After a scanning interval, the SS returns to the DL channel of the active BS. This behavior aims to keep the interruption as short as possible since no payload transmissions are possible during the scanning process. If no preferable BS could be detected on the scanned channel, it reinitiates the scanning mode after a  $P$  frames period (interleaving interval) to find a new BS. The total number of allowed repetitions of the scanning process is given with the parameter ( $T$ ).

The diagrams from Fig. 5 and Fig. 6 show the cumulative throughput at the end of simulation time, while the antenna gain and scanning threshold have been varied, again taking two MS speed values of 10 and respectively 50 m/s.

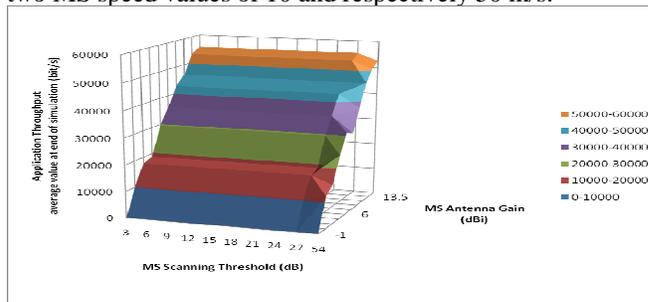


Figure 5. MS Antenna Gain - MS Scanning Threshold -- Decision Space for MS Speed 10m/s

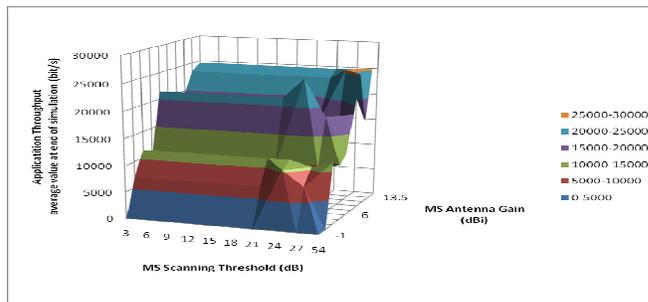


Figure 6. MS Antenna Gain - MS Scanning Threshold -- Decision Space for MS Speed 50m/s

Simulations set: A05B\_C\_D01 E03 F01\_S10 for low speed (s= 10m/s) and A05B\_C\_D01 E03 F01\_S50 for high speed (s= 50 m/s).

Scanning Method: N=4 P=240 T=10.

As expected, the diagrams illustrate the fact that antenna gain has a major effect on the throughput. On the contrary, in this scenario adjusting the scanning threshold for a given value of gain, does not change significantly the throughput, except the high speed scenario (s = 50 m/s) when taking a more sensible scanning threshold (> 21 dB) can bring some throughput raising with 10 – 20%. The explanation is that making the scanning threshold more sensible does not help if the major factor of antenna gain is low. The speed influence is again a worsening one: the maximum throughput that can be achieved for speed of 50 m/s versus 10 m/s is about a half.

These conclusions does not mean that scanning parameters are not important to be adjusted properly, but show that the context is important and in different contexts, scanning activity might be more or less critical.

The two sets of diagrams clearly show the dramatic effect of antenna gain increase on throughput increase, while the variation of scanning threshold is not so significant. For different MS speed values, the density of scanning has more effects that are visible.

C. Influence of MS scanning method and MS speed

The following set of simulations experimented different scanning methods (Table 2). N=scanning (frames), P=interleaving (frames), T=iterations.

Simulations set: A05 B05 Cx1-x6 D01 E03 F01\_S10-50 (it covers all speed range, from 10m/s to 50m/s).

Table 2 Scanning method parameters

MS scanning method	Cx1	Cx2	Cx3	Cx4	Cx5	Cx6
MS Scanning Threshold (dB)	10	20	30	40	50	60
N (frames)	30	25	20	15	10	5
P (frames)	50	100	150	200	250	250
T	10	10	10	10	10	10

It is seen that Cx1 represents a scanning method having relative large N/P ratio, i.e. the relative time spent with the scanning is larger versus the time spent to transmit the data payload. At the other end of the range, Cx6 has small N/P, i.e. the scanning relative time is less than the time spent for data transmission. On the other side, the scanning threshold has been adjusted as to compensate in a certain measure this scarcity of spanning activity, by taking a higher value of the scanning threshold (60 DB).

Fig. 7 shows an aggregated diagram in which the scanning threshold and scanning interleaving are varied on one dimension and the MS speed on the other dimension. It is seen that for low MS speed the scanning method is not so critical, therefore, a light scanning (small N/P) is sufficient to allow more relative time for data transmission. However, a dense scanning method is very effective for high speed, when mobile is quickly aware about the next BS available, and the low scanning method implies a slow reaction of mobile to communication condition changes.

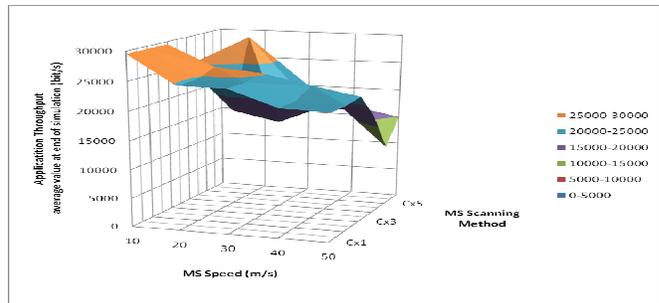


Figure 7. MS Scanning Method – MS Speed – Decision Space

Fig. 8 shows the throughput in two extreme cases in the range experimented: (Cx1- CX6 at s= 10 m/s – left and Cx1- Cx6- at s= 50 m/s – right). It is seen that at s= 50 m/s a more dense scanning is better (Cx1 and Cx3).

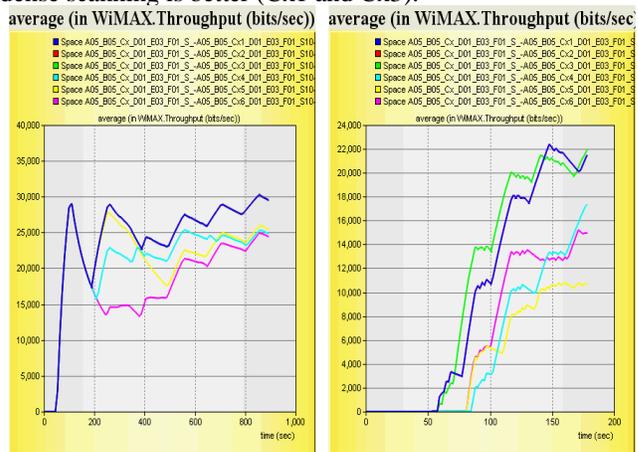


Figure 8. Influence of scanning density, scanning threshold variation and speed on throughput – sample cases

D. Influence of MS hysteresis threshold and MS speed

This section studies the effect of the hysteresis threshold values selection when the speed is also varied.

Simulations set: A05 B05 C08 D01-10 E03 F01\_S10-50

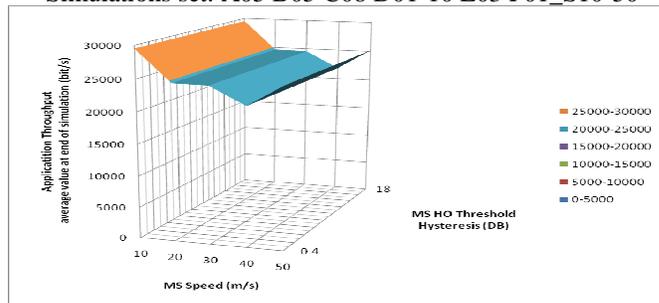


Figure 9. MS HO Threshold Hysteresis - MS Speed – Decision Space

The hysteresis can be used to select BSs as that are suitable candidates for the target BS in a HO. When finding the candidate BSs, the MS (or simulation) may compute the difference between the CINR of the serving BS and the CINR of the potential target BS. The value of this attribute (Hysteresis threshold) specifies the minimum amount by which the CINR of potential target BS must exceed that of

the serving BS. The value of new signal to noise ratio should be greater with the Handover Threshold Hysteresis than the current value in order that the HO can be triggered. An auxiliary Multi-target Handover Threshold Hysteresis (its value is less than Handover Threshold Hysteresis) can also be used to select among scanned possible target BSes before handover triggering.

As observed in Fig.9, there is no modification for different values of MS HO Threshold Hysteresis, so that parameter has no importance in an environment as that simulated one. That conclusion is applied for that context only; due different network topologies with different WiMAX parameter values could produce different results.

**E. Influence of MS Maximum TX Power and BS Antenna gain**

This section will study the effect on throughput of the MS Maximum TX Power and BS antenna gain.

Simulations set: A\_B05 C08 D07 E03 F\_S10 for MS speed 10m/s and A\_B05 C08 D07 E03 F\_S50 for MS high speed (s= 50 m/s).

A major conclusion, highlighted in both Fig. 10 and Fig.12, is that BS antenna gain has a major impact on MS observed throughput at all ranges of MS TX Power.

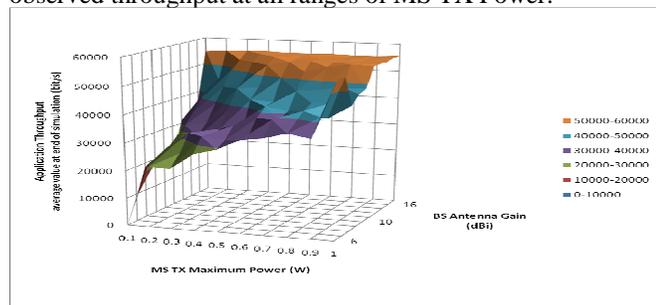


Figure 10. MS Maximum TX Power - BS Antenna Gain -- Decision Space for MS Speed 10m/s

For low speed, the application throughput grows with MS Maximum TX Power increase or with BS antenna gain. There is an important growth with MS Maximum TX Power for low values of BS Antenna gain, which will be less important as the BS antenna gain increase (see the 50000 – 60000) region in the figure.

Sample vertical sections of MS Maximum TX Power-BS Antenna Gain Decision Space for MS Speed 10m/s are presented in Fig. 11, describing the V2I system behavior under effect of BS antenna gain variance.

One can see that at low MS power (left-upper part diagram) the influence of the BS antenna gain is dramatic-which is normal in such low MS power condition. When the MS Maximum TX power is increasing this compensate a lower BS antenna gain and the throughput is better even for lower values of the BS antenna gain. The best results are obtained (bottom-right diagram) for sufficient power at MS (1.0W) and high BS antenna gain  $g = 16\text{dBi}$ . On the other side is to be observed that sufficient throughput can be obtained with less power at MS (i.e. for  $P > 0.6\text{W}$ ). This can give the possibility to apply policies in adjusting the maximum MS TX power in function of current condition.

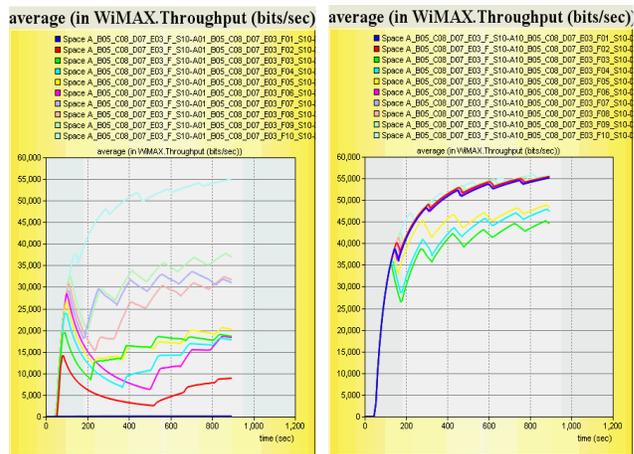


Figure 11. Throughput for different values of MS Maximum TX power and variation of the BS antenna gain (MS TX P = 0.1, 1.0W) s = 10 m/s

For high speed, the application behavior is more complex (Fig. 12).

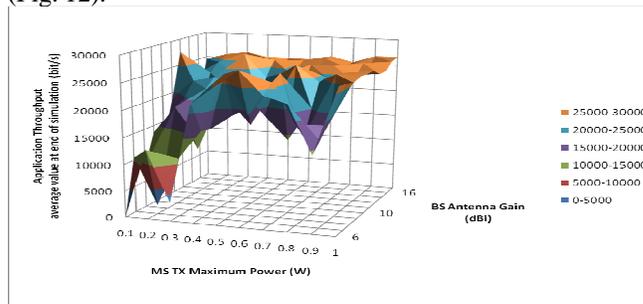


Figure 12. MS Maximum TX Power - MS Antenna Gain -- Decision Space for MS Speed 50m/s

In the simulated scenario (versus network topology and trajectory), there are optimal values for MS Maximum TX Power (ex. 0.5W for BS Antenna Gain -1dBi). These optimal values are different for each BS Antenna Gain value. The mobile could adjust its Maximum TX Power to the value indicated in Decision Space and it will obtain the maximum application throughput. There is no clear dependency between MS Maximum TX Power increasing or BS Antenna Gain and application throughput growth, as obtained for low speed.

The qualitative results for high speed are the same as for low speed, but the overall throughput is significantly lower (roughly twice), even for high MS TX Power and high BS antenna gain.

**F. Influence of MS and BS Antenna gains**

The MS and BS antenna gains are expected to have major influence on the overall performance.

Simulations set: A05 B\_C08 D07 E03 F\_S10 for low speed (s = 10 m/s).

For MS low speed (s = 10 m/s) the Fig. 13 diagram shows a rather monotonic increase of the throughput with both BS and MS antenna gain on both dimensions. However a “triangle” in the space (BS-gain, MS-gain) of saturation (region 50000- 60000) is seen (top-right side) where no

increasing in throughput is possible; in other words, the BS gain and MS gain “cooperates” and we obtain sufficient throughput if an empirical (approximated) relationship is fulfilled:

$$BS\_gain + MS\_gain > a$$

where limit  $a$  can be determined from the figure. This observation gives the possibilities to apply policies to cross-adjust the BS antenna gain and MS antenna gain.

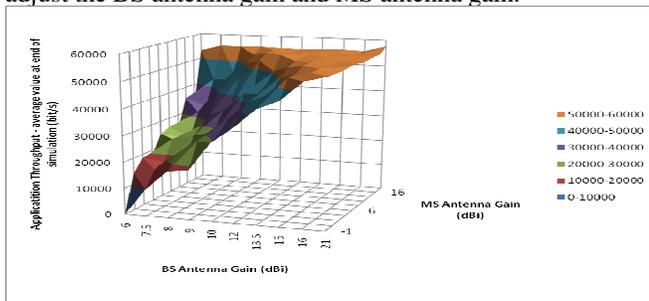


Figure 13. MS Antenna Gain - BS Antenna Gain -- Decision Space for MS Speed 10m/s

The simulations highlights that for high speed (Fig. 14), the application throughput has a approximately constant level at middle range of parameters, where the variance of MS Antenna Gain and BS Antenna Gain has no major effect, as opposite with low speed situations, where that constant level is reached near the highest range of parameters.

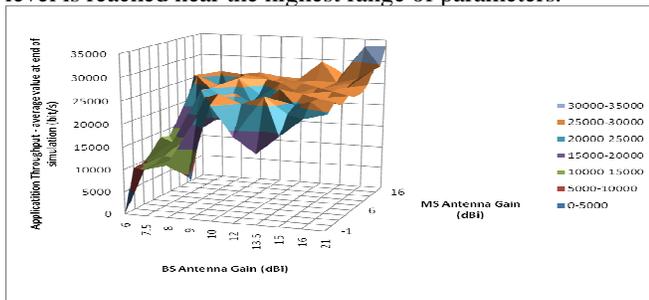


Figure 14. MS Antenna Gain - BS Antenna Gain -- Decision Space for MS Speed 50m/s

If BS antenna gain is  $g = 9\text{dBi}$  then after at  $MS\ g \geq 10\text{dBi}$ , we do not get further improvement (left diagram). A similar effect is seen in the right figure for  $g \geq 4\text{dBi}$ .

## V. CONCLUSIONS

### A. Conclusions related to simulations

One important conclusion of the simulation studies is that the antenna gain of MS and BS has major impact on throughput in a large range of other parameter variation (speed, power, scanning threshold, hysteresis threshold, etc.). Having a good antenna gain at both BS and MS is essential in all contexts. This is a reason to use directional antennas, which expose such properties.

As for other parameters (speed, scanning threshold, hysteresis threshold, etc.) the results obtained show that the behavior of throughput versus variation of parameter combinations is not a monotonic one and is context

dependent. This conclusion is however natural for a system that has tens of inter-dependent working parameters, but cross-optimizations are possible in the sense that the results of the decision spaces can be used in policies governing the ranging, scanning, HO, etc.

### B. Conclusions related to decision spaces method and cross-layer algorithm

The analysis of V2I system behavior on incremental variance of speed, antenna gain, maximum Tx power, and scanning threshold/methods under decision space matrix provides a data base usable for optimization methods/techniques. Each decision space allow not only the combined effect details of a pair of parameters variance, but allow to predict the system behavior for each specific parameter value (see decision space vertical section examples presented in Fig. 11).

Network operator could use such kind of extended simulations for different roads and highway, where the network topology and the road details are known. Vehicles provided with WiMAX terminal capabilities passing these roads could be helped to optimize the communications using cross layer-algorithm based on location and speed prediction from GPS information and optimal parameters set from network operator.

## ACKNOWLEDGMENT

The work on mobility presented in this paper has been performed within a European research project SMART-Net (SMART-antenna multimode wireless mesh Network), focused on the possible advantages if using directional antennas in mobility scenarios.

## REFERENCES

- [1] C. Chou, C. Li, W. Chien, and K. Lan, “A Feasibility Study on Vehicle-to-Infrastructure Communication: WiFi vs. WiMAX”, 2009 Tenth International Conference on Mobile Data Management: Systems, Services and Middleware, pp. 397-398.
- [2] N. Frangiadakis and N. Roussopoulos, “A Pegasus over WiFi and WiMax Demo: Connectivity at high speeds”, 2009 5th International Conference on Testbeds and Research Infrastructures for the Development of Networks & Communities and Workshops, pp. 1-3.
- [3] R. Matos, B. Sousa, P. Neves, S. Sargento, and M. Curado, “Advanced Mobility in Broadband Wireless Access Scenarios”, 2009 IEEE International Conference on Wireless and Mobile Computing, Networking and Communications, pp. 214-220
- [4] G. Zaggoulos, A. Nix, and D. Halls, “Novel Antenna Configurations for Wireless Broadband Vehicular Communications”, 2010 6th International Conference on Wireless and Mobile Communications, pp. 21-26.
- [5] M. Constantinescu and E. Borcoci, "Handover Performance Evaluation in WiMAX Mobility Scenarios", 7th International ICST Conference on Broadband Communications, Networks, and Systems, BROADNET-MOWAN 2010
- [6] M. Constantinescu, E. Borcoci, and T. Rashid, "WiMAX Handover Conditions Evaluation Towards Enhancement Through Cross-layer Interaction", 2010 Sixth Advanced International Conference on Telecommunications, pp. 422-427
- [7] M. Constantinescu and E. Borcoci, "A SIP-Based Cross-Layer Optimization for WiMAX Hard Handover", The 8th International Conference on Communications, COMM2010, pp. 375-378
- [8] [www.opnet.com](http://www.opnet.com), 01.02.2011

# Mobile Phone Security Awareness and Practices of Students in Budapest

Iosif Androulidakis

Jožef Stefan International Postgraduate School  
Jamova 39, Ljubljana SI-1000, Slovenia  
[sandro@noc.uoi.gr](mailto:sandro@noc.uoi.gr)

Gorazd Kandus

Department of Communication Systems  
Jožef Stefan Institute  
Jamova 39, Ljubljana SI-1000, Slovenia  
[gorazd.kandus@ijs.si](mailto:gorazd.kandus@ijs.si)

**Abstract**—The present paper presents the results of a survey about users' security practices regarding mobile phone usage that took place in 4 Universities of Budapest in February 2010. We targeted an extended pool of respondents reaching 959 answers. The general users' feeling is that mobile phone communication is secure and this possibly leads to a relaxation. As results indeed further showed, students are unaware of the necessary measures to avoid a possible unauthorized access and/or sensitive data retrieval from their phones and that they lack proper security education. There was also a statistically important difference in the answers, depending on the type of operating system (modern or not). Since users fail to secure their phones they should either be educated or preferably presented with transparent security features, built in their phones, in order to mitigate the dangers.

**Keywords**—*mobile phone security; security practices; user interface security; questionnaire survey; mobile phone usage*

## I. INTRODUCTION

Mobile devices are becoming a critical component of the digital economy, a style statement and useful communication device, a vital part of daily life for billions of people around the world. Modern mobile phones' enhanced capabilities allow them to be almost as versatile as a computer becoming a valuable business (mobile applications) and entertainment tool (mobile games, m-commerce). At the same time users store and process more data including sensitive information in their phones. A few years ago the only concern of a mobile phone user would be his communication privacy.

This is not the case anymore. Users have to be protected from unauthorized third party access to their data. Apart from the traditional security measures such as PIN (Personal Identification Number) usage and voice encryption, users have to take extra security measures and to follow new best practices. Unfortunately, as the survey revealed, users aren't adequately informed about security issues in regards to their mobile phones' options and technical characteristics and fail to follow proper security measures and practices. In Section II, related work is examined. The methodology used for the survey is described in Section III. Results are presented in Section IV, closing with conclusion and future work in Section V.

## II. RELATED WORK

Although there have been quite many theoretical studies concerning mobile services, a significant means for investigating and understanding users' preferences is asking their opinion via specific questioning techniques. The vast majority of these surveys indicate the growing importance of mobile phones in everyday life and the increased popularity of new features [1].

In any case, the security of mobile phones is proven not to be adequate in many research papers [2][3]. There also exist several survey studies in this direction. Some of these surveys studies focus on mobile phone's security issues [4][5] while others on mobile phone services, touching also security issues [6][7]. Modern smart phones, specifically, are open to more security risks [8].

A recent survey [9] published in November 2008 focused on mobile phones security issues and in which degree these issues concern the users. The conclusion was that a major part of the participants are extremely concerned about security and don't want any of their private data to be available to 3rd party unauthorized users.

It is interesting to note that according to other surveys [10] a major part of the participants is interested in mobile services adoption only if the prices are low and the security framework tight enough. At the same time, cyber security and safety education is left out from the educational system [11] and users do not know if their phones are secure or not [12]. Given the fact that mobile phones could be a dominant feature of future classroom, special security awareness and training courses, presenting the necessary guidelines, should definitely be implemented in schools. This is why the present paper tries to address users' security awareness and practices, as an enabler for greater mobile services market penetration.

## III. METHODOLOGY

A very useful evaluation method for surveying user's practices is the use of multiple-choice questionnaires (i.e. in person delivery or e-mail questionnaires) [13][14]. Our survey was conducted using in-person delivery technique, with a total of 959 respondents participating in this survey. This method was selected from other alternatives because is more accurate and has a bigger degree of participation from

the respondents (e-mail questionnaires usually treated as spam mail from the respondents or they might misunderstand some questions). Data entry took place using custom software [15]. Due to lack of financial resources the survey was limited to Europe. An interesting approach would be to use social networks such as Facebook to amend the results of the survey, especially targeting students from United States and other continents.

The target group of the survey was university students from ages mostly 18-26, incorporating both younger and older youth segments (24-26 years old percentage was 25.5%) because these ages are more receptive to new technologies. They also understand better the technological evolution than older people who use mobile phones mostly for voice calls.

In the analysis of the security feeling and the security knowledge a simple mathematical formula was developed to produce numerical values. We weighted the responses with the following weights: Very Much: 4, Much: 3, Moderately: 2, Not much: 1, Not at all: 0 and then divided by the number of occurrences, in order to get a mean value.

#### IV. RESULTS

The questionnaire was divided in two parts. In the first part participants were asked some demographic data including gender, age and field of studies as well as some economic data including mobile phone usage, connection type and budget spent monthly on phone service. In the second part we proceeded to our main contribution, the specific questions related with their practices and security perceptions regarding mobile phones' security issues.

##### A. Demographics

56.3% of the participants were females and 43.7% were males. Most of the respondents, in turn, were aged 18-26 (82.4%). The main body of respondents was studying Economics or Business Administration (30.1%) Following in the sample there were students of Humanities or Philology (22%), Engineering, Mathematics or Natural Sciences (13.7%), Medicine (13.2%), Law (10.8%) and other fields (10%).

Regarding mobile phone usage, 60.3% of them are using daily a single mobile phone, with some 21% using two phones regularly and even 10% using more than two phones. Nokia is the favourite brand, reaching one third of students (34.3%) followed by Sony-Ericsson (21.3%) and Samsung (17.6%) (Figure 1). Apple's iPhone (which is expected to have a higher percentage in the US market) has a very descent 7.8% of penetration given the generally low budgeted section of the population targeted. It is immediately apparent that focusing on Nokia and Sony-Ericsson phones a security awareness campaign would immediately target more than half of users yielding a very high return. Of course the brand itself is not enough to categorize attack vectors and practices, since there is also

the feature of the specific operating system running on each phone.

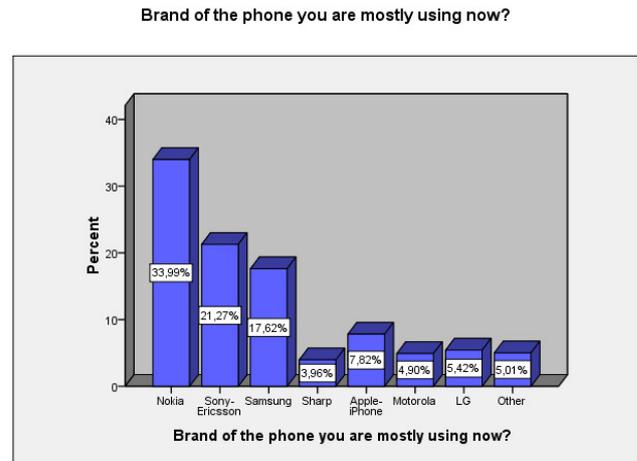


Figure 1. Favourite brands.

##### B. Economics

Proceeding to economics, participants were asked whether they are using a pre-paid or post-paid (contract) mobile phone connection. Half of students are using a contract based subscription, a rather high percentage, while 17.2% have both prepaid and postpaid SIMs (Subscriber Identity Module).

Answering how much money they spent monthly, student mobile phone users had a wide range of financial capabilities. The leading 25.7% spends 11-20 Euros (currency converted) monthly while almost equal parts of 20% spend 21-30, 31-40, or more than 40 Euros per month.

##### C. Security Specific Questions

The objective of this particular subsection and the main contribution of our research were to determine whether our participants acknowledge some security related features of their phone and what is their security feeling. The results are analysed in the following paragraphs.

Our fundamental research question was whether students are informed about how the options and the technical characteristics of their mobile phones affect the security of the latter and whether they are taking the necessary measures to mitigate the risks. The results that follow are totally in line with the initial response of students that only 29.6% believe they are much or very much informed while 42.6% state that they are not at all (a large 20.4%) or not much (Figure 2).

Are you informed about how the options and technical characteristics of your mobile phone affect its security?

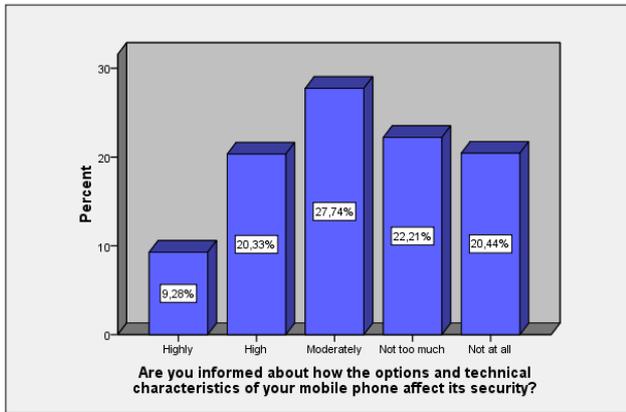


Figure 2. Knowledge of mobile phone security aspects.

Using the simple formula described in Section III (Methodology), the mean “security knowledge value” was 1.76, in the 0-4 scale (0 not at all, 4 very much). Further correlating their responses to the type of operating system–O/S (modern or not) proved that students owning phones without modern operating system have statistically (Pearson Chi-Square) better knowledge of security aspects than those who actually own a phone with modern O/S (Figure 3). As it was expected users that do not know the type of their O/S were the least informed about security.

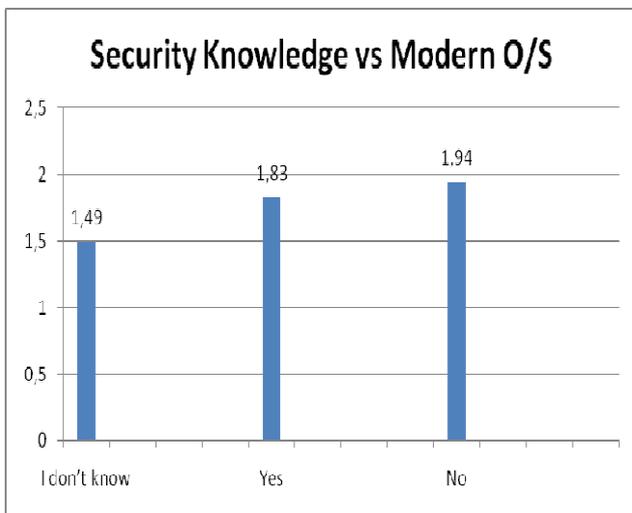


Figure 3. Security knowledge value vs. operating system.

Continuing with a general question about how “safe” mobile phone users feel, the majority (30%) replied “high (much)” followed by 26.7% “moderately” (Figure 4). On the other hand, some 27.9% felt not too much or not at all sure they are safe. This general feeling of security in turn

leads to an over-relaxation of students in regards to security practices as following answers reveal.

How safe do you consider communication through mobile phones?

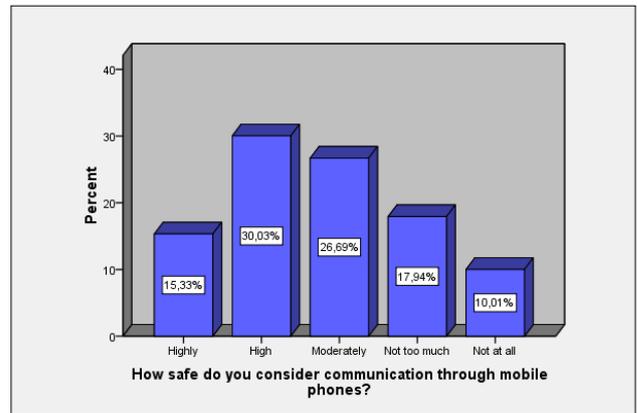


Figure 4. How safe do you consider communication through mobile phones?

Using the same methodology, the mean “security feeling” value was 2.22 in the 0-4 scale. The correlation to the operating system showed that users without modern O/S feel statistically (Pearson Chi-Square) the least secure while users that do not know the type of O/S are more “relaxed” (Figure 5).

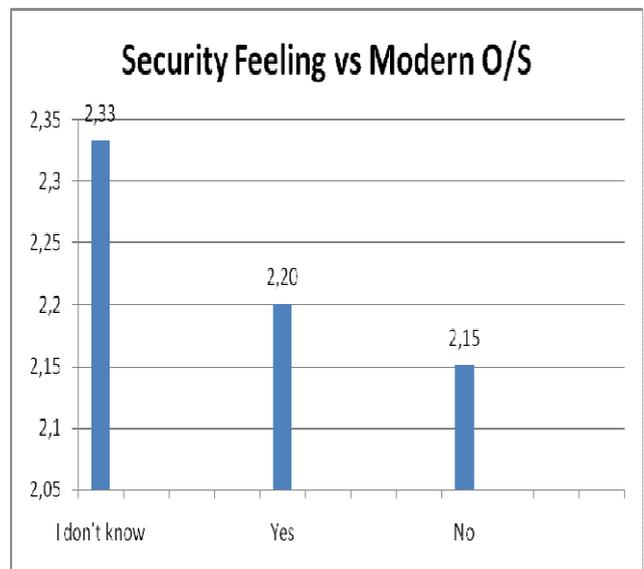


Figure 5. Security feeling value vs. operating system

In regards to operating system itself, a significant percentage of the participants (33.2%) doesn't know about the capabilities of his phone's operating system. Almost the same percentage (31.6%) of students is using mobile phones

with an advanced operating system. In any case, apart from the relaxation in security awareness that was previously shown, the ignorance of the type of operating system renders users more vulnerable to hacker attacks with the use of exploits specifically targeted for their phones.

Similarly, in Figure 6, only a very small percentage of the participants (less than 24%) knows his/her phone's IMEI (International Mobile station Equipment Identity) and has noted it somewhere. IMEI is very significant because if the phone is ever stolen, using this serial number the provider can block access to the stolen phone effectively mitigating stealing risks. Almost half of students are completely unaware of its existence. Knowledge of this feature would possibly help 41.1% of them who unfortunately had their phone stolen once or more (Figure 7). Similarly high percentages are noted by other studies too [16][17].

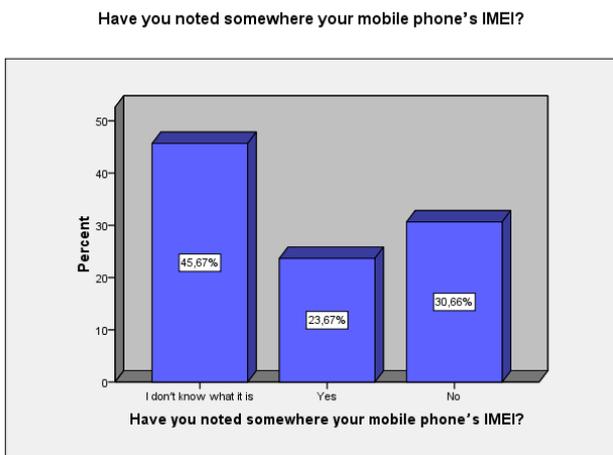


Figure 6. IMEI knowledge.

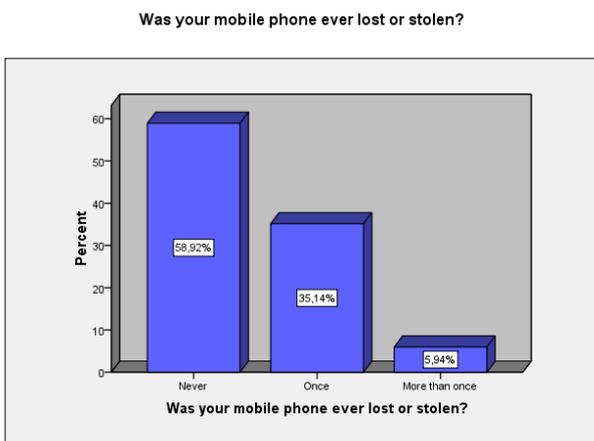


Figure 7. Lost or Stolen phone.

At the same time, 71% of users are not aware of the existence of the special icon that informs the user that his/her phone encryption has been disabled [3]. Ignorance of this security icon leaves users vulnerable to man in the

middle attacks since they can't recognize the attack taking place. This was probably the most expected result as even professionals are not aware of this feature and another hint that user interfaces should help and not obscure security.

Users, as expected, are actively (almost 70%) using SIM's PIN code. The negative finding that Figure 8 reveals is that only a small percentage (24.5%) uses screen-saver password while similar percentages do not know if their phone has such an option. That leaves 75% of users without a screen saver password, and their phones ready to be manipulated by "malicious" hands. An attack can take place in a few minutes by downloading specific software to the phone; this is why it is not enough to protect the phone only by PIN but also by a screen saver password.

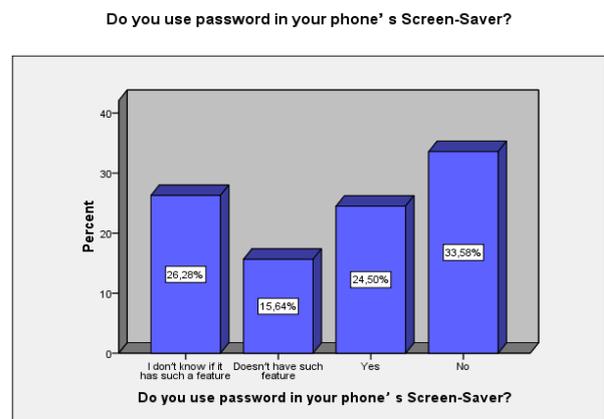


Figure 8. Screen-saver password.

A great attack vector of the past, Bluetooth, seems not to be the problem anymore (Figure 9). Just one out of five students has Bluetooth switched on and visible (leaving the phone vulnerable), while 42.3% of users have it switched off. It is not clear whether this is a security practice or a social practice that stemmed from the continuous harassments messages over Bluetooth caused upon users.

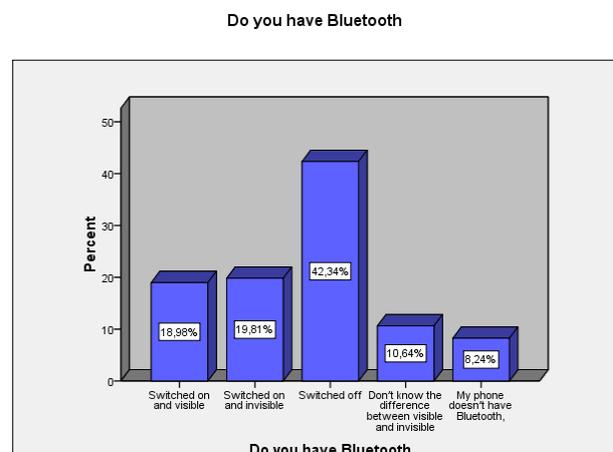


Figure 9. Bluetooth.

In a question that touches upon issues of politeness and openness, 44.7% of students are lending their phones, but only while they are present (Figure 10). This is a major factor that compromises the phone's security even if the participant is present, because a single minute is needed for someone to install malicious software in the phone. In that respect 36.2% of them refuse to lend their phone in any case being better safe and "impolite" than sorry.

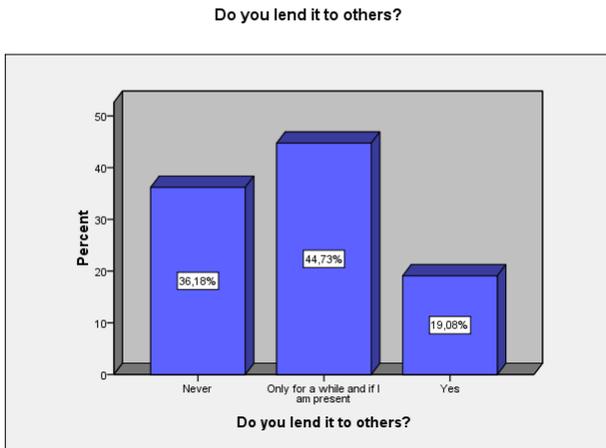


Figure 10. Phone lending.

Following, in Figure 11, with a question of both security and economic importance, almost 60% of participants don't download any software at all. There is also a 13% that actively downloads ringtones or logos, a 16% that tries applications and just 11% of "gamers". It is well interesting to note that security considerations is one of the hindering factors of mobile phone downloading [2]). In the antipode, getting familiar with downloading users are being more vulnerable to downloading and using unauthorised software that can harm their phone.

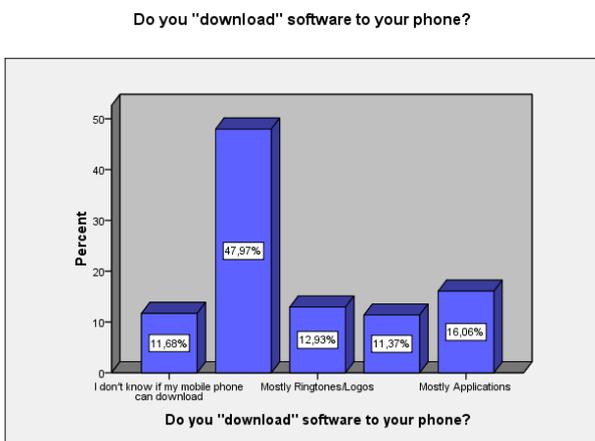


Figure 11. Software downloading.

This is where a mobile phone Antivirus would help. In our case (Figure 12), 19% of users acknowledge it exists such a product but don't use it, while 44% do not know whether such a product exists. That leaves 12.3% using it. Compared to PC users where nowadays everybody is using (at least) an antivirus shows a clear lack of security education and different mind-set. Organizations, in turn, show an increase in mobile phone antivirus tools usage [18]

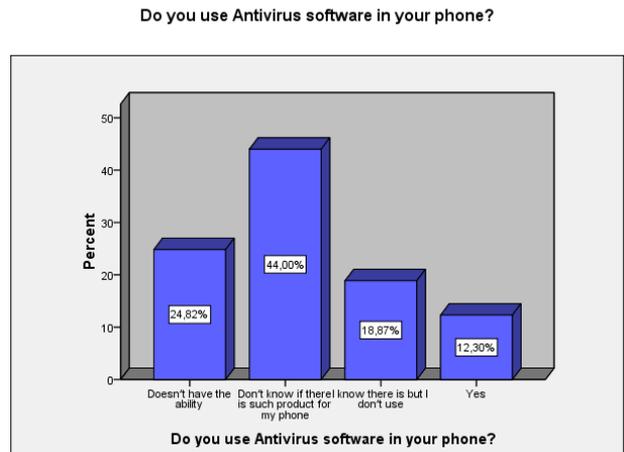


Figure 12. Anti-virus usage.

Being young, 57% of university students keep sensitive information into their mobile phones (Figure 13). It seems that we consider our mobile phone to be a very personal device and we save equally important and sensitive information there. Such kind of information should be protected but again, the results from our survey show that users fail to do so. The consequences from a breach of data of this type could be devastating for the life of the victim.

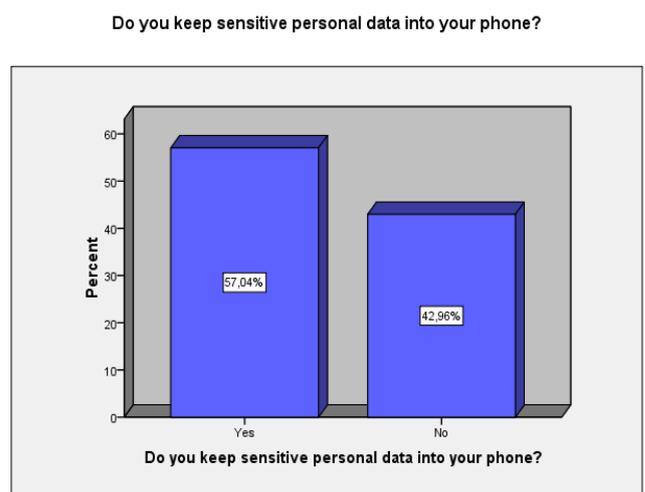


Figure 13. Sensitive information kept in phone.

In a rather alarming finding, 21.6% of users (Figure 14) keep passwords saved in plain in their mobile phone. At least, another 22% is using some form of encryption (i.e. letter scrambling). Since users generally follow the notion of encryption in these saved passwords, it is expected that they would be able to do the same with private information (i.e. photos) kept in the phone, should they be provided the necessary software. Once again, the issue of better designed user interfaces surfaces.

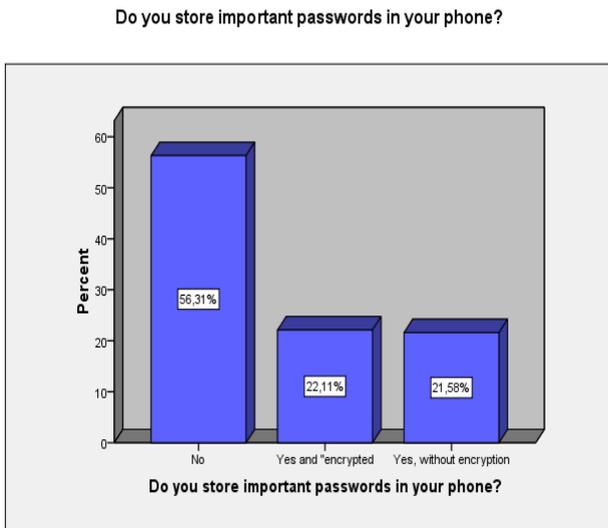


Figure 14. Important passwords kept in phone.

Closing our survey, the issue of backup was examined. As it can be seen in Figure 15, a large percentage of the participants reaching 47% never performs a backup of their phone's data. At least some 53% do backup up, although the majority (19%) less often than once per month.

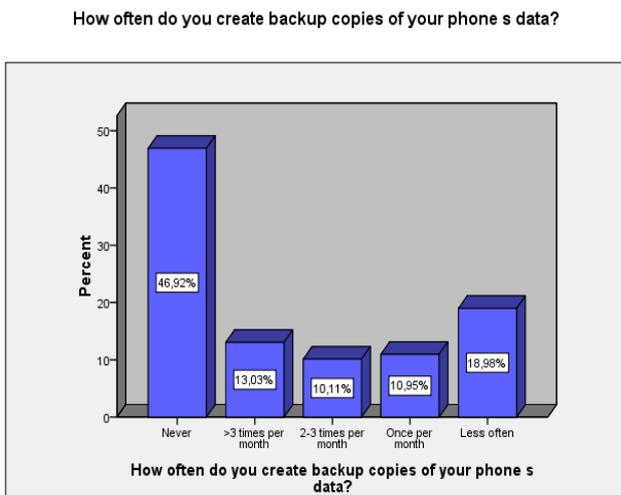


Figure 15. Backup frequency.

V. CONCLUSION AND FUTURE WORK

The majority of the respondents care about security issues and are concerned about data interception and the fact that an intruder could gain unauthorized access to their devices, as previous surveys have clearly showed. However, there is no culture of security and no advanced technical knowledge of their mobile phones.

A very high percentage of users didn't know there is an icon that informs them about the phone encryption status. Most of them don't take backups at all while at the same time would lend their phone that contains sensitive data and passwords to somebody else. Contributing to the problem, badly designed interfaces are an additional factor of hindering the development of security culture.

Students owning phones without modern operating system have statistically (Pearson Chi-Square) better knowledge of security aspects than those who actually own a phone with modern O/S. At the same time, they feel statistically the least secure while, on the other hand, users that do not know the type of O/S are more "relaxed".

In order to have comparative results, we have conducted a similar survey in more than 10 European countries reaching more than 7500 students and the results will soon be published. The preliminary findings however, show that users exhibit the same behaviour everywhere. Since students (who are young people and mostly receptive to technology and knowledge) do not actively follow most of security best practices then academia, phone manufacturers and operators must team up informing users, raising awareness level and building more secure systems and user interfaces with transparent security features.

REFERENCES

- [1] Synovate, "Global mobile phone survey shows the mobile is a 'remote control' for life", Synovate survey, <http://www.synovate.com>, 2009 [accessed: 09/10/2010]
- [2] Rahman, M. and Imai, H., Security in Wireless Communication, *Wireless Personal Communications*, vol. 22, issue, 2, pp. 218-228, 2002
- [3] I. Androulidakis, Intercepting Mobile Phones, Article in «IT security professional» magazine, Issue 8, pp. 42-28, Jan-Feb 2009
- [4] Trend Micro, "Smartphone Users Oblivious to Security", Trend Micro survey, <http://www.esecurityplanet.com>, 2009, [accessed: 09/10/2010]
- [5] Goode Intelligence, "Mobile security the next battleground", <http://www.goodeintelligence.com>, 2009, [accessed: 09/10/2010]
- [6] Androulidakis, N. and Androulidakis, I. Perspectives of Mobile Advertising in Greek Market, 2005 International Conference on Mobile Business (ICBM 2005), pp. 441-444, 2005
- [7] Vrechopoulos, A.P., Constantiou, I.D., and Sideris, I. Strategic Marketing Planning for Mobile Commerce

- Diffusion and Consumer Adoption, in Proceedings of M-Business 2002, pp. CD, July 8-9, 2002
- [8] comScore M:Metrics, “*Smarter phones bring security risks: Study*”, <http://www.comscore.com>, 2008 [accessed 09/10/2010]
- [9] I. Androulidakis and D. Papapetros, Survey Findings towards Awareness of Mobile Phones’ Security Issues, Recent Advances in Data Networks, Communications, Computers, Proceedings of 7th WSEAS International Conference on Data Networks, Communications, Computers (DNCOCO '08), pp 130-135, Nov. 2008
- [10] I. Androulidakis, C. Basios, and N. Androulidakis, Surveying Users’ Opinions and Trends towards Mobile Payment Issues, Frontiers in Artificial Intelligence and Applications - Volume 169, (Techniques and Applications for Mobile Commerce - Proceedings of TAMoCo 2008), pp. 9-19, 2008
- [11] National Cyber Security Alliance (NCSA), “*Schools Lacking Cyber Security and Safety Education*”, <http://www.staysafeonline.org>, 2009 [accessed: 09/10/2010]
- [12] McAfee, “*Most Mobile Users Don't Know if They Have Security*”, McAfee-sponsored research, <http://www.esecurityplanet.com>, 2008 [accessed: 09/10/2010]
- [13] Dillman, D. A., Mail and Internet Surveys: The Tailored Design Method, John Wiley & Sons, 2nd edition, November 1999
- [14] Pfleeger, S. L. and Kitchenham, B. A. Principles of Survey Research Part 1: Turning Lemons into Lemonade, ACM SIGSOFT Software Engineering Notes, vol. 26 (6), pp 16-18, November 2001
- [15] Androulidakis I., Androulidakis N. On a versatile and costless OMR system Wseas Transactions on Computers Issue 2, Vol 4, pp. 160-165, 2005
- [16] CPP, “*Mobile phone theft hotspots*”, CPP survey, <http://www.cpp.co.uk>, 2010 [accessed: 09/10/2010]
- [17] ITwire, “*One-third of Aussies lose mobile phones: survey*”, ITwire article, <http://www.itwire.com>, 2010 [accessed: 09/10/2010]
- [18] Darkreading, “*Survey: 54 Percent Of Organizations Plan To Add Smartphone Antivirus This Year*”, Darkreading article, <http://www.darkreading.com>. 2010 [accessed: 09/10/2010]
- 9) Have you noted somewhere your mobile phone’s IMEI? (A, I don’t know what it is, B yes, C no.)
- 10) Was your mobile phone ever lost or stolen? (A Never, B once, C more than once)
- 11) Are you aware of the existence of a special icon in your telephone which informs you for the encryption's deactivation? (A Yes, B No)
- 12) Do you have SIM card’s PIN activated? (A Yes, B No)
- 13) Do you use password in your phone's Screen-Saver? (A I don’t know if it has such a feature, B, doesn’t have such feature, C, Yes, D No)
- 14) Do you have Bluetooth: (A Switched on and visible, B Switched on and invisible, C Switched off, D don’t know the difference between visible and invisible, E My phone doesn’t have Bluetooth,
- 15) Do you lend it to others? (A Never, B Only for a while and if I am present, C Yes)
- 16) Do you "download" software to your phone? (A I don’t know if my mobile phone can download, B No, C mostly Ringtones/Logos, D mostly Games, E mostly Applications)
- 17) Do you use Antivirus software in your phone? (A Doesn’t have the ability, B Don’t know if there is such product for my phone, C I know there is but I don’t use D Yes)
- 18) Do you store important passwords in your phone (eg Credit cards passwords, ATM passwords)? (A No, B Yes and "encrypted", C yes, without encryption)
- 19) How often do you create backup copies of your phone's data? (A Never, B >3 times per month, B 2-3 times per month, C Once per month, D Less often)
- 20) Do you keep sensitive personal data into your phone (photos/videos/discussion recordings)? (A Yes, B No)
- 21) How safe do you consider communication through mobile phones? (A Very Much, B Much, C Moderately, D Not too much, E Not at all)
- 22) Are you informed about how the options and technical characteristics of your mobile phone affect its security? (A Very Much, B Much, C Moderately, D Not too much, E Not at all)

## APPENDIX

### The Questionnaire used

- 1) Male (A) or Female (B)?
- 2) Age? (A < 18, B 18-20 , C 21-23, D 24-26 , E >26)
- 3) Are you studying: (A: Humanities-Philology, B Medicine, C Law, D Engineering-Computer Science, E Maths-Natural Sciences, F Economics-Business Administration, G OTHER)
- 4) How many mobile phones do you use (daily)?  
A) 1            B) 2            C) >2    D) None
- 5) Are you a contract subscriber or a prepaid subscriber?  
A) Pre-paid (Card) B) Post-paid (Contract) C) Both
- 6) Your average monthly phone bill? (A up to 10 Euros, B 11-20 Euros, C 21-30 Euros, D 31-40 Euros, E >40 Euros)
- 7) Brand of the phone you are mostly using now? (A Nokia, B Sony-Ericsson, C Samsung, D Sharp, E Apple I-phone, F Motorola, G LG, H Other)
- 8) Does it have an advanced operational system (eg Symbian, Windows Mobile, Android)? (A I don’t know, B yes, C no.)

# Impulse Response and Generating Functions of $\text{sinc}^N$ FIR Filters

J. Le Bihan

Laboratoire RESO  
Ecole Nationale d'Ingénieurs de Brest (ENIB)  
CS 73862, 29238 BREST cedex 3, France  
E-mail: jean.le\_bihan@enib.fr

**Abstract** –  $\text{Sinc}^N$  finite impulse response (FIR) filters are built as a cascade of  $N$   $\text{sinc}$  filters, each of length  $M$ . They are used for digital signal processing applications, in various areas including telecommunications. Generating functions for the  $\text{sinc}^N$  FIR filters transfer functions are given. It is shown that z-Transform techniques can offer an efficient method to derive straightforward recurrence relations for fast computation of the impulse response. Moreover, a simple expression, valid for all the filter coefficients, is also obtained. It is new and general, compared to previously published formula.

**Keywords** – digital filter; impulse response; coefficient computation; z-Transform.

## I. INTRODUCTION

Different structures for digital filters have been proposed for applications as decimation, interpolation or noise suppression in Sigma-Delta A/D conversion [1, 2]. As suitable filters for high speed operation, cascaded-integrator-comb (CIC) filters [1] and  $\text{sinc}^N$  impulse finite response (FIR) filters [3, 4] are among the well known solutions. The use of  $\text{sinc}^N$  filters constituted of a cascade of  $N$   $\text{sinc}$  filters, each of length  $M$ , requires the calculation of all of the impulse response coefficients. So adequate recursions or expressions for rapid calculation of these coefficients have been the subject of numerous investigations, as in [5-7]. Interesting computation aids have already been published: a closed form expression for the first  $M$  coefficients has been given in [6], while a recurrence formula has been presented in [7].

It has been proved that using z-Transform techniques can be an efficient method to derive digital filter coefficients [8, 9]. Here, employing these techniques, new results are obtained. Simple effective recurrence relations are derived for the computation of the coefficients of  $\text{sinc}^N$  FIR filters. In addition, a simple expression is deduced, which is valid not only for the first coefficients of the impulse response but for all of them. In addition, generating functions for the transfer functions of  $\text{sinc}^N$  FIR filters are presented.

In this paper, the general form of the transfer function of the  $\text{sinc}^N$  FIR filters is given in section II, as well as its useful basic properties. In section III, it is shown how recurrence relations for the filter coefficients can be derived using z-Transform techniques applied to the expression of

the transfer function considered as a z-transform. Similarly, the way to obtain an expression for the filter coefficients, new and general compared to formula previously published by other authors, is described in section IV. Finally, in section V, forms of generating functions for the transfer functions of  $\text{sinc}^N$  FIR filters are expressed.

## II. TRANSFER FUNCTION OF $\text{sinc}^N$ FILTERS

The transfer function  $H^{M,N}(z)$  of  $\text{sinc}^N$  filters can be written as

$$H^{M,N}(z) = \left( \frac{1 - z^{-M}}{M(1 - z^{-1})} \right)^N \quad (1)$$

or

$$H^{M,N}(z) = M^{-N} (1 + z^{-1} + z^{-2} + \dots + z^{-(M-1)})^N \quad (2)$$

to which corresponds the following magnitude response

$$\left| H^{M,N}(e^{j\omega}) \right| = \left| \frac{\sin(M\omega/2)}{M \sin(\omega/2)} \right|^N \quad (\text{Fig. 1}).$$

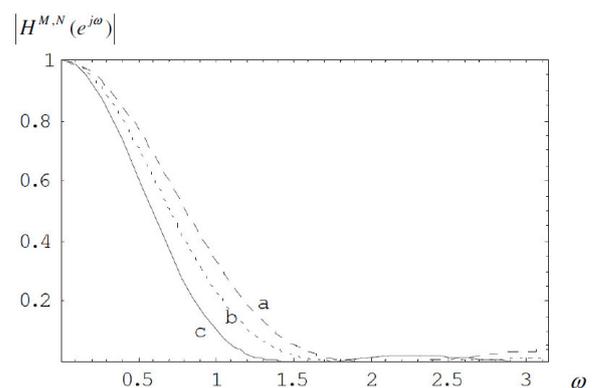


Fig. 1. Magnitude responses of  $H^{3,3}(e^{j\omega})$  (a),  $H^{3,4}(e^{j\omega})$  (b) and  $H^{4,3}(e^{j\omega})$  (c), versus normalized angular frequency  $\omega$

For simplicity, we will consider the scaled transfer function  $G^{M,N}(z) = M^N H^{M,N}(z)$

$$G^{M,N}(z) = \left( \frac{1 - z^{-M}}{1 - z^{-1}} \right)^N \quad (3)$$

$G^{M,N}(z)$  is a polynomial, with degree  $N(M-1)$ , of the form

$$G^{M,N}(z) = \sum_{j=0}^{N(M-1)} g_j^{M,N} z^{-j} \quad (4)$$

where  $g_j^{M,N}$  represent the coefficients of the impulse response of the  $\text{sinc}^N$  filters.  $h_j^{M,N}$  and  $g_j^{M,N}$  are related by  $h_j^{M,N} = M^{-N} g_j^{M,N}$ . Note that  $g_j^{M,N} = 0$  ( $\forall j < 0$ ),  $g_0^{M,N} = 1$  ( $\forall M, \forall N$ ),  $G^{M,0}(z) = 1$  ( $\forall M$ ),  $G^{1,N}(z) = 1$  ( $\forall N$ ) and  $g_{N(M-1)-j}^{M,N} = g_j^{M,N}$ .

In the following, we will propose simple recurrence relations useful for the computation of  $\text{sinc}^N$  FIR filter coefficients, and especially a new general expression of these coefficients.

### III. RECURRENCE RELATIONS FOR THE FILTER COEFFICIENTS

In this section, several relations are derived, according to different criteria such as simplicity, low order, varying index. Each of these independent relations can be useful, depending on the way chosen for computation, and implemented separately.

A. Calculation of  $g_j^{M,N}$  from 3 coefficients with the same values  $M$  and  $N$ , and lower indexes  $j-1$ ,  $j-M$  and  $j-(M+1)$

Differentiating  $\ln G^{M,N}(z)$  using (3) and multiplying by  $G^{M,N}(z)$  yields

$$\frac{dG^{M,N}(z)}{dz} = N \left( \frac{Mz^{-(M+1)}}{1-z^{-M}} - \frac{z^{-2}}{1-z^{-1}} \right) G^{M,N}(z) \quad (5)$$

which can be written

$$(1-z^{-1}-z^{-M}+z^{-(M+1)}) \left( -z \frac{dG^{M,N}(z)}{dz} \right) = N \left( z^{-1} - Mz^{-M} + (M-1)z^{-(M+1)} \right) G^{M,N}(z) \quad (6)$$

By using basic z-Transform techniques, (6) leads immediately to

$$g_j^{M,N} = \frac{1}{j} (N-1+j) g_{j-1}^{M,N} - [M(N+1)-j] g_{j-M}^{M,N} + [M(N+1)-(N-1)-j] g_{j-(M+1)}^{M,N} \quad (7)$$

which can also be written

$$g_j^{M,N} = g_{j-1}^{M,N} + g_{j-M}^{M,N} - g_{j-(M+1)}^{M,N} + \frac{1}{j} \left[ (N-1) (g_{j-1}^{M,N} - g_{j-(M+1)}^{M,N}) - M(N+1) (g_{j-M}^{M,N} - g_{j-(M+1)}^{M,N}) \right] \quad (8)$$

An advantage of (8) is the fact that  $M$  and  $N$  keep constant values. Relative drawbacks are that the number of terms is not the lowest possible and that shifts in  $j$  indexes increase with  $M$  and may be large. This recurrence relation is the same as (11) in [7], but here it has been straightforwardly derived using z-Transform techniques.

For example, for  $M=8$ ,  $N=4$  and  $j=9$ , the relation gives:

$$g_9^{8,4} = g_8^{8,4} + g_1^{8,4} - g_0^{8,4} + \frac{1}{9} [3(g_8^{8,4} - g_0^{8,4}) - 40(g_1^{8,4} - g_0^{8,4})]$$

$$\text{i.e., } g_9^{8,4} = 161 + 4 - 1 + \frac{1}{9} [3(161-1) - 40(4-1)] = 204$$

B. Calculation of  $g_j^{M,N}$  from 2 coefficients with the same value  $M$ , and indexes  $(j-1, N)$  and  $(j, N-1)$

Differentiating  $G^{M,N}(z)$  using (3) yields

$$\frac{dG^{M,N}(z)}{dz} = N \left[ \frac{Mz^{-(M+1)}}{1-z^{-1}} - \frac{z^{-2}(1-z^{-M})}{(1-z^{-1})^2} \right] \left( \frac{1-z^{-M}}{1-z^{-1}} \right)^{N-1} \quad (9)$$

which can be written

$$(1-z^{-1}) \left( -z \frac{dG^{M,N}(z)}{dz} \right) = N \left\{ -MG^{M,N-1}(z) + [M - (M-1)z^{-1}] G^{M,N}(z) \right\} \quad (10)$$

By using basic z-Transform techniques, it can be immediately deduced

$$g_j^{M,N} = \frac{1}{MN-j} \left[ (MN-j+1-N) g_{j-1}^{M,N} + MN g_j^{M,N-1} \right] \quad (11)$$

Advantages of (11) are a very low number of terms, the fact that  $M$  keeps a constant value and that shifts in  $N$  and  $j$  indexes equal only one and occur separately. This new and simple relation allows a fast recursive calculation of the impulse response coefficients. Therefrom, computation is quite easy, even permitting to fill up the start of the table of filter coefficients simply by hand.

For example, for  $M=8$ ,  $N=4$  and  $j=9$ , the relation gives:

$$g_9^{8,4} = \frac{1}{23} (20g_8^{8,4} + 32g_9^{8,3})$$

$$\text{i.e., } g_9^{8,4} = \frac{1}{23} (20 \times 161 + 32 \times 46) = 204$$

C. Calculation of  $g_j^{M,N}$  from 3 coefficients with the same value  $M$ , and indexes  $(j-1, N)$  and  $(j, N-1)$  and  $(j-M, N-1)$

Using (3),  $G^{M,N}(z)$  can be linked to  $G^{M,N-1}(z)$  as follows

$$G^{M,N}(z) = \frac{1-z^{-M}}{1-z^{-1}} G^{M,N-1}(z) \quad (12)$$

By using basic z-Transform techniques, it can be immediately deduced from (12)

$$g_j^{M,N} = g_{j-1}^{M,N} + g_j^{M,N-1} - g_{j-M}^{M,N-1} \quad (13)$$

Advantages of (13) are a low number of terms, with no multiplying factors, the fact that  $M$  keeps a constant value and that shifts in  $N$  equal only one. Relative drawbacks are that the number of terms is not the lowest possible and that shifts in  $j$  index increase with  $M$  and may be large. This simple relation allows a fast recursive calculation of the impulse response coefficients. Computation is quite easy, even permitting to fill up the start of the table of filter coefficients simply by hand.

For example, for  $M=8$ ,  $N=4$  and  $j=9$ , the relation gives:

$$g_9^{8,4} = g_8^{8,4} + g_9^{8,3} - g_1^{8,3}$$

$$\text{i.e., } g_9^{8,4} = 161 + 46 - 3 = 204$$

D. Calculation of  $g_j^{M,N}$  from  $N+1$  coefficients with the same lower value  $M-1$ , and indexes  $(j-k, k)$ ,  $k=0..N$

Using (3),  $G^{M,N}(z)$  can be expressed as follows

$$G^{M,N}(z) = \left[ 1 + z^{-1} \frac{1 - z^{-(M-1)}}{1 - z^{-1}} \right]^N \tag{14}$$

which can be written

$$G^{M,N}(z) = \sum_{k=0}^N \binom{N}{k} z^{-k} G^{M-1,k}(z) \tag{15}$$

By using basic z-Transform techniques, (15) yields

$$g_j^{M,N} = \sum_{k=0}^N \binom{N}{k} g_{j-k}^{M-1,k} \tag{16}$$

An advantage of (16) is the fact that  $M-1$  keeps constant in the right side of the equation. Drawbacks are that the number of terms, the number of sets of coefficients

involved in the relation, as well as shifts in  $j$  index increase with  $N$ .

For example, for  $M=8, N=4$  and  $j=9$ ,  $g_j^{M,N}$  can be easily computed using the values of the coefficients

$g_{9-k}^{7,k}$ , weighted by the binomial coefficients  $\binom{4}{k}$ :

$$g_9^{8,4} = \sum_{k=0}^4 \binom{4}{k} g_{9-k}^{7,k} = g_9^{7,0} + 4g_8^{7,1} + 6g_7^{7,2} + 4g_6^{7,3} + g_5^{7,4}$$

i.e.,  $g_{10}^{8,4} = 0 + 4 \times 0 + 6 \times 6 + 4 \times 28 + 56 = 204$

TABLE I  
Values of  $g_j^{M,N}$  for  $N = 1, \dots, 4$  and  $M = 1, \dots, 8$

	$M$	$N(M-1)$	$M^N$	$g_0^{M,N}$	$g_1^{M,N}$	$g_2^{M,N}$	$g_3^{M,N}$	$g_4^{M,N}$	$g_5^{M,N}$	$g_6^{M,N}$	$g_7^{M,N}$	$g_8^{M,N}$	$g_9^{M,N}$	$g_{10}^{M,N}$	$g_{11}^{M,N}$	$g_{12}^{M,N}$	$g_{13}^{M,N}$	$g_{14}^{M,N}$
N=1	1	0	1	1														
	2	1	2	1														
	3	2	3	1	1													
	4	3	4	1	1													
	5	4	5	1	1	1												
	6	5	6	1	1	1												
	7	6	7	1	1	1	1											
	8	7	8	1	1	1	1											
N=2	1	0	1	1														
	2	2	4	1	2													
	3	4	9	1	2	3												
	4	6	16	1	2	3	4											
	5	8	25	1	2	3	4	5										
	6	10	36	1	2	3	4	5	6									
	7	12	49	1	2	3	4	5	6	7								
	8	14	64	1	2	3	4	5	6	7	8							
N=3	1	0	1	1														
	2	3	8	1	3													
	3	6	27	1	3	6	7											
	4	9	64	1	3	6	10	12										
	5	12	125	1	3	6	10	15	18	19								
	6	15	216	1	3	6	10	15	21	25	27							
	7	18	343	1	3	6	10	15	21	28	33	36	37					
	8	21	512	1	3	6	10	15	21	28	36	42	46	48				
N=4	1	0	1	1														
	2	4	16	1	4	6												
	3	8	81	1	4	10	16	19										
	4	12	256	1	4	10	20	31	40	44								
	5	16	625	1	4	10	20	35	52	68	80	85						
	6	20	1296	1	4	10	20	35	56	80	104	125	140	146				
	7	24	2401	1	4	10	20	35	56	84	116	149	180	206	224	231		
	8	28	4096	1	4	10	20	35	56	84	120	161	204	246	284	315	336	344

In this section, using z-Transform techniques, efficient recursive formulae for computing the impulse response coefficients  $g_j^{M,N}$  of  $\text{sinc}^N$  filters have been derived, especially (11) and (13). The values of the coefficients  $g_j^{M,N}$ , for  $M \leq 8$  and  $N \leq 4$ , are given in Table I. For symmetry reasons, only the coefficients for  $j=0..INT\left[\frac{N(M-1)}{2}\right]$  are shown.

Further investigations should be undertaken in order to evaluate precisely the number of operations and the computation time when using such or such recurrence relation or a combination of them, and, in a way, the more efficient strategy.

#### IV. GENERAL EXPRESSION FOR THE FILTER COEFFICIENTS

Let us write  $G^{M,N}(z)$  under the form

$$G^{M,N}(z) = (1-z^{-M})^N Y^N(z) \tag{17}$$

with  $Y^N(z)$  defined as follows

$$Y^N(z) = \left( \frac{1}{1-z^{-1}} \right)^N \tag{18}$$

##### A. Explicit expression for the coefficients $y_j^N$ of $Y^N(z)$

Differentiating  $\ln Y^N(z)$  and multiplying by  $(1-z^{-1})[-zY^N(z)]$  gives

$$(1-z^{-1}) \left( -z \frac{dY^N(z)}{dz} \right) = Nz^{-1}Y^N(z) \tag{19}$$

Basic z-Transform techniques allow to obtain

$$y_j^N = \frac{N-1+j}{j} y_{j-1}^N \tag{20}$$

This relation permits easy recursive computation of  $y_j^N$  from  $y_{j-1}^N$ . Moreover, (20) leads to the following simple explicit relation for the coefficients  $y_j^N$

$$y_j^N = \binom{N-1+j}{j} \tag{21}$$

##### B. General expression for the coefficients $g_j^{M,N}$ of $G^{M,N}(z)$

Developing  $(1-z^{-M})^N$  in (17) leads to the following relation between  $G^{M,N}(z)$  and  $Y^N(z)$

$$G^{M,N}(z) = \sum_{k=0}^N (-1)^k \binom{N}{k} z^{-kM} Y^N(z) \tag{22}$$

Then using the z-Transform translation property, the coefficients  $g_j^{M,N}$  can be written

$$g_j^{M,N} = \sum_{k=0}^{[j/M]} (-1)^k \binom{N}{k} \binom{N-1+j-kM}{j-kM} \tag{23}$$

with  $j=0..N(M-1)$ , and where  $[j/M]$  denotes the integer quotient of  $j$  and  $M$ .

It is worth noting that this new expression of the coefficients  $g_j^{M,N}$  applies for all values of  $M, N$  and  $j$ .

This simple formula allows to calculate any coefficient numerically. It offers the obvious advantage, compared to previous results, to be general and valid for all the coefficients of the impulse response of the  $\text{sinc}^N$  filters. In fact, in [6] (cf. (19)) and in [7] (cf. (12) and (13)), explicit expressions are given only for the very first coefficients of the impulse response of the  $\text{sinc}^N$  filters. Note that here (23) leads immediately to these explicit expressions for  $j=0..(M-1)$  and for  $j=M$ , with  $k=0$  and  $k=1$  respectively.

For example, for  $M=8, N=4$  and  $j=9$  ( $[j/M]=1$ ),  $g_j^{M,N}$  can be easily computed using the products of

binomial coefficients  $\binom{N}{k} \binom{N-1+j-kM}{j-kM}$ , i.e.,

$$\binom{4}{k} \binom{12-8k}{9-8k}, \text{ multiplied by } (-1)^k :$$

$$g_9^{8,4} = \sum_{k=0}^1 (-1)^k \binom{4}{k} \binom{12-8k}{9-8k} = \binom{12}{9} - 4 \binom{4}{1},$$

$$\text{i.e., } g_9^{8,4} = 220 - 4 \times 4 = 204$$

In this section, a simple general expression (23) has been obtained for rapid computation of any of the coefficients of  $\text{sinc}^N$  filters.

As a summary, useful recurrence relations and formula (8), (11), (13), (16) and (23) have been derived for computing the  $\text{sinc}^N$  FIR filter coefficients  $h_j^{M,N} = M^{-N} g_j^{M,N}$ .

#### V. GENERATING FUNCTIONS FOR THE TRANSFER FUNCTIONS OF $\text{sinc}^N$ FIR FILTERS

Let us consider (12)  $G^{M,N}(z) = \frac{1-z^{-M}}{1-z^{-1}} G^{M,N-1}(z)$ , with  $G^{M,0}(z) = 1$ .

##### A. Ordinary generating function

Multiplying  $G^{M,N}(z)$  by  $x^N$  and summing with  $N$  varying from 0 to infinity leads to the following ordinary generating function

$$\Gamma_o^M(z, x) = \sum_{N=0}^{\infty} G^{M,N}(z) x^N = \frac{1}{1 - \frac{1-z^{-M}}{1-z^{-1}} x} \tag{24}$$

Conversely, developing  $\Gamma_o^M(z, x)$  into series expansion generates the expression of the filter transfer functions  $G^{M,N}(z)$  as the coefficients of  $x^N$  ( $N = 0, \dots, \infty$ ).

##### B. Exponential generating function

Multiplying  $G^{M,N}(z)$  by  $\frac{x^N}{N!}$  and summing with  $N$  varying from 0 to infinity leads to the following exponential generating function

$$\Gamma_e^M(z, x) = \sum_{N=0}^{\infty} G^{M,N}(z) \frac{x^N}{N!} = e^{\frac{1-z^{-M}}{1-z^{-1}} x} \tag{25}$$

Conversely, developing  $\Gamma_e^M(z, x)$  into series expansion generates the expression of the filter transfer functions  $G^{M,N}(z)$  as the coefficients of  $\frac{x^N}{N!}$  ( $N = 0, \dots, \infty$ ).

## VI. CONCLUSION

$\text{sinc}^N$  filters, constituted of a cascade of  $N$  *sinc* filters, each of length  $M$ , are useful for digital signal processing applications, in various domains including telecommunications. In this paper, z-Transform has been used as an efficient tool for deriving various formulae allowing to compute the impulse response coefficients of  $\text{sinc}^N$  FIR filters.

In particular, straightforward recursive relations have been demonstrated. Further investigations should be made to evaluate the computation time and the more efficient strategy when exploiting these relations separately or in combination.

Moreover, a simple general expression – new compared to previous published formula – has been given, valid for all the coefficients, whatever their rank  $j$  and the values of  $M$  and  $N$ .

In addition, generating functions for the  $\text{sinc}^N$  FIR filters transfer functions have been given.

## REFERENCES

- [1] E. B. Hogenauer, "An economical class of digital filters for decimation and interpolation", *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-29, 2, pp. 155-162, Apr. 1981
- [2] J. C. Candy, "A use of double integration in sigma delta modulation", *IEEE Trans. Commun.*, vol. COM-3, 3, pp. 249-258, March 1985
- [3] H. Meleis and P. Le Fur, "A novel architecture design for VLSI implementation of a FIR decimation filter", *Proc. ICASSP'85*, pp. 1380-1383
- [4] V. Friedman, D. M. Brinthampt, D.-P. Chen, and T. Deppa, J. P. Edward Jr., E. M. Fields, and H. Meleis, "A bit slice architecture for sigma delta analog-to-digital converters", *IEEE J. Select. Areas Commun.*, vol. 6, pp. 520-526, Apr. 1988
- [5] M. Shiraishi, "A simultaneous coefficient calculation method for  $\text{sinc}^N$  FIR filters", *IEEE Trans. Circuits Syst. I*, vol. 50, 4, pp. 523-529, Apr. 2003
- [6] S. C. Dutta Roy, "Impulse response of  $\text{sinc}^N$  FIR filters", *IEEE Trans. Circuits Syst. II*, 2006, 53, (3), pp. 217-219
- [7] E. E. Hassan, "Recurrence formula for impulse response coefficients of  $\text{sinc}^N$  FIR filter", *Electron. Letters*, 2006, 42, (15), pp. 850-851
- [8] J. Le Bihan, "Maximally linear FIR digital differentiators", *Circuits, Syst., Signal Process.*, vol. 14, 5, 1995, pp. 633-637
- [9] J. Le Bihan, "Coefficients of FIR digital differentiators and Hilbert transformers for midband frequencies", *IEEE Trans. Circuits Syst. II*, vol. 43, 3, March 1996, pp. 272-274

# Multi-Level Log-Likelihood Ratio Clipping in a Soft-Decision Near-Maximum Likelihood Detector

Sébastien Aubert<sup>\*†</sup>, Andrea Ancora<sup>\*</sup>

\* ST-ERICSSON;

635, route des Lucioles;

CP 06560 Sophia-Antipolis, France

{sebastien.aubert, andrea.ancora}@stericsson.com

Fabienne Nouvel<sup>†</sup>

†IETR;

20, avenue des Buttes de Coësmes;

CP 35043 Rennes, France

fabienne.nouvel@insa-rennes.fr

**Abstract**—Consider the MIMO detection background. While the hard-decision Sphere Decoder has been widely and recently considered as the most promising near-optimal detector, this perspective might fall down during the soft-decision extension through a List Sphere Decoder (LSD). Due to the finiteness of the LSD list output - that does not necessarily allow for generating explicitly the Log-Likelihood Ratios (LLRs), even through a max-log approximation - the issue of how to set the missing reliabilities has been addressed. This paper presents existing works concerning the main trend. In particular, it consists in setting the LLR to a pre-defined value, this operation being commonly referred as LLR Clipping. We discuss this choice that has a significant impact on the system performance, by providing a brief state of the art of the existing solutions. In addition in the presented work, a novel solution lies in the multi-level bit mapping. Despite of its simplicity, it allows for low distortion approximated LLR computation. By simulation, the superiority of our method over the existing solutions, is shown.

**Index Terms**—Log-Likelihood Ratio; bit clipping; List Sphere Decoder;

## I. INTRODUCTION

In order to achieve the 3GPP Long-Term Evolution (LTE) and 3GPP LTE-Advanced (LTE-A) requirements, Spatial Multiplexing Multiple-Input Multiple-Output (SM-MIMO) communication schemes have been implemented. In such a configuration and from the receiver point of view, a linear superposition of separately transmitted information symbols is observed, due to multiple transmit antennas that simultaneously send independent data streams. The interest of the detectors consists in recovering the transmitted symbols while approaching the channel capacity [1], and corresponds to an inverse problem with a finite-alphabet constraint.

As treated in several publications, the optimal - while exponentially complex in the number of transmit antennas and constellation size- Maximum Likelihood (ML) detector [2] can be efficiently approximated although avoiding an exhaustive search. In particular, some well-established techniques such as Sphere Decoding (SD) [3], Lattice Reduction (LR) [4] or a combination of both [5], have been shown to offer near-optimal performance. In the practical case of coded systems and due to their performance-complexity flexibility, the aforementioned detectors are straightforwardly modified in order to provide Soft-Output (SO). In the particular case of the classical  $K$ -Best, strongly presented in [6], a list  $\mathcal{L}$  - of size  $|\mathcal{L}|$  - of

candidate solutions is generated from a subset of lattice points. This detector is denoted as the List SD (LSD) and will be considered in the present article, unless otherwise specified.

In order to achieve the channel capacity, the acknowledged way of bit transmission lies in providing redundancy and interleaving, denoted as channel coding. By focusing on the performance-complexity optimization, modern capacity-approaching codes lie in probabilistic coding schemes [7]. In particular, convolutional codes [8] led to the widely employed turbo codes [1], that are considered in this paper. In such schemes, soft-decisions of coded symbols are typically produced from the detector output, plus any available side information, and passed on to the decoder in the form of bit-wise Log-Likelihood Ratios (LLR), with the sign representing the decision and the magnitude representing the reliability. In the present article, the logical  $\{0,1\}$  bit patterns are respectively mapped onto the - zero-mean - set of amplitude levels  $\{+1, -1\}$ . From the decoder point of view, a close-to-zero value corresponds to an unreliable bit.

By evoking first that the LLR approximation is expressed as a function of both the probability of any bit representation [9], given the data in reception, and since the LSD output contains at least one solution, it is clear that a correct LLR sign is robustly reached, leading to promising uncoded performance [3], [4], [5]. However, due to the finite nature of the output  $\mathcal{L}$  of the LSD that only offers a reduced-size list of candidates, its magnitude is regularly unknown, namely when the list does not contain both the hypothesis and its counter-hypothesis for a given bit. As a consequence, the performance in coded communication systems may be dramatically impacted.

While the exact LLR calculation is processed when possible, the main issue in the soft decision extension consists in how to estimate the missing LLR magnitude. To the best of the authors' knowledge, two distinct trends have been explored in the digital communications literature. Båro has early met this problematic aspect and proposed a path augmentation [10] that consists in considering a bit-wise granularity during the tree construction. However, such a scheme requires high computational complexity and poor performance is reached if no apriori information is available [11].

A simple yet efficient operation has been widely studied.

It is commonly referred as LLR Clipping (LC) [9], [12] and lies on setting the missing LLR values to a predefined value. However, the choice of the clipping level has a strong impact on the system performance [12], as addressed in the following, and any optimization is valuable.

*Contribution:* Our contributions can be summarized as follows:

- A multi-level bit mapping in the LC is presented, which leads to a significant coded performance improvement and offers convenient result improvement;
- The introduced solution preserves a general approach that makes it applicable to any LC.

*Outline of the paper:* In Section II, the notations are given and the problem statement is presented. In Section III, the necessity of introducing an accurate LC is explained and a detailed description of existing solutions is presented. The proposed solution is defended and introduced in Section IV. Section V aims at providing simulation results that show the superiority of the proposed solution. Finally, concluding remarks and perspectives are given in Section VI.

## II. SYSTEM MODEL AND PROBLEM STATEMENT

Consider a  $n_T$ -transmit and  $n_R$ -receive  $n_T \times n_R$  MIMO system model. Assuming narrow-band flat-fading, the receive symbols vector  $\mathbf{y} \in \mathbb{C}^{n_R}$  typically reads

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

where  $\mathbf{H} \in \mathbb{C}^{n_R \times n_T}$  is the complex channel matrix, assumed to be perfectly known at the receiver, and  $\mathbf{n}$  is a complex additive white Gaussian noise of variance  $\sigma_n^2$ . The entries of the transmit symbol vector  $\mathbf{x}$  are independently withdrawn from a constellation set  $\xi$ , containing  $|\xi|$  symbols, and such that  $\mathbf{x} \in \xi^{n_T}$ . Also, a layer  $\nu$  is defined as a spatial stream, the number of spatial multiplexing data streams being  $\min\{n_R, n_T\}$ . Each channel inputs  $\mathbf{b}_k^n \in \{\pm 1\}^{\nu \log_2\{|\xi|\}}$  is assigned to a symbol according to any encoding scheme, where  $n$  is any given layer and  $k$  is any bit position within the corresponding symbol. The whole symbols vector is mapped from a block of bit stream  $\mathbf{c}_l$ , by denoting  $l$  as the number of bits of the codeword with  $1 \leq l \leq \nu \log_2\{|\xi|\}$ . At this step, the uncoded or coded case is not distinguished. The block code may consider channel coding through the addition of redundancy and correlation, by introducing the code rate  $R \leq 1$  - where  $R = 1$  makes  $\mathbf{c}$  corresponding to uncoded bits -, and interleaving [13]. By applying an efficient modulation and code rate scheme, the channel capacity is almost achieved at any SNR point [1].

The extrinsic LLR of the bit  $\mathbf{b}_k^n$  is conditioned on the receive signal and is denoted  $\Lambda(\mathbf{b}_k^n | \mathbf{y})$ . Through the Bayes' theory and by using the *high SNR approximation*, which makes the classical max-log approximation accurate, it becomes advantageously rewrites [9]:

$$\Lambda(\mathbf{b}_k^n | \mathbf{y}) \approx -\frac{1}{2\sigma_n^2} \left( (d_{\mathbf{b}_k^n=+1}^2)_{\min} - (d_{\mathbf{b}_k^n=-1}^2)_{\min} \right), \quad (2)$$

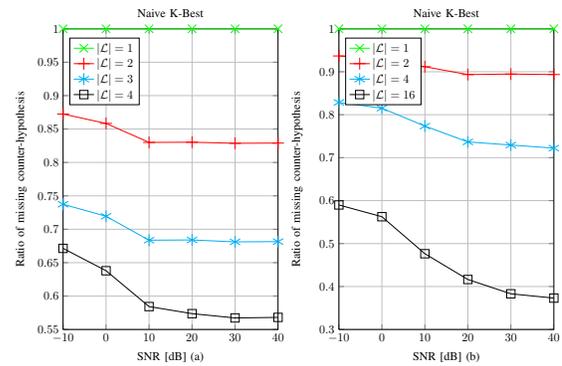


Fig. 1. Expected ratio of to-be-clipped values by considering the output of a naive K-Best, 4x4 complex MIMO system, QPSK (a) or 16QAM (b) modulation on each layer with  $|\mathcal{L}| = \{1, 2, 3, 4\}$  and  $|\mathcal{L}| = \{1, 2, 4, 16\}$ , respectively, 10,000 simulations.

where  $(d_{\mathbf{b}_k^n=+1}^2)_{\min}$  and  $(d_{\mathbf{b}_k^n=-1}^2)_{\min}$  denote the minimal square euclidean distance between  $\mathbf{y}$  and all the possibly transmit symbols vectors which are assigned to  $\mathbf{b}_k^n = +1$  and to  $\mathbf{b}_k^n = -1$ , respectively. Also, the layer index  $n$  within the transmit data and the bit index  $k$  within any symbol inside are such that  $1 \leq n \leq \nu$  and  $1 \leq k \leq \log_2\{|\xi|\}$ , respectively.

The importance of LC on both performance and complexity is highlighted. With this aim in view, the ratio of the occurrence of a missing counter-hypothesis as a function of  $|\mathcal{L}|$  is depicted in Figure 1, for a  $4 \times 4$  system, with QPSK or 16QAM modulations on each antenna.

It appears that the number of to-be-clipped LLR values depends on the SNR, on the list size  $|\mathcal{L}|$  and on the modulation scheme. In particular in the 16QAM case and for a high range SNR, the ratio is 73% and 38% for  $|\mathcal{L}|$  being 4 and 16, respectively.

Beyond this consideration, it is also clear that also the clipping value depends on the list size  $|\mathcal{L}|$ . As seen in Figure 1 and in the case of large  $|\mathcal{L}|$ , few missing counter-hypothesis remain. However, since a large neighbourhood study of lattice points has been processed, the corresponding bit sign is very improbable and should be clipped to a large value. In the case of a small  $|\mathcal{L}|$ , more missing counter-hypothesis are contained within the list output. Due to the partial nature of the study, the bit sign is not completely improbable, leading to choosing a small clipping value.

This idea has to be introduced in theory. It can be done through the introduction of the *mutual information*  $\mathcal{I}\{\mathbf{b}_k^n, \Lambda_{clip}(\mathbf{b}_k^n)\}$  of any bit amplitude and any corresponding LC value, denoted as  $\Lambda_{clip}(\mathbf{b}_k^n)$ . By assuming that the bit sign in the LC is correct and  $\Pr\{\mathbf{b}_k^n = +1\} = \Pr\{\mathbf{b}_k^n = -1\} = \frac{1}{2}$ , the mutual information reads:

$$\mathcal{I}(\mathbf{b}_k^n, \Lambda_{clip}(\mathbf{b}_k^n)) = 1 - \frac{1}{2} \left( \log_2 \left\{ 1 + e^{-\Lambda_{clip}(\mathbf{b}_k^n)} \right\} + \log_2 \left\{ 1 + e^{\Lambda_{clip}(\mathbf{b}_k^n)} \right\} \right). \quad (3)$$

Subsequently, the impact of the clipping value choice can be observed. In particular, a solution for determining the

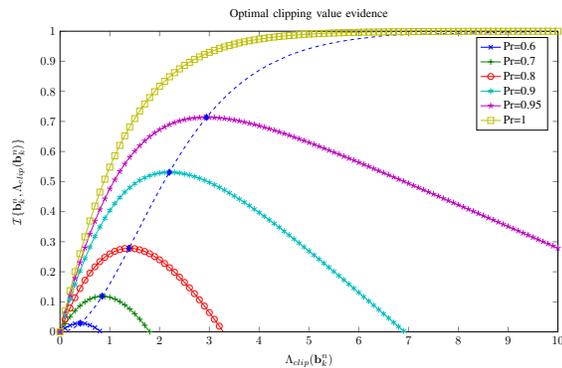


Fig. 2. Mutual information versus clipping value  $\Lambda_{clip}(\mathbf{b}_k^n)$  for multiple  $\Pr\{\Lambda_{clip}(\mathbf{b}_k^n)\}$  values.

LC setting relies on the mutual information maximization. Figure 2 plots the mutual information between  $\mathbf{b}_k^n$  and the clipped detector output  $\Lambda_{clip}(\mathbf{b}_k^n)$ , versus the clipping value  $\Lambda_{clip}(\mathbf{b}_k^n)$  and for different values of  $\Pr\{\Lambda_{clip}(\mathbf{b}_k^n)\}$ . Also, the maximal mutual information - given any  $\Pr\{\Lambda_{clip}(\mathbf{b}_k^n)\}$  value - is pointed out by diamond marks and indicates the optimal clipping value.

Consequently, it is clear that there exists an optimal value that is not a constant. At least, it depends on the number of missing counter-hypothesis. Its evolution is plotted in Figure 2 with a dashed style. Also, while an efficient balance - namely  $\Lambda_{clip}(\mathbf{b}_k^n) = 3$  [12] - can be found at intermediate clipping values, it has to be noticed that the LC choice has a significant impact on the coded performance. In particular, choosing the clipping level too high induces the decoder to assume a too high reliability for the bits with missing counter-hypothesis and consequently prevents error correction at these bit positions. Setting the clipping level too low limits the mutual information at the detector output and thus decreases its performance.

In order to address this problematic aspect, numerous techniques have been proposed and studied.

### III. SUMMARY OF EXISTING SOLUTIONS

In case both the hypothesis and its counter-hypothesis given any bit arise in  $\mathcal{L}$ , the LLR is calculated according to Equation (2). Naively, we can consider that the soft information about any given bit  $\mathbf{b}_k^n$  is essentially contained in  $\mathcal{L}$ . Indeed, if there are many entries in  $\mathcal{L}$  with  $\mathbf{b}_k^n = -1$ , then it can be concluded that the likely value for  $\mathbf{b}_k^n$  is indeed minus one, whereas if few entries occur, then the likely value is one. In the particular configuration of no arising counter-hypothesis within  $\mathcal{L}$ , Hochwald *et al.* proposed first a solution. It consists in setting its corresponding LLR value  $\Lambda_{clip}(\mathbf{b}_k^n)$  to an extremely large value [9], namely  $\pm 128$ , due to their high improbability. However, by proceeding this way, the importance given to unknown LLRs is too high and the channel coding gain is impaired.

#### A. Fixed LLR Clipping

Ideally,  $\Lambda_{clip}(\mathbf{b}_k^n)$  should be different for each channel bit, such that the mutual information is maximized [12]. Nevertheless, a simple while robust proposal has been introduced [9], [5]. For a missing  $\pm 1$  bit value on any of the  $\nu \log_2(|\mathcal{L}|)$  locations, the corresponding LLR is set to  $\pm 8$ , based on the argument that a missing bit value from the list of candidates makes it unlikely. As previously discussed, such a Fixed LC (FLC) value is efficient in the case of large  $|\mathcal{L}|$  only. It can anyway be considered as a convenient upper bound [11], [14].

Another reasonably balanced FLC is commonly employed. The FLC value can be set to  $\pm 3$  and offers convenient performance [12]. Again, it has been previously discussed that such a FLC value is efficient in the case of small  $|\mathcal{L}|$  only. In particular in [11], performance is depicted in the case of a naive LSD and  $\Lambda_{clip} \mathbf{b}_k^n = 3$ . Improved performance is offered compared to  $\Lambda_{clip} \mathbf{b}_k^n = 8$ , except with a very large study, namely with  $|\mathcal{L}| = 64$ . Consequently, this simple although efficient technique will be used as a reference in the simulation results.

By referring again to Figure 2, it must be highlighted that a constant clipping value that is much lower than the optimum value for a given bit position causes the channel decoder to largely ignore the clipped detector output values, which degrades its error-correction effectiveness. A clipping value that is much higher than its optimum value forces the channel decoder to assume that the clipped detector output values have the correct sign. In the case that this assumption is not correct, soft decisions on other bits must be compromised in order to meet the code constraints, leading to error propagation. Consequently, the optimal LC level strongly depends of the system configuration as well as the employed detector calibration.

#### B. Empirical LLR Clipping

Widdup *et al.* subsequently considered to use some information contained in the LSD output. In particular, since the costs of only the best  $|\mathcal{L}|$  solutions are known, the others must be estimated from the knowledge that their cost is at least as high as that of the worst known point, namely the current radius [15]. This solution is also denoted as the last list entry [11] and offers a significant coded performance gain compared to the naive solution. However, this solution does not match with the widely used K-Best LSD nature, that does not take advantage (without early termination condition) of the radius constraint. Namely, contrary a depth-first search, the radius is not shrunked during the process. Even if the largest Euclidean distance of the counter-hypothesis is considered, such a clipping does not give importance enough to the fact that the counter-hypothesis does not appear in  $\mathcal{L}$ . Consequently, it leads to a significant performance loss [11].

That is why Kawamoto *et al.* introduced a likelihood function generation for selecting an appropriate clipping value [16], [17]. Briefly, this solution lies in a statistical study of the radius and an empirical result. In particular, the expectation of the minimum squared Euclidean distance among the bits within

the list output that offers both the hypothesis and counter-hypothesis is calculated. This value is then grown up by 50% [16] in order to increase its weight. Thus, the LC value is obtained.

The presented solution lies in empirical results that offer neither strong theoretical results nor convincing performances.

### C. SNR-aware LLR Clipping

As previously mentioned, the LC value is expected to depend on the Signal-to-Noise Ratio (SNR), on  $|\mathcal{L}|$  and on the modulation scheme (not of the code rate, as clearly introduced in [18], Theorem 1). However, the main issue relies on how to take these factors into account.

Milliner *et al.* recently proposed an analytical expression, especially detailed in [18] in the particular case of BPSK, that is claimed to provide the optimal clipping value in the AWGN case for any arbitrary code rate. Starting from the exact LLR definition, an approximation is proposed through the introduction of the channel state information based bit error probabilities of a bit to be +1 and -1.

The Symbol Error Rate (SER) for the ML detector in the case of a Pulse Amplitude Modulation (PAM) transmission over an AWGN SISO channel with effective SNR [11] is:

$$P_{S,PAM} = 2 \left( 1 - \frac{1}{\sqrt{|\xi|}} \right) Q \left( \sqrt{\frac{3}{|\xi|-1}} \sqrt{|\mathcal{L}|} \left( \frac{E_s}{N_0} \right)_i \right), \quad (4)$$

where  $Q(\cdot)$  denotes the Q-function and  $\left( \frac{E_s}{N_0} \right)_i$  is the instantaneous SNR for the  $i$ -th detection layer ( $i$ -th component of the transmit signal) [11]. Following classical QAM extensions of the PAM SER expression yields the QAM SER, from which the QAM BER can be easily obtained. The predicted error probability  $P_b(|\mathcal{L}|, \left( \frac{E_s}{N_0} \right)_i)$  then yields the SNR-aware  $\Lambda_{clip,i}(\mathbf{b}_k^n)$  value for bits in the  $i$ -th detection layer:

$$\Lambda_{clip,i}(\mathbf{b}_k^n) \approx -\ln P_b(|\mathcal{L}|, \left( \frac{E_s}{N_0} \right)_i). \quad (5)$$

This solution offers convenient results. In particular, it outperforms any FLC solution [11]. However, a drawback remain. Due to the AWGN assumption, the technique is optimal only in mean over multiple Rayleigh channel realizations. Also, the whole bit sequence is considered. In particular, the additional knowledge about the bit position is not taken into account.

By considering the aforementioned techniques, a simple while efficient optimization is introduced in the following.

## IV. PROPOSED SOLUTION

An original approach lies in taking into account the multi-level bit mapping nature of Quadrature Amplitude Modulation (QAM) which is a multi-level bits-to-symbol mapping. Every symbol correspond to a codeword. They are each characterized by a different mean Euclidean distance, and hence a different level of protection against noise and amplitude impairments. This aspect depends on the bit position within the bit sequence. In Figure 3, the example of 16QAM case is shown. In particular for this case, two levels exist comprising Most

Significant Bits (MSB), with a bold representation, and Least Significant Bits (LSB) for each signal point. For 64QAM, three levels of protection exist while only one exists for 4QAM. The latter being actually not strictly a multi-level mapping since the protection level is the same among the codeword. In the following, for the sake of simplicity and without loss of generality, only the 16QAM case will be considered.

In the LTE-A downlink case [19], QAM modulations with a multilevel Gray mapping can be partitioned into square subsets with minimum mean intrasubset Euclidean distance. In Figure 3, the MSBs of a signal point determine in which subset it is located.

This idea is novel in this context. In [12], the multilevel

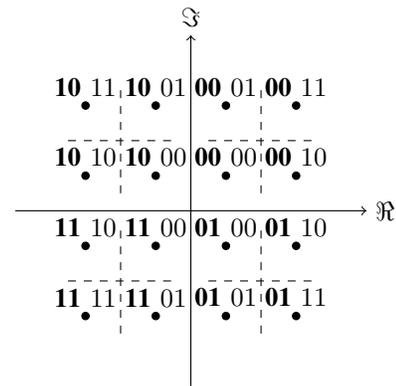


Fig. 3. 16QAM modulation constellation for LTE-A downlink [19]

bit mapping is considered in order to reduce the detector complexity, which becomes nearly independent of the signal constellation size. However, there is no link in the LLR calculation, and in particular in the LC.

Analytical expressions could be introduced from [20] and with required updates. However, Figure 4 exhibits more clearly the multilevel impact on LLRs in the particular case of a trivial  $4 \times 4$  AWGN MIMO channel with 16QAM modulations on each layer. As it is shown, the LLRs are distributed differently depending on the bit index and the SNR. In particular, there is no maximal value for MSB, while there is a maximal value for positive LSB, from a Soft-Decision ML output. Also, consistently with previous discussions, this maximal value is different according to the SNR. The idea presented here consists in exploiting this additional knowledge in order to apply different clipping values to generate lower-distortion approximated LLR.

Through a SNR normalization, the positive LSB are shown in Figure 5 to be still upper bounded by a constant value in the Rayleigh channel case, this upper bound is marked by diamonds. For multiple SNR, the distribution of LLRs normalized by SNR is depicted in a  $4 \times 4$  complex MIMO system. Similarly to [11], the proposed technique is SNR-aware and consider the modulation type as well. It does not consider the list size and offers an optimization by considering the additional knowledge of the bit position.

In particular, the positive LSB LLRs have been shown to

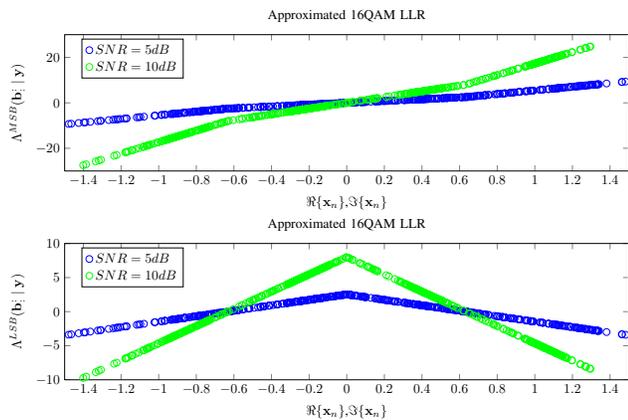


Fig. 4. Separated  $\Lambda^{LSB}(\mathbf{b}_k^n | \mathbf{y})$  and  $\Lambda^{MSB}(\mathbf{b}_k^n | \mathbf{y})$  representation of  $\Lambda(\mathbf{b}_k^n | \mathbf{y})$  as a function of both the real and imaginary parts of any transmit symbol, for multiple SNR values,  $4 \times 4$  complex MIMO system, 16QAM modulation on each layer, 800 distinct values per bit weight.

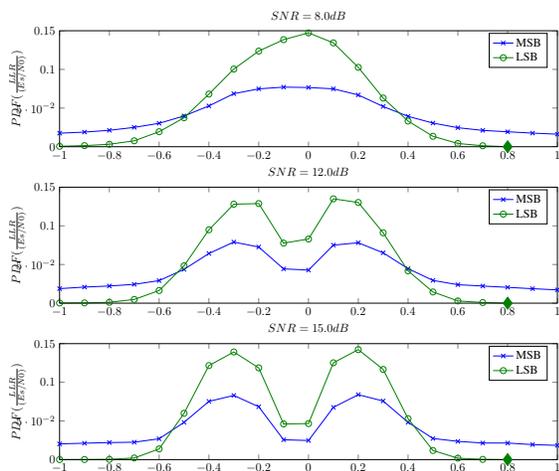


Fig. 5. Probability Density Function (PDF) of LLRs normalized soft ML output,  $4 \times 4$  complex MIMO system, 800 distinct values per bit weight.

offer a constant maximal magnitude, through a SNR normalization step. Concerning the negative LSB and the MSB, the LLRs still depend on the SNR. This point is more clearly illustrated in Figure 6, where the maximum normalized LLRs from the soft-decision ML output are plotted as a function of SNR for different bit positions. From Figure 6, an efficient  $\Lambda_{clip}^{LSB+}(\mathbf{b}_k^n)$  is obtained. Also, the authors highlight that it is constant over the SNR range and independent of the number of clipped bits. This is not the case for the other bit positions which depend on the SNR and on the number of clipped bits. Nevertheless, further note that the absolute clipping values for MSB are the same:

$$\Lambda_{clip}^{MSB-}(\mathbf{b}_k^n) = -\Lambda_{clip}^{MSB+}(\mathbf{b}_k^n). \quad (6)$$

The efficiency of the proposed solution is presented through coded performance comparisons in the next section.

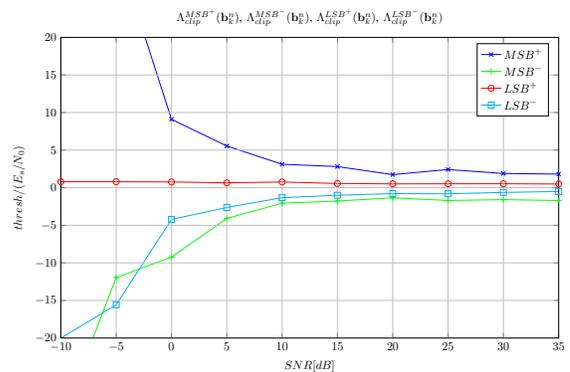


Fig. 6. Thresholds evolution as a function of SNR,  $4 \times 4$  MIMO Rayleigh channel,  $\frac{1000|\xi|}{2}$  simulations per bit weight and per SNR value.

## V. SIMULATION RESULTS

For the sake of comparability of the provided simulation results, the employed solution lies on considering a reference Soft-Decision ML detectors, which provides all the needed LLR values through an explicit computation. A certain percentage of these values are arbitrary discarded. Such a scheme is denoted as the *eclipsed* ML detector, and strictly corresponds to a LSD. The interest lies in keeping under control the percentage of missing LLR. The missing values are then clipped according to the proposed solution. The position of the discarded value is done randomly in order to make the ratio uniformly distributed between all the bit positions. Consequently, the real impact of the proposed solution is shown. In particular in the presented simulations, the ratio is defined as the proportion of forced clipped value from the soft-decision ML output. In the case of a ratio of  $p$ :  $p\%$  of the provided LLRs are clipped according to any clipping technique. The  $100 - p\%$  other LLR values remain explicitly computed. All the simulation results are compared to the ML detector with max-log approximation, used as a reference, and to the constant  $\pm 3$  clipping value, which has been previously shown to be a very efficient empirical result.

By considering the simulation conditions below, Figure 7 shows the BER and BLock Error Rate (BLER) error performance of the *eclipsed* ML detector in a  $4 \times 4$  SM-MIMO system. At the transmitter, binary information data bits are first serial-to-parallel-converted into four data streams and are segmented into blocks containing a selected number of bits per packet frame according to the employed Modulation and Coding Scheme (MCS). The information data sequence is encoded by Turbo coding with memory 2 code and  $1 + D + D^2$  and  $1 + D^2$  feedback and feedforward polynomials, respectively. Also, the data is interleaved from a Look-Up Table. The original coding rate  $R = \frac{1}{3}$  and then punctured according to the coding rate of  $R = \frac{1}{4}$  when the resultant encoded sequence is data-modulated. The considered modulation format is 16QAM only with multilevel LTE-A bit mapping. Blocks of information bits are fed to the channel encoder, and subsequently transmitted over a Rayleigh fading

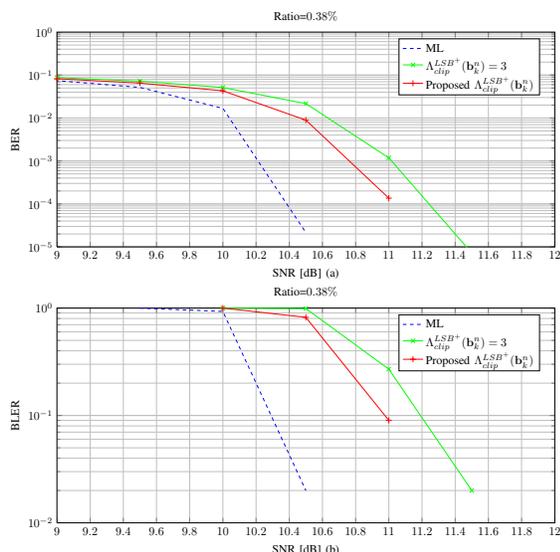


Fig. 7. BER (a) and BLER (b) performance plots,  $4 \times 4$  complex MIMO system, 16QAM modulation and  $R = \frac{3}{4}$  on each layer, 100 simulated blocks (a),  $3 \times 10^5$  simulated bits (b).

MIMO channel, by forcing the 1,049 transmit symbols per block, independently of the employed modulation. It presently corresponds to 1008 transmit redundant bits per block, to within one rounding. For each block, 20 half iterations within the turbo decoder have been performed. We consider it is sufficient for achieving the convergence of the turbo decoder and consequently for reaching the maximal performance.

Both the BER and BLER performances confirm the efficiency of our proposed technique for 38% of clipping values. This proportion matches with the ratio obtained through a naive LSD with  $|\mathcal{L}| = 16$  in a  $4 \times 4$  MIMO system with 16QAM modulation on each layer, as depicted in Figure 1. Such a LSD calibration may be shown to allow near-optimal performance in the hard-decision case [6]. As it is shown in Figure 7, the proposed technique outperforms the clipping of  $\pm 3$  for every bit position proposed in [12] by 0.26 dB for a BER of  $10^{-3}$  and by 0.21 dB for a BLER of  $10^{-1}$ .

## VI. CONCLUSION

This paper proposed a novel technique that allows for low distortion approximated LLR computation at the output of a soft-decision near-ML detector. Moreover, such approximation is judiciously applied depending on the actual bit-mapping indexing used by QAM constellations. To the best of the authors' knowledge, no method presented so far exploited this information to solve the problem of LC for the class of receivers considered and such a technical solution offers a significant performance gain. Also, the presented technique is general and may be applied to any more advanced technique.

## ACKNOWLEDGMENT

This work was supported by ST-ERICSSON.

## REFERENCES

- [1] C. Berrou, A. Glavieux and P. Thitimajshima. Near Shannon limit error-correcting coding and decoding: Turbo-codes (1). *Communications, IEEE International Conference on*, 2:1064–1070, May 1993.
- [2] W. Van Etten. Maximum likelihood receiver for multiple channel transmission systems. *Communications, IEEE Transactions on*, pages 276–283, Feb. 1976.
- [3] E. Agrell, T. Eriksson, A. Vardy and K. Zeger. Closest point search in lattices. *Information Theory, IEEE Transactions on*, 48(8):2201–2214, Aug. 2002.
- [4] D. Wübben, R. Böhnke, V. Kühn and K.-D. Kammeyer. Near-maximum-likelihood detection of MIMO systems using MMSE-based lattice reduction. *Communications, IEEE International Conference on*, 2:798–802, June 2004.
- [5] X.F. Qi and K. Holt. A Lattice-Reduction-Aided Soft Demapper for High-Rate Coded MIMO-OFDM Systems. *Signal Processing Letters, IEEE*, 14(5):305–308, May 2007.
- [6] K.-W. Wong, C.-Y. Tsui, S.-K. Cheng and W.-H. Mow. A VLSI Architecture of a K-Best Lattice Decoding Algorithm For MIMO Channels. *Circuits and Systems, IEEE International Symposium on*, 3:273–276, May 2002.
- [7] D.J. Jr Costello and G.D. Jr Forney. Channel Coding: The Road to Channel Capacity. volume 95, pages 1150–1177, Feb. 2007.
- [8] P. Elias. Channel Coding: The Road to Channel Capacity. pages 37–46, March 1955.
- [9] B.M. Hochwald and S. ten Brink. Achieving Near-Capacity on a Multiple-Antenna Channel. *Communications, IEEE Transactions on*, 51(3):389–399, Mar. 2003.
- [10] S. Båro, J. Hagenauer and M. Witzke. Iterative detection of MIMO transmission using a list-sequential (LISS) detector. *Communications, IEEE International Conference on*, 4:2653–2657, May 2003.
- [11] D.L. Milliner, E. Zimmermann, J.R. Barry and G. Fettweis. Channel State Information Based LLR Clipping in List MIMO Detection. *Personal, Indoor and Mobile Radio Communications International, IEEE International Symposium on*, pages 1–5, Sept. 2008.
- [12] Y.L.C. de Jong and T.J. Willink. Iterative Tree Search Detection for MIMO Wireless Systems. *Communications, IEEE Transactions on*, 53(6):930–935, June 2005.
- [13] S. Benedetto, D. Divsalar, G. Montorsi and F. Pollara. Serial Concatenation of Interleaved Codes: Performance Analysis, Design, and Iterative Decoding. *Information Theory, IEEE Transactions on*, 44(3):909–926, May 1998.
- [14] M. Myllyla, J. Antikainen, M. Juntti and J.R. Cavallaro. The effect of LLR clipping to the complexity of list sphere detector algorithms. *Signals, Systems and Computers, Asilomar, Conference on*, pages 1559–1563, Nov. 2007.
- [15] B. Widdup, G. Woodward and G. Knagge. A Highly-Parallel VLSI Architecture for a List Sphere Detector. *Communications, IEEE International Conference on*, 5:2720–2725, June 2004.
- [16] J. Kawamoto, H. Kawai, N. Maeda, K. Higuchi and M. Sawahashi. Investigations on likelihood function for QRM-MLD combined with MMSE-based multipath interference canceller suitable to soft-decision turbo decoding in broadband CDMA MIMO multiplexing. *Spread Spectrum Techniques and Applications, IEEE International Symposium on*, pages 628–633, Aug. 2004.
- [17] K. Higuchi, H. Kawai, N. Maeda, M. Sawahashi, T. Itoh, Y. Kakura, A. Ushirokawa and H. Seki. Likelihood function for QRM-MLD suitable for soft-decision turbo decoding and its performance for OFCDM MIMO multiplexing in multipath fading channel. *Personal, Indoor and Mobile Radio Communications, IEEE International Symposium on*, 2:1142–1148, Sep. 2004.
- [18] E. Zimmermann, D.L. Milliner, J.R. Barry and G. Fettweis. Optimal LLR Clipping Levels for Mixed Hard/Soft Output Detection. *Global Telecommunications Conference, IEEE*, pages 1–5, Dec. 2008.
- [19] 3GPP. Technical Specification Group Radio Access Network; Spreading and modulation (FDD) v9.1.0. TR 25.213, 3rd Generation Partnership Project (3GPP), 2009.
- [20] C.C. Wang. A Bandwidth-Efficient Binary Turbo Coded Waveform Using QAM Signaling. *Communications, Circuits and Systems and West Sino Expositions, IEEE International Conference on*, 1:37–41, July 2002.

# Analysis of the Least-Squares Adaptive Algorithms in Interference Cancellation Configuration

Silviu Ciochină, Constantin Paleologu, and Andrei Alexandru Enescu

Department of Telecommunications  
University Politehnica of Bucharest  
1-3 Iuliu Maniu Blvd., Bucharest, Romania  
{silviu, pale, aenescu}@comm.pub.ro

**Abstract**— An analysis of the interference cancellation configuration working based on the least-squares (LS) adaptive algorithms is presented in this paper. The existence of a “residual leakage” phenomenon of the useful signal, to the output of the adaptive filter, through the error signal is demonstrated. As a consequence, the useful signal can be attenuated, up to complete cancellation. This process is important for low values of the weighting parameter  $\lambda$  and is practically absent for  $\lambda$  close to unit. The simulations performed in an echo cancellation configuration support the theoretical findings.

**Keywords**— adaptive filters, least-squares (LS) algorithms, interference cancellation, echo cancellation

## I. INTRODUCTION

A lot of applications in the telecommunications field require cancelling an unknown interference that corrupts a useful signal. Such a problem can be solved using an adaptive filter working in an interference cancellation configuration. The goal of this system is to produce an estimate of the interference that will be subtracted from the received signal [1], [2]. One of the most common applications of this configuration type is echo cancellation [3], [4]. In this case, the principle is to synthesize a replica of the echo and to subtract it from the returned signal.

Besides convergence rate, an important aspect of an echo canceller is its performance during “double-talk” (i.e., near-end speech) [3], [4]. In the case of the normalized least-mean-square (NLMS) algorithm [1], the presence of near-end signal considerably disturbs the adaptive process. To eliminate the divergence of echo cancellers the standard procedure is to inhibit the weight updating during the double-talk. The presence of double-talk is detected by a double-talk detector (DTD). The DTD acts with a delay since it requires a number of samples to detect the double-talk presence. However, this very small delay can be enough to generate a considerable perturbation of the echo estimate.

Therefore, it will be desirable to implement fast converging and double-talk robust adaptive algorithms [5] in future echo cancellers. Based on convergence performance alone, a least-squares (LS) algorithm [1] is clearly the algorithm of choice.

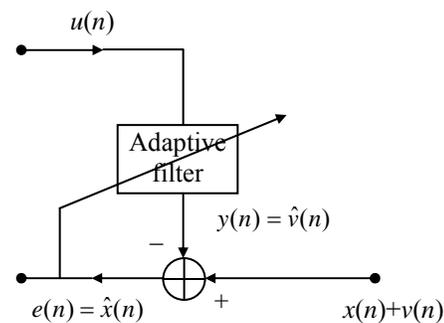


Figure 1. Interference cancellation configuration.

In this paper we will prove the existence of an interesting phenomenon that appears when an LS adaptive algorithm is used in an interference cancellation configuration. The phenomenon consists of a “residual leakage” of the useful signal, through the error signal into the output of the adaptive filter, even if it is uncorrelated with the input signal of the adaptive filter. The process depends on the algorithm weighting parameter  $\lambda$ . As a consequence, not only the perturbation (i.e., the echo), but also the useful signal can be severely attenuated, up to complete cancellation. Controlling this phenomenon in case of an echo canceller could solve the “double-talk” problem without any DTD system.

The paper is organized as follows. In Section II we briefly review the adaptive interference cancellation configuration, working in ideal conditions. In Section III, the real behavior of the configuration is considered, working based on an LS algorithm. Replacing the statistical averages by temporal estimators proves to lead to the “residual leakage” phenomenon, in the case of the low memory algorithms (i.e., using small  $\lambda$ ). Simulations performed in the context of echo cancellation are presented in Section IV. Finally, Section V concludes this work.

## II. IDEAL BEHAVIOR OF THE ADAPTIVE INTERFERENCE CANCELLATION CONFIGURATION

In the case of the interference cancellation applications (Fig. 1), an adaptive filter is used to cancel an unknown interference,  $v(n)$ , that corrupts a useful signal,  $x(n)$ .

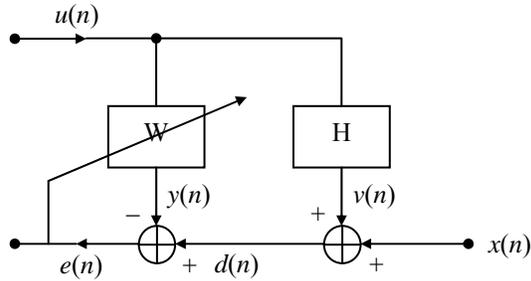


Figure 2. Echo cancellation configuration.

This scheme needs two inputs. A so-called primary signal consists of the corrupted signal,  $x(n) + v(n)$ , and plays the role of the “desired signal” in the adaptive configuration. The second one,  $u(n)$ , is correlated with the perturbation  $v(n)$  and uncorrelated with  $x(n)$ . It is applied to the input of the adaptive filter. In addition,  $x(n)$  and  $v(n)$  are mutually uncorrelated. The output of the adaptive filter  $y(n)$  is expected to be an estimate of  $v(n)$  and, consequently, the error signal  $e(n)$  should be an estimate of  $x(n)$ .

A specific application of this configuration is echo cancellation [3], [4]. In this case (Fig. 2),  $v(n)$  is the echo generated by a system characterized by the impulse response

$$\mathbf{h} = [h_0, h_1, \dots, h_{N-1}]^H, \quad (1)$$

where the superscript  $H$  denotes Hermitian transposition (transposition and complex conjugation) and  $W$  is an adaptive filter, having the coefficients

$$\mathbf{w} = [w_0, w_1, \dots, w_{N-1}]^H. \quad (2)$$

The sequence  $u(n)$  is the far-end signal and  $x(n)$  is the near-end. We suppose that  $x(n)$  and  $u(n)$  are uncorrelated, i.e.,

$$E\{u(n)x^*(n-k)\} = 0, \forall k \in Z, \quad (3)$$

where  $E$  is the expectation operator and superscript  $*$  denotes the complex conjugation.

Defining the vector:

$$\mathbf{u}(n) = [u(n), u(n-1), \dots, u(n-N+1)]^T, \quad (4)$$

where superscript  $T$  denotes the transposition operation, we have:

$$v(n) = \mathbf{h}^H \mathbf{u}(n), \quad (5)$$

and

$$e(n) = v(n) - y(n) + x(n). \quad (6)$$

In case of the Wiener filter the optimal coefficients are given by normal equation [1]:

$$\mathbf{R} \mathbf{w}_{opt} = \mathbf{p}, \quad (7)$$

with

$$\mathbf{R} = E\{\mathbf{u}(n)\mathbf{u}^H(n)\} \quad (8)$$

and

$$\mathbf{p} = E\{\mathbf{u}(n)d^*(n)\}. \quad (9)$$

Taking into account the relations (3) and (6) we obtain:

$$\begin{aligned} \mathbf{p} &= E\{\mathbf{u}(n)d^*(n)\} = E\{\mathbf{u}(n)(x^*(n) + v^*(n))\} = \\ &= E\{\mathbf{u}(n)x^*(n)\} + E\{\mathbf{u}(n)v^*(n)\} = \\ &= E\{\mathbf{u}(n)v^*(n)\} = E\{\mathbf{u}(n)\mathbf{u}^H(n)\mathbf{h}\} = \mathbf{R}\mathbf{h} \end{aligned} \quad (10)$$

so that, according to (7) and (10), the value of optimal coefficients results as

$$\mathbf{w}_{opt} = \mathbf{h}. \quad (11)$$

In this case,

$$y(n) = \mathbf{w}_{opt}^H \mathbf{u}(n) = \mathbf{h}^H \mathbf{u}(n) = v(n) \Rightarrow e(n) = x(n), \quad (12)$$

so that the separation of signals  $x(n)$  and  $v(n)$  is correctly performed.

### III. THE RESIDUAL LEAKAGE PHENOMENON

Let us consider the real case of the classical recursive least-squares (RLS) adaptive algorithm [1]. In this situation, the statistical expectation is replaced by a weighted sum as follows:

$$E\{\bullet\} \rightarrow \sum_{i=1}^n \lambda^{n-i} \{\bullet\}, \quad (13)$$

where  $\lambda$  is the exponential weighting factor (also known as the “forgetting factor”) of the RLS algorithm. Consequently, the normal equation from (7) becomes

$$\begin{aligned} \sum_{i=1}^n \lambda^{n-i} \mathbf{u}(i)\mathbf{u}^H(i) \mathbf{w}_{opt} &= \sum_{i=1}^n \lambda^{n-i} \mathbf{u}(i)(v^*(i) + x^*(i)) = \\ &= \sum_{i=1}^n \lambda^{n-i} \mathbf{u}(i)v^*(i) + \sum_{i=1}^n \lambda^{n-i} \mathbf{u}(i)x^*(i). \end{aligned} \quad (14)$$

For values of the exponential weighting factor  $\lambda$  very close to unit and for a value of  $n$  high enough we may write

$$\lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=1}^n \lambda^{n-i} \mathbf{u}(i) x^*(i) \cong E \{ \mathbf{u}(n) x^*(n) \} = \mathbf{0} \quad \Rightarrow$$

$$\sum_{i=1}^n \lambda^{n-i} \mathbf{u}(i) \mathbf{u}^H(i) \mathbf{w}_{opt} \cong \sum_{i=1}^n \lambda^{n-i} \mathbf{u}(i) v^*(i) = \sum_{i=1}^n \lambda^{n-i} \mathbf{u}(i) \mathbf{u}^H(i) \mathbf{h}$$

$$\Rightarrow \quad \mathbf{w}_{opt} \cong \mathbf{h} \quad (15)$$

so that  $e(n) = x(n)$  like in the ideal case.

On the other hand, for a value of the exponential weighting factor small enough, so that  $\lambda^k \ll 1$  for  $k > n_0$ , we may use the following approximation:

$$\sum_{i=1}^n \lambda^{n-i} \{ \bullet \} \cong \sum_{i=n-n_0+1}^n \lambda^{n-i} \{ \bullet \}. \quad (16)$$

According to the orthogonally principle, the normal equation becomes

$$\sum_{i=n-n_0+1}^n \lambda^{n-i} \mathbf{u}(i) e^*(i) = \mathbf{0}. \quad (17)$$

This is a homogeneous set of  $N$  equations with  $n_0$  unknown parameters,  $e(i)$ . If  $n_0 < N$  then the set of equations has the unique solution

$$e(i) = 0 \quad \text{for } i = n - n_0 + 1, \dots, n \quad (18)$$

leading to

$$y(n) = \mathbf{w}^H(n) \mathbf{u}(n) = x(n) + v(n). \quad (19)$$

Consequently, there is a “residual leakage” of  $x(n)$ , through the error signal into the output of the adaptive filter. In this situation the useful signal,  $x(n)$  is suppressed together with the perturbation  $v(n)$ . A small value of  $\lambda$  or a high value of  $N$  intensifies this phenomenon.

Concluding, in the real case of any LS adaptive algorithm used in an interference cancellation configuration two effects appear:

- $\mathbf{w}$  differs from  $\mathbf{h}$  in a certain extent and this may be viewed as a divergence of the algorithm;
- $y(n)$  will contain a component proportional to  $x(n)$ , that will be subtracted from the total received signal; this phenomenon is in fact a leakage of the  $x(n)$  in  $y(n)$ , through the error signal  $e(n)$ ; the result consists of an unwanted attenuation of the desired signal.

If we consider an echo cancellation application where the echo path is significantly long, leading to a high value for the length of the adaptive filter, we have to use a value of  $\lambda$  very close to unit in order to reduce unwanted attenuation of the near-end signal.

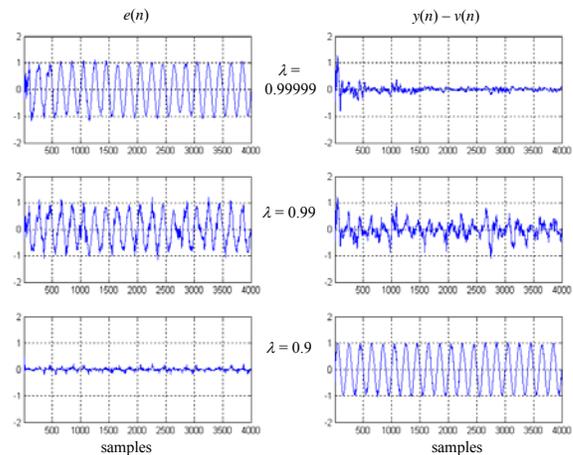


Figure 3. The filter length is  $N = 32$  and different values of  $\lambda$  are used.

However, the requirements of an echo canceller are both rapid convergence and a low computational cost. Thus, a highly desirable algorithm is a “low cost” LS algorithm (in terms of the computational complexity), i.e., the fast LS algorithms [1].

#### IV. SIMULATION RESULTS

For the first set of simulations we analyze a simple interference cancellation application. Using the configuration from Fig. 2 we choose  $u(n)$  uniformly distributed in  $[-1;1]$  and  $x(n)$  as a low frequency sine wave ( $\omega_0 = 0.01\pi$ ). The adaptive filter has  $N$  coefficients. In the first experiment we fixed  $N = 32$  and we take different values for the exponential weighting factor  $\lambda$  (Fig. 3).

In a second case we used a fixed weighting factor  $\lambda = 0.99$  and different values for the length of the adaptive filter  $N$  (Fig. 4). In order to outline the “residual leakage” phenomenon we plot the signals  $e(n)$  (the recovered sine wave) and the difference  $y(n) - v(n)$  [the component from  $x(n)$  leaked into the output of the adaptive filter].

We can notice that similar effects appear when the value of the exponential weighting factor  $\lambda$  decreases or the length of the adaptive filter  $N$  increases. The desired signal  $x(n)$  leaks into the output of the adaptive filter, leading to an unwanted attenuation of the recovered signal  $e(n)$ .

In order to observe the phenomenon in the spectral domain we repeat the previous experiments using a higher frequency sine wave ( $\omega_0 = 0.5\pi$ ) as desired signal. The corresponding spectral results are given in Fig. 5 and Fig. 6. The conclusions are practically the same. The recovered sine wave is strongly attenuated when the value of  $\lambda$  decreases. A similar effect appears when the length of the adaptive filter  $N$  increases.

In order to approach the context of echo cancellation, a second set of simulations was performed using speech sequences for both  $u(n)$  and  $x(n)$  signals.

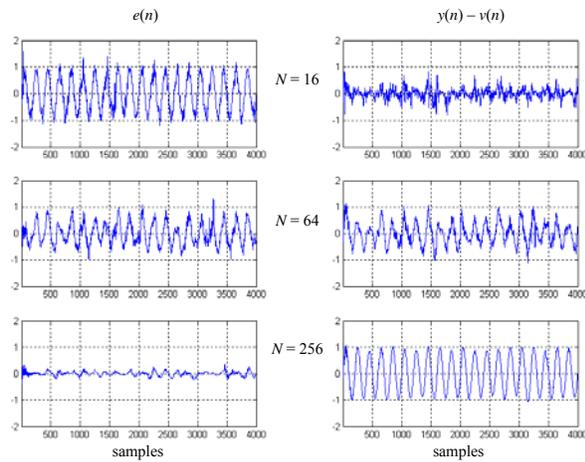


Figure 4. The exponential weighting factor is  $\lambda = 0.99$  and different values of  $N$  are used.

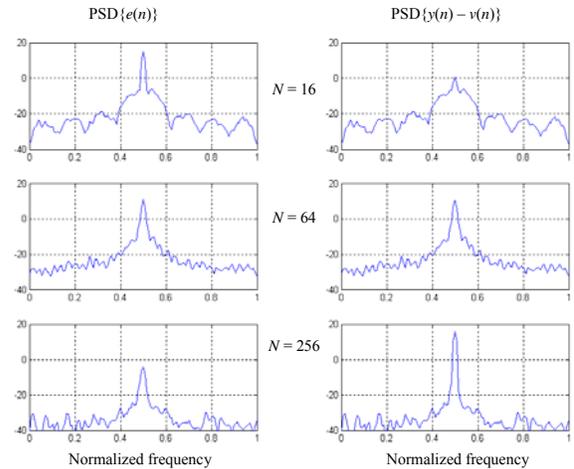


Figure 6. Power spectra [dB]; the exponential weighting factor is  $\lambda = 0.99$  and different values of  $N$  are used.

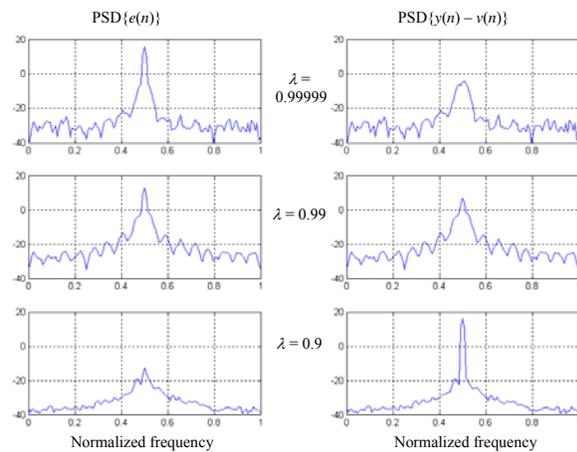


Figure 5. Power spectra [dB]; the filter length is  $N = 32$  and different values of  $\lambda$  are used.

Moreover, the H filter is a particular echo path according with ITU-T G.168 Recommendation for digital echo cancellers [6]. It is convenient to subtract out the direct near-end component from the error signal  $e(n)$  [7]. The residual error  $r(n) = e(n) - x(n)$  cumulates the undesired attenuation of the near-end signal  $x(n)$  and the imperfect rejection of the echo path response  $v(n)$ . In a real application such a subtraction can never be done because the signal  $x(n)$  is not available.

In the first experiment we choose a value of the exponential weighting factor very close to unit ( $\lambda = 0.99999$ ) and a 32 msec. echo path (corresponding to  $N = 256$ ). One can see that the near-end signal  $x(n)$  is recovered in  $e(n)$  with slight distortions (Fig. 7). Next, in order to point out the leakage process, we decrease the value of  $\lambda$  (Fig. 8). Because of the lower value of  $\lambda$  the “residual leakage” phenomenon is significant. The adaptive filter rejects not only the far-end signal but also the near-end signal. The transmitted signal remains in the absence of the far-end signal (where there is no input signal for the adaptive filter).

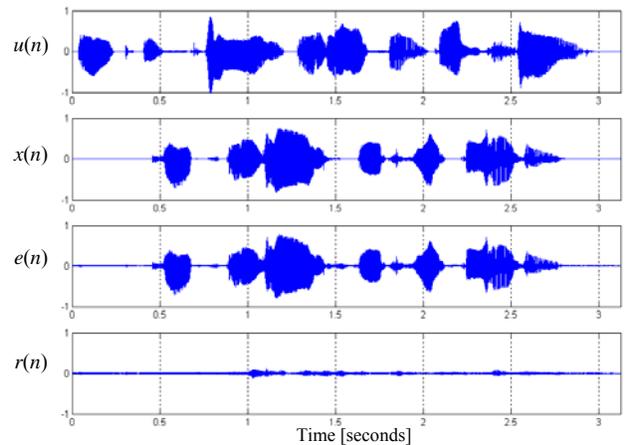


Figure 7. Double-talk situation. The exponential weighting factor is  $\lambda = 0.99999$  and the filter length is  $N = 256$ .

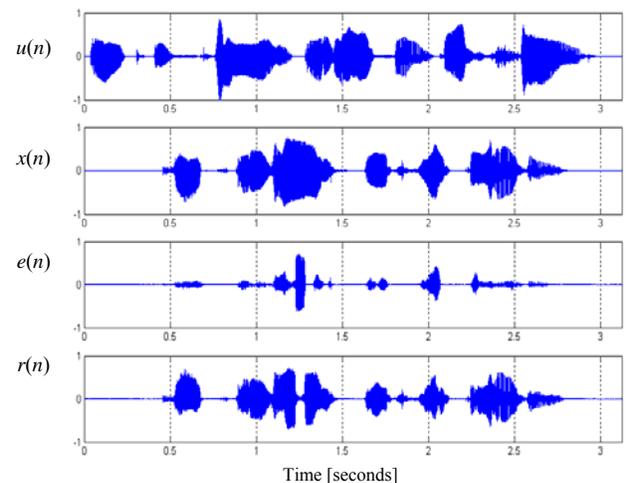


Figure 8. Double-talk situation. The exponential weighting factor is  $\lambda = 0.99$  and the filter length is  $N = 256$ .

## V. CONCLUSIONS

In this paper we have demonstrated the existence of a phenomenon that appears when a LS adaptive algorithm is used in an interference cancellation configuration. It consists of a “residual leakage” of the useful signal, through the error signal into the output of the adaptive filter that leads to an unwanted attenuation of the useful signal. This process strongly depends on the algorithm weighting parameter  $\lambda$  (it is important for low  $\lambda$  and is practically absent for  $\lambda \cong 1$ ) and it is influenced by the length of the adaptive filter  $N$  (it is amplified by a high value of  $N$ ).

Both the theoretical and experimental developments lead to the conclusion that the “residual leakage” phenomenon can be avoided for a memory of the algorithm significantly higher than the adaptive filter length.

Controlling this phenomenon in case of an echo canceller we can solve the “double-talk” problem without any DTD system. Taking into account that the length of an echo path is in general high enough, we have to use a value of  $\lambda$  very close to unit in order to reduce unwanted attenuation of the near-end signal.

## ACKNOWLEDGMENT

This work was supported by the UEFISCSU Romania under Grant PN-II-RU-TE no. 7 / 05.08.2010 and Grant PN-II-RU-PD no. 39 / 12.08.2010.

## REFERENCES

- [1] S. Haykin, *Adaptive Filter Theory*, Fourth Edition, Prentice Hall International, Inc., Upper Saddle River, New Jersey, 2002.
- [2] J. Benesty and Y. Huang, Eds., *Adaptive Signal Processing—Applications to Real-World Problems*. Berlin, Germany: Springer-Verlag, 2003.
- [3] S. L. Gay and J. Benesty, Eds., *Acoustic Signal Processing for Telecommunication*. Boston, MA: Kluwer Academic Publisher, 2000.
- [4] J. Benesty, T. Gaensler, D. R. Morgan, M. M. Sondhi, and S. L. Gay, *Advances in Network and Acoustic Echo Cancellation*. Berlin, Germany: Springer-Verlag, 2001.
- [5] T. Gänslér, St. L. Gay, M. M. Sondhi, J. Benesty, “Double-Talk Robust Fast Converging Algorithms for Network Echo Cancellation”, *IEEE Trans. Speech and Audio Processing*, vol. 8, no. 6, pp. 656-663, Nov. 2000.
- [6] *Digital Network Echo Cancellers*, ITU-T Rec. G.168, 2002.
- [7] D.L. Duttweiler, “Proportionate Normalized Least-Mean-Squares Adaptation in Echo Cancelers,” *IEEE Trans. Speech and Audio Processing*, vol. 8, no. 5, pp. 508-518, Sept. 2000.

## An Evaluation Of Multiwavelet Families For Stereo Correspondence Matching

Pooneh Bagheri Zadeh

Department of Engineering, Faculty of Technology  
De Montfort University  
Leicester, UK  
e-mail: pbz@dmu.ac.uk

Cristian V. Serdean

Department of Engineering, Faculty of Technology  
De Montfort University  
Leicester, UK  
e-mail: cvs@dmu.ac.uk

**Abstract**—This paper presents an evaluation of different types and families of multiwavelets in stereo correspondence matching. Different multiwavelet families with different filter types such as balanced versus unbalanced, symmetric-symmetric versus symmetric-antisymmetric are used. Normalized cross correlation is employed to find the best correspondence points and generate a disparity map. In the case of balanced multiwavelets, due to similar spectral content of the four generated low frequency subbands, they are shuffled to form a single baseband and then this baseband is used to generate a disparity map. However, in the case of unbalanced multiwavelets, the resulting basebands are used to form four disparity maps and then these maps are combined using a Fuzzy algorithm to generate a single disparity map. Middlebury stereo test images are used to generate experimental results. Results show that the unbalanced multiwavelets produce a smoother disparity map with less mismatch errors compared to balanced multiwavelets.

**Keywords**—Multiwavelets; stereo correspondence matching; normalized cross correlation;

### I. INTRODUCTION

Stereo correspondence is an issue of great importance in the field of computer vision and 3D reconstruction. It concerns the matching of points between a pair of stereo images of the same scene. The disparity is calculated as the distance of the correspondence points when one of the two stereo image pairs is projected onto the other. The disparity map along with the stereo camera parameters are then used to calculate the depth map and produce a 3D view of the scene. Nevertheless, a number of problems such as occlusion, ambiguity, illumination variation and radial distortion limit the accuracy of the disparity map, which is crucial in generating a precise 3D view of the scene [1].

Over the past years much research has been done to improve the performance of correspondence matching techniques. Multiresolution-based stereo matching algorithms have received much attention due to the hierarchical and scale-space localization properties of the wavelets [2][3]. This allows for correspondence matching to be performed on a coarse-to-fine basis, resulting in decreased computational costs. Sarkar and Bansal [3] presented a

multiresolution-based correspondence technique using a mutual information algorithm. They showed that the multiresolution technique produces significantly more accurate matching results compared to non-multiresolution based algorithms, at much lower computational cost.

Research has shown that multiwavelets (unlike scalar wavelets) can possess orthogonality (preserving length), symmetry (good performance at the boundaries via linear-phase), and a high approximation order simultaneously [4], which could potentially increase the accuracy of correspondence matching techniques. Bhatti and Nahavandi [5] introduced a multiwavelet based stereo correspondence matching algorithm. They use the wavelet transform modulus maxima to generate a disparity map at the coarsest level. This is then followed by the coarse-to-fine strategy to refine the disparity map up to the finest level. Bagheri Zadeh and Serdean [6] proposed another multiwavelet based stereo correspondence matching technique. They used a global error energy minimization technique to find the best correspondence points between the same multiwavelet's lowest frequency subbands of the stereo pair, followed by a fuzzy algorithm to form a dense disparity map.

In spite of their highly desirable advantages compared to scalar wavelets, the application of different types and families of multiwavelets in stereo correspondence matching has been little investigated in the literature so far.

This paper studies the application of different types and families of multiwavelets in stereo correspondence matching. A multiwavelet is first applied to the input stereo images to decompose them into a number of subbands. Normalized cross correlation is used to generate a disparity map at the coarsest level. In the case of balanced multiwavelets, as the four low frequency subbands have similar spectral content, they are shuffled to generate one baseband, while in the case of unbalanced multiwavelets, the resulting basebands are used to form four disparity maps and then a Fuzzy algorithm is used to combine the four maps and generate one disparity map.

The rest of the paper is organized as it follows. Section II introduces a brief review of the multiwavelet transform. The proposed stereo matching technique for both balanced and

L <sub>1</sub> L <sub>1</sub>	L <sub>1</sub> L <sub>2</sub>	L <sub>1</sub> H <sub>1</sub>	L <sub>1</sub> H <sub>2</sub>
L <sub>2</sub> L <sub>1</sub>	L <sub>2</sub> L <sub>2</sub>	L <sub>2</sub> H <sub>1</sub>	L <sub>2</sub> H <sub>2</sub>
H <sub>1</sub> L <sub>1</sub>	H <sub>1</sub> L <sub>2</sub>	H <sub>1</sub> H <sub>1</sub>	H <sub>1</sub> H <sub>2</sub>
H <sub>2</sub> L <sub>1</sub>	H <sub>2</sub> L <sub>2</sub>	H <sub>2</sub> H <sub>1</sub>	H <sub>2</sub> H <sub>2</sub>

Figure 1. One level of 2D Multiwavelet decomposition.

unbalanced multiwavelets is discussed in Section III. Experimental results are presented in Section IV and the paper is concluded at Section V.

## II. MULTIWAVELET TRANSFORM

In many respects multiwavelet transforms are very similar to scalar wavelet transforms. In contrast to the wavelet transform, which supports one wavelet and one scaling function, multiwavelets have two or more scaling and wavelet functions. A multiwavelet with two scaling and wavelet functions can be defined as [7]:

$$\Phi(t) = \sqrt{2} \sum_{k=-\infty}^{k=\infty} H_k \Phi(mt - k) \tag{1}$$

$$\Psi(t) = \sqrt{2} \sum_{k=-\infty}^{k=\infty} G_k \Psi(mt - k)$$

where  $\Phi(t)$  and  $\Psi(t)$  are the multi-scaling function and multiwavelet function and  $H_k$  and  $G_k$  are  $r \times r$  matrix filters ( $r$  is the number of scaling- and wavelet functions). To date, most multiwavelets have  $r = 2$  [4,7].

One level of decomposition for a 2D multiwavelet with multiplicity 2 produces sixteen subbands as shown in Figure 1, where  $L_x L_y$  represent the approximation subbands and  $L_x H_y$ ,  $H_x L_y$  and  $H_x H_y$  are the detail subbands, with  $x = \overline{1, 2}$ .

The major advantage of multiwavelets over scalar wavelets is their ability to possess symmetry, orthogonality and higher order of approximation simultaneously, which is impossible for scalar wavelets. Furthermore, the multichannel structure of the multiwavelet transform is a closer approximation of the human visual system than what wavelets offer. In the case of unbalanced multiwavelets, the resulting approximation subbands carry different spectral content of the original image (both high- and low-frequencies), while for balanced multiwavelets, the approximation subbands contain similar spectral content of the original image [8]. This feature of unbalanced multiwavelets has the potential to increase the accuracy of the calculated disparity maps and to reduce the number of

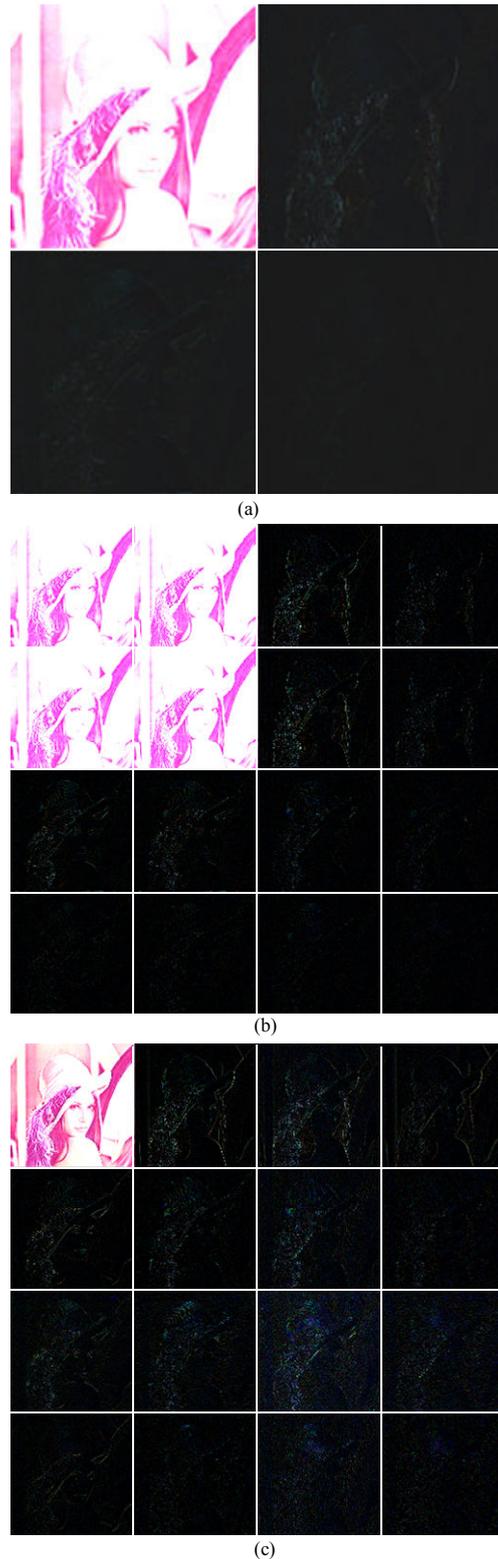


Figure 2. Single level decomposition of Lena test image (a) Antonini 9/7 wavelet transform, (b) balanced bat01 multiwavelet transform and (c) unbalanced GHM multiwavelet transform.

erroneous matches compared to that of balanced multiwavelets.

Figures 2(a) to 2(c) give a visual comparison of the resulting subbands for the Antonini 9/7 scalar wavelet, as well as for the balanced bat01 and unbalanced GHM multiwavelets applied to the Lena test image. As it can be seen from Figure 2, multiwavelets generate four subbands instead of each subband that wavelets create. The resulting unbalanced multiwavelet subbands carry different spectral content of the original Lena test image, while the balanced multiwavelet subbands produce similar spectral content of the original image. More information about the generation of multiwavelets, their properties and their applications can be found in [4-7].

### III. EVALUATION OF MULTIWAVELETS' FAMILY IN STEREO CORRESPONDENCE MATCHING

The proposed stereo correspondence matching system is based on multiwavelets and normalized cross correlation. Figures 3(a) and 3(b) show block diagrams of the proposed system for balanced and unbalanced multiwavelets respectively. A pair of stereo images is input to the stereo matching system. The images are first rectified to suppress the vertical displacement. A multiwavelet transform is then applied to each input stereo image. A number of different types and families of multiwavelets are evaluated. Since, the information in the approximation subbands is less sensitive to the shift variability of the multiwavelets, these subbands are used for correspondence matching purposes. In the case of balanced multiwavelets (Figure 3(a)), since their basebands contain similar spectral information, it is possible to use the shuffling technique proposed in [9] to rearrange the multiwavelet coefficients and generate a single low frequency subband. Figure 4 shows how four multiwavelet basebands are shuffled and a single baseband is formed. Figure 4(a) shows the four multiwavelet basebands with eight pixels (two from each baseband) highlighted and given a unique numeric label. Figure 4(b) shows the same set of pixels after shuffling, where coefficients corresponding to the same spatial locations in different basebands are placed together and one baseband is generated. Normalized cross correlation is then employed to find the best correspondence points between the two basebands of the stereo images and a disparity map is generated.

Figure 3(b) shows a block diagram of the unbalanced multiwavelet based stereo matching system. The shuffling technique works very well for balanced multiwavelets but it is not suitable for unbalanced multiwavelets due to their different spatio-frequency subband content. The unbalanced multiwavelets basebands contain both high and low frequency information with  $L_1L_1$  (top left baseband) containing most of the image energy. For correspondence matching purposes, the same basebands from the two views are input to the normalized cross correlation block, generating four disparity maps as a result. As most of the image energy is concentrated in  $L_1L_1$ , its output disparity map is more reliable than the other three disparity maps

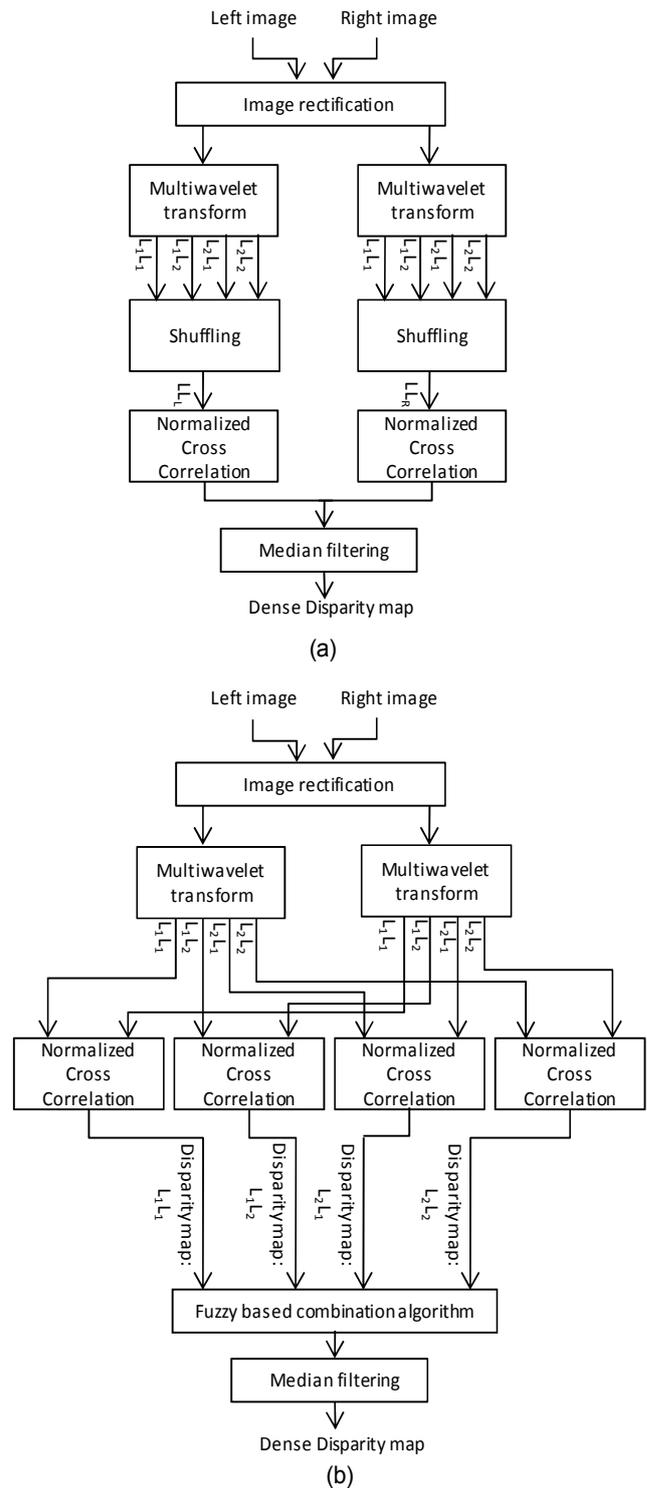


Figure 3. Block diagram of multiwavelet based stereo matching technique, (a) balanced- and (b) unbalanced-multiwavelets.

generated from other basebands,  $L_1L_2$ ,  $L_2L_1$ ,  $L_2L_2$ . Based on this property of unbalanced multiwavelets, a Fuzzy

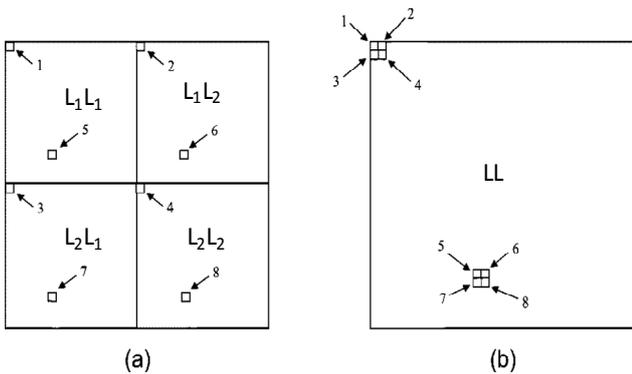


Figure 4. Shuffling method for multiwavelet baseband coefficients; selected pixels are numbered to indicate correspondence (a) before shuffling and (b) after shuffling.

algorithm is employed to combine the four disparity maps. This algorithm gives a higher weight to the disparity values of the  $L_1L_1$  disparity map, while the disparity values of the other three disparity maps are used to refine the final disparity map. A median filter is then applied to further smooth the resulting disparity map.

IV. SIMULATION RESULTS

The performance of different types and families of multiwavelets in stereo correspondence has been evaluated using 'Teddy' and 'Cones' stereo test images from the Middlebury stereo database [10]. Figure 5 shows the left image and the ground truth of these test images. The experimental results were generated using a number of multiwavelets, i.e. balanced versus unbalanced and symmetric-symmetric (SYM-SYM) versus symmetric-antisymmetric (SYM-ASYM) multiwavelets (listed in Table I). Table I shows the percentage of "bad pixels" at which the disparity error is larger than 1, for all regions (all). To give a visual comparison, the resulting disparity maps for balanced GHM and unbalanced BIGHM multiwavelets, applied to 'Cones' and 'Teddy' test images are shown in Figures 6(a) and 6(b), respectively. In these figures areas with intensity zero represent unreliable disparities. As can be seen from the results presented in Table I, generally unbalanced multiwavelets give better results compared to the balanced multiwavelets. The symmetric-symmetric multiwavelets seem to produce slightly better results compared to symmetric-antisymmetric multiwavelets (SA4). However, the symmetric-symmetric and symmetric-antisymmetric property of multiwavelets doesn't seem to have much effect on the resulting disparity map. From Figure 6, it is clear that the unbalanced multiwavelet based algorithm produces more accurate and smoother disparity maps compared to the balanced multiwavelet case. This can be explained by the fact that the approximation subbands of the unbalanced multiwavelet carry different spectral content of the input images, which enables the matching algorithm to generate more reliable matches.

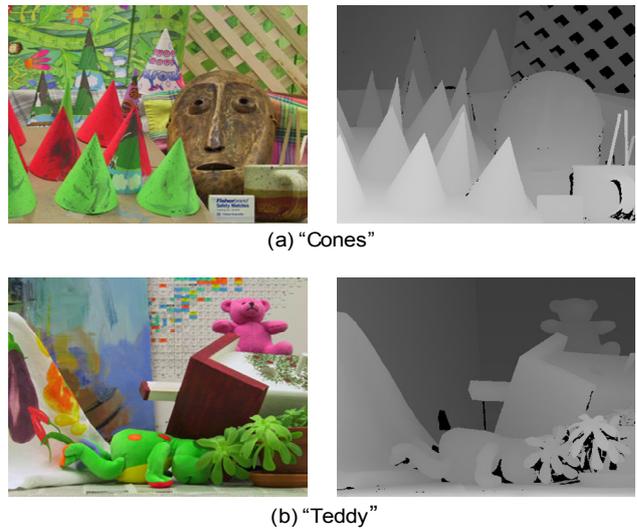


Figure 5. Left image and the ground truth of (a) 'Cones' and (b) 'Teddy'.

TABLE I. EVALUATION RESULTS OF DIFFERENT MULTIWAVELETS IN STEREO CORRESPONDENCE MATCHING.

'Teddy' (All)			
Balanced Multiwavelets		Unbalanced Multiwavelets	
CARDBAL2	9.84	BIH32S	8.92
CARDBAL 3	9.52	BIH52S(SYM-SYM)	8.91
BAT 01	10.37	BIH34N	8.92
BAT02	9.66	BIH54N (SYM-SYM)	8.99
GHM (SYM-SYM)	10.48	BIGHM	9.02
	9.84	SA4 (SYM-ASYM)	9.91
'Cones' (All)			
Balanced Multiwavelets		Unbalanced Multiwavelets	
CARDBAL2	8.84	BIH32S	8.75
CARDBAL 3	9.28	BIH52S(SYM-SYM)	8.85
BAT 01	9.34	BIH34N	8.74
BAT02	9.81	BIH54N (SYM-SYM)	8.91
GHM (SYM-SYM)	9.69	BIGHM	8.54
		SA4 (SYM-ASYM)	9.39

V. CONCLUSIONS

This paper has investigated the application of different types and families of multiwavelets in stereo correspondence matching. For this purpose, two correspondence matching algorithms were designed to deal with both balanced and unbalanced multiwavelets. In the case of balanced multiwavelets, due to the similar frequency content of the four multiwavelet low frequency subbands, they were shuffled to generate a single baseband and then normalized cross correlation was used to generate a disparity map. In the case of unbalanced multiwavelets, the four generated basebands and normalized cross correlation was used to generate four disparity maps. These maps were then combined using a Fuzzy algorithm to form a single disparity map. The results generated using Middlebury stereo test images show that unbalanced multiwavelets work better than balanced ones in stereo correspondence matching, while the symmetric-symmetric and symmetric-antisymmetric

property of multiwavelets doesn't have a significant effect in reducing erroneous matches.

#### ACKNOWLEDGMENT

This work was supported by EPSRC under the EP/G029423/1 grant.

#### REFERENCES

- [1] D. Scharstein and R. Szeliski, "A Taxonomy and Evaluation of Dense Two-Frame Stereo Correspondence Algorithms," *Int. Journal of Comp. Vision*, vol. 47, pp. 7-42, April 2002.
- [2] S. Mallat, *A Wavelet Tour of Signal Processing*, 3rd ed., Academic Press, 2009.
- [3] I. Sarkar and M. Bansal, "A wavelet-based multiresolution approach to solve the stereo correspondence problem using mutual information," *IEEE Trans. on system, man, and cybernetics*, vol. 37, pp. 1009-1014, August 2007.
- [4] V. Strela and A.T. Walden, "Signal and image denoising via wavelet thresholding: orthogonal and biorthogonal, scalar and multiple wavelet transforms," *In Nonlinear and Nonstationary Signal Processing*, pp. 124-157, 1998.
- [5] A. Bhatti and S. Nahvandi, "Depth estimation using multi-wavelet analysis based stereo vision approach," *International Journal of Wavelets, Multiresolution and Information Processing*, vol. 6, pp. 481-497, 2008.
- [6] P. Bagheri Zadeh and C. Serdean, "Stereo Correspondence Matching: Balanced Multiwavelets versus unbalanced Multiwavelets," *The 2010 European Signal Processing Conference (EUSIPCO-2010)*, pp. 1509-1513, August 2010.
- [7] V. Strela, "Multiwavelets: theory and applications," PhD thesis, MIT, 1996.
- [8] L. Ghouti, A. Bouridane, M. K. Ibrahim, and S. Boussakta, "Digital Image Watermarking Using Balanced Multiwavelets," *IEEE Trans. on Signal Pro.*, Vol. 54, no. 4, pp. 1519-1536, 2006.
- [9] M.B. Martin, and A.E. Bell, "New image compression techniques multiwavelets and multiwavelet packets," *IEEE Trans. Image Process.* Vol. 10, no. 4, pp. 500-510, 2001.
- [10] <http://vision.middlebury.edu/stereo/>

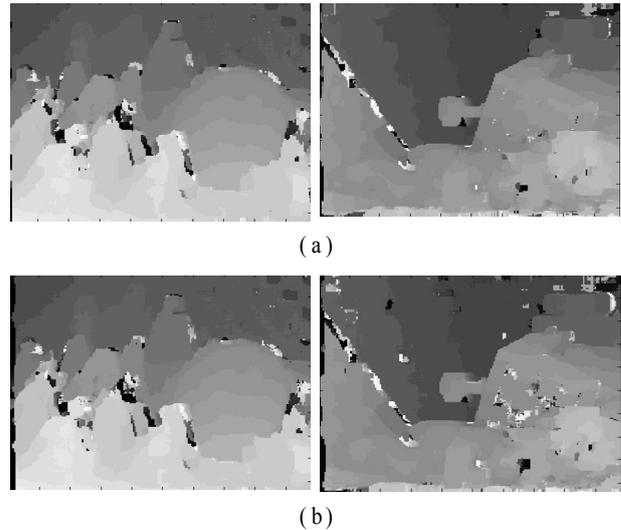


Figure 6. Disparity maps for 'Cones' and 'Teddy' stereo test image (a) unbalanced BiGHM and (b) balanced GHM multiwavelets.

# New Adaptation Method Using Two-dimensional PCA for Speaker Verification

Chunyan Liang, Xiang Zhang, Lin Yang, Li Lu, Yonghong Yan

ThinkIt Speech Lab, Institute of Acoustics, CAS

Beijing, P.R. China

Email: {chyliang, xzhang, lyang, llu, yonghong.yan}@hccl.ioa.ac.cn

**Abstract**—In this paper, a new adaptation method based on two-dimensional principal component analysis is introduced into speaker recognition. In the method, mixture and dimension of mean vectors based on the Gaussian Mixture Models (GMMs) are differentiated, and the covariance matrix is computed dimension-wisely. The experiments are carried out on the core conditions of NIST 2008 speaker recognition evaluation data. The experimental results indicate that the 2DPCA-based method can achieve comparable performance to the conventional eigenvoice approach. Besides, the fusion of the two different systems can make significant performance improvement compared to the eigenvoice system alone, achieving relative reduction on EER between 7% and 25% for different test conditions.

**Keywords**-speaker recognition; 2DPCA; eigenvoice; SVM

## I. INTRODUCTION

State-of-the-art speaker verification systems are based on statistical generative models such as Gaussian Mixture Models (GMMs). In this case, one needs to create a generative model for each client, as well as a generative model for a corresponding anti-client, often replaced by a universal background model (UBM) [1]. The support vector machines (SVMs) [2] have also proved to be effective for speaker recognition. A commonly used method for combining GMM and SVM is to concatenate GMM mean vectors as super-vectors for SVM design [3].

For speaker verification, the client model is often derived by adapting the parameters of the UBM using the speaker's training speech. Some adaptation methods have been proven to be successful [4], such as eigenvoice [5]. Eigenvoice speaker adaptation has been shown to be effective for speaker recognition in recent years. The eigenvoice approach involves three steps. First, an eigenspace is established with many speaker dependent (SD) models from training speakers via principal component analysis (PCA). Each of the SD models is represented as a column vector, with the mixture and dimension treated without distinction. Then a group of eigenvoice coefficients is determined for each testing speaker. Finally, we obtain the client models which are expressed as a linear combination of bases in the eigenspace.

In this study, we adopt the speaker adaptation method based on two-dimensional PCA (2DPCA). In 2DPCA, each training SD model is represented as a matrix (the mixture and dimension of mean vectors are represented in separate directions) rather than as a vector which is the case for

eigenvoice. Thus, more compact bases with lower dimension than those of eigenvoice can be obtained from 2DPCA, and the speaker adaptation formula using these bases can have a dimension-wise speaker weight. For speech recognition, the speaker adaptation method using 2DPCA has been shown to perform competitively [6]. In this paper, we introduce the new adaptation method into speaker recognition to update the GMM mean vectors of the client models and concatenate them as supervectors for SVM.

The remainder of this paper is organized as follows. In Section II, we give a brief overview of eigenvoice. Section III describes the 2DPCA method and the application of 2DPCA-based method in GMM framework for the task of speaker recognition. The details of the performed experiments and results are presented in Section IV. Finally, we conclude this paper in Section V.

## II. EIGENVOICE

The underlying hypothesis of eigenvoice adaptation is that all voices represented in a space of a large dimension could in fact be well represented in a low-dimensional linear subspace [5][7]. The most commonly used tool to select the low-dimensional subspace is the well-known principal component analysis (PCA) [8].

Given a set of  $T$  speaker dependent models (SD models) already adapted by Bayesian Maximum A Posteriori (MAP), PCA is used to compute the  $K$  leading eigenvectors of the covariance matrix of the  $T$  parameter vectors. The model of a new speaker  $c$  can then be represented as a linear combination of the  $K$  eigenvectors:

$$\mu(c) = Vx + \mu(ubm). \quad (1)$$

where  $\mu(ubm)$ , consisting of  $M \times D$  elements ( $M$  is the number of Gaussian components, and  $D$  represents the feature dimension for each Gaussian component), is the concatenated mean supervector of all the mixture component means of the UBM model, and  $\mu(c)$  is the adapted supervector of the new speaker  $c$ .  $V = [v_1, v_2, \dots, v_K]$  represents the eigenspace, and it is the concatenated matrix of the  $K$  eigenvectors with the  $K$  largest eigenvalues.  $x$  is the eigenvoice coefficients of client  $c$ .

### III. METHODS

#### A. Two-Dimensional Principal Component Analysis

In 2DPCA [6][9], the mapping from an input 'matrix' into the feature space is performed by

$$\omega = X \cdot \phi. \quad (2)$$

where  $X \in \mathbf{R}^{D \times N}$  is the input matrix,  $\phi$  is the base vector which is unitary ( $\phi^T \cdot \phi = 1$ ), and  $\omega$  is a D-dimensional feature vector (which is a scalar in PCA). A set of such bases,  $\{\phi_k\}_{k=1}^K$ , can be obtained by maximizing the following criterion:

$$J(\phi) = \text{tr}(E[(\omega - E[\omega])(\omega - E[\omega])^T]). \quad (3)$$

where  $\text{tr}(\cdot)$  denotes the trace of a matrix. The criterion can be expressed in terms of  $X$  and  $\phi$  as

$$J(\phi) = \phi^T \cdot \{E[(X - E(X))^T(X - E(X))]\} \cdot \phi. \quad (4)$$

The covariance matrix  $G$  is defined as

$$G = E[(X - E(X))^T(X - E(X))]. \quad (5)$$

Using a set of example matrices,  $\{X_s\}_{s=1}^S$ , the sample covariance matrix can be obtained by

$$G = \frac{1}{S} \sum_{s=1}^S (X_s - \bar{X})^T (X_s - \bar{X}). \quad (6)$$

where  $\bar{X} = 1/S \sum_{s=1}^S X_s$  is the average of the example matrices. Thus, the criterion becomes

$$J(\phi) = \phi^T \cdot G \cdot \phi. \quad (7)$$

A set of orthonormal projection vectors,  $\{\phi_k\}_{k=1}^K$ , can be obtained as the  $K$  leading eigenvectors of the covariance matrix  $G$  to maximize the above criterion.

The set of such bases project an input matrix  $X_s$  into the feature space by

$$W(s) = (X_s - \bar{X}) \cdot \Phi. \quad (8)$$

where  $\Phi = [\phi_1 \cdots \phi_k \cdots \phi_K]$  and  $W(s) = [\omega_1(s) \cdots \omega_k(s) \cdots \omega_K(s)]$ . Then, the low-rank approximations of the input matrix can be obtained as

$$X_s \approx \bar{X} + W(s) \cdot \Phi^T = \bar{X} + \sum_{k=1}^K \omega_k(s) \cdot \phi_k^T. \quad (9)$$

In (9),  $X_s$  can be exactly reconstructed when  $K=S$ .

#### B. Application of 2DPCA-based method in GMM framework for the task of speaker recognition

In this section, we will discuss the application of 2DPCA to speaker adaptation in the GMM framework for the task of speaker recognition. Here, let  $\mu_m(s) \in R^{D \times 1}$  be the mean vector of the  $m$ -th Gaussian component of speaker  $s$  from the total  $S$  training speaker models. Adapted from the UBM model using MAP adaptation, the SD mean model of speaker  $s$  is viewed as a matrix:

$$\mu(s) = [\mu_1(s) \cdots \mu_m(s) \cdots \mu_M(s)]. \quad (10)$$

where  $M$  is the number of the Gaussian components. In the expression above, column corresponds to mixture and row corresponds to the dimension of mean vectors.

We apply 2DPCA to the training examples  $\{\mu(s)\}_{s=1}^S$  as follows [6]. First, denoting  $\tilde{\mu}(s) = \mu(s) - \mu(\text{ubm})$ , we obtain the covariance matrix as:

$$G = \frac{1}{S} \sum_{s=1}^S \tilde{\mu}(s)^T \tilde{\mu}(s). \quad (11)$$

When denoting  $\tilde{\mu}_d(s) \in R^{1 \times M}$  as the  $d$ -th row vector of  $\tilde{\mu}(s)$ , the covariance matrix can be shown as

$$G = \frac{1}{S} \sum_{s=1}^S \sum_{d=1}^D \tilde{\mu}_d(s)^T \tilde{\mu}_d(s). \quad (12)$$

As such, 2DPCA is equivalent to line-based PCA, which partitions a matrix into lines and each line is treated as a sample data in standard PCA framework [10].

Then, the eigenvectors corresponding to the  $K$  largest eigenvalues ( $K \leq S - 1$ ) of  $G$  can be found as the bases  $\{\phi_k^{M \times 1}\}_{k=1}^K$ , corresponding to the bases  $\{\phi_k^{M \times D}\}_{k=1}^K$  in eigenvoice. Thus, 2DPCA produces a more compact set of eigenvectors by the factor of  $D$ . Using the bases, we update the model for a new speaker by

$$\mu(\text{new}) = \mu(\text{ubm}) + W_{\text{new}} \cdot \Phi^T. \quad (13)$$

where  $\Phi = [\phi_1 \cdots \phi_k \cdots \phi_K]$ , and  $W_{\text{new}}$  is the speaker weight which can be derived in a maximum-likelihood estimation (MLE) framework as follows [6]: given the observation data  $O = \{o_1 \cdots o_t \cdots o_T\}$ , the auxiliary function is defined as

$$Q(\lambda, \hat{\lambda}) = -\frac{1}{2} P(O|\lambda) \times \sum_{m=1}^M \sum_{t=1}^T \gamma_m(t) (D \cdot \log(2\pi) + \log|C_m| + h(o_t, m)). \quad (14)$$

where  $\lambda$  is the current model parameter and  $\hat{\lambda}$  is the re-estimated model parameter,  $\gamma_m(t)$  denotes the occupation probability of being in mixture  $m$  at time  $t$  given  $O$ , and  $C_m$  is the covariance matrix for the  $m$ -th Gaussian, which

is diagonal in our work. The last term in (14) contains the model parameter:

$$h(o_t, m) = (o_t - (\mu_m(ubm) + W_{new} \cdot \Phi_m^T))^T \cdot C_m^{-1} \cdot (o_t - (\mu_m(ubm) + W_{new} \cdot \Phi_m^T)). \quad (15)$$

We can derive the following equation to find the weight  $W_{new}$ :

$$\begin{aligned} & \sum_{m=1}^M \sum_{t=1}^T \gamma_m(t) C_m^{-1} \cdot (o_t - \mu_m(ubm)) \cdot \Phi_m \\ & = \sum_{m=1}^M \sum_{t=1}^T \gamma_m(t) C_m^{-1} \cdot W_{new} \cdot \Phi_m^T \cdot \Phi_m. \end{aligned} \quad (16)$$

The above equation can be solved for  $W_{new}$  using the same procedure in [11].

### C. Feature extraction and SVM modeling

After the speaker model are updated as (1) and (13), the parameters from  $\mu(s)$  are concatenated into a single supervector consisting of  $D \times M$  elements according to Kullback-liebler divergence [3] and modeled using SVMs, where  $M$  is the number of Gaussians in UBM model and  $D$  is the dimension of mean vectors in each Gaussian component.

An SVM is trained for each target speaker by regarding the target speaker's training supervector as positive examples, and the supervectors from a background training set as negative examples. Our experiments are implemented using the SVMLight with a linear inner-product kernel function.

## IV. EXPERIMENT

### A. Experiment setup

The experiments for different systems based on the two kinds of speaker adaptation methods (eigenvoice and 2DPCA) are carried out on the NIST 2008 speaker recognition evaluation corpus. The NIST SRE2008 evaluation tasks are distinguished by including in the training and test conditions not only conversational telephone speech but also interview speech recorded with different microphones involving an interview scenario. We carry out the experiments on three types of trials: telephone-telephone, interview-interview and interview-telephone. The performance is measured in terms of equal error rate (EER) and DET curves [12].

The input speech utterance is first converted to a sequence of 36-dimensional feature vectors including 18 MFCC coefficients and their first order derivatives over 5 frames. To reduce channel effects, feature warping to a Gaussian distribution, CMN, CVN are performed to the feature vectors.

The gender dependent UBM models with 1024 mixture components are trained using the NIST SRE 2004 1side training corpus. The background data for SVM system are selected from the data form NIST SRE2004 and NIST SRE2005. Eigenvectors for both eigenvoice and the 2DPCA

Table I  
SVM SYSTEMS BASED ON DIFFERENT ADAPTATION METHODS ACROSS ALL MALE SPEAKERS IN THE TEST CORPUS. THE VALUE IN EACH TABLE CELL IS THE EER (%).

task	eigenvoice	2DPCA	fusion
telephone-telephone	5.62	5.54	4.98
interview-interview	2.38	2.97	2.04
interview-telephone	4.74	5.21	3.59

are also gender dependent. 600 eigenvoices and 600 eigenvectors for 2DPCA for both male and female are trained using the Switchboard II, Switchboard Cellular corpus as well as the data from NIST SRE2005 and NIST SRE2006. For eigenchannel compensation in feature domain, telephone and microphone data from NIST SRE2004, NIST SRE2005 and NIST SRE2006 are used.

The raw score are speaker-normalized by means of gender-dependent ZTnorm. For Znorm and Tnorm, telephone and interview utterances are drawn from the NIST SRE2006 corpus.

We use the linear fusion for the two systems, with the weight of 0.5 for each system.

### B. Experiment results

In this subsection, we list the results of the systems based on both eigenvoice and the 2DPCA-based method as well as the fusion of them on the three test conditions in NIST SRE 2008. The DET curves are also given below.

Table I lists the performance of the SVM system based on eigenvoice and 2DPCA-based method on the three trial conditions across all male speakers. From Table I, we can see that the system based on 2DPCA can achieve comparable performance to the conventional eigenvoice system for male speakers. As well, the fusion of the two systems makes significant performance improvement compared to the eigenvoice system alone, yielding 11.4% improvement on EER for the telephone-telephone condition, 14.3% for the interview-interview condition and 24.3% for the interview-telephone condition.

Table II shows the results of the SVM system based on the two different adaptation methods across all female speakers. It can also be seen that the performance of the 2DPCA-based system is comparable to the eigenvoice system. Compared to the single system based on eigenvoice, the fusion of the two systems achieve relative reduction of 7.14% on EER for the telephone-telephone condition, 11.9% for the interview-interview condition and 7.3% for interview-telephone.

Figure 1 and Figure 2 show the DET curves of the systems based on different speaker adaptation methods for male and female speakers respectively.

Table III summarizes the approximate average training time per file. The training time mainly consists of two parts, the time of estimating the parameters of feature vector and the time of SVM training. As we can see, our method uses

Table II  
SVM SYSTEMS BASED ON DIFFERENT ADAPTATION METHODS ACROSS ALL FEMALE SPEAKERS IN THE TEST CORPUS. THE VALUE IN EACH TABLE CELL IS THE EER (%) .

task	eigenvoice	2DPCA	fusion
telephone-telephone	7.14	8.11	6.63
interview-interview	4.53	6.22	3.99
interview-telephone	6.46	11.12	5.99

Table III  
AVERAGE TRAINING TIME PER FILE FOR EIGENVOICE AND 2DPCA.

Systems	Training time cost(sec)
eigenvoice	2.88
2DPCA	7.86

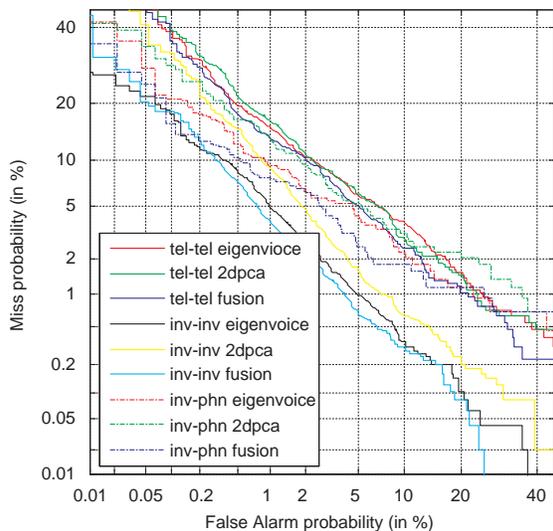


Figure 1. DET curves comparing systems based on eigenvoice and 2DPCA as well as the fusion system for male speakers

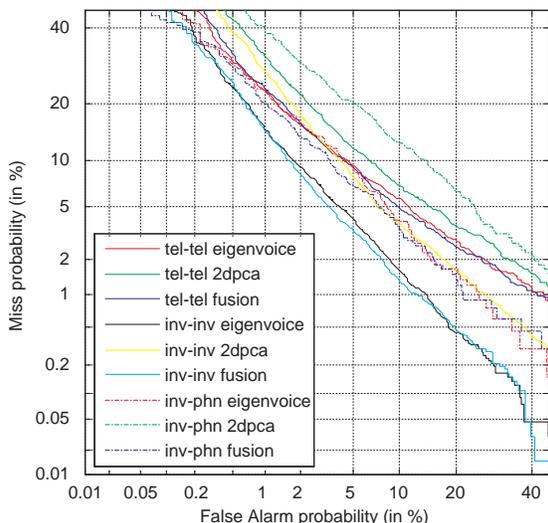


Figure 2. DET curves comparing systems based on the eigenvoice and 2DPCA as well as the fusion system for female speakers

more time cost than the eigenvoice system, which implies

its limits in real time work.

### V. CONCLUSION

In this paper, we have introduced the new adaptation method using 2DPCA into speaker recognition. The 2DPCA of training models produces the more compact bases whose dimension is lower than that of eigenvoice, and the speaker weight consists of dimensional elements. Experiments show that the system based on 2DPCA can achieve comparable performance to the conventional eigenvoice system and the fusion of the two systems can further improve the performance, yielding 7%-25% improvement on EER for different tasks, which indicates that the 2DPCA-based method and eigenvoice are complementary to each other to some extent when used in speaker recognition. Future work include generalizing this approach to other PCA-based modeling methods such as eigenspace-based MLLR [13].

### ACKNOWLEDGEMENT

This work is partially supported by The National Science and Technology Pillar Program (2008BAI50B03), National Natural Science Foundation of China (No. 10925419, 90920302, 10874203, 60875014).

### REFERENCES

- [1] D. A. Reynolds, T. F. Quatieri and R. B. Dunn, "Speaker verification using adapted Gaussian mixture models", *Digital Signal Processing*, vol. 10, pp. 19-41, 2000.
- [2] N. Cristianini and J. Shawe-Taylor, "Support vector machines", Cambridge University Press, Cambridge, UK, 2000.
- [3] W. Campbell, D. Sturim, D. Reynolds, and A. Solomonoff, "SVM based speaker verification using a GMM supervector kernel and NAP variability compensation", *IEEE ICASSP 2006*, vol. 1, 2006.
- [4] J. Mariethoz and S. Bengio, "A comparative study of adaptation methods for speaker verification", In *Seventh International Conference on Spoken Language Processing*. 2002.
- [5] B. Mak, J. Kwok, and S. Ho, "Kernel eigenvoice speaker adaptation", *IEEE Transactions on Speech And Audio Processing*, 13(5):984, 2005.
- [6] Y. Jeong and H.S. Kim, "New speaker adaptation method using 2-D PCA", *IEEE Signal Process. Lett.*, vol. 17, no. 2, pp. 193-196, 2010.
- [7] H. Wang, Q. Zhao and Y. Yan, "Using Eigenvoice Coefficients as Features in Speaker Recognition", *2009 International Conference on Electronic Computer Technology*, pp. 262-266. 2009.
- [8] I. Jolliffe, "Principal component analysis", Springer New York, 2002.

- [9] J. Yang, D. Zhang, A.F. Frangi, and J.Y. Yang, "Two-dimensional PCA: A new approach to appearance-based face representation and recognition", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 26, no. 1, pp. 131-137, Jan. 2004
- [10] L. Wang, X. Wang, X. Zhang, and J. Feng, "The equivalence of two-dimensional PCA to line-based PCA", *Pattern Recognition. Lett.*, vol. 26, no. 1, pp. 57-60, Jan. 2005.
- [11] C.J. Leggetter and P.C. Woodland, "Maximum likelihood linear regression for speaker adaptation of continuous density hidden Markov models", *Comput. Speech Lang.*, vol. 9, no. 2, pp. 171-185, Apr. 1995.
- [12] "The NIST Year 2008 Speaker Recognition Evaluation Plan", <http://www.nist.gov/speech/tests/spk/2008/index.html>, 2008.
- [13] K.T. Chen, W.W. Liau, H.M. Wang, and L.S. Lee, "Fast speaker adaptation using eigenspace-based maximum likelihood linear regression", *ICSLP 2000*, vol. 3, pp. 742-745, Oct. 2000.

## Language Recognition With Locality Preserving Projection

Jinchao Yang, Xiang Zhang, Li Lu, Jianping Zhang, Yonghong Yan

ThinkIT Speech Lab, Institute of Acoustics  
Chinese Academy of Sciences  
Beijing, P.R.China

{yangjinchao, xzhang, llu, jzhang, yonghong.yan}@hccl.ioa.ac.cn

**Abstract** - In this paper, we introduce locality preserving projection (LPP) to language recognition under the support vector machine (SVM) framework. The success of the use of total variability in language recognition shows that the global structure and linear manifold preserve discriminative language dependent information. The proposed LPP language recognition system believes the local structure and nonlinear manifold also contain discriminative language dependent information. Experiment results on 2007 National Institute of Standards and Technology (NIST) language Recognition Evaluation (LRE) databases show LPP language recognition system combining total variability language recognition system gains relative improvement in EER of 11.7% and in minDCF of 9.6% comparing to total variability language recognition system in 30-second tasks, and further improvement is obtained combining with state-of-the-art systems. It leads to gains of 13.8% in EER and 20.2% in minDCF compare with the performance of the combination of the MMI and the GMM-SVM systems.

**Index Terms**— language recognition, language total variability, PCA, LDA, LPP, SVM,

### I. INTRODUCTION

The aim of language recognition is to determine the language spoken in a given segment of speech. Phoneme recognizer followed by language models (PRLM) and parallel PRLM (PPRLM) approaches that use phonotactic information have shown very successful performance [1][2]. In PPRLM, several tokenizers are used to transcribe the input speech into phoneme strings or lattices [3][4], which are scored by n-gram language models. It is generally believed that phonotactic feature and spectral feature provide complementary cues to each other [1]. The spectral features of speech are collected as independent vectors. The collection of vectors can be extracted as shifted-delta-cepstral acoustic features, and then modeled by Gaussian Mixture Model (GMM). The result was reported in [5]. The approach was further improved by using discriminative train that named Maximum Mutual Information

(MMI).

Several studies using SVM in language recognition to form GMM-SVM system [6][7]. SVM as a classifier maps the input feature vector into high dimensional space then separate classes with maximum margin hyperplane. It is important to choose an appropriate SVM feature expansion.

Recently total variability approach has been proposed in speaker recognition [8][9], which uses the factor analysis to define a new low-dimensional space that named total variability space. In this new space, the speaker and the channel variability are contained simultaneously. In our previous work, we introduce the idea of total variability to language recognition and propose total variability language recognition system. The success of the use of total variability in language recognition show that most of the discriminative language dependent information is captured by low-dimensional subspace.

Actually, total variability method is a classical application of the probabilistic principal component analysis (PPCA) [10]. In our previous work about language recognition with language total variability, we can say that PCA+LDA is used to reduce the dimension of GMM supervector before SVM model. Locality preserving projection (LPP) [11][12] that gains an embedding that preserves local and linear information is different from PCA and LDA which effectively preserve global structure and linear manifold.

In this paper, LPP algorithm is carried out after PCA to a conversation to get the supervector that contain discriminative language dependent information by the local structure and nonlinear manifold. We can call laplacian supervector extraction method.

SVM classifiers are employed to model the laplacian supervector and LDA and diagonal covariance gaussians are used as backend in Language Score Calibration.

This paper is organized as follows: In Section 2, we give a simple review of Support Vector Machines and total variability language recognition system. Section 3 shows the laplacian algorithmic procedure. In Section 4, the proposed language recognition system is presented in detail. corpora and evaluation are given in Section 5. Section 6 gives out experi-

mental result. Finally, we conclude in Section 7.

## II. BACKGROUND

### A. Support Vector Machines

An SVM [13] is a two-class classifier constructed from sums of a kernel function  $K(\cdot, \cdot)$ :

$$f(x) = \sum_{i=1}^N \alpha_i t_i K(\mathbf{x}, \mathbf{x}_i) + d \quad (1)$$

where  $N$  is the number of support vectors,  $t_i$  is the ideal output,  $\alpha_i$  is the weight for the support vector  $x_i$ ,  $\alpha_i > 0$  and  $\sum_{i=1}^N \alpha_i t_i = 0$ . The ideal outputs are either 1 or -1, depending upon whether the corresponding support vector belongs to class 0 or class 1. For classification, a class decision is based upon whether the value,  $f(x)$ , is above or below a threshold.

The kernel  $K(\cdot, \cdot)$  is constrained to have certain properties (the Mercer condition), so that  $K(\cdot, \cdot)$  can be expressed as

$$K(\mathbf{x}, \mathbf{y}) = \phi(\mathbf{x})' \phi(\mathbf{y}) \quad (2)$$

where  $\phi(x)$  is a mapping from the input space (where  $\mathbf{x}$  lives) to a possibly infinite dimensional SVM expansion space. We refer to the  $\phi(x)$  as the SVM features.

### B. Language Total Variability

In total variability speaker recognition, the factor analysis is used to define a new low-dimensional space that named total variability space and contains the speaker and the channel variability simultaneously. Then, the intersession compensation can be carried out in low-dimensional space. We define language total variability space.

#### a. Language Total Variability Space Estimation

There is only one difference between total variability space  $T$  estimation and eigenvoice space estimation in speaker recognition [9]. All the recordings of speaker are considered to belong to the same person in eigenvoice estimation, however, in total variability space estimation, a given speaker's entire set of utterances are regarded as having been produced by different speakers. We suppose that different conversation from one language is produced by different languages.

For a given conversation, the language and variability dependent supervector is denoted in equation (4).

$$M = m_{ubm} + Tw \quad (3)$$

where  $m_{ubm}$  is the UBM supervector,  $T$  is language total variability space, and the member of the vector  $w$  are total factor. We could call  $w$  language total factor vector. We can think the language total factor vector model a new feature extractor that project a conversation to a low rank space  $T$  to get a language and variability dependent language total factor vector  $w$ .

#### b. Intersession Compensation

After the feature extractor, the intersession compensation can be carried out in low-dimensional space. We use the Linear Discriminant Analysis approach (LDA) to intersession compensation. All the language total factor vector of the same language are think as the same class.

$$w^* = Aw \quad (4)$$

By LDA transformation in equation (4), the language total factor vector  $w$  is projected to new axes that maximize the variance between language and minimizing the intra-class variance. The matrix  $A$  is trained using the dataset show in session 5 and is contained of the more larger eigenvectors of equation (5).

$$S_b \nu = \lambda S_w \nu \quad (5)$$

where  $\lambda$  is the diagonal matrix of eigenvalues. The matrix  $S_b$  is the between class covariance matrix and  $S_w$  is the within class covariance matrix.

## III. LAPLACIAN ALGORITHM PROCEDURE WITH LOCALITY PRESERVING PROJECTION

Since total variability method is a classical application of the probabilistic principal component analysis (PPCA). In our previous work about language recognition with language total variability, we can say that PCA+LDA is used to reduce the dimension of GMM supervector before SVM model. As a type of PCA, the total variability method does not need language information. And PCA seeks directions that are efficient for representation. LDA seeks directions that are efficient for discrimination. PCA+LDA that aims to preserve the global structure and linear manifold is successful for language recognition, then the local structure and nonlinear manifolds may be useful to language recognition.

Though LPP is still a linear technique, it seems to reserve important aspects of the intrinsic nonlinear manifold structure by preserving local structure. The algorithmic procedure in this paper is formally stated below.

### A. PCA projection

In session 2.2, a conversation is projected to language total variability space to get a language and variability dependent language total factor vector  $w$ . Actually, it is a PCA projection. In this paper, LPP is used after PCA projection.

### B. Constructing the nearest-neighbor graph

Let  $G$  denote a graph with  $m$  nodes. We put an edge between nodes  $i$  and  $j$  while  $i$  is among  $k$  nearest neighbors of  $j$  or  $j$  is among  $k$  nearest neighbors of  $i$ .

### C. choosing the weights

If nodes  $i$  and  $j$  are connected, let

$$S_{ij} = e^{-\frac{(w_i - w_j)^2}{\tau}} \quad (6)$$

The justification for this choice of weights can be traced back to [14].

### D. eigenmap

Compute the eigenvectors and eigenvalues for generalized eigenvector problem:

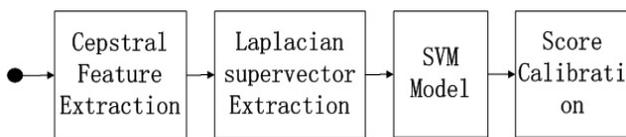
$$WLW^T a = \lambda WDW^T a \quad (7)$$

where  $D$  is a diagonal matrix whose entries are column sums of  $S$ ,  $D_{ij} = \sum_j S_{ji}$ .  $L = D - S$  is the Laplacian matrix. The  $i$ th row of matrix  $W$  is  $w_i$ . Let  $a_0, a_1, \dots, a_{l-1}$  be the solution of (7), ordered according to their eigenvalues,  $0 \leq \lambda_0 \leq \lambda_1 \leq \dots \leq \lambda_{l-1}$ . Thus, the embedding is as follows:

$$w_i \longrightarrow y_i = A^T w_i, A = (a_0, a_1, \dots, a_{l-1}) \quad (8)$$

where  $y$  is a  $l$  dimensional vector, and  $A$  is a transformation matrix.

## IV. THE PROPOSED LANGUAGE RECOGNITION SYSTEM



Input Utterance

**Fig. 1.** The Proposed Language recognition System

Figure 1 show the frame of the proposed Language recognition system.

### A. MSDC Feature Extraction

The MSDC feature in the system is 7 MFCC coefficient concatenated with SDC 7-1-3-7 feature, which are in total 56 dimension coefficients each frame. MSDC feature refers to this 56 demension feature in my system. Nonspeech frames are eliminated after speech activity detection, then 56 dimension MSDC feature are Extraction. Then feature warping and cepstral variance normalization are applied on the previously extracted MSDC feature which results that each feature is normalized to mean 0 and variance 1.

### B. Laplacian supervector Extraction

In our system, we use MSDC feature after compensation of channel factors. Firstly, total variability spaces are estimated as session 2.2. MSDC feature, UBM and Language-independent Total variability space  $T$  are need as equation (3) in language total factor vector extraction (actually it is a PCA Projection). Then LPP transformation matrix is learned as session 3. The embedding is as follows to each GMM supervector  $x$ :

$$x \longrightarrow y = A^T x \quad (9)$$

$$A = A_{PCA} A_{LPP} \quad (10)$$

where  $A_{PCA}$  denote the transformation matrix of PCA as session 2.2. And  $A_{LPP}$  denote the transformation matrix of LPP, while the algorithmic procedure is in session 3. We call  $A$  Laplacian transformation matrix.

### C. SVM Model and Language Score Calibration

Our experiments are implemented using the SVMTool [15] with a linear inner-product kernel function.

Calibrating confidence scores in mutiple-hypothesis language recognition has been studied in [16]. We should estimate the posterior probability of each hypotheses and make a maximum a posterior decision. In standard SVM-SDC system [7], log-likelihood ratios (LLR) normalization is applied as a simple backend process and is useful. Suppose  $S = [S_1 \dots S_L]^t$  is the vector of  $L$  relative log-likelihoods from the  $L$  target languages for a particular message. Considering a flat prior, a new log-likelihood normalized score  $S'_i$  is denoted as:

$$S'_i = S_i - \log\left(\frac{1}{L-1} \sum_{j \neq i} e^{S_j}\right) \quad (11)$$

A more complex full backend process is given [7] [17], LDA and diagonal covariance gaussians are used to calculate the log-likelihoods for each target language and achieve improvement in detection performance.

In this paper, the two backend processes are used in language recognition system. Experiments also show the similar conclusion that the LDA and diagonal covariance gaussians backend process is superior over log-likelihood ratios normalization.

## V. CORPORA AND EVALUATION

The experiments are done using the NIST LRE 2007 evaluation database. There are 14 target languages in corpora used in this paper: Arabic, Bengali, Chinese, English, Farsi, German, Hindustani, Japanese, Korean, Russian, Spanish, Tamil, That and Vietnamese. The task of this evaluation was to detect the presence of a hypothesized target language for each test utterance. The Training data was primarily from Callfriend corpora, Callhome corpora, mixer corpora, OHSU corpora, OGI corpora and LRE07Train. The development data

consist of LRE03, LRE05, LRE07Train. We use equal error rate (EER) and the minimum decision cost value (minDCF) as metrics for evaluation.

## VI. EXPERIMENTS

Firstly, total variability Language recognition System (PCA+LDA) is experimented, then export to Laplacian Language recognition System (PCA+LPP).

**Table 1.** Results of the MMI system, GMM-SVM system, the total variability and proposed Laplacian language recognition systems on the NIST LRE07 30s corpus.

System	EER	MinDCF
MMI (a)	3.62	3.78
GMM-SVM (b)	2.65	2.61
total variability (c)	3.15	2.61
Laplacian (d)	3.29	2.83

In Table 1, we give the performance of the MMI, the GMM-SVM, the total variability and proposed Laplacian language recognition systems on NIST 2007 language recognition Evaluation 30s corpus after score backend. EER and minDCF are observed. With the performance comparison, we can see that total variability and proposed Laplacian language recognition systems achieve performance comparable to that obtained with state-of-the-art approaches, which shows my proposed systems are effective and Laplacian language recognition system indeed contains language information.

**Table 2.** Score Fusion and Super Join Results of the total variability and proposed Laplacian language recognition systems on the NIST LRE07 30s corpus.

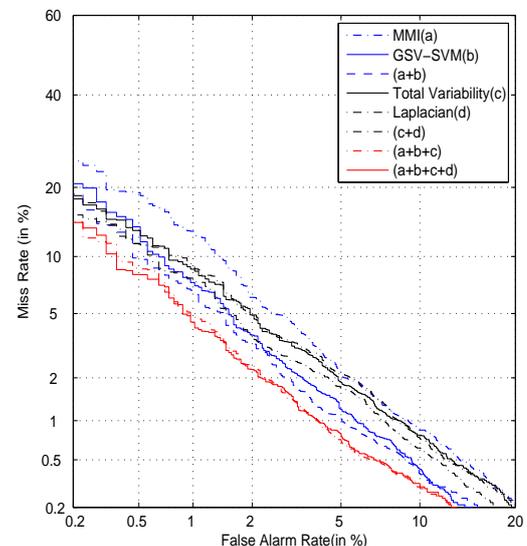
System	EER	MinDCF
total variability (c)	3.15	2.61
Laplacian (d)	3.29	2.83
c+d Score Fusion	2.78	2.36
c+d Super Join	2.87	2.51

Table 2 shows the score fusion and super join results of the total variability and proposed Laplacian language recognition systems. The score fusion leads to gains of 11.7% on EER and 9.6% minDCF compare with the performance of the total variability language recognition systems, And super join gains 8.9% on EER and 3.8% minDCF. The result show Laplacian language recognition system that preserves local and nonlinear information includes different language information comparing to total variability language recognition system that preserves global and linear information.

**Table 3.** Results of the combination of MMI system and GMM-SVM system, and the combination of the MMI system, GMM-SVM system, total variability system, and Laplacian system on the NIST LRE07 30s corpus.

System	EER	MinDCF
Score Fusion (a+b)	2.47	2.42
Score Fusion(a+b+c)	2.18	2.06
Score Fusion(a+b+c+d)	2.13	1.93

Table 3 shows the results of the combination of the MMI system, the GMM-SVM system, the total variability system and the Laplacian system. EER and minDCF are observed. In language recognition evaluation, MMI and GMM-SVM are primary acoustic system. Usually the combination of the MMI system and the GMM-SVM system is the given performance of acoustic system. It leads to gains of 13.8% on EER and 20.2% minDCF compare with the performance of the combination of the MMI and GMM-SVM systems. Lastly, figure 1 gives DET curves for each system and fusion of each system.



**Fig. 2.** DET curves for each system and fusion of each system

## VII. CONCLUSIONS

In this paper, we propose a new language recognition system by introducing LPP to language recognition. while our previous propose total variability language recognition system show discriminative language dependent information is contained by global structure and linear manifold, the new

language features of Laplacian supervector that preserve local structure and nonlinear manifolds also contain discriminative language dependent information. SVM classifiers are employed to model the new language features and LDA and diagonal covariance gaussians are used as backend in Language Score Calibration. Experiments show that combining two systems LPP and total variability can achieve relative improvement in EER of 11.7% and in minDCF of 9.6% compare to only total variability in 2007 NIST language Recognition Evaluation databases 30-second tasks. Further improvement of relative improvement of 13.8% in EER and 20.2% in minDCF is obtained combining with state-of-the-art systems, comparable with the performance of the combination of the MMI and GMM-SVM systems.

## Acknowledgments

This work is partially supported by The National Science and Technology Pillar Program (2008BAI50B00), National Natural Science Foundation of China (No. 10925419, 90920302, 10874203, 60875014).

## References

- [1] M.A. Zissman, "Language identification using phoneme recognition and phonotactic language modeling," in *IEEE International Conference On Acoustics Speech And Signal Processing*. Institute Of Electrical engineers INC (IEE), 1995, vol. 5, pp. 3503–3503.
- [2] Y. Yan and E. Barnard, "An approach to automatic language identification based on language-dependent phone recognition," in *icassp*. IEEE, 1995, pp. 3511–3514.
- [3] J.L. Gauvain, A. Messaoudi, and H. Schwenk, "Language recognition using phone lattices," in *Eighth International Conference on Spoken Language Processing*. ISCA, 2004.
- [4] W. Shen, W. Campbell, T. Gleason, D. Reynolds, and E. Singer, "Experiments with lattice-based PPRLM language identification," in *IEEE Odyssey 2006: The Speaker and Language Recognition Workshop, 2006*, 2006, pp. 1–6.
- [5] P.A. Torres-Carrasquillo, E. Singer, M.A. Kohler, R.J. Greene, D.A. Reynolds, and J.R. Deller Jr, "Approaches to language identification using Gaussian mixture models and shifted delta cepstral features," in *Seventh International Conference on Spoken Language Processing*. Citeseer, 2002.
- [6] H. Li, B. Ma, and C.H. Lee, "A vector space modeling approach to spoken language identification," *IEEE Transactions on Audio, Speech, and Language Processing*, vol. 15, no. 1, pp. 271–284, 2007.
- [7] W.M. Campbell, J.P. Campbell, D.A. Reynolds, E. Singer, and P.A. Torres-Carrasquillo, "Support vector machines for speaker and language recognition," *Computer Speech & Language*, vol. 20, no. 2-3, pp. 210–229, 2006.
- [8] N. Dehak, R. Dehak, P. Kenny, N. Brummer, P. Ouellet, and P. Dumouchel, "Support vector machines versus fast scoring in the low-dimensional total variability space for speaker verification," in *Proc. International Conferences on Spoken Language Processing (ICSLP)*, 2009, pp. 1559–1562.
- [9] N. Dehak, P. Kenny, R. Dehak, P. Dumouchel, and P. Ouellet, "Front-end factor analysis for speaker verification," *submitted to IEEE Transaction on Audio, Speech and Language Processing*.
- [10] M.E. Tipping and C.M. Bishop, "Probabilistic principal component analysis," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 61, no. 3, pp. 611–622, 1999.
- [11] X. He and P. Niyogi, *Locality preserving projections*, Citeseer, 2005.
- [12] X. He, S. Yan, Y. Hu, P. Niyogi, and H.J. Zhang, "Face recognition using laplacianfaces," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pp. 328–340, 2005.
- [13] N. Cristianini and J. Shawe-Taylor, "Support Vector Machines," *Cambridge University Press, Cambridge, UK*, 2000.
- [14] M. Belkin and P. Niyogi, "Laplacian eigenmaps and spectral techniques for embedding and clustering," *Advances in neural information processing systems*, vol. 1, pp. 585–592, 2002.
- [15] R. Collobert and S. Bengio, "SVMtorch: support vector machines for large-scale regression problems," *The Journal of Machine Learning Research*, vol. 1, pp. 143–160, 2001.
- [16] N. Brummer and D.A. van Leeuwen, "On calibration of language recognition scores," in *IEEE Odyssey 2006: The Speaker and Language Recognition Workshop, 2006*, 2006, pp. 1–8.
- [17] E. Singer, P.A. Torres-Carrasquillo, T.P. Gleason, W.M. Campbell, and D.A. Reynolds, "Acoustic, phonetic, and discriminative approaches to automatic language identification," in *Eighth European Conference on Speech Communication and Technology*, 2003.

# Speech Recognition and Text-to-speech Solution for Vernacular Languages

## Free software and community involvement to develop voice services

**James K. Tamgno**

ESMT Dakar, Sénégal.  
james.tamgno@esmt.sn

**Claude Lishou**

UCAD Dakar, Sénégal  
claudelishou@ucad.edu.sn

**Aristide T. Mendo'o**

ESMT Dakar, Sénégal  
mendoo.aristide@gmail.com

**Morgan Richomme**

Orange Labs Lanion, France  
morgan.richomme@orange-ftgroup.com

**Seraphin D. Oyono Obono**

DUT Durban, South Africa  
eyonoobonosd@dut.ac.za

**Pascal U. Elingui**

ESMT Dakar, Sénégal  
elinguiuriel@gmail.com

**Abstract** — This paper summarizes the work performed to study and develop a model Automatic Speech Recognition (ASR) system and a speech synthesis or Text-To-Speech (TTS) system on keywords of the vernacular language Wolof, respectively based on the open source software toolkits Julius and Festival. Much research has been developed in this area. Our goal is to be the first to develop a model for speech recognition and synthesis in Wolof, and also to create different lexicons and knowledge bases of phonetic, acoustic and linguistic features in order to introduce other languages.

**Keywords** – *Speech Recognition; Speech Synthesis; Wolof.*

### I. INTRODUCTION

Speech technologies – such as automatic speech recognition (ASR) and text-to-speech (TTS) systems – can play a significant role in bridging the “Digital Divide”, which is currently preventing the vast majority of developing-world citizens from participating in the Information Age [1]. Most importantly, these technologies can lower the level of sophistication required to access information services, and thereby contribute towards the establishment of a fully inclusive information society. By circumventing language barriers and lessening the impact of illiteracy or disability, these technologies address real needs. Also, given the central role of language in cultural matters, speech technologies can play a significant role in guiding a diverse set of cultures towards the use of Information Technology.

The availability of these technologies creates new opportunities, but in order to realize the full potential of mobile ICT services, important challenges and obstacles must be overcome [2]. Most importantly, speech technologies need to be tailored to the properties of each new language in which they are to be used.

Thus, conscious of the dominant status of the Wolof language in Senegal, and in order to address services in local languages for the customers of its African subsidiary companies, France Telecom through Orange Labs signed a teaching partnership contract with the ESMT. Through this partnership, a research project was initiated, relating to the «Study and integration of a free software based speech recognition and text-to-speech solution for vernacular languages, within the framework of a simple grammar».

The current paper describes how this project was completed; it is organized as follows: Section II reviews some relevant theory and describes the methodology and technologies we used to analyze and to create new speech recognition in Wolof based on the Julius toolkit. Section III presents some theory on speech synthesis and describes the methodology and technologies we used to analyze and to create new speech synthesis in Wolof using Festival, within the framework of a simple grammar. We also describe how the development and the deployment were made. Section IV concludes the paper and outlines our future work.

#### A. Motivations

The United Nations Conference on Trade and Development (UNCTAD) Information Economy Report 2009 presents Africa as the fastest growing mobile market in the world (UNCTAD 2010). The number of mobile subscribers worldwide are expected to grow to 5.5 billion by the end of 2013 and 70% of them will be in developing countries. Mobile phones are contributing to unprecedented social and economical development on the continent with pioneering initiatives led by international agencies, Non-Governmental Organizations (NGO) and the private sector in agriculture, health, education, banking, citizen media, disaster and humanitarian relief. There is a widespread agreement that Information and Communications Technology (ICT) services, especially mobile ones, have the potential to play a major role in furthering social and rural development in developing economies such as Africa.

#### B. Vernacular Language Specificities in Senegal

According to National African Language Resource Center (NALRC), the local African language Wolof is understood and spoken by about 90% of Senegalese, while the official language is French. Wolof is a typical African oral language. It means that no formal grammar has been defined. Very few dictionaries [17] have been produced and pronunciation of the same grapheme may be very different thus a standard linguistic approach cannot be used here. Moreover, the

vocabulary is poor and lots of words are directly imported from other languages such as English, French or Arabic. Illiteracy is counted among the problems that limit the development and progress of ICT in some undeveloped countries. By not writing, reading or understanding certain languages may hinder access to ICTs by some African populations who only communicate in their local language. With an aim of promoting the African languages and dialects, various initiatives get busy to return the contents and the accessible software in African languages. By not using this language in the development of mobile applications and ICT simply denies access to hundreds of thousands of people to certain telecommunications services and jobs [1].

The weight of vernacular languages in African societies is very important; some countries deal with hundreds of local languages. One of the main problem with these languages or dialects is that, most of the time, there are no written languages (no formal grammar, limited number of dictionaries, few linguists) and have to deal with lots of import from other languages (French, English, Dutch, Portuguese,..). Therefore the procedures for speech recognition and text synthesis have to be adapted and the engineering work must be completed with a complex linguistic work. Additionally industrial actors of the sector have the technology to provide speech recognition and synthesis in any language, but this is not economically profitable. The cost of this technology is too high regarding the potential revenues that are too low. The choice of free software as sustainable and open technology was a possible answer to take into account the specificities of vernacular languages.

C. Approach

In this study, according to the architecture (Figure 10), core issue for investigation may include:

1. Collecting information on theoretical models of platform ASR/TTS and system providing similar services.
2. Analysis of various aspects of the modeling language :
  - o Captures words (feature analysis)
  - o Converts digital signal of voice into phonemes
  - o Attempts to Recognize grapheme/phoneme
  - o Finds match in acoustic model database
3. Attempts to make sense of what we are saying
4. Find a Typical architecture of a conversational agent

II. SPEECH RECOGNITION CONCEPTION

An utterance leading to interact with machines is the main idea behind speech recognition. Whatever language one may speaks, it has been progressively made possible. Looking the way to do so, researches have felt on recognition systems using Hidden Markov Model (HMM). The approach based on HMM represents the current art of the state in open-source field. A new approach has emerged to deal with the challenging problem of conversational speech recognition [3][4].

This part concerning Speech recognition describes theoretical aspect, presents new languages creation

methodology, and leads us to a Julius ASR development for Wolof.

A. Review of theory

The matter of the problem is to get one's utterance transcripts into text format. The goal can be for a given acoustic observation  $X=X_1, X_2, \dots, X_n$ , find a the corresponding sequence of words  $\hat{W}=w_1, w_2, \dots, w_m$ , with de maximum a posteriori probability  $P(W/X)$ . Using Bayes decision rules, this can be expressed by [5]:

$$\hat{w} = \underset{w}{\operatorname{argmax}} P(X/Y)P(W) \quad (1)$$

Since the acoustic observation X is fixed, (1) is equal to  $P(X/W)$  is the probability to observing acoustic observation X given a specific word sequence W.  $P(W/X)$  is determined by an acoustic model.  $P(W)$  is the probability of observing W independent of the acoustic observation. It is referred to as a language model.

B. Methodology to create a new language recognized by Julius

B.1. Inside Julius: System Architecture

The language model consists of a word pronunciation dictionary and a syntactic constraint. Various types of language model are supported: word N-gram model (with arbitrary N), rule-based grammars and a simple word list for isolated word recognition.

$$P(w_1/w_2^N) = \frac{P(w_1, w_2, \dots, w_N)}{P(w_2, \dots, w_N)} = \frac{\prod_{i=1}^N P(w_i/w_1^{i-1})}{\prod_{i=2}^N P(w_i/w_2^{i-1})}$$

N-gram Bayes rule

Given a language and an acoustic model, Julius performs speech recognition task, whether embed in applications or in server-client architecture.

An overview of Julius base system is presented in Figure 1. Acoustic models should be HMM defined for sub-word units.

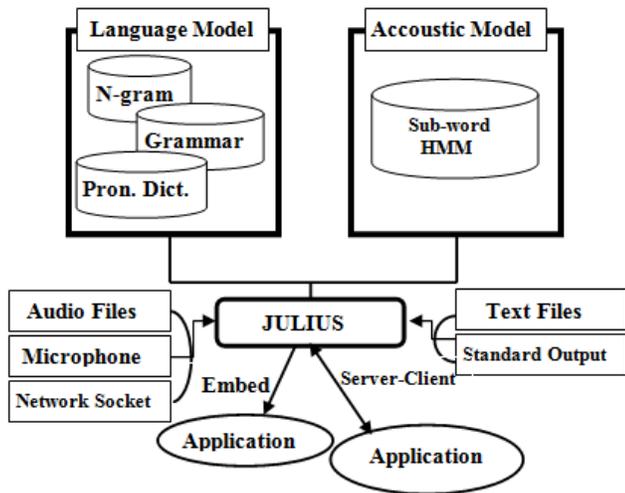


Figure 1. Julius overview [7].

The structure of Julius is illustrated in Figure 2. The top-level structure is the “engine instance”; it contains all the modules required for a recognition system: audio input, voice detection, feature extraction, Language Model (LM), Acoustic Model (AM) and search process.

An “AM process instance”, holds an acoustic HMM and work area for acoustic likelihood computation. The “MFCC (Mel-Frequency Cepstral Coefficients) instance” is generated from the AM process instance to extract a feature vector sequence from speech waveform input. The “LM process instance” holds a language model and work area for the computation of the linguistic likelihoods. The “Recognition process instance” is the main recognition process, using the AM process instance and the LM process instance [17].

- building, we use the HTK toolkit [10] to create a new language in three main steps, as illustrated in Figure 3.:
- Data preparation: All needed data such as lexicon, audio files are collected and compile.
- Training: HMM parameters (mean, variance, etc.) are estimated using HTK tool suite.
- Evaluation: testing and analyzing acoustic model performances.

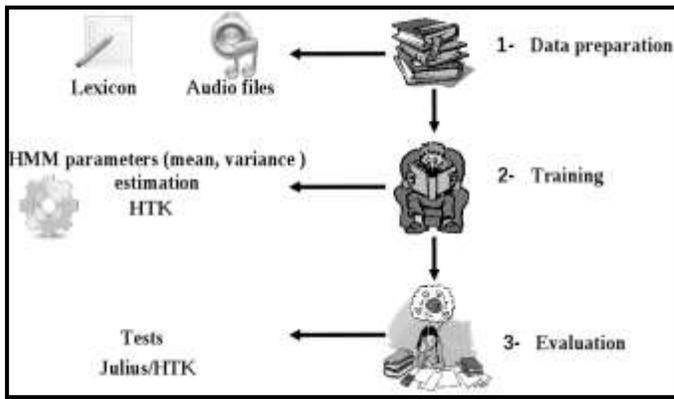


Figure 3. Steps for a new language creation with HTK [10].

C. Developments

New language creation following the three phases is performed by acoustic model maker (Make\_AM.sh), which is a program built around Perl and Bash script, running HTK commands. It provides options for data compilation, acoustic model training and evaluation.

Make\_AM.sh helps us to create an acoustic model in a chosen language (illustrated by Figure 4).

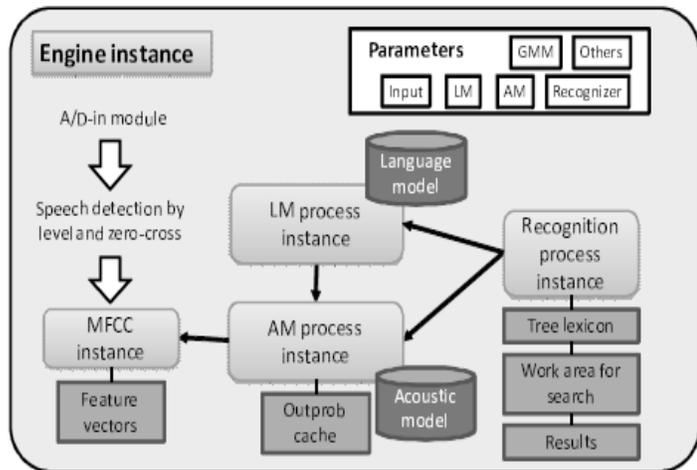


Figure 2. Internal structure of Julius [7].

B.2 Creating a new language

The language model in this context is rule-based grammar. This type of language model is defined in Backus-Naur Form (BNF) and compiled with “makdfa.pl”, a Perl Julius tool [9]. Since Julius performs rule-based grammar

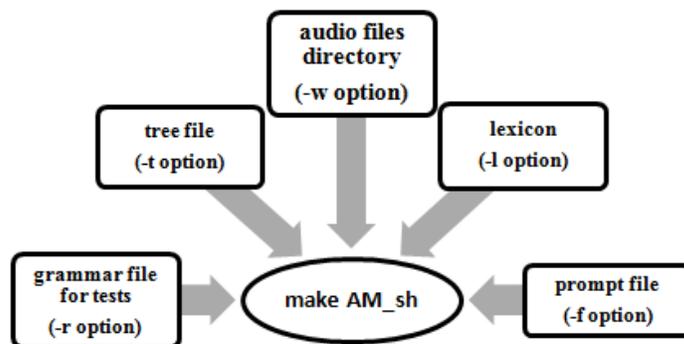


Figure 4. Running Make\_AM.sh

Make\_AM.sh architecture running on Linux operating system is presented in Figure 5.

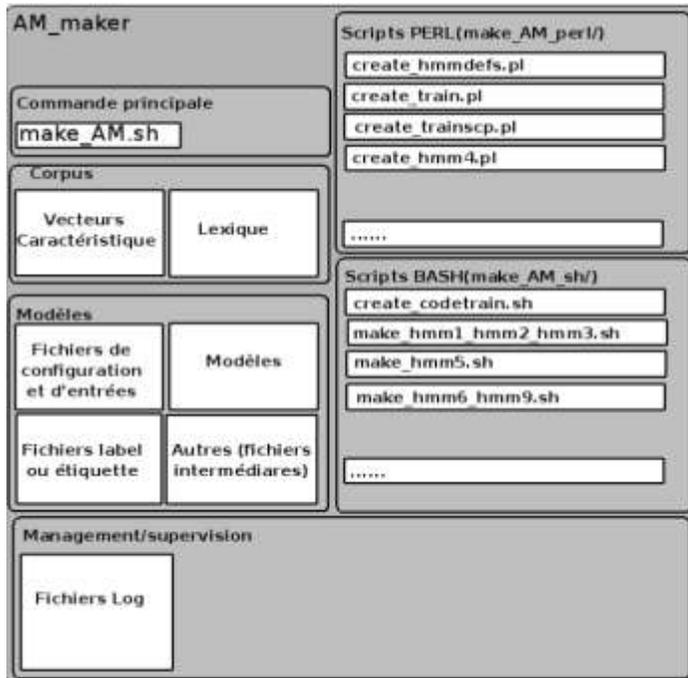


Figure 5. Make\_AM.sh architecture

```

Algorithm: Make_AM.sh.sh
Input: dictionary ,speech files,
          prompt file, tree file, grammar file
Output: acoustic model files.
1. if required input then
2.   pronunciation dictionary
3.   transcription file
4.   MFCC extraction
5.   flat HMM
6.   silence model
7.   switch model (-p option) do
8.     case monophone:
9.       monophone model
10.    case triphone :
11.      triphone model
12.    end switch
13.    if adaptation (-a option) then
14.      acoustic model independent of speaker
15.    end if
16.    if test with HTK(-T option) then test & analyze
17.    else
18.      if live test with HTK(-H option) then
19.        test with HTK
20.      else
21.        if live test with Julius (-J option) then
22.          test with Julius
23.        end if
24.      end if
25.    end if
26.    end if
27.    end if
28.    return acoustic model
29.  else return exception
30. end if
31. end if
32. end
    
```

### C.1 Data preparation

For best results, the corpus of acoustic data used for learning (Wolof language) must be performed in a good quality recording studio. The records were made by several Indigenous from Senegal and other regions, for the model to capture most of the nuances of phones [17], to these records, we use Handy Recorder with four channels (H4n).

The age of the speakers selected depends on the intended target of the service. The age of the speakers was between 18 and 55. To proceed we use Make\_AM.sh (-f option).

### C.2 Training

Make\_AM.sh passes through all training steps specified in HTK manual (htkbook). Before that, it compiles data in previous phase to prepare feature vectors MFCC and needed files.

### C.3 Evaluation and results

We perform Make\_AM.sh (-T/-J option) to evaluate the acoustic model built with HTK recognition tool. We did Wolof recordings of about 123 sentences.

Parameter	Description
Testing audio files	Reusing training audio files
SOURCERATE	parameter not specified configuration files

Table 1. Test environment parameters

Evaluation provides following results:

Recordings carried out at the frequency of	Percentage of sentences recognized	Ajusting SOURCERATE parameter
16 KHz	100%	Yes/No
48 KHz	88,57%	No
48 KHz	91,43%	Yes

Table 2. Test results

Because Speech Recognition Engines need Acoustic Models trained with speech audio that has the same sampling rate and bits per sample as the speech it will recognize, several reasons can justify the results obtained in Table 2. The different speech mediums have limitations that affect speech recognition, telephony bandwidth limitations, desktop sound card and processor limitations, some VOIP PBX's, such as asterisk, actually represent audio data internally at 8kHz/16-bits sampling rates. Speech Recognition Engines work best with Acoustic Models trained with audio recorded at higher sampling rate and bits per sample. So to made adjustments we decided to collect speech recorded at the highest sampling rate your audio card support, and then downsample the 48kHz/16-bit audio to 16kHz/16-bit audio,

that can be supported by the speech medium as indicated by VoxForge [9], and create Acoustic Models from this. This approach permits us to be backward compatible with older Sound Cards that may not support the higher sampling rates/bits per sample, and also permit us to look to the future so that any submitted audio at higher sampling rates/bits per sample will be usable down the road when Sound Cards that support higher sampling rates/bits per sample will become more common, and processing power increases.

### III. TEXT-TO-SPEECH CONCEPTION

#### A. Reminders on theory

One goal of TTS is to be able to provide textual information to people via voice messages. TTS provides voice output for all types of information that can be stored in databases and information services. Most speech synthesizers are built with modules that perform the 5 steps [15][16] shown in Figure 1.

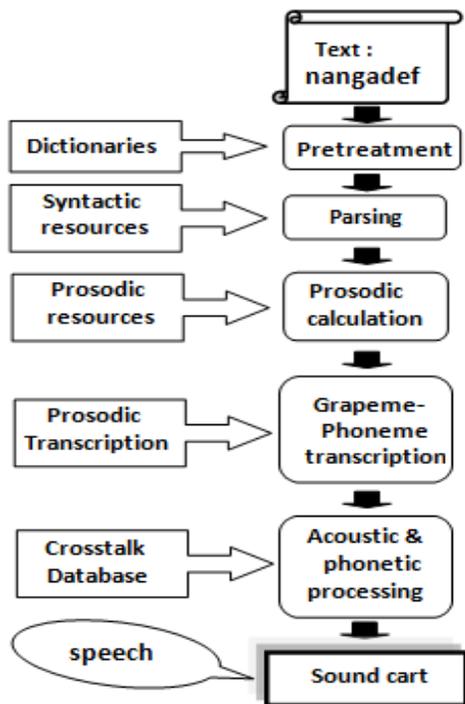


Figure 6. Five steps of Speech Synthesis

#### B. Methodology to create a new voice in festival

When creating a new voice in Festival [13][14], two scenarios are possible. The language to which the voice will be created is already supported by Festival, or that language is not yet supported. Our study was based on the latter, which requires the providing of:

- Phone set list
- Token processing rules (number etc)
- Prosodic phrasing method
- Word pronunciation (lexicon and/or letter to sound rules)

- Intonation (accents and FO contour)
- Durations
- Waveform synthesizer

But if the language is already supported by festival, we need to consider:

- Waveform synthesis
- Speaker specific intonation
- Speaker specific duration

Another possible solution for building a new voice in festival is to do voice conversion, as is done at Oregon Graduate Institute (OGI) and elsewhere.

For the construction of new voice in Wolof, we did not completely follow the steps described in the tutorial of the festival [17]. Instead, we have, due to a problem matching accented characters (UTF-8) of the Wolof alphabet, bypassed all the steps before the grapheme-phoneme transcription by an additional module written in Perl, composed of different lexicons and Perl functions which do all the steps from preprocessing to grapheme-phoneme transcription. This module produces output correspondence in ARPABET to a Wolof text input.

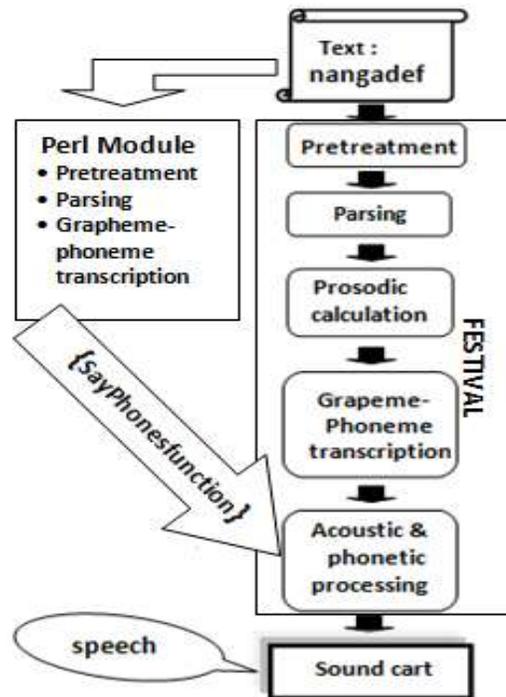


Figure 7. Grapheme-phoneme transcription

The connection with festival takes place using the *festival\_client* command, with as parameter the text transcript in ARPABET alphabet by the Perl module. *festival\_client* executes the *SayPhones* command on festival server that can synthesize data in ARPABET alphabet, conforming to the phone set list of the chosen language, namely English for our case.

```

Algorithm: transcription
Input:word list to transcribe=LAenrg()
List of Wolof characters=LCW()
Hash tables, matching graphemes Wolof
<--> graphemes ARPABET
Hash tables special cases, matching graphemes Wolof
<-->graphemes ARPABET
fonction épurer() /*correct or delete all non orthographic elements
and no Wolof's character of the word which is crossed to him in
parameter */
fonction transcrire() /*give the equivalent ARAPBET of Wolof
word which is it given in paramètre also by taking into account
cases individual of transcription linked to the position of phonemes in
words */
Output: transcription of the text in ARPABET
1. Begin:
2. valeurRetour = "";
3. Silence = "pau";
4. for all word of LAenrg() $mot1 € LAenrg()
5. $mot2 = epurer($mot1)
6. $mot3 = transcrire ($mot2)
7. valeurRetour = valeurRetour . Silence /* add a silence at the end
of the valeurRetour */
8. valeurRetour = valeurRetour . $mot3 /* add a $mot3 at the end
of the valeuRetour */
9. end for
10. end
    
```

The advantage here is that although working in a language not yet supported by festival, we could with the available diaphone database create a voice in Wolof (or other vernacular language), without having to make recordings. It can thus build several voices without making great effort.

The diaphone database is extracted from audio recordings of a real person voice. It is important to choose a speaker with a distinctive voice for a better synthesized voice quality.

The recordings are made in the tutorial [17] with the stack of software *speech\_tools* of CSTR; but it is completely in command line, then more advanced utilities such as *Audacity* [16] with a Graphic User Interface (GUI) recorder, make records and treatments faster and easier.

**C. Building a new voice in vernacular language:**

Vernacular languages (Wolof, in this case) are often very poor in vocabulary [17], and are often not formalized. These languages fill their vocabulary gaps by borrowing words in others more advanced languages. This makes it necessary to take into account all the languages of borrowing during the construction of the synthesizer. The construction of vernacular general synthesizer thus requires an abundant lexicon grouping non orthographic characters, exceptions, acronyms and words borrowed from others advanced languages. Figure 8 illustrates the applications talking-Web width festival in a Web page [18].

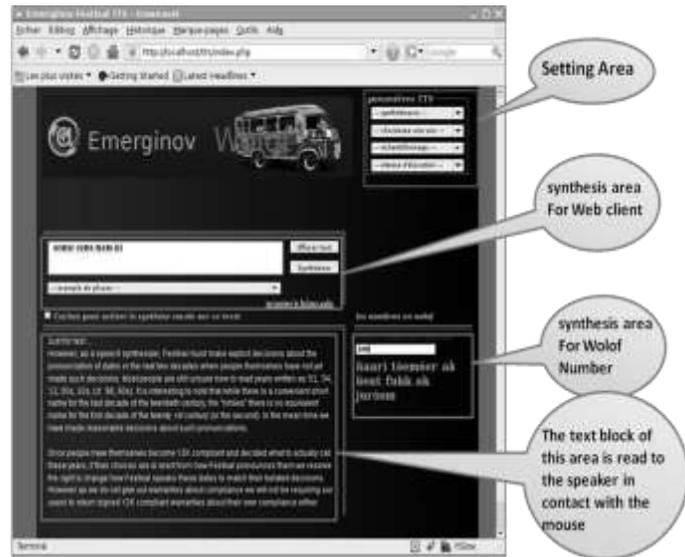


Figure 8. Talking web based on festival [18].

**D. The talking web based on festival**

The high level of illiteracy is the factor that leads indigenous to mystify Information Technology and Communication (ICT). We took advantage of open source software festival, to build a JavaScript library, named *ajaxForFestival*. This library based on Asynchronous JavaScript and XML (Ajax) allows to easily implement Web-talking at any website, this by a set of 12 functions.

Thus, we have described the approach that we used to build a voice based on an African vernacular language (Wolof) on the open source Festival TTS. We also describe the Web service built from speaking of TTS and *ajaxForFestival* JavaScript library that we built. This helps make Web content accessible to illiterate populations.

**E. Web API to build lexicons database**

Figure 9. describes the open API allows web users to create lexicons and their transcription in any language of their chose. This helps us build a database of lexicons in any other vernacular language [18].

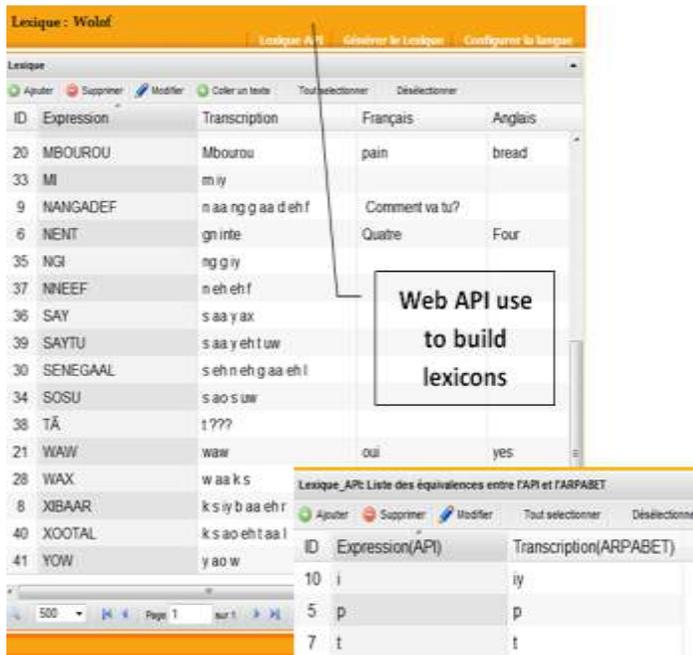


Figure 9. GUI to build lexicons database [18].

#### IV. CONCLUSION

The objective of our work was to conduct a theoretical study in the areas of recognition and speech synthesis in the Wolof language on the one hand and implementation of results obtained in the open source software for Julius Speech Recognition and festival for speech synthesis on the other. Today most of our results are already hosted by the platform Emerginov of France Télécom [18].

The filling of different lexicons and knowledge bases was the most time consuming task and requires some mastery of phonetic features, the acoustic and linguistic of Wolof, and a high degree of concentration. Recognizing this enormous challenge, we have grouped all these characteristics, then we have introduced various web APIs, finally we wrote scripts to automate the filling of the lexicons. This makes the self-evolving platform, because even in production environments lexicons can continue to be met.

Despite the problems we encountered, we are relieved that the end customer France Telecom (Orange Labs) appreciated the positive results we have obtained so far. In our conceptual approach, we have as much as possible left the door open to other languages that may in future be incorporated into the platform without major problems. There are many possibilities for future development of the system. That is why we believe that the work produced will open doors to a large development in Senegal.

#### REFERENCES :

- [1] P Nasfors. "Efficient voice information services for developing countries". Master's thesis, Department of Information Technology, Uppsala University, May 2007.
- [2] E. Barnard, L. Cloete and H. Patel. Language and technology literacy barriers to accessing government services. Lecture Notes in Computer Science. vol. 2739, pp. 37-42, June 2003.
- [3] J. Bridle, L. Dengand and J. Picone, "An investigation of segmental hidden dynamic models of speech coarticulation for automatic speech recognition". Final Report for the 1998 Workshop on Language. Engineering, Center for Language and Speech Processing at Johns Hopkins University, pp. 161-168, July 1998.
- [4] J. Ma and L. Deng, "Target-directed mixture linear dynamic models for spontaneous speech recognition. IEEE transactions on speech and audio processing", vol. 12, pp. 31-38, January 2004.
- [5] M. A. Spaans, "On Developing Acoustic Models Using HTK". Faculty of Electrical Engineering, Mathematics and Computer Science Delft University of Technology, pp. 13-20, December 2004.
- [6] T. Kawahara, H. Nanjo, T. Shinozaki and S. Furui, "Benchmark Test for Speech Recognition Using the Corpus of Spontaneous Japanese". ISCA & IEEE Workshop on Spontaneous Speech Processing and

All the work done in this project is to ultimately get the architecture shown in Figure 10, where module TTS-ASR implemented, will provide interfaces for dialogue between clients, the administrator and open source tools Julius and Festival. The proposed interfaces should allow the creation of services based on speech recognition and speech synthesis. Customers can access services via various media (SIP, Web, Mobile phone...). The concept behind the services to implement is sending requests to the voice system, which responds with speech synthesis.

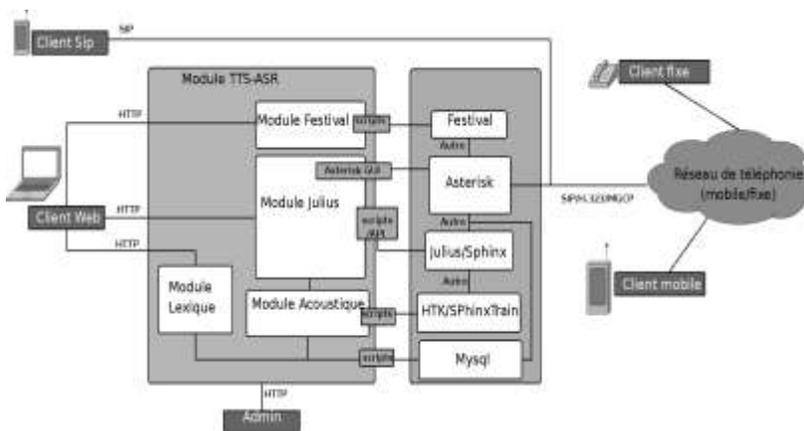


Figure 10. Final architecture

- Recognition (SSPR), Proc. SSPR2003, pp. 135–138, May 2003.
- [7] "Julius architecture", <http://julius.sourceforge.jp>, last accessed on the 8<sup>th</sup> of March 2011.
- [8] A. Lee and T. Kawahara, "Recent Development of Open-Source Speech Recognition Engine Julius". Nagoya Institute of Technology & Kyoto University, In Proc. APSIPA ASC, pp.131-137, October 2009.
- [9] "Collection of transcribed speech for use with Free and Open Source Speech Recognition Engines", <http://www.voxforge.org/>, last accessed on the 8th of March 2011.
- [10] "HTK website and resources", <http://htk.eng.cam.ac.uk/>, last accessed on the 8th of March 2011.
- [11] M. Morel and A. Lacheret-Dujour, "Synthèse vocale à partir du texte de la conception à la mise en œuvre" Laboratoire CRISCO, Université de Caen, vol. 42, pp. 193–221, September 2006.
- [12] S. Baloul, "Développement d'un système automatique de synthèse de la parole à partir du texte arabe standard voyélé". Université du Maine, May 2003.
- [13] A. W. Black P. Taylor and R. Caley, "The Festival Speech Synthesis System", Edition 1.4, for Festival 1.4.3, December 2002.
- [14] A. W. Black and K. A. Lenzo, Festvox: "Building Synthetic Voice", Voices version 2.1. <http://www.festvox.org/bsv/>, last accessed on the 8th of March 2011.
- [15] "Festival website and resources", <http://festival.sourceforge.com/documentation/1.96-beta/files.html>, last accessed on the 8th of March 2011.
- [16] "Audacity. The Free, Cross-Platform Sound Editor ", <http://audacity.sourceforge.net/>, last accessed on the 8th of March 2011.
- [17] J. Diouf "Dictionnaire Wolof-Français et Français-Wolof", Edition KARTHALA, August 2003.
- [18] "Links to integrated platforms", [http://projects.emerginov.org/esmt\\_Wolof/tts/](http://projects.emerginov.org/esmt_Wolof/tts/), last accessed on the 8th of March 2011.

# Linguistic Text Compression

Ondřej Kazík  
 Charles University  
 Faculty of Mathematics and Physics  
 Prague, Czech Republic  
 Email: kazik.ondrej@gmail.com

Jan Lánský  
 The University of Finance and Administration  
 Department of Computer Science  
 Prague, Czech Republic  
 Email: zizelevak@gmail.com

**Abstract**—Compression of texts written in natural language can exploit information about its linguistic structure. We show that separation of coding of part-of-speech tags of a sentence (so called sentence types) from the text and coding this sentence types separately can improve resulting compression ratio. For this purpose the tagging method NNTagger based on neural networks is designed. This article is focused on a specification and formalization of a compression model of texts written in Czech. Language with such a complicated morphology contains a great amount of implicit grammatical information of a sentence and it is thus suitable for this approach. We propose methods of constructing of initial dictionaries and test their influence on resulting compression ratio.

**Keywords**—text compression; part-of-speech tagging; neural networks

## I. INTRODUCTION

It was pointed out in several works concerning text compression (e.g., [22]) that knowledge of the structure of a coded message can be exploited for more successful compression of this message. It turned out that a set of basic elements (words, syllables, characters) as well as a structure of documents is strongly dependent on the used language. This motivates us to exploit implicit linguistic information contained in an input text.

We use during the compression knowledge about structures larger than words, so called sentence types – a sequence of words' tags in a sentence. We assume that the separate coding of a sentence type and words of a sentence with knowledge of their tag is a suitable supplement of common compression methods based on words as the elements.

A tagging method has to precede the compression itself. Since compression is not sensitive to small errors, we will seek for approximative method of assigning tags to the words in a sentence.

Our goal it then to propose such model of sentences which comprehends the sentence type. This model will serve for an effective coding of document's sentences.

In Section II, we provide a brief overview of tagging and compression methods, other than the approaches described in this article. In Section III we introduce the NNTagger, an approximative part of speech tagger based on neural networks. In Section IV we describe the model of a text,

sentence and other subelements and coding and decoding algorithms which are based on them. In Section V we compare results of our method with other compression techniques. Finally, in Section VI we summarize our findings and discuss possible future work.

## II. RELATED WORK

Many nature language processing applications, e.g., syntactic analysis, require as a preprocessing phase assigning definite morphological information to each word in an input text. This procedure is tagging and consists of a morphological analysis, i.e., assigning of all possible tags to a word alone, and disambiguation; i.e., resolving tag ambiguity by the context of a word. Both these phases are chosen with regard to the used language and tag set, which is the range of the tagging method. The Czech language in comparison with the English language is a highly inflecting language, which needs a more sophisticated processing of suffixes and in many applications it requires a tag set, which contains more detailed morphological information. Also it is highly ambiguous. In the English language, the relative small tag sets are used (e.g., Brown, Penn, CLAWS, London-Lund [23]). By contrast, the Czech tag sets (as the Prague Dependency Treebank positional tag set [11] or the set used in morphological analyzer Ajka [31]) represent broad morphological information containing several thousands of possible tags. For the Czech language there are correspondent morphological analyzers, e.g., the tools Free Morphology [9] and Ajka [31]. There is a large number of disambiguation methods suitable for different languages and tagging sets, for instance visible [23] or hidden Markov models [6], taggers based on a decision tree TreeTagger [30], transformations [3], maximum entropy model [26], exponential tagger [10], averaged perceptron-based model ([5] or MORČE [40] for the Czech language) and artificial neural networks based taggers (NetTagger [29]) which is close to our approach, especially the context net of NNTagger described later.

The text file compression is usually employed to save hard drive space or bandwidth during their transfer. The compression of texts can be classified according to elements of coding, e.g., letters, words or syllables. The methods

based on words require dividing the input document into a stream of words and non-words [24], [25]. Overview of word methods is provided by [36]. There are word-based variants of all main groups of compression algorithms: e.g., word-based Huffman encoding [13], [28], word-based LZW [7], word-based Burrows-Wheeler Transformation [7], [15], and word-based PPM [2], [24]. Syllables as source units for compression can be successfully used in languages, where words are naturally divided into syllables (e.g., Czech, Russian, and German). This approach was firstly used in [39]. Knowledge of regularities of the language, in which an encoded message is written, can be exploited to improve text compression. For English texts, there are several methods of such improvement: e.g., specific order of letters in the alphabet for lexicographic sorting [1], [4], replacing common clusters (n-grams) of letters with one symbol [38], static initialization of PPM method [35]. In order to manage multilingual text files, the methods need to recognize the language of a message and its encoding, like in modifications of Word Replacing Transformation, TWRT [36]. Separation of the parts of speech of each word is used as additional information for compression of English texts and compared with character and word based compression programs also in [38]. Compression of small files is a separate subject of research. One of such approaches is focused on compression of short text files for mobile phones. The low-complexity static partial matching model (PPM) is described in [27]. A sequence of pruned suffix trees is used in statistic model in [17].

Lansky and Zemlicka [22] propose theoretical basis of syllable-based compression, LZWL and HuffSyll algorithms. Right configuration of the characteristic words and syllables dictionaries used by LZWL and HuffSyll algorithms for initialization is emphasized in [21]. Various approaches to small and medium text file compression by means of Burrows-Wheeler Transformation (BWT) are compared in [19]. Suitable inflation of the characteristic syllables and words sets, which are used to initialize the HuffSyll and LZWL methods is examined in [18]. XBW Project [32], [33] studies the compression of large, non-well formed XML files. Its basis is BWT and it implements many other coding algorithms, which we utilize in this paper [20].

### III. NNTAGGER (APPROXIMATIVE PART OF SPEECH TAGGING)

Linguistics as a science investigates a language from many viewpoints, e.g., syntax, morphology, lexicology etc. The language we are interested in, the Czech language can be described from these perspectives. The Czech language has a free word order, a rich inflective morphology which lays grammatical morphemes at the end of the word. Also the sets of words belonging to different parts of speech have different properties. For example, the inflective and autosemantic parts of speech (e.g., nouns, verbs etc.) tend to

have longer words and greater variation both in a document and among different documents than the non-inflective and synsemantic ones (prepositions, conjunctions etc.). These characteristics can be exploited in our approximative tagging method NNTagger.

Every tagging algorithm has to solve three main problems: tokenization, i.e., division of the input text into basic elements (words, sentences), morphological analysis, i.e., assigning of possible tags to a word alone, and disambiguation which determines the tag according to a particular use of the word in a sentence.

#### A. Pre-processing

Two questions must be solved before the tagging. First, we focus on a definition of words, which will be the basic element of processing. Second, we should consider a width of the sequence of words, which will be available for the tagging procedure, i.e., how to separate single sentences from an input text.

1) *Words*: We are extending the definition of words from [22], where authors, with regard to separability, introduced these classes of words: words consisting of small letters (small), those consisting of capital letters (capital), with first letter capital (mixed), numeric and other words. We have taken into account characteristics of training data from the PDT corpus [12]. There not only numeric and alpha words but also punctuation marks are considered as the basic elements on the word layer. This is consistent with the Manning and Schütze [23], where the information about macrostructure of sentence contained in the punctuation is emphasized.

2) *Sentences*: Now we need to decide, how long blocks of words will be passed on to a tagging algorithm. Annotated available training data (the PDT corpus [12]) contain texts tagged divided into whole sentences. We gather an inspiration from description of the algorithm in [23].

We are passing the sequence of non-special words from left to right and inspecting them as candidates on a sentence boundary. We use this heuristics:

- If the inspected word is “;” or “...”, it is marked as the last word in the sentence.
- In the case of “?” or “!”, it is marked, if in the same time is not followed by another “?” or “!”.
- The inspected word is “.”. It is marked only, if the following word is capitalized but the preceding word does not consist of one capital letter (abbreviation of a first name) or this word is not in a list of abbreviations which usually is not placed at the end of a sentence (“mgr”, “např”, “tzn” etc.).
- If after the inspected word is end of line and next word is not small or if after this word is more than one end of line, we put there the sentence boundary. This is the case of e.g., a heading.

- If immediately after the supposed sentence boundary is a quotation mark, shift the boundary after it.

In the next section we will show that it is unsuitable to pass these large segments to the compression algorithm directly and thus they will be segmented to shorter sequences.

### B. Tag Set

The decision about a tag set was limited by available data. In our case it is the annotated corpus PDT with a positional morphological tag system, where each position corresponds to some morphological category. So the tag set should be equivalent to a subset of these positions. The decision was based on a presupposition that the tag should carry an information about characteristics of the given word form, its statistical features and its relation to other words in a sentence. In the same time the sets of word forms belonging to these tags should have minimal overlapping, which disqualify some detailed categories (as the cases). Hence we can focus on the position *Part of Speech* (POS, 12 values) and the position *Detailed Part of Speech* (SUBPOS, 75 values), from which we can also derive the part of speech category. Due to various criteria for subsumption under a part of speech, words from the same set can play various roles in a sentence (e.g., Numerals or Pronouns). The detailed part of speech category would partially reduce this ambiguity. But we decided only for the first category (Part of Speech) because of the following reasons:

- The larger is the tag set, the more difficult is to determine the correct value of the tag. This is due to ambiguity of the division criteria and to shortcomings of the chosen tagging method.
- With an amount of information extracted from the text complexity of the model of language and number of its parameters increase.
- The larger alphabet of symbols, i.e., tag set, corresponds to the larger set of words, which are compound of this alphabet, i.e., sentence types, and also the higher probability of occurrence of a new sentence type. This would complicate the compression.

The selected set of part of speech tags contains thus 11 values: nouns (also marked by N), adjectives (A), pronouns (P), numerals (C), verbs (V), adverbs (D), prepositions (R), conjunctions (J), particles (T), interjections (I) and punctuations (Z). This category carries according to the entropy estimation about 25 % of the information of a word form.<sup>1</sup>

### C. Morphological Phase

The input of this phase is a non-special word  $w$  with the length  $m$ :

$$w = \alpha_1 \alpha_2 \dots \alpha_m$$

<sup>1</sup>The entropy was measured on the PDT corpus. Shannons entropy of the part of speech is 2.9 bits and of the whole word forms 11.9 bits.

The output is then the vector  $\mathbf{y}^m(w)$  with the length equal to the size of the tag set  $n = |\mathcal{T}|$ . It holds for elements of this vector:

$$y_t^m(w) \in [0, 1]$$

This value can be interpreted as the likelihood that the word  $w$  has the tag  $t \in \mathcal{T}$  on this level.

We define the auxiliary vector  $\hat{\mathbf{y}}(t)$  which can be considered as a certainty that the word has the tag  $t$ .

$$\begin{aligned} \hat{y}_t(t') &= 1 && \text{if } t = t' \\ &= 0 && \text{otherwise} \end{aligned}$$

First, it is possible to tag the non-alpha words. If  $w$  is a word containing a symbol from  $\Sigma_Z$ , then  $w$  is a punctuation:  $\mathbf{y}^m(w) = \hat{\mathbf{y}}(Z)$ . The procedure is then terminated.

Similarly, if the word  $w$  contains a symbol from  $\Sigma_C$ , it is a number and thus:

$$\mathbf{y}^m(w) = \hat{\mathbf{y}}(C)$$

Next only alpha words are processed, so the small, the capital and the mixed are distinguished. Lets the function  $\psi : \Sigma_M \cup \Sigma_V \rightarrow \{1, 2, \dots, 41\}$  assigns to every letter a natural number which corresponds to some basic Czech small letter.<sup>2</sup> A size of letters is irrelevant for this mapping, thus:

$$\forall \alpha \in \Sigma_V : \psi(\alpha) = \psi(\mu(\alpha))$$

For the letters which are not included to the basic Czech alphabet (which have small probability in Czech texts) it assigns some value from this alphabet. For example:  $\psi('ä') = \psi('a')$

We decided to use neural nets for the next processing. This decision was motivated by the amount of information, which contains regular structures of word forms in such inflecting languages as is the Czech language. Another advantage is the absence of any dictionary. On the other side, we expect a lower accuracy than by other methods of the morphological analysis. This should not have a major influence on the compression.

Due to a speed of learning and a number of parameters, we decided to limit relevant letters for a morphological analysis to a fixed window. Symbols of the word  $w$  will be inserted to this window justified to the right. This reflects characteristic of the Czech language, where the major part of morphological information carries an end of a word, i.e., suffix and ending. We choose the size of the window to eight characters.

The morphological net is based on the Back-propagation net [37]. It includes input layer units, which form vector  $\mathbf{x}^m$  of the length 331.

<sup>2</sup>We consider the basic Czech alphabet as a set of small letters together with their possible variants (with the accent, wedge or circle), total 41 letters. The mapping  $\mu$  is a bijection after a restriction on this set.

- In a first part of the layer, every unit corresponds to a possible symbol of the input window. It thus consists of  $8 \cdot 41 = 328$  symbols and these are set as follows:

$$\begin{aligned} \forall i \in \{1, \dots, 8\}, j \in \{1, \dots, 41\} : \\ x_{41 \cdot (i-1) + j}^m = 1 \quad \text{if } j < |w| \text{ and } \psi(\alpha_{|w|-i+1}) = j \\ x_{41 \cdot (i-1) + j}^m = 0 \quad \text{otherwise} \end{aligned}$$

- The next two units are coding a size of the word  $w$ :

$$\begin{aligned} x_{359}^m = 0 \text{ and } x_{360}^m = 0 \text{ if } w \text{ is small} \\ x_{359}^m = 1 \text{ and } x_{360}^m = 0 \text{ if } w \text{ is mixed} \\ x_{359}^m = 1 \text{ and } x_{360}^m = 1 \text{ if } w \text{ is capital} \end{aligned}$$

- The last symbol contains a length of the word:

$$x_{361}^m = |w|$$

The output vector has the length 10, which is a length of  $\mathbf{y}^m$  shortened by the position for punctuation (Z) handled before. On an output of the morphological phase value 0 is being appended on this position. During the training, an input word from the corpus is transformed according the previous definition to the input vector, which is presented to the morphological net and a result is computed. Next, this result is compared to the desired output, i.e., vector  $\hat{\mathbf{y}}(t)$ , where  $t$  is the right tag assigned to the word in the corpus. Weights of the net are subsequently adjusted according to the Backpropagation algorithm. [37]

We transformed the data from PDT [12] to a suitable form, which includes beside word forms and their part of speech tags the information about ends of sentences. The data containing 670528 words were split to training, with which the net was trained, and testing, on which an error rate was measured in an approximate ratio 2 : 1.

We set in all cases the learning rate of the algorithm  $\alpha = 0.2$ . Momentum was not used, because there was no or negative influence on a learning. Training data were passed in six cycles by the net. This procedure was performed on nets with different configurations. For nets with one hidden layer it were 50 (i.e., 331 in the input, 50 in the hidden and 10 in the output layer), 100, 150, for nets with two hidden layers then 50-25, 100-25, 100-50, 150-100.

On the testing data is then with estimated the tag of each word as a maximal value in the output vector of the morphological phase:

$$\hat{t}^m(w) = \arg \max_t y_t^m(w)$$

We see that larger nets do not automatically lead to better performance. Greater amount of parameters makes probably the learning harder and longer. The relative simple net has the best result with only one hidden layer containing 100 units.

#### D. Context Phase

The goal of the context phase is a processing of information about the surrounding words, the actual context of a word, in order to estimate its real tag. We took up a neural net-based disambiguation method Net-Tagger, described in [29]. Nevertheless there are several differences between this and our approach. It is namely:

- Inputs of context net are not probabilities of tags assigned to the word, which was looked up in a dictionary of word forms, but an output vector of the morphological phase.
- Context net contains a hidden layer. Computational power of the nets without this layer is limited. [37] On the other side, multilayer nets can model more complicated functions.

We take into account for the actual word its left context with size  $p$  and right context with size  $f$ . Every word in this context is represented by  $|\mathcal{T}|$  (i.e., 11) units expressing likelihood that the word has a corresponding tag, based on all available information. This means for the actual word  $w_i$  and words in the right context, for which context phase was not yet processed, the outputs of the morphological phase, i.e., vectors  $w_{i+1} \dots w_{i+f}$ . For words in the left context, previous outputs of the context layer are available. The input layer thus contains  $11 \cdot (p+1+f)$  neurons. If  $|S|$  is a length of the sentence  $S$ , we set the input vector  $\mathbf{x}^c$  for the word  $w_i$  as follows:

$$\begin{aligned} \forall j \in \{-p, \dots, f\}, t \in \{1, \dots, 11\} : \\ x_{11 \cdot (j+p) + t}^c = y_t^c(w_{i+j}) \quad \text{if } j < 0 \text{ and } 1 \leq i+j \\ x_{11 \cdot (j+p) + t}^c = y_t^m(w_{i+j}) \quad \text{if } 0 \leq j \text{ and } i+j \leq |S| \\ x_{11 \cdot (j+p) + t}^c = 0 \quad \text{otherwise} \end{aligned}$$

The output of the net is the vector  $\mathbf{y}^c(w_i)$  with a length  $|\mathcal{T}|$ .

The input of the words from the left context brings a recursion into the computation. This can complicate the learning procedure, especially in the moment, when the outputs are not yet correct. Hence in the time of learning we replace the vector  $\mathbf{y}^c(w_{i-j})$ , analogously to [29], by weighted average of the previous output vector  $\mathbf{y}^c(w_{i-j})$  of the context net and the output of the morphological phase  $\mathbf{y}^m(w_{i-j})$ :

$$(1 - e_\tau) \cdot \mathbf{y}^c(w_{i-j}) + e_\tau \cdot \mathbf{y}^m(w_{i-j})$$

Here  $e_\tau$  is a coefficient, which in the beginning, when the error of the net is high, is near to 1 and with successful learning falls. We used there for this value an exponential moving average of an error of the net, i.e., ratio of wrong assigned tags, if the error is greater than 0.1. In the other case 0 is assigned to  $e_\tau$ .

The inputs from the training data were put to the net during the learning according to the previous definitions and

Table I  
ERROR RATE OF THE NNTAGGER DEPENDING ON A CONTEXT SIZE. IN  
THE ROWS ARE SIZES OF A LEFT CONTEXT  $p$ , IN THE COLUMNS ARE  
SIZES OF A RIGHT CONTEXT  $f$ .

	0	1	2	3
0	4.72 %	4.74 %	4.67 %	4.61 %
1	4.72 %	4.53 %	4.42 %	4.45 %
2	4.72 %	4.51 %	4.40 %	4.41 %
3	4.72 %	4.48 %	4.43 %	4.37 %

the outputs were vectors  $\hat{y}(t)$ , where  $t$  is a correct tag of the word. The output of the best morphological net was used as the input of the context net, i.e., the net 331-100-10. The same data source as by the morphological phase was used for the training. The training was processed six times with a learning rate equal to 0.2 and without a momentum.

The output of the NNTagger method is thus a sequence of tags induced from values of the vectors  $y^c(w)$ . Formally speaking, the tag belonging to a word  $w$  is:

$$\hat{t}(w) = \arg \max_t y_t^c(w)$$

We show in the table I achieved error rates on the remaining test data depending on sizes of the left ( $p$ ) and the right context ( $f$ ).

The lowest relative accuracy of the tagger was by interjections which were never tagged. This part of speech occurs in Czech texts only rarely. Greater error rate was also by particles which have disputable morphological characteristics. The highest absolute number of mistakes was between nouns and adjectives in both directions.

#### IV. COMPRESSION OF SENTENCES

Any natural language utterance contains a great amount of implicit information. Now we will show use of the information about parts-of-speech in an input text for its compression.

We define a *type of a sentence* (or of a sentence part), which is determined by its non-special words  $w_1, \dots, w_n$ , as a sequence of part-of-speech tags of this sentence. ( $T(S) = t_{1,n}^S$ )

##### A. XBW Project

We adopt the platform XBW [20] and exploit some of its features in our work. The goal of the XBW project is creation of an interface to test of different compression methods with an accent on compression of texts in the XML-format. It implements some of the elementary algorithms for a number coding and decoding (Elias, arithmetic, Huffman coding), string compression (Burrows-Wheeler, PPM, MTF, RLE, LZ), XML files processing, dictionary compression, some of auxiliary data structures (trie) and suggests their collaboration.

##### B. Sentence Parts

We performed an experiment in order to find out duplicities of sentence types. We split a language corpus between training and testing data in an approximate proportion 2:1. A dictionary of the sentence types was created from the training data. Then there were only 16 % overlapping of the sentence types from the test data and those in the dictionary. Variation of types of whole sentences is thus too high for our purpose. Hence, we propose using shorter segments of a sentence, called *sentence parts*. Due to a heuristic character of the definition, sentence parts do not have to match with sentence structure.

The parsing of an input sentence proceeds as follows. After skipping of an initial sequence of punctuations, the sentence is being passed from the left to the right and divided in these cases:

- Put a boundary of a sentence part after occurrence of punctuation marks “;”, “:”, “””, “(” or “)”. If it is immediately followed by some other punctuation marks, move the boundary after them. If there is not a whitespace between the punctuation and a next word (so this word immediately follows), move boundary of sentence part before the punctuation.
- Put the boundary after “-”, if between it and the next word is a white space (unlike the case of a composed word or the Czech particle -li).
- Put the boundary before the Czech conjunctions “a”, “i”, “ani”, “nebo” or “či”. These conjunctions usually have no commas before them.

After division of the sentence part the parsing continues on the rest of the sentence until the end of sentence is reached.

Types of above defined sentence parts have the duplicity in the dictionary higher, approximately 37 %. Nevertheless the number of new emerging types is relatively high. Thus we will seek for methods of an effective coding of new sentence types.

After the pre-processing we have input text decomposed in sentence parts in this form:

$$S^P = w_1 \cdot s_1 \cdot w_2 \cdot s_2 \cdot \dots \cdot w_n \cdot s_n$$

where  $n$  is length of the sentence,  $w_i$  (for  $i = 1, \dots, n$ ) is a non-special word (alpha word, numeric word or punctuation) and  $s_i$  is a special word (even with zero length), which always follows  $w_i$ . There is also a sentence type ( $T(S_P) = t_{1,n}^{S_P}$ ) for every sentence part  $S_P$ .

##### C. Models

Lets define a model of text generation which will be used for text compression and decompression.

Four models will be created for these layers of the text: sentence part, sentence type, word, and symbol. Except the first one we can all of them subsume under a template. If we have  $\mu$  as a basic element (type, word or symbol), then this model (template  $M$ ) includes:

- *Dictionary of the basic elements*  $\Delta$ . This is the injective function, whose domain is a set of the known elements and the range is the set of natural numbers.
- *Probability distribution*  $\rho$  of the known elements and reserved value *esc* for elements, which are not found in dictionary  $\Delta$ . These structures are adaptive. If a new element occurs, new values are added to the  $\Delta$  and  $\rho$ . After the occurrence of an element  $\mu$ , relative probability  $\rho(\Delta(\mu))$  increases. The implementation of  $\rho$  depends on concrete compression method (in this case arithmetic coding). Dictionaries and distributions of elements could be initialized from frequent symbols based on a text corpus. Details about the initialization of dictionaries will be provided in the next section.
- *Model of unknown elements*  $M^{sub}$ . If the element is not found in the dictionary, it has to be coded on a lower layer (the models of lengths, symbols, or parts-of-speech).

For coding of the element  $\mu$  to a bit string with knowledge of the model  $M$  we will use this notation:

$$K(\mu|M)$$

In most cases we abstract from the concrete coding algorithm. So the code of symbol with index  $i$  and distribution  $\rho$  (i.e.,  $K(i|\rho)$  can be realized by any adaptive algorithm).

We introduce an operator of concatenation:

$$K = K_1 \cdot K_2$$

which means that the bit string  $K$  consists of the code  $K_1$  followed by the code  $K_2$ .

We do not describe the decoding explicitly. It follows directly from the coding procedure, which is proposed as a bijection between a set of possible texts and a set of possible codes.

For every layer we have thus its own model structure: a model of sentence parts  $M^S$ , a model of sentence types  $M^T$ , model of words  $M^W$  and model of symbols  $M^C$ . In addition, we have a model for lengths of words or sentences  $M^{L_k}$ , which describes generation of non-zero lengths. Some of these structures can be multiple instantiated, as is the case of the word models for each part-of-speech.

We will describe the compression procedure from above, i.e., from higher structures as files and sentence parts to the single characters.

The input file is decomposed after pre-processing into the sequence of sentence parts  $S_1, \dots, S_m$ . These sentence parts are coded one by one and the code of empty sentence  $\lambda$  is appended at the end.

#### 1) Sentence parts:

*Model:* As we said, a sentence part with length  $n$  is determined by a sequence of non-special words  $w_1, \dots, w_n$ , a sequence of special words  $s_1, \dots, s_n$ , which follow them, and its type, i.e., a sequence of tags,  $t_{1,n}^{SP}$ . All of these

sequences have the lengths equal to  $n$ . Empty sentence part ( $\lambda$ , zero length sequence) has a special meaning and it represents an end of file.

Model of sentence parts consists of these submodels:

- *Model of sentence types*  $M^T$ .
- For every part-of-speech one *model of words*  $M^{W_N}, \dots, M^{W_I}$ . We have not find any reason for separate models of symbols for each part-of-speech, because the basic distribution of letters differs in the Czech words only slightly. So there is shared model of symbols for each of them. On the other hand, there are differences in average lengths of each part-of-speech, so we proposed independent models of lengths.
- *Model of punctuations*  $M^{Cz}$ .
- *Model of numeric words*  $M^{W_{num}}$ .
- *Model of special words*  $M^{W_S}$ .

The last two have their own models of lengths and symbols.

*Coding:* Code of a sentence part consists of a code of the sentence type and codes of each word of this sentence part. Thus:

$$K(S|M^S) = K(t_{1,n}^S|M^T) \cdot K(w_1|M_1) \cdot K(s_1|M^{W_S}) \cdot \dots \cdot K(w_n|M_n) \cdot K(s_n|M^{W_S})$$

where  $K(w_i|M_i) = K(w_i|M^{W_{t_i}})$  if  $t_i$  is a tag of the word  $w_i$ , different from punctuation or numeral. If the word is numeral, the alpha numeral must be distinguished from numbers. It is accomplished by a one-bit token. If it is an alpha numeral we use its model:

$$K(w_i|M_i) = 0 \cdot K(w_i|M^{W_C})$$

For a number the word is encoded by the model of numbers:

$$K(w_i|M_i) = 1 \cdot K(w_i|M^{W_{num}})$$

On the other hand, a punctuation is a one-character word  $w_i = \alpha$ , thus the coding is provided directly by the symbol model of punctuations.

$$K(w_i|M_i) = K(\alpha|M^{Cz})$$

#### 2) Sentence types:

*Model:* The model of sentence types contains a dictionary  $\Delta^T$  of sentence types which were in an initialization set or appeared in the input text. The dictionary is represented by a trie (which maps a type to a natural number) over the alphabet of tags and by a table (which maps a value to a type). The next part of this model is a probability distribution  $\rho^T$  of items of the dictionary  $\Delta^T$ , which is used for coding and decoding. The dictionary and the distribution include an element corresponding with the type of an empty sentence  $\lambda$ .

A model of lengths  $M^{L_{30}}$  is then needed because of the coding of new types. Lengths of sentence types are at the

same time the lengths of whole sentence parts, so we do not need to code this information on the higher level. Finally we have to consider a model of generation of the unknown type tags. The basic variant is to have an unconditional probability distribution of tags occurrence  $\rho^P$ . However, we decided to include the information about right context of the actual tag. So, the resulting model is a bigram model. Instead of one distribution, there is a distribution  $\rho^{P_t'}$  for every possible tag  $t'$  which is on the right side of the actual position  $t$ . Together with them there is a distribution  $\rho^{P_\emptyset}$  of the tags at the end of the type, which have no right context. We used the entropy of tags and estimated that the improvement on the new types would be approximately 21 %.

*Coding:* We have to divide the coding of sentence types to two cases. In the first case, the type  $t_{1,n}$  is in the dictionary  $\Delta^T$ . It is then possible to code corresponding value according to the distribution  $\rho^T$ :

$$K(t_{1,n}|M^T) = K(\Delta^T(t_{1,n})|\rho^T)$$

The distribution is subsequently adjusted.

Otherwise, the type is not in the dictionary. Thus, the escape symbol is generated and the new type must be coded. This consists of coding of its length  $n$  and all of its tags from the right to the left with knowledge of the right context.

### 3) Words:

*Model:* There are several instances of a model of words in the general model of texts. There are models corresponding to all part-of-speeches except punctuations (alpha words:  $M^{W_N}, \dots, M^{W_r}$ ), model of numbers (numeric words  $M^{W_{num}}$ ) and model of special words ( $M^{W_S}$ ).

Each model  $M^W$  is responsible for coding of a word  $w$  (the string of symbols  $\alpha_1 \dots \alpha_m$ ). The knowledge about the position of the alpha word in a sentence part is essential for coding of its size. From a measuring of word sizes on the corpus we determined the most probable sizes of a word at the beginning and on following positions of a sentence part. The implicit value on the first position is a mixed word (i.e., word with capital the first letter and small other letters), on the following positions are implicit small words.

Regarding this, there are again the dictionary  $\Delta^W$  and the adaptive probability distribution  $\rho^W$ , both of them include except the *esc* symbol for new word these function symbols:

- $esc_{M \rightarrow C}$ , i.e., the token which means that the word is capital at the beginning of a sentence part.
- $esc_{M \rightarrow S}$  for small words at the beginning of the sentence part.
- $esc_{S \rightarrow C}$  for capital words on the second or another following position.
- $esc_{S \rightarrow C}$  for mixed word which is not at the beginning.

The dictionary and the distribution are moreover able to code the empty word  $\lambda$ , as for instance the non-existent white space between a word and the next punctuation. The

dictionary is represented by a pair of a trie over the alphabet of UNICODE characters and a table which maps from the values to the words. In order to code new words, the  $M^W$  contains a model of length  $M^{L_{20}}$  and a model of symbols  $M^C$  (in the case of alpha words representing small letters).

*Coding:* In the first place, if the word is an alpha word, we have to compare its size (small, mixed or capital) with the predicted one and, if necessary, to emit an escape symbol. Then the small variant of the word  $w' = \mu_g(\alpha_1) \dots \mu_g(\alpha_n)$  is taken into account.

- If the word is capital on the first position:

$$K(w|M^W) = K(esc_{M \rightarrow C}|\rho^W) \cdot K'(w'|M^W)$$

- If it is a small word on the first position:

$$K(w|M^W) = K(esc_{M \rightarrow S}|\rho^W) \cdot K'(w'|M^W)$$

- If it is a capital word on next positions:

$$K(w|M^W) = K(esc_{S \rightarrow C}|\rho^W) \cdot K'(w'|M^W)$$

- If it is a mixed word on next positions:

$$K(w|M^W) = K(esc_{S \rightarrow M}|\rho^W) \cdot K'(w'|M^W)$$

- In all other cases (including the non-letter words) the size of word is known:

$$K(w|M^W) = K'(w'|M^W)$$

Next, the  $w'$  is coded. If the  $w'$  is in the dictionary  $\Delta^W$ , the code can be emitted directly:

$$K'(w'|M^W) = K(\Delta^W(w')|\rho^W)$$

Otherwise, we generate the codes of the escape symbol, the length of the word  $w'$  and all of its symbols. Then the dictionary and the distribution are adjusted.

### 4) Lengths:

*Model:* Model of lengths ( $M^{L_k}$ ) contains only a probability distribution  $\rho^L$  of  $k + 1$  values. The values  $1, \dots, k$  are the most frequent lengths. The value  $k + 1$  is reserved as a special symbol for higher lengths. As usually the distribution is adaptive and after (de)coding of a length  $l$  is the relative probability of occurrence of  $l$  increased. The range of  $\rho^L$  is, on the contrary, fixed and new values are not added into the distribution.

*Coding:* There are two possibilities. First, if  $1 \leq l \leq k$ , we code it directly:

$$K(l|M^{L_k}) = K(l|\rho^L)$$

Otherwise  $k < l$ , the auxiliary overflow symbol is emitted, followed by the alpha code of a surplus value:

$$K(l|M^{L_k}) = K(k + 1|\rho^L) \cdot \alpha(l - k)$$

### 5) Symbols:

Table II  
COMPARISON OF VARIOUS METHODS COMPRESSION RATIOS  $T_B$ .

Method	Compression Ratio
<b>A</b> <sub>1</sub>	4.60
<b>A</b> <sub>2</sub>	4.41
<b>A</b> <sub>3</sub>	4.19
<b>A</b> <sub>4</sub>	3.85
<b>A</b> <sub>5</sub>	3.34
gzip	4.43
bzip2	4.38

*Model:* Model of UNICODE-symbols generation is in the sentence part model included in these instances:

- Model of small letters  $M^{CP}$  shared by models of all alpha part-of-speeches ( $M^{WN}, \dots, M^{Wt}$ )
- Model of characters in numbers  $M^{Cnum}$  (digits, decimal separators, symbol “-”) which is part of model of numbers (numeric words).
- Model of punctuations  $M^{Cz}$
- Model of special words symbols  $M^{Cw}$  (e.g., white space)

Each of these models  $M^C$  contains a symbol dictionary  $\Delta_C$ , consisting of two tables mapping frequent symbols to natural numbers and back, and an adaptive probability distribution  $\rho^C$ . This distribution includes among others a special symbol *esc* for new characters.

*Coding:* The coding procedure of a symbol  $\alpha$  has two alternatives. If the symbol is in the dictionary  $\Delta^C$ , we emit simply the corresponding code:

$$K(\alpha|M^C) = K(\Delta^C(\alpha)|\rho^C)$$

with increase of a relative frequency of  $\Delta^C(\alpha)$ .

If  $\alpha$  is new, we put *esc* symbol on the output followed by full 32-bit beta code of the character:

$$K(\alpha|M^C) = K(esc|\rho^C) \cdot \beta^{32}(\alpha)$$

New element is subsequently added into the  $\Delta^C$  and the  $\rho^C$ .

## V. RESULTS

We test there our method in various configurations of the initial dictionaries and the coding algorithm of basic elements.

We tested the mixtures of the initial dictionaries (**A**<sub>1</sub>...**A**<sub>5</sub>) on the testing set  $T_B$  containing 100 shorter documents (as newspaper articles). We show the results in table II in bpc compared with commonly used methods bzip2 and gzip.

We see the expected improvement of a compression rate for small documents. The variant **A**<sub>3</sub> already overcame both methods.

Next we present dependency of a compression rate on a size of a compressed file. We split the set  $T_A$  by size

Table III  
COMPARISON OF COMPRESSION RATIOS FOR VARIOUS FILE SIZES IN  $T_A$ .

Method	10–100 kB	100–500 kB	500–1000 kB
<b>A</b> <sub>1</sub>	4.21	3.78	3.69
<b>A</b> <sub>2</sub>	4.00	3.61	3.51
<b>A</b> <sub>3</sub>	3.76	3.42	3.34
<b>A</b> <sub>4</sub>	3.47	3.22	3.16
<b>A</b> <sub>5</sub>	3.10	2.95	2.93
gzip	3.69	3.46	3.47
bzip2	3.35	2.71	2.57

into three categories and observed results within them. The results are in table III.

There is again revealed that variant **A**<sub>5</sub> achieves best results for all sizes of an input file. Due to this initialization the method overcame successful algorithm bzip2 based on a Burrows-Wheeler transform.

We performed the compression and decompression of the set  $T_A$  concatenated into one file in order to find out time demands of the algorithm. The compression of such concatenated file with size 4.76 MB with the variant **A**<sub>5</sub> took about 82 seconds on the computer with processor unit AMD Sempron 2800+, 1.60 GHz and 960 MB RAM. The decompression took about 4 seconds. Here we can observe the asymmetry between compression and decompression phases.

## VI. CONCLUSIONS AND FUTURE WORKS

In this paper we have introduced the method of extraction of information from an input text which would allows better modelling of this text and improve our ability to compress it. Such knowledge has been for us the sequence of tags of words in the text which represent the words' belonging to the parts of speech. The sequence of tags for a sentence is called a sentence type. This makes possible for us to introduce knowledge about the structure of a sentence and functions of particular words in an utterance.

In order to get the sentence type, we have proposed the approximative part-of-speech tagger for the Czech language, the NNTagger, based on neural networks. Both procedures – the morphological analysis and disambiguation – were implemented by Backpropagation nets. Such an approach was motivated by a great amount of information contained in the regular structures of words. The NNTagger was trained on the corpus of tagged texts and tested in various configurations. The slightly lower accuracy of this method is compensated by the absence of any dictionary in both phases and justified by the observation that a compression ratio is not sensitive to small errors.

We have constructed model structures of texts on different layers (a symbol, word, sentence type, sentence part) with respect to the additional knowledge (part of speech, kind of symbol). The coding and decoding procedures have been

implemented based on these models and for that purpose a novel notation has been introduced.

We proposed the methods of creating databases of frequent words and sentence types from a set of Czech documents in order to set initially each of the models. Such initialization can improve the resulting compression ratio especially in domain of small documents. However, each of the models contributes to this improvement with another proportion due to its different function in the language. Therefore, several possible mixtures of the initial dictionaries have been compiled with respect to the improvement of a compression ratio and the size of the databases.

The method has been tested in various configurations of the initialization databases and results compared with both commonly used compression programs and similar purely word-based compression methods. These results proved an advantage of coding of the sentence type as information about syntactic structures separately from the words of a sentence. The creation of the initialization mixtures shows that the sets containing words of a part of speech have different properties, e.g., the variability both in a document and among different documents, the distribution of word lengths etc., and their models must be initialized differently. With such an initialization, the algorithm overcame commonly used compression algorithms *gzip* and *bzip2* in compression of small files up to 100 kB.

Features of our solution can be summarized as follows:

- Approximative tagger NNTagger had slightly lower accuracy than usually used Czech tagging algorithms. On the other hand, it does not need any additional dictionary of word forms or lemmas.
- The method of exploitation of part-of-speech tags sequences as a linguistic information in a sentence has been proposed which allows us better to grasp model of text generation.
- Our text compression algorithm gained 17–27 % better compression ratio in comparison with algorithms based only on words with comparable settings and initialization.

The experiments with the linguistic compression in the Czech language have shown some possible challenges for a future work. The tagging method NNTagger has recorded good results regarding the low costs on additional data; however it has lower accuracy than other methods, particularly on non-inflective words. Thus some of the more sophisticated methods should be used and the effect of a tagger's accuracy on the compression should be tested. Also the frequency of unknown sentence types has been relatively high during the compression, so we plan to implement other string compression method, e.g., PPM [20]. Finally, an open question is an application of analogous compression method on languages with different characteristics. For instance, due to less stress on the word morphology in the English language, appropriate tagger has to be used, the morpholog-

ical phase of NNTagger should be replaced by dictionary-based method, and another corresponding tag set has to be chosen. On the other side, fixed word order can cause an increase in the significance of sentence types and possibly their better compression ratio. The symbol-based languages, like Chinese, raise other challenges. In such languages, the word segmentation is a hard problem which has to be solved by non-trivial algorithm. Also the model of words would be modified with respect to different relation between words and characters.

#### Acknowledgement

This work has been supported by the Internal Grant Agency of VŠFS under project IGA VŠFS 7721.

#### REFERENCES

- [1] Abel J. and Teahan W. Universal Text Preprocessing for Data Compression. In: *IEEE Transactions on Computers*, **54** (5), pp. 497–507, (2005)
- [2] Adiego J. and Fuente P. On the Use of Words as Source Alphabet Symbols in PPM. In: Storer JA, Cohn M (Eds.). *Proceedings of 2006 IEEE Data Compression Conference*, IEEE Computer Society Press, Los Alamitos, California, USA (2006) 435.
- [3] Brill E. *A Corpus-Based Approach to Language Learning*, dissertation, Department of Computer and Information Science, University of Pennsylvania; Philadelphia, 1993.
- [4] Chapin B. and Tate SR. Higher Compression from the Burrows-Wheeler Transform by Modified Sorting. In: Storer JA, Cohn M (Eds.). *Proceedings of 1998 IEEE Data Compression Conference*, IEEE Computer Society Press, Los Alamitos, California, USA (1998) 532.
- [5] Collins M. Discriminative Training Methods for Hidden Markov Models. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*. Morristown, NJ, 2002; 1–8.
- [6] Cutting D., Kupiec J., Pedersen J., and Sibun P. A Practical Part-of-Speech Tagger. In *Proceedings of the Third Conference on Applied Natural Language Processing*. Trento, 1992.
- [7] Dvorský J., Pokorný J., and Snášel V. Word-based Compression Methods for Large Text Documents. In: Storer JA, Cohn M (Eds.). *Proceedings of 1999 IEEE Data Compression Conference*, IEEE Computer Society Press, Los Alamitos, California, USA (1999) 523.
- [8] Gailly J. *The gzip home page*. <http://www.gzip.org/> [1 March 2010].
- [9] Hajič J. *Disambiguation of Rich Inflection (Computational Morphology of Czech)*, Karolinum; Prague, 2004.
- [10] Hajič J. and Hladká B. Tagging Inflective Languages: Prediction of Morphological Categories for a Rich, Structured Tagset. In *Proceedings of COLING-ACL Conference*. Montreal, 1998; 483–490.

- [11] Hana J. and Zeman D. *Manual for Morphological Annotation*. Revision for the Prague Dependency Treebank 2.0. ÚFAL Technical Report No. 2005–27. Charles University; Prague, 2005.
- [12] Honetschläger V et al. *The Prague Dependency Treebank*, ÚFAL MFF UK; Prague, 2006. <http://www.ufal.mff.cuni.cz/pdt2.0/> [1 March 2010].
- [13] Horspool RN. and Cormack GV. A general purpose data compression technique with practical applications. *Proceedings of the CIPS Session 84*, (1984) 138-141.
- [14] Horspool RN. and Cormack GV. Constructing WordBased Text Compression Algorithms. In: Storer JA, Cohn M (Eds.). *Proceedings of 1992 IEEE Data Compression Conference*, IEEE Computer Society Press, Los Alamitos, California, USA (1992) 62-71.
- [15] Isal RYK. and Moffat A. Word-based Block-sorting Text Compression. In: *Proceedings of the 24th Australasian conference on Computer science*, Gold Coast, Queensland, Australia, (2001) 9299.
- [16] Kochánek J., Lánský J., Uzel P., and Žemlička M. Multi-stream Compression. In: Storer JA, Marcellin MW (Eds.). *Proceedings of 2008 Data Compression Conference*, IEEE Computer Society Press, Los Alamitos, California, USA, (2008) pg. 527.
- [17] Korodi G., Rissanen J., and Tabus I. Lossless Data Compression Using Optimal Tree Machines. In: Storer JA, Cohn M (Eds.). *Proceedings of 2005 IEEE Data Compression Conference*, IEEE Computer Society Press, Los Alamitos, California, USA (2005) 348357.
- [18] Kuthan T., and Lánský J. Genetic algorithms in syllable based text compression. In: Pokorný J, Snášel V, and Richta K (Eds.). *Proceedings of the DATESO 2007 Annual International Workshop on Databases, Texts, Specifications and Objects*. CEUR-WS, Vol. 235, (2007) 21-34.
- [19] Lánský J., Chernik K., and Vlčková Z. Comparison of Text Models for BWT. In: Storer JA, Marcellin MW (Eds.). *Proceedings of 2007 IEEE Data Compression Conference*, IEEE Computer Society Press, Los Alamitos, California, USA (2007) 389.
- [20] Lánský J., Šesták R., Uzel P., Kovalčín S., Kumičák P., Urban T., and Szabó M. XBW - *Word-based compression of non-valid XML documents*. <http://xbw.sourceforge.net/> [1 March 2010].
- [21] Lánský J. and Žemlička M. Compression of Small Text Files Using Syllables. In: Storer JA, Cohn M (Eds.). *Proceedings of 2006 IEEE Data Compression Conference*, IEEE Computer Society Press, Los Alamitos, California, USA (2006) 458.
- [22] Lánský J. and Žemlička M. Text Compression: Syllables. In: Richta K, Snášel V, Pokorný J (Eds.). *Proceedings of the DATESO 2005 Annual International Workshop on Databases, Texts, Specifications and Objects*. CEUR-WS, Vol. 129, (2005) 32-45.
- [23] Manning C. and Schütze H. *Foundations of Statistical Natural Language Processing*, MIT Press; Cambridge, 1999.
- [24] Moffat A. Word based text compression. In: *Software-Practice and Experience* **19** (2), (1989) 185198.
- [25] Moura ES., Navarro G., Ziviani N., and Baeza-Yates R. Fast searching on compressed text allowing errors. In: *Proceedings of the 21st annual international ACM SIGIR conference on Research and development in information retrieval*. ACM Press, New York, (1998) 298-306.
- [26] Ratnaparkhi A. A Maximum Entropy Model for Part-of-Speech Tagging. In *Proceedings of the First Empirical Methods in Natural Language Processing Conference*. Philadelphia, 1996; 133-141.
- [27] Rein S., Gühmann C., and Fitzek F. Compression of Short Text on Embedded Systems. *Journal of Computers*, Vol. 1, No. 6, (2006).
- [28] Ryabko BY. Data compression by means of a book stack. *Prob. Inf. Transm.*, **16**(4), (1980). In Russian.
- [29] Schmid H. Part-of-speech tagging with neural networks. In *Proceedings of the International Conference on New Methods in Language Processing*. Manchester, 1994; 44-49.
- [30] Schmid H. Probabilistic Part-of-Speech Tagging Using Decision Trees. In *Proceedings of the 15th COLING Conference*. Kyoto, 1994; 172-176.
- [31] Sedláček R. *Morfologický analyzátor češtiny*, diploma thesis, FI MU; Brno, 1999. In Czech.
- [32] Šesták R., and Lánský J. Compression of Concatenated Web Pages Using XBW. In: Geffert V et al. (Eds.). *SOFSEM 2008*, LNCS 4910, Springer-Verlag Berlin, Heidelberg, Germany, (2008) pg. 743-754.
- [33] Šesták R., Lánský J., and Žemlička M. Suffix Array for Large Alphabet. In: Storer JA, Marcellin, MW (Eds.). *Proceedings of 2008 Data Compression Conference*, IEEE Computer Society Press, Los Alamitos, California, USA, (2008) pg. 543.
- [34] Seward H. *bzip2: Documentation*. <http://www.bzip.org/1.0.5/bzip2-manual-1.0.5.pdf> [1 March 2010].
- [35] Shkarin D. *Durilca Light and Durilca 0.4b*. <http://compression.graphicon.ru/ds/> [3 February 2007].
- [36] Skibinski P. *Reversible data transforms that improve effectiveness of universal lossless data compression*. Doctor of Philosophy Dissertation, University of Wrocław, (2006).
- [37] Šíma J., and Neruda R. *Teoretické otázky neuronových sítí*, Matfyzpress; Prague, 1997. In Czech.
- [38] Teahan W. *Modelling English text*. Ph.D. dissertation, University of Waikato, New Zealand, (1998).
- [39] Üçölük G., and Toroslu H. A Genetic Algorithm Approach for Verification of the Syllable Based Text Compression Technique. *Journal of Information Science*, Vol. 23, No. 5, (1997) 365-372.
- [40] Votrubec J. *Volba vhodné sady rysů pro morfologické značkování češtiny*, diploma thesis, MFF UK; Prague, 2005. In Czech.

## Dynamic Codec Selection Algorithm for VoIP

Maja Sulovic

Faculty of Electrical Engineering  
Department of Telecommunications  
Sarajevo, Bosnia and Herzegovina  
maja.sulovic@etf.unsa.ba

Mesud Hadzialic

Faculty of Electrical Engineering  
Department of Telecommunications  
Sarajevo, Bosnia and Herzegovina  
mhadzialic@etf.unsa.ba

Darijo Raca

Faculty of Electrical Engineering  
Department of Telecommunications  
Sarajevo, Bosnia and Herzegovina  
draca@etf.unsa.ba

Nasuf Hadziahmetovic

Faculty of Electrical Engineering  
Department of Telecommunications  
Sarajevo, Bosnia and Herzegovina  
nhadziahmetovic@etf.unsa.ba

**Abstract**—This paper gives a proposal on algorithm for adaptive adjustment of VoIP sources transmission rate based on voice quality estimated at the receiver. This adjustment is achieved through the appropriate use of differing voice codecs, as changes in network conditions occur, in order to maintain an efficient utilization of the available resources. We propose a simple algorithm for dynamic selection of voice codec, depending on network conditions during the on-going voice session. Algorithm is embedded in the source code of the programming environment OPNET Modeler 14.5. Simulation results show that the proposed algorithm makes better use of the available bandwidth, achieving superior performance in comparison to the situation without implementation of algorithm for adaptive codec selection.

**Keywords**—dynamic codec selection; speech codec; packet delay; packet loss; MOS

### I. INTRODUCTION

Nowadays, VoIP is widely accepted technology used for transmission of voice over IP networks. Packets belonging to the same VoIP session may travel independently through different paths in the network. In this paper various codecs such as G.711, G.729 and G.723 will be discussed. Each of these codecs use a different technique for sample coding and different levels of compression that directly affects the voice session quality. Selection of the right codec represents a compromise between desired performance and available resources in the network.

In this paper, as criteria for determining the quality of the voice session we use E-model, which estimates user's satisfaction through R value, on the basis of the used codec, delay and loss in the network. Since metrics values on the network level in IP network are variable, this will practically mean that during the session, same codec may not be optimal in all moments. There are different techniques used for solving this problem, e.g., AMR (Adaptive Multi Rate) codecs, implementation of additional access mechanisms at application layer, the dynamic selection of codec during the sessions and others. This paper addresses the last mentioned technique, and introduces a

simple algorithm, which estimates the optimal codec during voice session, based on current and past values of average delay and current value of packet loss.

As authors of this paper, we are familiar with research work dealing with this problem. In [1], codec selection algorithm is based on average delay value which is then compared with current delay value. Algorithm verification is based on comparison between different amount of available bandwidth scenarios. It has been shown that growth of available bandwidth as a consequence, increases MOS value when using adaptive codec selection method. Similar results are also shown in [2, 3]. In [4], authors proposed adaptive multi-rate VoIP control scheme that adapts the voice encoding rate and packetization interval in relation to transmission rate in the PHY layer. In [5] authors proposed algorithm which, based on information extracted from RTCP packets and MAC layer, dynamically adapt codec for ongoing VoIP calls. In [6] authors described an end-to-end based adaptation, which adjusts application parameters to changing network conditions in order to achieve better bandwidth utilization and QoS, by employing adaptive codec switching techniques to further enhance QoS.

In this paper, the called party of a voice session is measuring the average packet delay and loss. It periodically sends reports to the calling party that is responsible for session management. Based on the received reports, algorithm that is implemented on the calling party learns about the network by memorizing the minimum and the maximum value of average packet delay and packet loss for each codec that has been used in the session. Memorized values are then compared with reported average delay and packet loss. If the algorithm determines that there are improvements of network conditions (e.g., reported average delay is less than the minimum average delay for currently used codec that algorithm has knowledge of until the moment of observing), the algorithm makes a decision about switching to a new codec with lower compression ratio. Based on the feedback in the following report algorithm

decides whether the previous decision on codec change was correct or not. Incorrect decisions will occur in situations when despite new minimum average delay being detected by the algorithm, available resources in the network are not sufficient enough to use codec with lower compression ratio. All incorrect codec switching decisions will then reflect in the following report, through reported average delay that is higher than the maximum delay allowed for the voice session. The algorithm then concludes that value of the minimum delay, previously memorized, should not be used as threshold value for switching to a better codec and the session codec will switch to a codec with a higher compression ratio, thus enabling the algorithm to recognize similar situations in the future and prevent incorrect codec switch decisions from occurring.

The proposed algorithm was implemented in simulation package OPNET Modeler 14.5. Used version of the Modeler does not have similar algorithm for dynamic codec selection that has been implemented. To the best author knowledge there is no similar solution to this problem.

## II. OVERVIEW OF SPEECH CODECS

There are various codecs specified by the ITU-T. Codecs have different performance and impact on the voice quality due to different degree of compression. High degree of compression results with higher compression delay and increases loss sensitivity compared to codecs with low or no compression. Contrary to this, codecs with high degree of compression have less bandwidth requirements, and thus have better performance in network congestion situations. Therefore, it is necessary to select the appropriate codec to obtain best quality of voice with the lowest bandwidth requirements [7].

The G.711 codec does not use any compression; it has 8-kHz sampling rate, requires 64 kbit/s of audio bandwidth and provides very good quality level. The G.729 codec is computationally complex, but provides significant bandwidth savings. It has 8:1 compression and requires just 8 kbit/s of audio bandwidth. The maximum achievable MOS is about 3.9. The G.723.1 codec is mostly used in VoIP applications due to its low bandwidth requirement. There are two versions of this codec, with bit rates at 5.3 kbit/s and 6.3 kbit/s. Every codec adds additional delay to the total packet transmission delay due to signal encoding, decoding, compression and decompression. Main characteristics of the codecs mentioned are shown in Table 1 [8, 9, 10].

TABLE 1. CHARACTERISTIC OF CODECS

Codec	Bit Rate (kbit/s)	Link Utilization (kbit/s)	Delay (ms)	Loss (%)	MOS
G.711	64	87.2	0.125	7-10	4.10
G.729	8.0	31.2	15	< 2	3.92
G.723.1	5.3	20.8	37.5	< 1	3.65

## III. IMPACT OF NETWORK LAYER METRICS ON QUALITY OF SPEECH

QoS concept is observed as layered model and is defined at user, application and system layer. Quality of service at observed layer of QoS layered model represents characterization of expected quality, which should be achieved during transfer of data units. Between any two layers, it is important to determine mapping between expected performance at lower layer and its impact on QoS parameters on higher layer. The main question is how given QoS guarantees, at layer N-1, impact performance metrics at layer N [11, 12].

ITU-T E-model represents an analytical model of voice quality defined in the ITU-T recommendation G.107. E-model provides a framework for real-time on-line quality estimation from network performance measurement (e.g., delay and loss characteristics) and application level factors (e.g., low bit rate codecs). The result of the E-model is the calculation of the R-factor. The R-factor can be further translated into MOS scale through these expressions:

$$MOS = \begin{cases} 1 & R < 0 \\ 1 + 0.035R + 7 \cdot 10^{-6} R(R - 60)(100 - R) & 0 < R < 100 \\ 4.5 & R > 100 \end{cases} \quad (1)$$

The R-factor is defined as:

$$R = R_0 - I_s - I_d - I_e + A \quad (2)$$

$R_0$  represents the basic signal-to-noise ratio.  $I_s$  reflects the impairments occurring simultaneously with the voice signal due to quantization. It is a function of several parameters, none of which are related to the underlying packet transport.  $I_d$  models the impairments caused by one-way delay. Voice quality degrades more rapidly when this delay exceeds 177.3 ms. This effect is modeled using following expression:

$$I_d = 0.024d + 0.11(d - 177.3)H(d - 177.3) \quad (3)$$

where  $d$  is the one-way delay (in milliseconds) and  $H(x) = 0$  za  $x < 0$  i  $H(x) = 1$  when  $x \geq 0$ .  $I_e$  is the equipment impairment factor that covers the distortion of the original voice signal due to low-rate codec and packet loss in both, network and playout buffer.  $I_e$  value is codec dependable. The advantage factor  $A$  represents the measure of the willingness of a VoIP user to trade call quality for convenience. [13]

VoIP application is very sensitive to delay. Acceptable one-way delay according to ITU-T G.114 recommendation is 150 ms. Delay between 150 ms and 400 ms makes the conversation possible, but considerably annoying. Delay over 400 ms is unacceptable according to ITU-T G.1010. Packet loss also must be managed or controlled in VoIP,

since its effect on VoIP is treated as noise. Unlike delay, VoIP can tolerate packet loss to some extent. [14]

#### IV. PROPOSED ALGORITHM FOR DYNAMIC SPEECH CODEC SELECTION

After review of most commonly used codecs and QoS metrics which evaluate performance and dependencies between QoS metrics from different layers of layered QoS model, this chapter presents details of proposed algorithm for dynamic codec selection.

##### A. Algorithm Design

Next steps present design of proposed algorithm in detail:

1. Define initial minimum delay values for codecs G.729A and G.723.1 5.3K
2. Define initial maximum delay values for codecs G.711 and G.729
3. Define initial maximum loss values for codecs G.711, G.729A and G.723.1 5.3K
4. Store the value of currently used codec and its average delay from report in an array that stores up to three previously used codecs and their corresponding delays from previous reports.
5. If codec changes occurred at least three times until current moment of the call session, then determine new minimum average delay if possible
  - 5.1. If current and penultimate used codecs are the same, go to step 5.1.1, else go to step 6
    - 5.1.1. If previously stored average delay is less than the current average delay, go to step 5.1.2, else go to step 6
    - 5.1.2. store average delay from the previous report into variable minimum average delay for codec used during the receiving of the previous report
6. Compare received average delay and loss with minimum average delay, maximum average delay and maximum packet loss
  - 6.1. If average delay is less than minimum average delay, change codec according to:
    - 6.1.1. If G.711, then G.729A
    - 6.1.2. If G.729, then G.723.1 5.3K
  - 6.2. If average delay is higher than maximum average delay, change codec according to:
    - 6.2.1. If G.723.1 5.3K, then G.729A
    - 6.2.2. If G.729A, then G.711
  - 6.3. Return to step 4

Step 1 defines initial minimum average delay for G.723.1 5.3K and G.729A codecs. Algorithm compares average delay from the report and minimum average delay value to assume whether network conditions have been improved. If so, algorithm assumes that there has been a release of resources in the network and switches to the codec with higher bandwidth requirements in order to

improve speech quality that is expressed through the MOS value.

Steps 2 and 3 are analogous to step 1. They define maximum value of packet delay and packet loss for codecs G.711 and G.729. If average delay and loss from the currently received report are higher when compared to stored maximum values, algorithm concludes that there is network congestion occurrence, and that is necessary to use a codec that requires less bandwidth.

In step 4, algorithm learns about conditions in the network. Algorithm stores the information about three last used codecs and its resulting average delays in the array.

Step 5 is the crucial step of the algorithm. It will be explained using the following example. If voice application uses codec G.711 and if average delay from current report is higher than the maximum allowed delay for G.711, algorithm switches the session codec to one that requires less bandwidth (G.729A). Next statistical report will indicate latency reduction due to the fact that G.729A codec has lower bandwidth consumption and thus lower contribution on total network load on bottleneck. Upon receiving this report, algorithm compares last three stored delays and may assume that there was a release of resources in the network, and try to re-use the codec G.711. However, if there has been no release of resources in the network, delays in the forthcoming report for the codec G.711 will again have a high value. This way, algorithm calculates delay value for G.729A which should not trigger codec change and updates the variable that stores the minimum delay for codec G.729A. Without this step, algorithm would trigger codec change upon receiving any statistical report that is indicating latency reduction. Application would switch session codec to a codec with lower compression ratio and that would again worsen the situation in the bottleneck of the network. Described situation resulting effect would be a constant switching from codec to codec.

Step 6 includes decision making on the optimal codec for current conditions in the network. The algorithm compares average delay from last received report with minimum and maximum delay and maximum packet loss defined for currently used codec. If current average delay exceeds maximum delay or maximum packet loss, algorithm switches the session codec to one with lower bandwidth requirements. Contrary to this, if current average delay is below minimum delay, algorithm switches the session codec to one with higher bandwidth requirements.

##### B. Implementation in OPNET Modeler 14.5

This section provides a brief description of the proposed algorithm implementation in the OPNET Modeler. To implement dynamic codec selection algorithm in the OPNET Modeler, we had to modify process models, which implement the calling party as well as the called party. Process in OPNET Modeler is modeled as a finite state machine (FSM) that allowed us to specify the C/C++ code that implements a process.

For sending received statistics report from the called party to the calling party, firstly received statistics (average delay and loss) on the called party is collected and afterwards collected data is forwarded to the calling party. Statistical data are collected in “receive” state in the called party process editor, see Figure 1. Upon receiving voice packet, “receive” state calculates current average packet delay and packet loss. Report sending interval is set to 5 seconds. Once this interval is ended, process model that represents called party switches its state and enters into state responsible for sending collected data to calling party. If average delay exceeds 300 ms before 5 s interval is completed, called party sends additional early report for fast adaptation to network conditions change. For forwarding statistical data to the calling party a new “send report” state is created, within called party process. Newly created state will be responsible for sending periodic reports that are sent via RTP.

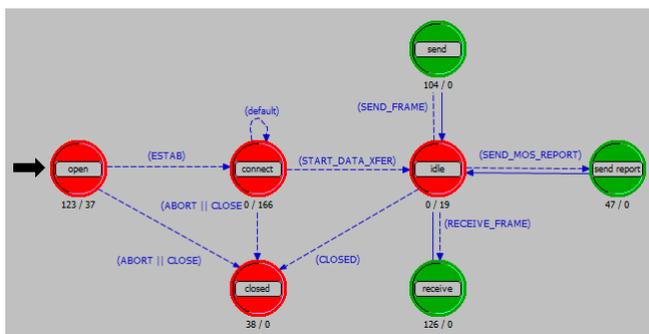


Figure 1. Called party process editor

Algorithm described in Section 4.1 is implemented in “receive” state in the calling party process editor. If algorithm determines that the codec in use needs to be changed, calling party changes currently used session codec. From that moment on, every voice packet sent to the called party includes information about new session codec, which is stored in the packet header. [15]

V. SIMULATION RESULTS

Simulation of simple network containing two end nodes is created for the purpose of verification of the proposed algorithm. Our goal was to create three cases where available bandwidth is optimized for voice traffic transmission with only one of the mentioned codecs. Nodes support voice and video applications. The voice application is set to default OPNET setting IP Telephony. Video applications settings are set to achieve mentioned goals, and can be seen in Table 2. Thus we created three profiles, VoIP Profile, Video High Load Profile and Video Low Load Profile. Video High Load Profile bitrate is 20 kbit/s; Video Low Load Profile bitrate is 10 kbit/s.

Elements presented in Table 3 are used in the simulation. Network consists of two core routers. Link between these routers has a capacity of 160 kbit/s, and represents a bottleneck of the topology. Each core router has 10 Mbit/s connection with a switch. Each switch has 10 Mbit/s link connection with workstation. Link between routers has FIFO queuing implemented, where maximum number of packets in the buffer is 100. In order to reduce simulation duration, we choose to simulate simple network topology. Therefore, although our network topology is very simple, our simulation results are valid and can, with proper equipment, be tested in more realistic case studies. Measured metrics in this simulation are average delay at the network level, sent/received number of bytes and MOS value.

TABLE 2. ELEMENTS USED IN SIMULATION

Device Name	Device Description
CS 7609	Core routers in simulation
CS 6509	Access switch in simulation
ethernet_wkstn	This node model represents a workstation with client-server applications running over TCP/IP and UDP/IP
10BaseT	Connection between workstation and switch
ppp_adv	Data Rate is 160000; Connection between core routers

Figure 2 shows a comparison between delays in the cases when using a fixed codec and when using a dynamic codec selection. Initial codec in the last case is codec G.729A. It can be seen that during the Video High Load profile codecs G.711 and G.729A have unacceptable delays which are approximately 1.1 s and 0.7 s, respectively. After the termination of the Video High Load profile, delay values in the case when using fixed G.711, G.729A and G.723.1 5.3K codecs have acceptable delay values below 100 ms. When new video profile Video Low Load starts, delay value for G.711 codec increases its value to approximately 0.9 s, while delay values for G.723.1 3.5K and G.729A codecs increase to 200 ms at most. In fourth case, which relates to the application of the proposed algorithm for dynamic codec selection, Figure 2 clearly shows moments of codec change and adjustment to network conditions. It may be noted that in this case, delay during Video High Load profile is approximately equal to the delay when using fixed G.723.1 5.3K codec. This is due to the fact that the algorithm chose G.723.1 5.3K codec as the optimal codec. At the time of termination of the Video High Load profile, there is a gradual change of codecs, first with the change from G.723.1 5.3K to G.729A and afterwards from G.729A to G.711. After Video Low Load profile starts, slight increase in delay value can be observed, after which the algorithm concludes that there might be congestion in the network, and switches codec from G.711 to G.729A.

Figure 3 shows the MOS value in cases analogous to those with previous images. One can clearly see the advantage of using an algorithm for dynamic codec selection, where the MOS value in all the session moments

is approximately equal to the maximum MOS values in the other three cases.

If we observe the percentage of lost packets shown in Table IV, the worst results are obtained by using the G.711 codec, where packet loss is 7%. G.729A and G.723.1 5.3k codecs both achieve packet loss of 1%. Application of the proposed algorithm for dynamic codec selection results with no voice packet loss. Average values for delay and MOS shown in Table 4 will provide a more complete picture of the benefits of the proposed algorithm. Using proposed algorithm, we obtain best results in terms of MOS values 2.44, compared to other values obtained for G.711, G.729A and G.723.1 5.3k, respectively. The same conclusion applies to average delay, where using the algorithm proposed results with 189.9 ms delay, whereas other cases result with higher delay values.

TABLE 3. AVERAGE DELAY AND PACKET LOSS

	G.711	G.729A	G.723.1 5.3K	Dynamic Codec Selection
Delay (ms)	728.8	262.0	206.2	189.9
Packet loss (%)	7	1	1	0
MOS	1.5120	2.1488	1.9419	2.4346

## VI. CONCLUSIONS AND FUTURE WORK

In this paper, we presented standard voice codecs used for transmission of voice traffic over the Internet. Basic characteristics of each of the discussed codec, its advantages and disadvantages are expressed through QoS metrics and average delay and loss at the network level, which ultimately results in the uniform assessment of quality, represented by the MOS value at user's level. Main goal for a voice over Internet session is to meet QoS recommendations and to simultaneously achieve the highest possible MOS value. In network congested condition, low rate codec has better performance than high rate codec. In contrast, when there is no network congestion, high rate codec has better performance. This means that the best network performance could be achieved through trade-off between codec bandwidth requirements and desired quality. In other words, voice session will adapt sending rate to available bandwidth. Since usage of fixed codec during one voice session cannot achieve this trade-off, we proposed a simple algorithm for dynamic selection of the codec depending on the current conditions in the network. We showed that use of a proposed algorithm maintains high level of MOS value in cases of network congestion.

In [3], main goal is to reduce delay when congestion occurs, and algorithm is based on TCP Vegas-like congestion avoidance technique, for the rate and loss control of VoIP flows over the WLAN. The idea for our algorithm proposal is also based on delay and packet loss similar to [3], but, unlike aforementioned, it does not make difference between congestion loss and error loss of packets (due to wireless environment). Also, in [3], authors do not analyse impact of decision of algorithm on final MOS values. The

algorithm in [3] is tested with more than one parallel calls, and also takes into account fairness between VoIP calls. Because of nature of our simulation, which is simple, we do not take this into account. Our algorithm is similar to one proposed in [5], but with one main difference. All of these algorithms have memoryless property that means that they do not memorize previous network conditions states. Our algorithm takes this parameter into account and based on that makes further decisions.

The next step is to examine the influence of report sending interval length on the adaptation process, which actually represents a trade-off between desired quality of conversation and adaptation rate to network changes. In future work, we intend to investigate impact of proposed algorithm in wireless environment and create real application based on proposed algorithm. Also we intend to simulate more than one call in the same time and analyse fairness impact on flows as it is done in [3].

## REFERENCES

- [1] A. Ossipov, "Dynamic Voice Codec Determination Mechanism for VoIP," MS Thesis, Wichita State University, Kansas, USA, May 2006.
- [2] A. Servetti and J.C. De Martin, "Adaptive interactive speech transmission over 802.11 wireless LANs," Proc. IEEE Int. Workshop on DSP in mobile and Vehicular Systems, Nagoja, Japan, April 2003., pp. 1-4.
- [3] A. Trad, Q. Ni, and H. Affi, "Adaptive VoIP Transmission over Heterogeneous Wired/Wireless Networks," 2nd International Workshop MIPS 2004, LNCS, Vol. 3311, Grenoble, France, November 2004, pp. 25-36.
- [4] T. Kawata and H. Yamada, "Adaptive Multi-Rate VoIP for IEEE 802.11 Wireless Networks with Link Adaptation Function," Global Telecommunications Conference, GLOBECOM '06. IEEE, San Francisco, USA, December 2006, pp. 1-5.
- [5] A. Sfairopoulou, C. Macián, and B. Bellalta, "VoIP codec adaptation algorithm in multirate 802.11 WLANs: distributed vs. centralized performance comparison," LNCS, 2007, Volume 4606/2007, Barcelona, Spain, pp. 52-61.
- [6] See Leng Ng, S. Hoh, and D. Singh, "Effectiveness of adaptive codec switching VoIP application over heterogeneous networks," IEEE 2nd International Conference on Mobile Technology, Applications and Systems, Guangzhou, China, November 2005, pp. 1-7.
- [7] H. Kazemitabar, S. Ahmed, K. Nisar, A.Said, and H. Hasbullah, "A comprehensive review on VoIP over Wireless LAN networks," ISSR Journal, Vol. 2, No. 2, September 2010, pp. 1-16.
- [8] ITU-T Recommendation G.711, "Pulse code modulation (PCM) of voice frequencies," 1988.
- [9] ITU-T Recommendation G.729, "Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear prediction (CS-ACELP)," 1996.
- [10] ITU-T Recommendation G.723.1, "Dual rate speech coder for multimedia communications transmitting at 5.3 and 6.3 kbit/s," 2006.
- [11] Kun I. Park, "QoS in Packet Networks," Springer, October 2004.
- [12] T. Braun, M. Diaz, J. Enriquez-Gabeiras, and T. Staub, "End-to-End Quality of Service Over Heterogeneous Networks," Springer, October 2008.
- [13] ITU-T Recommendation G.107, "The E-Model, a Computational Model for Use in Transmission Planning," March 2005.
- [14] K. Perlicki, "Simple analysis of the impact of packet loss and delay on voice transmission quality," Journal of Telecommunications and Information Technology, February 2002, pp. 53-56.

[15] "OPNET Modeler Documentation Set," Version 14.5, 2010.

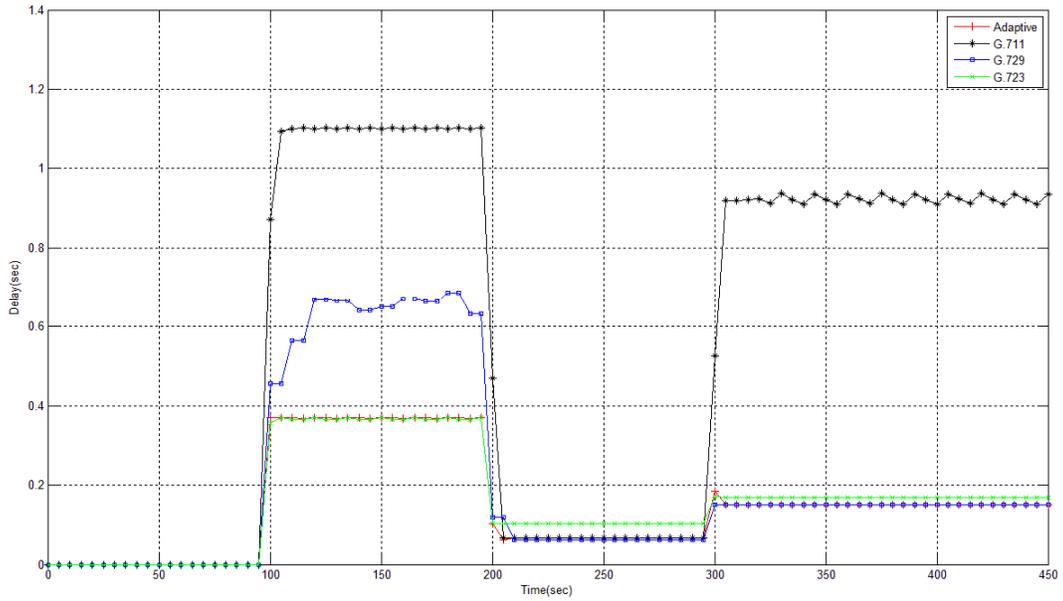


Figure 2. Delay comparison between using fixed codec and adaptive selection of speech codec

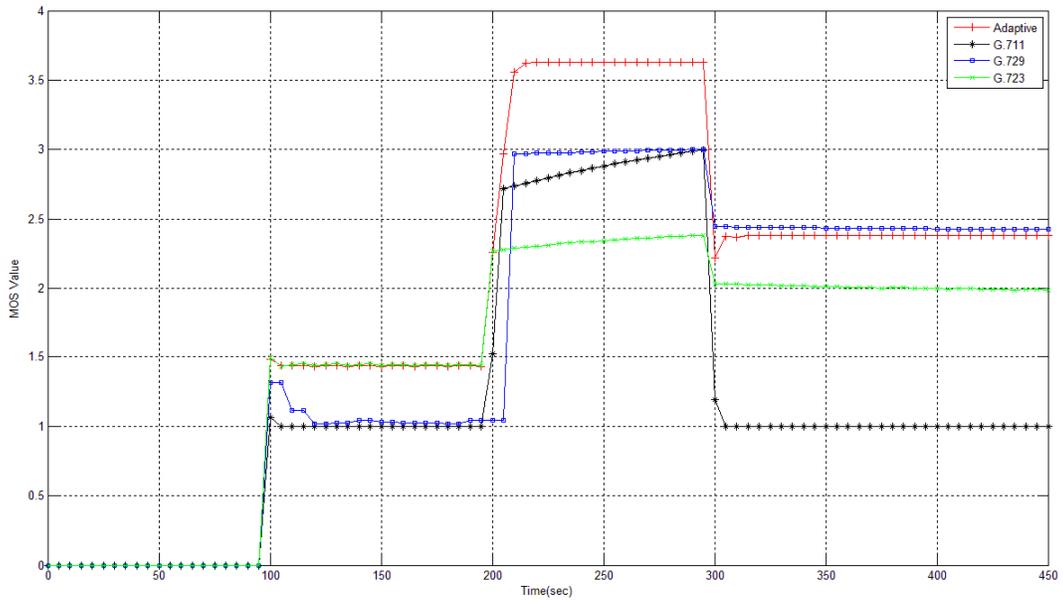


Figure 3. Obtained MOS value

## E-Model based Adaptive Jitter Buffer with Time-Scaling Embedded in AMR decoder

Liyun Pang

Audio Technology Team  
Huawei European Research Center  
Munich, Germany  
liyun.pang@huawei.com

Laszlo Böszörményi

Department of Information Technology  
University Klagenfurt  
Klagenfurt, Austria  
office-lb@itec.uni-klu.ac.at

**Abstract**—The premier factors affecting voice quality in packet networks are latency, jitter and packet loss. Jitter buffers are commonly used to counter jitter introduced by queuing in packet switched networks so that a continuous play-out of voice transmitted over the network can be ensured. In this paper, a new adaptive jitter buffer algorithm is proposed. The algorithm consists of an adaptive play-out algorithm based on the extended E-Model with spike detection and a time scaling technique relying on a speech classification mechanism embedded in the AMR decoder. Simulation results show that the proposed algorithm outperforms the best existing algorithms for random packet loss under various network scenarios.

**Keywords** - adaptive jitter buffer; E-Model; AMR; time-scaling

### I. INTRODUCTION

In recent years, the transport of Voice over IP (VoIP) has gained popularity and is becoming widely used. VoIP can be seen as a replacement to traditional circuit switched telephony with the advantages of cost reduction, simplified network and simplified network management. Voice quality is a critical parameter in the acceptance of VoIP services. Despite the amount of research and development work done in this area, it is still hard to guarantee the same Quality of Service (QoS) as that of traditional telephony. Among the various parameters affecting voice quality, packet losses, latency and delay jitter are the key factors, inevitable in a packet network, contributing to speech quality degradation. In order to balance jitter introduced by queuing in packet switched networks, a jitter buffer mechanism is required at the receiver for ensuring continuous play-out.

When a jitter buffer is applied, received packets are buffered for a while after arrival, and played out sequentially at scheduled time. If some packets arrive after their scheduled play-out time, they are discarded. A late play-out time reduces such kind of packet loss, but introduces unnecessarily long end-to-end delay. The problem of delay jitter is thereby converted into end-to-end delay and packet loss. Previous work mainly focused on designing jitter buffer based solely on the trade-off between end-to-end delay (play-out delay) and packet loss rate due to late arrival. The play-out delay is adjusted either at the beginning of a talk-spurt

[1][2] (called per-talk-spurt), or within the speech talk-spurt using time-scale modification to ensure continuous play-out of voice data [3][4] (called per-packet). Although such designs can achieve a minimum average end-to-end delay for a specified packet loss rate, they do not take into account the overall perceived speech quality. Recently, some quality-based approaches have been proposed. Instead of achieving a compromise between delay and packet loss, these approaches adjust the jitter buffer with the objective of optimizing the perceived speech quality given by the Mean Opinion Score (MOS) [5][6]. To develop such quality based approaches, the ITU-T E-Model [13] is one of the most well-known methods. The output of the E-Model, the so called  $R$  value, can be easily mapped onto a corresponding MOS value using a transformation given in Appendix I of [13].

Although the ITU-T E-Model has been initially developed for network planning purposes, there has been proposals to extend it not only for evaluating speech quality in conversational communication, but also for monitoring VoIP performance during transmission. In [7][8][9] quality-based play-out scheduling approaches were proposed to maximize perceived speech quality using the  $R$  value of the E-Model as cost function. These approaches rely only on adjusting the play-out delay on per-talk-spurt basis. When talk-spurts are long and the network delay varies significantly within them, the performance of these so called talk-spurt-based methods is limited. In [10], a per-packet quality-based jitter buffer algorithm is described. The play-out delay estimation is based on maximizing the  $R$  value (or equivalently maximizing the MOS value) and is designed as an unconstrained optimization problem. However, since time-scale modification is required in all per-packet jitter buffer algorithms and a speech frame normally can only be time-scaled within a certain range to avoid degrading voice quality [3], a constrained optimization problem is more suitable.

When designing a quality based algorithm for jitter buffer management, an estimate of network delay distribution is required. Some works assume a certain parametric model to estimate the Cumulative Distribution Function (CDF) of the network delay distribution. For instance, Pareto in [10], Weibull in [7] and Gamma in [11] were used. The use of a certain type of distribution to model

delay behavior in a network is arguable. In fact, delay and jitter in a VoIP session are non stationary and have a high degree of variability even within a single session.

In this paper, a new adaptive jitter buffer system is proposed, implementing per-packet scheduling based on the extended E-Model. In addition, the proposed system contains a spike detection mechanism and a classifier based time scaling technique similar to that proposed in [4]. The time scaling technique is implemented directly inside the speech decoder ( AMR decoder [16] ) which is advantageous for the quality and makes it possible to use the internal parameters of the codec. For instance, pitch values and gains are particularly useful parameters for time scaling.

The paper is organized as follows. In Section II , the extended E-Models used in the proposed jitter buffer system are introduced. In Section III, the E-model based play-out algorithm is proposed. Section IV presents the modified time-scaling embedded AMR decoder. Finally, simulation results illustrating the performance of the proposal and conclusion are presented in Section V and Section VI.

## II. EXTENDED E-MODEL

The ITU-T E-Model is a computational model for the prediction of the expected voice quality which combines different impairments due to codec, echo and other transmission parameters. The underlying assumption of the E-Model is that all impairment factors contributing to speech quality degradation are additive on a psychological scale, and summed to form a rating factor  $R$ . The rating factor lies in the range of 0 to 100. A rating of '0' represents a MOS value '1' (bad quality) and '100' of  $R$  represents MOS value '4.5' (high quality). The output  $R$  value is obtained by subtracting impairment factors from a basic quality measure [13]:

$$R = R_0 - I_s - I_d - I_{e,eff} + A \quad (1)$$

where  $R_0$  represents the basic signal-to-noise ratio;  $I_s$  is the Simultaneous Impairment Factor;  $I_d$  represents the Delay Impairment Factor;  $I_{e,eff}$  is the Effective Equipment Impairment Factor.  $A$  is an advantage Factor which has accordingly no relationship with all other parameters and normally can be neglected. All input parameters and their recommended ranges can be found in [13]. For those parameters which are not available at the time of planning, the default values from the ITU [14][15] are recommended. If we only focus on an IP network, the expression of E-Model in (1) can be simplified to the transport layer [12]

$$R = 93.2 - I_d - I_{e,eff} \quad (2)$$

with  $I_d$  referring to impairments only due to end-to-end delay  $d$ .  $I_d$  can be derived by curve fitting as described in [12]

$$I_d = 0.024 \cdot d + 0.11 \cdot (d - 177.3) \cdot H(d - 177.3) \quad (3)$$

where  $H(x)$  is the step function ( $H(x) = 0$  if  $x < 0$  ;  $H(x) = 1$  else). In this paper, we consider only random

packet losses for the AMR codec, therefore,  $I_{e,eff}$  is obtained either by applying provisioning values from [15]

$$I_{e,eff} = 5 + 90 \cdot (p_n + p_b)/(p_n + p_b + 10) \quad (4)$$

or from the empirical formula [7]

$$I_{e,eff} = 14.96 + 16.68 \cdot \ln(1 + 30.11 \cdot (p_n + p_b)) \quad (5)$$

where  $p_n$  is the packet loss rate in the network and  $p_b$  is the late packet loss rate dropped by the jitter buffer. Since packets are discarded when they arrive after their scheduled play-out time, the late loss rate  $p_b$  is calculated as

$$p_b = (1 - p_n)(1 - P_r(X \leq d)) = (1 - p_n)(1 - F(d)) \quad (6)$$

with  $F(d)$  being the CDF of network delay which is obtained in this paper from histogram statistics of previous network delay. If we define the sum of  $I_d$  and  $I_{e,eff}$  as a new impairment factor  $I$

$$I = I_d + I_{e,eff} \quad (7)$$

then (2) is simplified as

$$R = 93.2 - I \quad (8)$$

This formulation of  $R$  (8) is used as the cost function in our jitter buffer management to estimate the play-out delay by maximizing  $R$  which is equivalent to minimizing  $I$ . Equations (4) and (5) for modeling  $I_{e,eff}$  are used both and their performance is compared.

## III. PROPOSED PLAY-OUT ALGORITHM

The proposed receiver includes an adaptive jitter buffer algorithm and the time-scaling embedded in the decoder, as shown in Fig. 1. The Adaptive play-out algorithm is the main control unit. A spike, i.e., the sudden and very high increase of network delay, is very common in VoIP transmission. For this reason, spike detection in [2] is used to switch between NORMAL mode and SPIKE mode. In SPIKE mode, the scheduled play-out time follows current network condition. In NORMAL mode, the scheduled play-out time is based on the delay estimation implementing the extended E-Model, mentioned in Section II.

The proposed play-out algorithm will be described using the basic notations listed in Table I.

When a packet arrives at the receiver before its scheduled time, it can be played out without packet loss. Before playing out the current speech frame, the play-out delay of the next coming packet has to be estimated to obtain the modified current frame length. The play-out delay is chosen to maximize the perceived speech quality in terms of  $R$ . As discussed in Section II,  $R$  depends on the end-to-end delay  $d$ , network loss rate  $p_n$  and buffer loss rate  $p_b$ . The buffer loss rate is determined by the play-out buffering algorithm, thus by the end-to-end delay (play-out delay). Therefore (8) can

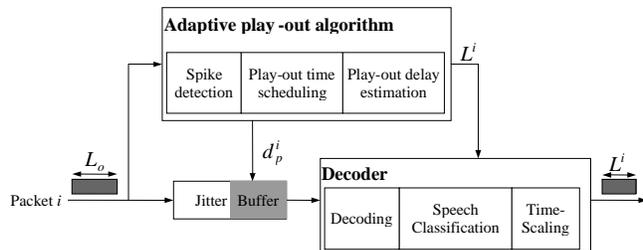


Figure 1. proposed adaptive jitter buffer at the receiver

TABLE I. BASIC NOTATIONS

symbol	Definition
$d_n^i$	network delay of packet $i$
$d_p^i$	actual play-out delay of packet $i$
$\hat{d}_p^i$	estimated play-out delay of packet $i$
$L_o$	original frame length, 160 for AMR
$L^i$	modified frame length of packet $i$
$\Delta^i$	frame length difference of packet $i$
$L_{max}$	possible maximum time-scaled frame length
$L_{min}$	possible minimum time-scaled frame length

be expressed as a function of play-out delay, and applied as the cost function in the play-out buffering algorithm to predict the voice quality.

The play-out delay for each packet is estimated based on maximizing the expected  $R$  value. The operation of the jitter buffer is based on the statistics of the delay and packet loss of the previous received packets.

The algorithm works as follows

1. Receive a new  $packet^i$ , obtain network delay information  $d_n^i$  from the RTP header information.
2. Spike Detection: check the current network condition, and switch between SPIKE mode and NORMAL mode.
3. Play-out time Scheduling
  - a) If this is the first packet of the talk-spurt, follow network delay  
 $d_p^i = d_n^i$
  - b) Otherwise, use the estimated play-out delay  
 $d_p^i = \hat{d}_p^i$
4. Play-out delay Estimation
  - a) SPIKE: follow the current network delay  $\hat{d}_p^{i+1} = d_n^i$ , and skip step 5.
  - b) NORMAL: estimate play-out delay based on the E-Model.
5. E-Model based play-out delay estimation in NORMAL mode
  - a) Update delay statistics of the most recent received  $W$  (history window size) packets only in NORMAL mode
  - b) Find the optimal play-out delay for  $packet^{i+1}$   
 $\hat{d}_p^{i+1}: I_m(\hat{d}_p^{i+1}) = \min_{d_{min} \leq d \leq d_{max}} I_m(d)$

where  $d_{min}$  and  $d_{max}$  are the constraints specified by the time-scaling to make the artifacts less audible:

$$d_{min} = d_p^i - (L_o - L_{min})$$

$$d_{max} = d_p^i + (L_{max} - L_o)$$

6. Calculate the new length of packet<sup>i</sup>

$$\Delta^i = \hat{d}_p^{i+1} - d_p^i$$

$$L^i = L_o + \Delta^i$$

7. Send  $packet^i$  and expected length  $L^i$  to the decoder.

#### IV. TIME-SCALING EMBEDDED IN THE DECODER

The E-Model based play-out scheduling algorithm described in section III is applied specifically to AMR codec. The standard 3GPP AMR decoder is modified to embed a time scaling technique based on speech classification. According to the evaluated frame type, different time-scaling (extension or suppression) operations are applied to the excitation frame which is segmented to four sub-frames.

##### A. Speech classification

Speech is categorized into silence and talk-spurt, which is further subdivided into voiced/unvoiced. Special frames such as plosive or over-voiced frames are also differentiated from others by using the internal parameters inside the AMR decoder

- Silence/Talk-spurts: The classification between silence and talk-spurts is realized by Voice activity detection (VAD). When operated in DTX mode, Silence and talk-spurts are distinguished from each other by checking if the frame type is SID frame.
- Voiced/Unvoiced: Considering the speech classification implemented in the VMR codec [17], our voiced/unvoiced decision is based on three parameters: The Voicing Factor  $F_v$ , Spectral Ratio  $e_{tilt}$  and Energy Variation  $V_e$ .  $F_v$  is calculated as an averaged normalized correlation over four sub-frames of speech with the pitch lags  $T_o$ .  $e_{tilt}$  is estimated as the ratio between the low and high frequency energy.  $V_e$  is applied to evaluate the variance of energy inside a frame.  $F_v$ ,  $e_{tilt}$  and  $V_e$  are then compared to predefined thresholds to identify the frame type. The Voiced/Unvoiced classification on the word "success" is illustrated in Fig. 2.
- Other Speech classes: Besides the voiced, unvoiced and silence classifications described above, some specific frames must be distinguished from the voiced/unvoiced frames to avoid speech quality degradation due to time scaling. The average pitch gain of some unvoiced frames is higher than 0.45, while the maximum pitch gain for some voiced frames is below 0.5. These unvoiced and voiced frames are termed as over\_voiced and under\_voiced frames respectively in this paper. It is suggested not applying time scaling on these frames as well as the plosive frames in order to prevent quality degradation.

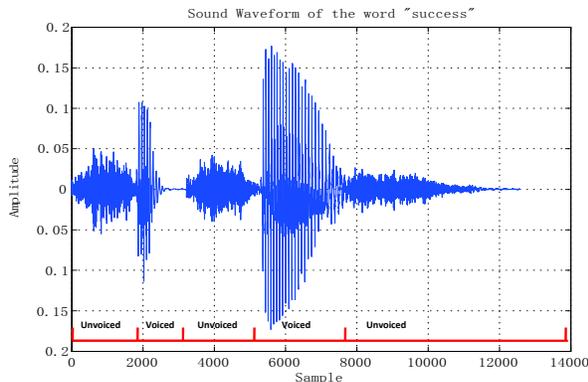


Figure 2. Voiced/Unvoiced classification of the word “success”

**B. Time Scale Modification inside AMR decoder**

Since pitch lag and pitch gain are internal parameters used by the AMR decoder, it is also advantageous to scale the speech inside the decoder, directly in the excitation domain. According to the speech classification, voiced and unvoiced frames are processed differently. Moreover, some frames are not modified to prevent quality degradation, as proposed in [3]. The different processing operations based on the result of speech classification are summarized in TABLE II.

Voiced frames are extended by repeating the pitch cycle preceding the minimum energy point. The number of added pitch cycles is determined by the difference between the original frame length and the expected frame length, combined with the pitch lag of the sub-frame. The voiced frames are suppressed by removing some pitch cycles just before the minimum energy point in the last sub-frame backwards. The number of subtracted cycles depends on the pitch lag of the last sub-frame and the length difference. Time scaling on unvoiced frames is much simpler. A certain number of zeros are uniformly inserted in the unvoiced frame for extension, while zeros are removed from the frame for compression. The number of zeros inserted or removed relies on the expected new length. In unvoiced frames, the samples can be removed from the beginning if the previous frame is unvoiced or from the end of the frame if the previous frame isn't unvoiced. The original and modified signals are illustrated in Fig. 3.

**C. Modified AMR decoder**

The modified AMR decoder is illustrated in Fig. 4. The generated excitation is formed by the fixed and adaptive codebooks with their corresponding gains. The excitation is classified into voiced/unvoiced/silence and other specific frames. According to the frame type decision, different time scaling techniques are applied to the excitation signal. The reconstructed speech is obtained by feeding the scaled excitation of new length into the LP Synthesis Filter. In order to keep the synchronization between encoder and decoder, the adaptive codebook is updated before time scaling.

TABLE II. TIME SCALING

Frame Type	Time Scale Modification	
	Extension	Suppression
Silence	Comfort Noise	Comfort Noise
Voiced	Duplicate some pitch cycles	Remove selected pitch cycles
Unvoiced	Insert zeros between the excitation samples	Remove samples from the excitation signal
Under voiced Over voiced Positive Onsets	No time scaling	

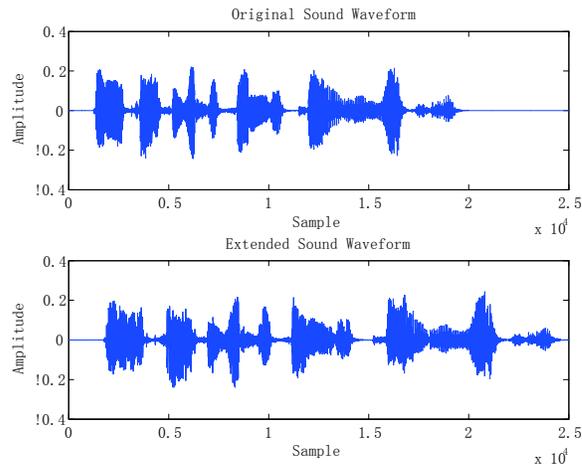


Figure 3. Original and extended sound waveform

**V. EXPERIMENTAL RESULTS**

In the experiment, we implemented three other most promising algorithms to compare with our proposed jitter buffer algorithm, denoted as Algorithm 1 [1], Algorithm 2 [7] and Algorithm 3 [3]. Our algorithm referred to as Algorithm 4. The results are shown in Table III for five traces. Each trace contains 7500 packets and the window size  $W$  is set to 300. During the experiment, we implement the proposed play-out delay estimation both on the  $I_{e,eff}$  model in (4) and (5). The performance comparison is shown in Fig.5. We observed that both (4) and (5) lead to quite similar results when used in the jitter buffer management algorithm. The maximum length after extension  $L_{max}$  is limited to twice of the original length (320 ms) and the minimum length after suppression  $L_{min}$  must not be shorter than half the original length (80 ms), as suggested in [3]. The maximum allowable end-to-end delay is 400ms.

From Table III, it can be seen that Algorithm 4 achieves the highest MOS scores (which are obtained from the impairment factor  $R$ ) among all tested traces. Algorithm 1 and Algorithm 2 apply both talk-spurt based jitter buffer management. For Algorithm 1, play-out time is defined with the help of statistical estimation of the play-out delay based on network characteristics of several previous talk-spurts.

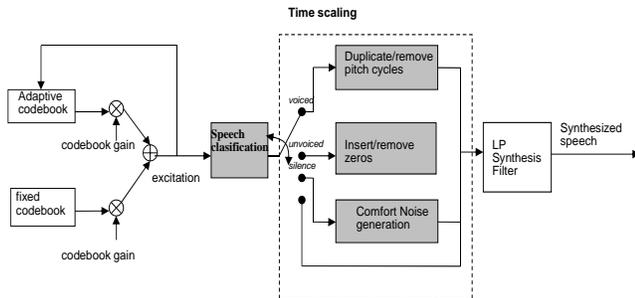


Figure 4. Modified architecture of the AMR decoder

Algorithm 2 implements an extended E-Model for estimation with the assumption of Weibull delay distribution. Both talk-spurt based algorithms are not efficient when the talk-spurts are very long and for cases where the network delay varies significantly such as in cases of spikes. Although a spike detection mechanism is adopted in per-talk-spurt scheduling, the play-out delay cannot be updated until the next talk-spurt. The scheduled play-out time cannot follow such spikes within a talk-spurt and results in more discarded packets due to late arrival, as in trace 1 and trace 5.

Both Algorithm 3 and our proposed algorithm (Algorithm 4) schedule play-out delay on a per-packet basis, thus adjust the play-out time in a highly dynamic way and adapt more quickly to the network conditions even during speech activity (talk-spurt). Algorithm 3 is based on achieving an optimal trade-off between packet loss rate and end-to-end delay, but it does not provide a direct access to the perceived speech quality, which is exactly the goal of the optimization. Our proposed algorithm estimates the play-out delay based on maximizing the MOS value derived from the impairment factor  $R$ , therefore achieving best performance in all trace files.

The information of network delay, delay jitter and network loss rate of five trace files are also listed in Table III. The optimal  $R$  is calculated by assuming no buffer delay and no late loss rate. The optimal  $R$  can only be achieved with full knowledge about the network condition before transmission, thus it cannot be realized in real time QoS monitoring. The difference between the Optimal MOS and the result of our jitter buffer algorithm is partly due to the constraints  $L_{min}$  and  $L_{max}$  required by time-scaling.

The performance of play-out delay estimation of trace 1, trace 2 are illustrated in Fig.6. Both the results from Algorithm 3 and from our proposed algorithm are shown. Both algorithms adapt play-out delay quite well to the varying network delay. In the cases of spikes, our algorithm reduces the packet loss rate at the expense of additional delay.

## VI. CONCLUSION AND FUTURE WORK

In this paper, we focused on impairment of random packet loss and end-to-end delay for the AMR codec. We proposed an adaptive jitter buffer algorithm based on the extended E-Model with spike detection and a time scaling technique embedded directly in the AMR decoder. The simulation results show that the proposed method achieve

better perceived speech quality compared to other existing algorithms under various network scenarios. Moreover, these results are not specific to AMR. As the time scaling algorithm is closely connected to the CELP coding scheme, the proposed jitter buffer management can be extended to other codecs, in particular to CELP based codecs being most advanced form of speech codecs.

For future activities, subjective listening tests are planned in order to validate the proposed method. We will also extend our work under bursty packet loss conditions.

## ACKNOWLEDGEMENT

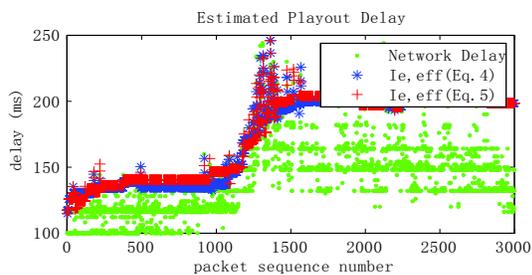
We greatly appreciate the support and guidance of Dr. Herve Taddei, Dr. Christophe Beaugeant and Dr. Imre Varga. We would also like to thank Dr. Anisse Taleb and Mr. David Virette for their helpful comments and advice.

## REFERENCES

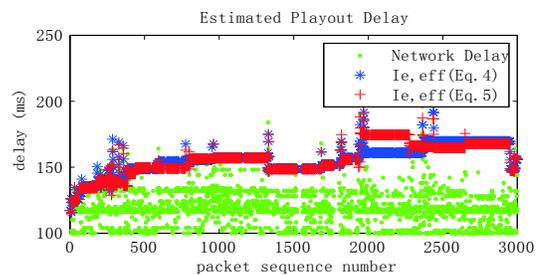
- [1] S. Moon, J. Kurose, and D. Towsley, "Packet audio play-out delay adjustment: Performance bounds and algorithms," *Multimedia Systems*, vol.6, no.1, Jan.1998, pp.17-28.
- [2] M. Narbut and L. Murphy, "VoIP Playout Buffer Adjustment using Adaptive Estimation of Network Delays," *Proc. 18th Int. Teletraffic Congress (ITC-18)*, 2003, pp.1171-1180.
- [3] Y. Liang, N. Färber, and B. Girod, "Adaptive Play-out Scheduling and Loss Concealment for Voice Communication Over IP Networks," *IEEE Trans. on Multimedia*, vol.5 2003.
- [4] P. Gournay and K. Anderson, "Performance Analysis of a Decoder-Based Time Scaling Algorithm for Variable Jitter Buffering of Speech Over Packet Networks," *Proc. IEEE Int. Conf. on Acoustics*, May 2006.
- [5] K. Fujimoto, S. Ata and M. Murata, "Adaptive playout buffer algorithm for enhancing perceived quality of streaming applications," *Proc. IEEE GLOBECOM'02*, vol. 3, Nov.2002.
- [6] L. Atzori and M. Lobina, "Playout Buffering of Speech Packets Based on a Quality Maximization Approach," *IEEE Trans. on Multimedia*, vol.8 2006.
- [7] L. Sun and E. Ifeachor, "Voice quality prediction models and their application in voip networks," *IEEE Trans. on Multimedia*, vol.8 2006.
- [8] Z. Li, S. Zhao, and J. Kuang, "An improved speech playout buffering algorithm based on a new version of E-Model in VoIP," *ChinaCom'08*, 2008.
- [9] L. Huang, Y. Chen and T. Yaw, "Adaptive VoIP Service QoS Control based on Perceptual Speech Quality," *Proc. IEEE ICACT'07*, 2007, pp.885-890.
- [10] C. Wu and W. Chang, "Perceptual Optimization of Playout Buffer in VoIP Applications," *ChinaCom'06*, 2006.
- [11] C. Wu, K. Chen, Y. Chang, and C. Lei, "An Empirical Evaluation of VoIP Playout Buffer Dimensioning in Skype, Google Talk, and MSN Messenger," *Proc. ACM NOSSDAV'2009*, 2009, pp.97-102.
- [12] R.Cole and J.Rosenbluth, "Voice over ip performance monitoring," *ACM SIGCOMM*, 2001.
- [13] ITU-T rec G.107, "E-model, a computational model for use in transmission planning," April 2009.
- [14] ITU-T rec G.109, "Definition of categories of speech transmission quality," Sep. 1999.
- [15] ITU-T rec G.113, "Transmission impairments due to speech processing," Nov. 2007.
- [16] 3GPP TS 26.071, AMR Speech Codec; General description.
- [17] 3GPP2 C.S0052-0, Source-Controlled Variable-Rate Multimode Wideband Speech Codec(VMR-WB).

TABLE III. COMPARISON OF DIFFERENT ALGORITHMS AND NETWORK DELAY TRACES

Trace	Algorithm	Average play-out delay(ms)	Late loss rate(%)	MOS	Optimal MOS	Average network delay(ms)	STD of network delay(ms)	Average delay jitter(ms)	Maximum jitter(ms)	Network loss rate (%)
1	Algorithm 1	180.5	7.5	2.0	3.5	136.7	25.0	36.7	146	2.4
	Algorithm 2	191.0	4.0	2.5						
	Algorithm 3	165.2	3.5	2.6						
	Algorithm 4	173.9	2.2	2.9						
2	Algorithm 1	153.3	2.3	3.4	4.1	119.7	12.4	19.7	120	0.24
	Algorithm 2	178.5	0.9	3.8						
	Algorithm 3	150.4	1.5	3.6						
	Algorithm 4	160.0	0.8	3.9						
3	Algorithm 1	148.6	6.0	2.5	4.1	126.8	19.9	26.8	134	0.51
	Algorithm 2	180.6	0.9	3.7						
	Algorithm 3	154.7	1.2	3.7						
	Algorithm 4	158.9	0.7	3.8						
4	Algorithm 1	133.7	0.3	4.1	4.2	112.3	8.8	12.3	48	0
	Algorithm 2	170.0	0.1	4.1						
	Algorithm 3	134.7	0.4	4.1						
	Algorithm 4	134.8	0.3	4.1						
5	Algorithm 1	147.6	2.6	3.4	4.2	116.5	44.9	16.5	305	0
	Algorithm 2	164.4	2.1	3.5						
	Algorithm 3	146.0	1.2	3.8						
	Algorithm 4	148.0	1.0	3.9						

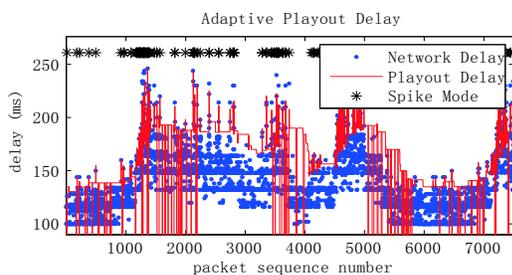


(a)

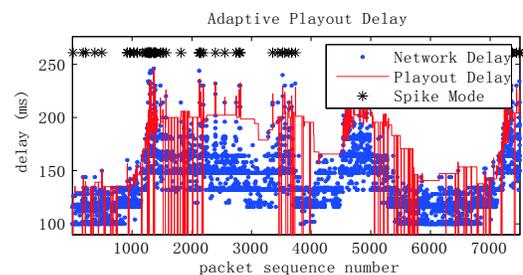


(b)

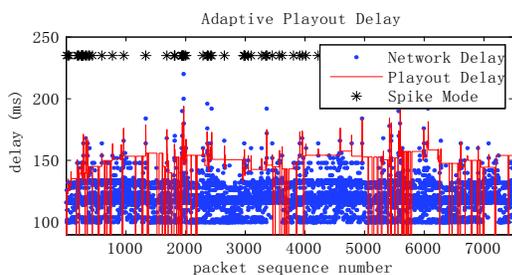
Figure 5. Playout delay estimation based on different  $I_{e,eff}$  Models (a) Trace1 (b) Trace2



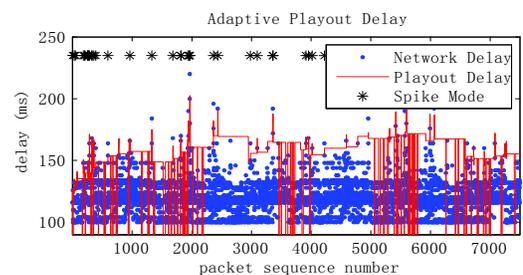
(a)



(b)



(c)



(d)

Figure 6. Play-out delay estimation: (a)Algorithm3 for Trace 1 (b)Algorithm4 for Trace1 (c)Algorithm3 for Trace2 (d) Algorithm4 for Trace2

## Service Strategies for Eliminating Digital Divide in Korea

Hyongsoon Kim, Sang-Ug Kang, Kwang-Taek Ryu  
 Digital Infrastructure Division, IT Convergence Service Department  
 National Information Society Agency  
 Seoul, Korea  
 {khs, sukang, ryukt}@nia.or.kr

Eunyoung Lee  
 Dept. of Computer Science  
 Dongduk Women's University  
 Seoul, Korea  
 elee@dongduk.ac.kr

**Abstract**—Republic of Korea is one of the countries which possess the top-class network infrastructure in the world. However, the network facilities are concentrated in densely populated areas, and this causes the data access speed of less populated areas, such as rural areas, to become relatively slow. The low speed of network makes broadcast-communication convergence services in rural areas difficult, and it leads to the digital divide between cities and rural areas which will emerge as a social issue. In the paper, we analyze the rural BcN project of the Korea government, which aims at accelerating the network infrastructure of rural areas. We also propose the service strategies for boosting the usage of the network infrastructure in rural areas.

*Keywords*-digital divide; BcN; convergence services; policies.

### I. INTRODUCTION

For the last few decades, the Korea government (i.e., the government of Republic of Korea) has executed a series of projects to build the network infrastructure of high speed. As the result of successful projects, Republic of Korea (in short, Korea in this paper) has become one of the top Information Technology (IT) countries in the world [1]. The Korea government executed the high-speed information and communication network project in the 1990's, and has constructed the broadcast-communication convergence network (BcN) infrastructure since the mid 2000s [2], [3], [4].

The development of the network infrastructure of high speed was government driven, so it made it possible to build a network infrastructure timely and effectively. Fast network infrastructure built by the government has been provided to IT industries in Korea with low costs or almost for free, and it enables IT companies to concentrate on developing new IT services and technologies without worrying building their privately-owned networks. The networks developed by the government-driven projects have been served as the infrastructure and the testbed for commercial services and academic researches. It has been pointed out by many professionals that the excellent network infrastructure is one of the main mile stones which lead Korea to the top class of IT technology in the world.

However, the investment of the Korea government and nation-wide network operators is focused on only densely

populated areas. According to the survey of Korea Telecom, the current nation-wide BcN coverage are up to about 80%, and would be increased to 90% by 2010 [5]. But, small farming and fishing villages, which lack profitability from the perspective of service providers, are not included in the operators' BcN expansion plans. Although the residents in small farming and fishing villages can access the Internet, the data rate remains at 1–3Mbps (xDSL type). As a result, various convergence services cannot be used in these areas, unlike in metropolitan-, medium-sized, or small-sized cities.

Voluntary investment to these less populated areas by communication service providers cannot be expected, because communication service providers generally pursue short-term profits. This situation will cause the gap between people with effective access to digital and information technology (those in city areas) and those with very limited or no access at all (those in rural areas), called the *digital divide*. The digital divide could wide the knowledge divide between city areas and rural areas, which is already considered as a serious social problem. The governmental support is inevitably necessary to alleviate the digital divide. The Korea government has set up the BcN establishment project to provide high-speed network convergence services such as IPTV and VoIP in sparsely populated rural areas, and has conducted the project since 2010.

In this paper, we analyze the rural BcN project of the Korea government, and propose the service strategies for boosting the usage of the resulting network infrastructure. In Section II, we outline the motivation of rural BcN and the development plan of it. After the brief introduction of rural BcN, we analyze and discuss the strategies for service applications to stimulate the usage of rural BcN in Section III. We draw a conclusion in Section V.

### II. RURAL BCN DEVELOPMENT PLAN

The Korea government has set up the BcN establishment project to provide high-speed network convergence services in sparsely-populated rural areas. The outline of the rural BcN project was made in 2009 [6], and the full version of the rural BcN project has been made and executed since 2010 [7], [8].



Figure 1. Strategy Deployment Circle

### A. Motivation

The construction project of high-speed networks in rural areas of over 50 households had been performed from 2002 to 2005. As the result of the success of the project, high-speed Internet service came to be provided to 3.53 million households, that is, 94 percent of all rural households. Through the construction of high-speed networks in rural areas of less than 50 households from 2006 to 2008, 3.76 million households, in other words, 99.7 percent of all rural households came to enjoy high-speed Internet services.

However, the speed of the network in rural areas was not fast enough for emerging broadcast-communication convergence services. The early version of the rural BcN project was made in 2009, and a feasibility study on the project was conducted. The result of the feasibility study was reflected in the plan called “Mid- and Long-term Development Plan for Broadcast-Communication Networks” in 2009 [6]. In 2010, the panel on the construction of smart infrastructure in rural areas has been formed, and the project has begun.

The broadcast-communication convergence network in rural areas is needed to lay a foundation for universal broadcast-communication convergence services by expanding broadband broadcast-communication networks, which are concentrated in cities, into rural areas. Constructing BcN in rural areas will help solve rural areas’ economic, social and cultural issues and will vitalize the economy and industry through the diffusion and spread of specialized services based on broadband broadcast-communication net-

works. It is also required to establish a national cooperation system integrating ministries, local governments and carriers in order to construct broadband broadcast-communication networks and expand broadcast-communication convergence services.

### B. Phases of the development plan

In order to construct broadband networks and create an environment for increasing the use of broadcast-communication convergence services, executive plans for rural BcN have four sub-goals: network infrastructure, service accessibility, customized services, and political and economical support. The plans for network infrastructure and political support are presented briefly in this section, and we will make a close analysis and discussion of service plans in Section III.

To construct broadband networks in rural areas, broadband networks in small rural villages (with less than 50 households) will be built by 2016. The expenses are partially covered by the central government. The central government, local governments, and carriers share the costs in the ratio of 1:1:2. For the broadband network construction in medium-sized rural villages (with more than 50 households), the efficiency of the broadband network will be increased. The construction of broadband networks in medium-sized areas will be completed by 2012. This construction will be supported through efficient policies, such as inducing to carriers to divide construction projects among themselves

or to construct networks jointly. Broadband networks will be built in remote islands and isolated areas, too. Remote islands and isolated areas are considered as places where new wired networks cannot be constructed. For these areas, a foundation to provide broadcast-communication convergence services will be made by considering and applying various alternatives.

The use of wireless networks such as 4G will be tested for rural areas. Test beds for wireless ultra wide-band networks will be built in small rural areas that have no broadband networks, and the possibility that wireless ultra wide-band network could replace fixed broadband networks will be verified. The government is trying to rent the optic fiber cables from Korea Electric Power Corporation in rural and isolated areas. The rented optic fibers can be considered as another option, but it requires the governmental involvement in the process of mediation and negotiation. Pilot projects will be promoted to verify other wireless technologies such as WiFi Mesh, WiBro. Services through a two-way satellite will be provided for areas where other alternatives cannot be applied.

In order to encourage service carriers to construct broadband networks efficiently, a supportive legislation for carriers will be enacted. Carriers will be encouraged to voluntarily invest in building broadband networks in non-competitive rural areas by reforming laws and regulations that deal with the construction of broadcast-communication infrastructure and developing a support system.

Policy and institutional supports for facilitating the use of services will be provided. The use of services will be facilitated by organizing and operating an effective body that promotes and supports the projects, and by developing a reasonable and affordable tariff plan for specialized services for rural areas. The ecology of broadcast communication industry for boosting the local industry in rural areas will be established. A basis for sustainable development of rural areas will be founded, and the opportunity to grow together will be offered to the related industries by establishing an ecological system that promotes a virtuous cycle.

Figure 1 shows how each direction of deployed strategies is related to other directions and makes influence in a circular way. The first direction of the rural BcN plan begins with building an economical and efficient network. In this part, the voluntary competition of the private sector and the governmental support are combined. In order to avoid duplicate investment, the communication resources will be used as much as possible.

The second direction of the rural BcN plan will be to promote the pilot projects on the basis of users' actual demand. In this part, the actual demand of rural residents will be identified before providing services. Analyzing the actual demand will create a service environment where the services reflect the characteristics of rural communities, and the most frequently used services.

The last direction of the rural BcN plan is to implement a nation-wide cooperation system. There exists a need for a cooperative body which involves all government agencies, local governments and carriers. The cooperative body can effectively foster the ecological environment such as convergence among related industries.

### III. STRATEGIES FOR SERVICE APPLICATIONS

To stimulate the usage of BcN in rural areas, the project should be conducted in two directions: making people to access the convergence services more easily and developing killer services to satisfy the needs of rural residents. We suggest service strategies in these two directions in the following subsections.

#### A. Enhancing user accessibility

To enhance the accessibility of users in rural areas, terminal guidelines for rural residents should be authored and provided. This will facilitate the use of broadcast-communication convergence services in rural areas by eliminating the complexity of using the terminals. The guideline of IT terminals and interfaces will enable the aged rural residents to use the IT terminals and interfaces more easily. As well as authoring the guideline, testing whether or not the IT terminals meet guidelines should be performed through rural BcN pilot projects. Using the terminals with guidelines should be also encouraged through the pilot projects.

The user interfaces (UI) should be enhanced for rural residents, and for this purpose, UI design solutions should be developed and provided with the governmental support. The study of UI structures and UI deployment should be conducted first, so that rural residents can easily access programs, contents, and information through BcN. The developed UI will be tested and verified for its usability by applying them to pilot services. For example, a new user interface of IPTV would be implemented and provided to rural residents with guidelines, and the effectiveness will be scored by the real users. Easy and accessible user interfaces will encourage rural residents to use broadcast-communication convergence services.

For customer support, the remote support solutions can be considered for the candidate. Concerning the use of broadcast-communication convergence services, it is important to resolve the inconvenience of rural residents. To support customers in rural areas, various measures should be invented. One of the feasible ways is developing terminals and solutions which can be controlled remotely. When telephone consultation is not sufficiently convenient to customers, a video-based remote support system like IPTV would be an attractive substitute for resolving customers' inconvenience. The terminals and solutions with remote control function will be developed in the pilot projects, and will be distributed to rural residents. Customer satisfaction surveys will be conducted several times during the period

of pilot projects, and the remote customer support will be tested and verified based on user feedback.

There must exist a system which connects and integrates IT services for rural areas. The system aims at enhancing existing services into broadcast-communication network-based services specialized for rural areas, and maximizing synergy. The goals can be achieved by establishing an inter-ministerial cooperation body and by promoting the connection and convergence among different IT services for rural areas. It is important to establish a nationwide cooperation body committed to exploring ways of integrating and connecting IT services for rural communities, although those IT services have been promoted separately by different government agencies. The unified governmental body can integrate the developed public services for rural areas, and make effective support to future development. A centralized service center will be a part of the unified governmental body, and it will operate an open platform, develop and connect specialized services, and offer a test environment.

#### *B. Developing specialized service applications*

There have been a few services for rural areas, but those services scatter among several governmental bodies. It is urgent and important to integrate all these scattered services to avoid duplicate investment and to maximize the synergy. The IT services for rural areas have performed one-way communication from service providers to users, and there is no connection among these IT services though their duplicate functionalities. These services have Web-based user interfaces, and are delivering only simple information to users. To supply broadcast-communication convergence services effectively, it is important to integrate similar IT services and upgrade it into a two-way communication service with higher quality.

Integration is not an easy job, so is integrating the existing IT services for rural areas. To support and expedite the integration of IT services, it is necessary to develop a software platform to support the integration. First of all, the open platform must be developed, and be applied for the integration of IT services through a couple of pilot projects. The government should be in charge of the pilot projects since the projects are not beneficiary enough for IT industry to invest their money to.

After integrating the duplicated IT services of various governmental ministries or bodies, it is needed to find good broadcast-communication convergence services, and to provide a test bed for the developed new services. At least one research center should be dedicated for these purposes. The research center will be in charge of operating the integration platform of rural IT services, and be a base station if some rural IT services, e.g., open agricultural market, are expanded nation-wide.

The specialized services for rural areas must be identified, implemented and tested through the pilot projects.

The demand of rural residents can be categorized in five fields: welfare, culture, education, economy and safety. The services should be identified based on these five fields, and the service provision models should be aligned with existing services offered by other ministries and local governments.

For the welfare of the rural residents, a u-health system will be implemented for senior residents and children based on IPTV and video telephone. Local weather, traffic information and local news will be delivered through IPTV. IPTV can be also used for education: home tutoring for students in rural areas, Korean language classes for multicultural families and educational lectures of new agricultural technologies. To stimulate the economy of rural areas, IPTV could offer e-commerce services which connect consumers in cities to the providers in rural areas directly. The history of products can be traced by the consumers in cities through IPTV. This will give more credibility of the agricultural products, and guarantee more profit to the providers in rural areas. For regional safety, guarding services using CCTV will be implemented through IPTV and video telephone. Smart phones as well as IPTV will be used for an emergency alert system.

Executing pilot services should begin with offering the service of the highest demand. The feasibility of each service candidate must be verified and evaluated based on the benefits of the service. The provision of pilot services will be given in steps. At the first step, a framework of service development will be come up with, and only a few services will be implemented. In the following step, these pilot services will be spread around widely in order to stimulate the usage of rural BcN and its related services. At the last step, the services will be provided nation-widely after the usefulness of the services is verified.

To increase the use of services, promotional and educational activities should be organized and executed. For effective educational and promotional activities, existing organizations and facilities owned by other institutions and carriers can be utilized. A regular education system for rural residents should be organized by training at least one educator per community. There must a regular re-education program for the educators at least once in a year, which is run by a dedicated organization.

As well as educational activities, regular surveys should be conducted on the use of services and customer satisfaction. The feedback from the users will increase the use of services in rural areas through a virtuous cycle of demand analysis, service development and provision, customer satisfaction study, and increase of usage rate. To achieve this goal, it is necessary to develop indicators and indices for evaluating customers' usage statistics and satisfaction level objectively. A survey on the usage status and satisfaction level of broadband networks can be conducted on a yearly basis.

#### IV. SUGGESTIONS

We have designed and developed the BcN establishment policy in rural areas with assisting the Korea government. We want to make a few suggestions to whom are considering deploying an enhanced IT infrastructure in isolated areas of their countries.

First, the economic efficiency of network facilities should be considered. The feasible options for our BcN establishment in rural areas would be narrowed into two candidates: one is the wired networking using PONs and the other is the wireless networking using wireless LANs or wireless mesh networks (WMNs). Using PONs as the network infrastructure of rural areas costs more than using wireless network techniques. But, using PON has an advantage when upgrading the speed of the network to over 1 Gbps. It is much easier and less costly to accelerate PONs than to upgrade wireless networks. The wireless networking is much cheaper than the wired networking using PONs, while the coverage of wireless networks is smaller. The wireless networking has the strength in the case that subscribers inhabit densely in a relatively small area. The network deployment method should be determined with considering the characteristics of a given area and the budget of the local government.

For the successful deployment of the BcN in rural areas, good services should be also developed and provided. Good services on BcN will stimulate the network usage. At the same time, the provided services should reflect the unique features of rural areas. For example, u-Health service can be one of promising killer services, when considering the fact that the population of rural areas becomes older than the population of urban areas on average. U-Learning would be another killer service for rural areas in order to narrow the academic gap between rural and urban areas. The services can be implemented in cooperation with private service providers and local government, and during the development phase, the opinion of local residents, who are the future beneficiaries, must be feed-backed.

Local governments must have plans for diminishing the burden of communication charge, i.e., the fee of using the BcN network and its related services in rural areas. Service providers should supply cheaper rate plans for rural inhabitants, and local government has to consider to support a part of communication charge as a way of eliminating the digital divide. One of the feasible ways would be to designate the BcN services as universal services like electricity or telephone services. Another feasible way would be lowering the rate by a long-term contract between local government and private service providers.

#### V. CONCLUSION AND FUTURE WORK

Korea is one of the countries which possess the top-class network infrastructure in the world. However, the network facilities are concentrated in densely populated areas, which

causes the data access speed of less populated areas, such as rural areas, to become relatively slow. The low speed of network makes broadcast-communication convergence services in rural areas difficult, and it leads to the digital divide between cities and rural areas, which will emerge as a social issue. The Korea government has set up the BcN establishment project to provide high-speed network convergence services, such as IPTV and VoIP, to sparsely-populated rural areas, and has conducted the project since 2010.

In the paper, we analyzed the rural BcN project of the Korea government, which aims at accelerating the network infrastructure of rural areas. The service strategies for boosting the usage of the network infrastructure in rural areas were proposed. Boosting network usage is as important as speeding up the network itself. Fast network infrastructure is useless unless it is used actively. The suggested boosting strategies can go to two directions: enhancing user accessibility to the services and developing specialized services.

The rural BcN project of the Korea government is under progress now, and it will become complete by 2016. We believe that we can make another further discussion of the rural BcN project after the project itself is over with some analyzable result.

#### REFERENCES

- [1] National Information Society Agency of Korea, "Broadband convergence Network annual report 2007," Ministry of Information and Communication of Korea, Tech. Rep., 2007.
- [2] Ministry of Information and Communication of Korea, "Basic blueprint for building the Broadband convergence Network," Ministry of Information and Communication of Korea, Tech. Rep., 2004.
- [3] Ministry of Information and Communication of Korea, "Second phase plan for establishing BcN," Ministry of Information and Communication of Korea, Tech. Rep., 2006.
- [4] Korea Communication Council, "Third phase plan for establishing BcN," Korea Communication Council, Tech. Rep., 2008.
- [5] Korea Communication Council, "UBcN deployment policy," Korea Communication Council, Tech. Rep., January 2009.
- [6] Korea Communication Council, "Directions for BcN infrastructure of rural areas," Korea Communication Council, Tech. Rep., 2009.
- [7] Korea Communication Council, "Project plans for promoting BcN construction in rural areas," Korea Communication Council, Tech. Rep., March 2010.
- [8] Korea Communication Council, "Long-term plan for constructing broadband networks in rural areas," Korea Communication Council, Tech. Rep., June 2010.

# UEP for Wireless Video Streaming Using Spatially Multiplexed MIMO System With a Suboptimal Joint Detection

Sanghyun Bak and Jaekwon Kim  
 Computer and Telecommun. Engineering Division  
 Yonsei University, Wonju, Republic of Korea  
 Email: jaekwon@yonsei.ac.kr

**Abstract**—Recently, unequal error protection (UEP) techniques for video transmission over multiple-input multiple-output (MIMO) wireless channel have been actively researched. Spatially multiplexed (SM) MIMO system with joint detection achieves a good bit error rate (BER) performance as well as a high data rate, however, UEP can not be implemented because all the physical (PHY) layer signals are jointly detected. In transmitting video streams, peak signal to noise ratio (PSNR) is important rather than BER, and it is well known that PSNR is improved by UEP. In this paper, we propose an UEP technique for wireless video streaming using SM MIMO system with a suboptimal joint signal detection method. Computer simulations demonstrate that the PSNR performance of QR-LRL-based UEP, that is computationally very efficient when compared with the conventional optimal maximum likelihood (ML) detection, is slightly better than that of the optimal ML detection that is considerably more computation intensive.

**Keywords**-UEP, H.264/AVC, MIMO, Joint Detection, ML, Spatial Multiplexing.

## I. INTRODUCTION

Video streaming service in wireless environments is a challenging task that requires both high data transmission speed and high transmission reliability [1]. Wireless channel is characterized by limited spectral bandwidth, limited transmit power, and unstable channel gains [2]. Consequently, error-resilient video streaming at a high speed using these limited resources is a demanding task in wireless video communications.

A popular application (APP) layer H.264/AVC was designed to enable network-friendly video streaming as well as enhanced compression efficiency [1]. Various error resiliency schemes can be used for H.264/AVC video streaming such as semantics, syntax error detection, data partitioning (DP), slice interleaving, flexible macro-block ordering (FMO), parameter set sharing, and error concealment techniques [3]. However, all these error resiliency schemes are APP layer schemes that do not consider physical (PHY) layer systems. There have been several previous works addressing APP-PHY cross layer design.

In the APP-PHY cross layer design approaches in [4] and [5], an orthogonal space-time block code (OSTBC) was adopted for multiple-input multiple-output (MIMO) systems

[6]. The OSTBC MIMO system, however, does not increase data transmission speed when compared with the traditional single antenna PHY system [7]. Due to the low transmission speed, MIMO system using an OSTBC seems not suitable for real time video streaming.

In [8], an implicit unequal error protection (UEP) technique for video streaming over MIMO wireless channel was proposed. Spatially multiplexed (SM) MIMO system provides  $N_{min} = \min(N_T, N_R)$  virtual subchannels, where  $N_T$  is the number of transmit antennas and  $N_R$  is the number of receive antennas, thereby  $N_{min}$  times higher data transmission speed than single antenna PHY system. When linear signal detection is used as in [8], the virtual subchannels show different transmission reliability, i.e., signal-to-noise ratio (SNR). In [8], UEP is achieved implicitly by assigning a APP video stream of higher priority to a PHY subchannel with higher reliability. In [9], a transmission power for multiple transmit antennas was also controlled to further benefit from MIMO systems. In these two previous works, however, linear signal detection methods were considered that offer severely degraded error performance when compared to a joint signal detection method.

In this paper, we propose an UEP technique for prioritized video streaming, assuming SM MIMO system with a suboptimal joint signal detection method QR-decomposition-least-reliable-layer (QR-LRL). There are various joint detection methods such as sphere decoding (SD) [10], QR-decomposition-M-algorithm-maximum likelihood detection (QRM-MLD) [11], lattice reduction aided detection (LRAD) [12], and optimal ML signal detection [13]. From the perspective of video streaming, the main problem of using joint signal detection is that all the signals from multiple antennas are jointly detected, thus the transmission reliability of all PHY subchannels are not differentiated. Consequently, UEP for prioritized video streaming exploiting the reliability information of subchannels is not allowed. In this paper, we show that the suboptimal joint detection method QR-LRL in [14] can be used for UEP. The QR-LRL is a computationally efficient but suboptimal joint detection method, thus the average BER performance is inferior to that of ML signal detection. The PSNR performance of QR-LRL, however,

is slightly superior to that of the ML detection as will be demonstrated in the simulation section.

## II. SYSTEM MODEL

The developed APP-PHY cross layer system, illustrated in Fig. 1, is described in this section. The APP layer of the system is based on the H.264/AVC encoder, which delivers network abstraction layer (NAL) unit streams. In Fig. 1,  $X_{f,s,m}$ ,  $f = 1, 2, \dots, F$ ,  $s = 1, 2, \dots, S$ ,  $m = 1, 2, \dots, M$  stands for the  $m$ -th macroblock in the  $s$ -th slice of the  $f$ -th frame, where  $F$  is the number of frames,  $S$  is the number of slices in a frame, and  $M$  is the number of macroblocks in a slice. The notation  $\hat{X}_{f,s,m}$  in Fig. 1 denotes the compressed macroblock that is different from  $X_{f,s,m}$  due to quantization in the process of compression. In this paper, a slice or a NAL packet was considered as a single PHY packet. Each PHY packet is 31-bit cyclic redundancy check (CRC) encoded. Then, when SM MIMO system is assumed, NAL units or PHY packets are assigned to transmit antennas of appropriate reliability. At the receiver side,  $N_T$  PHY packets are jointly detected. Then, the CRC parity bits are used to determine if each packet is received safely. The acknowledgement (ACK) or not ACK (NACK) information is passed to the APP layer. If NACK for a slice is received, the APP layer does not try to decode the slice but perform an error concealment for the slice. Transmission of sequence parameter set (SPS) and picture parameter set (PPS) NAL units are assumed error-free.

We consider MIMO systems with  $N_T$  transmit antennas and  $N_R (\geq N_T)$  receive antennas. Let  $\mathbf{x} = [x_1 \ x_2 \ \dots \ x_{N_T}]^T$  denote the transmit signal vector, where  $x_i$ ,  $i = 1, 2, \dots, N_T$  is the transmitted signal from the  $i$ th transmit antenna;  $\mathbf{y} = [y_1 \ y_2 \ \dots \ y_{N_R}]^T$  denotes the received signal vector, where  $y_j$ ,  $j = 1, 2, \dots, N_R$  is the received signal at the  $j$ th receive antenna;  $\mathbf{H}$  with dimension of  $N_R \times N_T$  denotes the channel gain matrix, of which entry  $h_{j,i}$ ,  $j = 1, 2, \dots, N_R$ ,  $i = 1, 2, \dots, N_T$  is the channel gain between the  $i$ th transmit antenna and the  $j$ th receive antenna;  $\mathbf{z} = [z_1 \ z_2 \ \dots \ z_{N_R}]^T$  denotes the noise vector, where  $z_j$ ,  $j = 1, 2, \dots, N_R$  is assumed to be zero mean complex white Gaussian with variance of  $\sigma_z^2$ . Then the MIMO system can be described as follows.

$$\mathbf{y} = \sqrt{\frac{E_x}{N_T}} \mathbf{H} \mathbf{x} + \mathbf{z}. \quad (1)$$

In this paper, we assume independent and identically distributed (i.i.d.) Rayleigh fading channel gains, and ideal channel estimation at the receiver side.

## III. UNEQUAL ERROR PROTECTION USING MIMO PHY SYSTEMS

In this section, two previous UEP techniques are described, and we propose a novel UEP technique exploiting SM MIMO systems with a suboptimal joint detection.

### A. UEP Using OSTBC MIMO System

In [4] [5], the orthogonal space-time block code  $\mathbf{G}_4$  in [6] was used assuming 4 transmit antennas, i.e.,  $N_T = 4$ . Then, assuming 4 receive antennas, i.e.,  $N_R = 4$ , and collecting the received signals during 4 symbol periods, the received signal matrix  $\mathbf{Y} \in \mathbb{C}^{4 \times 4}$  is expressed as

$$\mathbf{Y} = \sqrt{\frac{E_x}{3}} \mathbf{H} \mathbf{X}^{\text{OSTBC}} + \mathbf{Z}. \quad (2)$$

where  $y_{j,n}$ ,  $j = 1, 2, 3, 4$ ,  $n = 1, 2, 3, 4$ , the  $(j, n)$ th entry of  $\mathbf{Y}$ , denotes the received signal at the  $j$ th receive antenna during the  $n$ th symbol period; the noise matrix  $\mathbf{Z} \in \mathbb{C}^{4 \times 4}$  is composed of entries  $z_{j,n}$ ,  $j = 1, 2, 3, 4$ ,  $n = 1, 2, 3, 4$ , that denotes the noise at the  $j$ th receive antenna during the  $n$ th symbol period; the space-time coded signal is given as

$$\mathbf{X}^{\text{OSTBC}} = \begin{bmatrix} x_1 & -x_2^* & -x_3^* & 0 \\ x_2 & x_1^* & 0 & -x_3^* \\ x_3 & 0 & x_1^* & x_2^* \\ 0 & x_3 & -x_2 & x_1 \end{bmatrix}. \quad (3)$$

We note that the number 3 in (2) is due to the fact that signals are transmitted from only 3 antennas simultaneously, although  $N_T = 4$ . From the received signal matrix  $\mathbf{Y}$  in (2), optimal ML signal detection of  $x_1$  is performed via the following simple linear processing.

$$\begin{aligned} \hat{x}_1 &= \sum_{j=1}^4 \{h_{j,1}^* y_{j,1} + h_{j,2} y_{j,2}^* + h_{j,3} y_{j,3}^* + h_{j,4}^* y_{j,4}\} \\ &= \sqrt{\frac{E_x}{3}} \|\mathbf{H}\|_F^2 x_1 + \\ &\quad \sum_{j=1}^4 \{h_{j,1}^* z_{j,1} + h_{j,2} z_{j,2}^* + h_{j,3} z_{j,3}^* + h_{j,4}^* z_{j,4}\} \end{aligned} \quad (4)$$

Dividing both sides of (4) with  $\sqrt{\frac{E_x}{3}} \|\mathbf{H}\|_F^2$ , we have the estimate  $\hat{x}_1 = x_1 + \hat{z}_1$ . The three noise term in (4) can be shown to be a complex Gaussian noise with the same variance of  $\|\mathbf{H}\|_F^2 \sigma_z^2$ . With similar linear processing produces  $\hat{x}_s$ ,  $s = 2, 3$  with the same noise variance. The SNR of the signal  $\hat{x}_s$  or  $\hat{x}_s$  is given as

$$\text{SNR}^{\text{OSTBC}}_s = \frac{E_x \|\mathbf{H}\|_F^2}{3\sigma_z^2}, \quad s = 1, 2, 3. \quad (5)$$

From (5), it can be seen that the SNR of the three signals are identical, consequently the SNR information can not be used to implement UEP. In [4][5], various combinations of QP, channel coding rate, and constellation size were assigned to video streams of different priorities. However, it was shown that the data transmission speed of OSTBC MIMO system is almost the same as or even lower than that of the traditional single input single output (SISO) systems [7]. Note that the data transmission speed for the space-time code

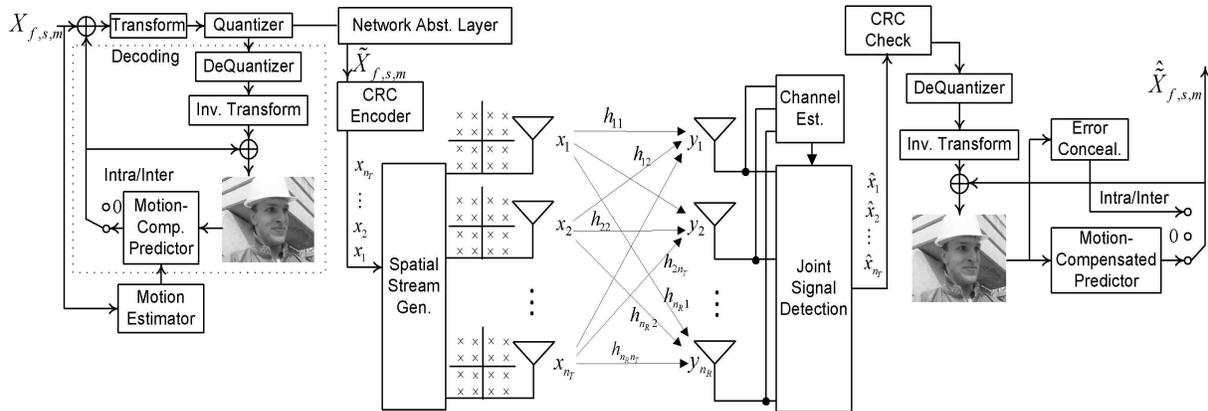


Figure 1. APP-PHY cross layer system.

$G_4$  is 3/4 symbols per channel use and that of SISO system is 1 symbol per channel use. We argue that the low data transmission speed of OSTBC MIMO system renders itself not suitable for video streaming service, which necessitates the SM MIMO system that is addressed in the following section.

### B. UEP Using SM MIMO Systems with Linear Signal Detection

In [8][9], SM MIMO systems were used to provide UEP for prioritized video streaming. The considered detection methods in [8][9] are linear equalizers such as zero-forcing (ZF) or minimum mean squared error (MMSE). If ZF or MMSE equalizer is used and when  $N_T = N_R = 4$ , the corresponding weight matrices are

$$\mathbf{W}_{ZF} = (\mathbf{H}^H \mathbf{H})^{-1} \mathbf{H}^H = \mathbf{H}^{-1} \quad (6)$$

$$\mathbf{W}_{MMSE} = \left( \mathbf{H}^H \mathbf{H} + \frac{4\sigma_z^2}{E_x} \mathbf{I}_4 \right)^{-1} \mathbf{H}^H \quad (7)$$

where  $(\cdot)^H$  denotes the Hermitian transpose. The reliability of the  $s$ th signal (or the reliability of the corresponding subchannel),  $s = 1, 2, 3, 4$ , by the above two methods are derived as

$$\text{SNR}_{ZF}^s = \frac{E_x}{\|\mathbf{w}_{s,ZF}\|^2 4\sigma_z^2} \quad (8)$$

$$\text{SINR}_{MMSE}^s = \frac{|\mathbf{w}_{s,MMSE} \mathbf{h}_s|^2}{\sum_{m \neq s} |\mathbf{w}_{m,MMSE} \mathbf{h}_m|^2 + \|\mathbf{w}_{s,MMSE}\|^2 \frac{4\sigma_z^2}{E_x}} \quad (9)$$

From (8) and (9), it can be seen that the reliability of the 4 subchannels are different. Let subch( $r = s$ ) denote the  $s$ th most reliable subchannel, i.e., subchannel with the  $s$ th highest SNR or SINR. Exploiting this relationship, in [8][9], a video stream of higher priority was assigned to a

PHY subchannel with higher reliability to implement UEP implicitly. Note that the symbol transmission rate of the SM MIMO system is 4 symbols per channel use, that is much higher than that of single antenna system or OSTBC MIMO system.

Although the high transmission rate of SM MIMO systems with linear equalization is desirable, its BER performance is harshly degraded when compared to OSTBC MIMO systems. Fig. 2 compares the PER performance of OSTBC MIMO and SM MIMO with the linear MMSE detection. The linear MMSE detection was used rather than the ZF detection because the MMSE detection outperforms ZF detection in general. The size of PHY packet transmitted over each virtual subchannel of SM MIMO system is 5,000 bits. Considered constellation is a 16-QAM, thus a PHY packet is transmitted from its assigned transmit antenna for  $5,000/4 = 1,250$  symbol periods. The PHY packet size of OSTBC MIMO system is 5,004 bits, thus a PHY packet is transmitted from 4 transmit antennas for  $5,004/[4(3/4)] = 1,668$  symbol periods.

From Fig. 2, it can be observed that the packet error rate (PER) performance of SM MIMO systems with linear detection suffers from significantly degraded performance when compared to OSTBC MIMO system, and that the 4 subchannels in SM MIMO with linear detection show differentiated PER performance, i.e., differentiated reliabilities.

In order to retain the high transmission speed of SM MIMO systems and to achieve a good PER performance simultaneously, we propose to use SM MIMO system with joint signal detection that is discussed in the next section.

### C. Proposed UEP Technique Using SM MIMO Systems with a Suboptimal Joint Detection

In this subsection, we propose a novel UEP technique for video streaming over SM MIMO systems with a suboptimal joint signal detection QR-LRL.

Maximum likelihood signal detection method that achieves the optimal PER performance is described as

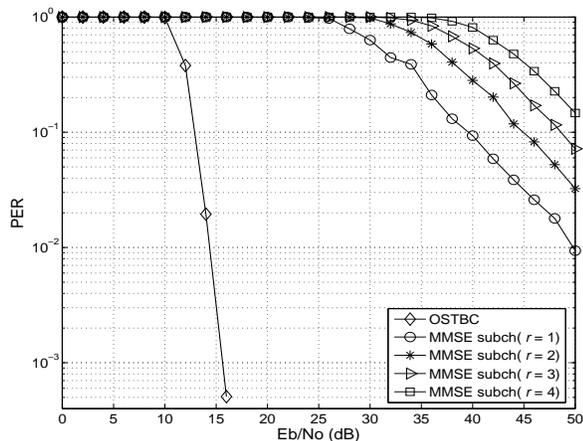


Figure 2. PER performance of OSTBC MIMO and SM MIMO with the linear MMSE detection.

$$\hat{\mathbf{x}}_{ML} = \arg \min_{\mathbf{x} \in \mathcal{C}^4} \left\| \mathbf{y} - \sqrt{\frac{E_x}{4}} \mathbf{H} \mathbf{x} \right\|^2 \quad (10)$$

where,  $\mathcal{C}$  is the set of constellation points. The brute force implementation of the above ML signal detection searching over  $|\mathcal{C}|^4$  vectors is almost impossible due to its required high computational complexity [13], where  $|\mathcal{C}|$  denotes the cardinality of the argument set. SD method [10] is well known to achieve the ML performance with a significantly reduced average complexity, however, its complexity depends on the channel conditions and the worst case complexity is still high. There are also suboptimal signal detection methods such as QRM-MLD [11] and LRAD [12] methods that achieve near-ML performance. Although the implementation complexity of the aforementioned methods are significantly reduced and the optimal or a near-ML performance is achieved, PHY subchannels are not differentiated in terms of transmission reliability, not allowing the implementation of UEP.

In QR-LRL [14], the signals are ordered and detected sequentially, and the transmission reliability of the subchannels depend on the detection order. In QR-LRL, the most unreliable subchannel is selected using the SNR criterion (8).

$$s_{LRS} = \arg \min_{s \in \{1,2,3,4\}} \text{SNR}_{s}^{\text{ZF}}. \quad (11)$$

All the constellation points are tried as the above selected symbol, and for each point as the selected symbol value, the remaining symbols are detected by the conventional VBLAST [15], producing a candidate vector set  $\mathcal{S}_{\text{QR-LRL}} = \{\mathbf{x}_c\}_{c=1}^{|\mathcal{C}|}$ . The ML metrics are calculated for the candidate vectors to choose the most likely symbol vector as follows.

$$\hat{\mathbf{x}}_{\text{QR-LRL}} = \arg \min_{\mathbf{x} \in \mathcal{S}_{\text{QR-LRL}}} \left\| \mathbf{y} - \sqrt{\frac{E_x}{4}} \mathbf{H} \mathbf{x} \right\|^2 \quad (12)$$

We note that ML metrics are calculated for  $|\mathcal{S}_{\text{QR-LRL}}| = |\mathcal{C}|$  candidate vectors in (12), while  $|\mathcal{C}|^4$  vectors are considered in (10). Trying all the constellation points as the first symbol, QR-LRL achieves almost the same BER performance as the conventional ML signal detection. Without loss of generality, we assume that the 3rd symbol has the minimum SNR. Obviously, the true signal vector has the smallest ML metric with high probability, hence the true signal vector is detected once it is included in the candidate vector set  $\mathcal{S}_{\text{QR-LRL}}$ . Therefore, with high probability, the error happens when the true symbol vector is not contained in the candidate vector set. Since all the symbols are tried as  $x_3$ , error happens when true  $x_3$  is tried but signal detection of the following system equation by VBLAST [15] is erroneous.

$$\underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix}}_{\tilde{\mathbf{y}}} - \sqrt{\frac{E_x}{4}} \mathbf{h}_3 x_3 = \sqrt{\frac{E_x}{4}} \underbrace{[\mathbf{h}_1 \ \mathbf{h}_2 \ \mathbf{h}_4]}_{\tilde{\mathbf{H}}} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_4 \end{bmatrix}}_{\tilde{\mathbf{x}}} + \mathbf{z}. \quad (13)$$

where  $\mathbf{h}_i$  is the  $i$ th column of the matrix  $\mathbf{H}$  in (1).

It was shown that the condition number of above  $\tilde{\mathbf{H}}$  is significantly reduced compared to  $\mathbf{H}$  by choosing the least reliable symbol (LRS) by (11). Detection ordering of (13) is decided in the decreasing order of the following SNR.

$$\widetilde{\text{SNR}}_s^{\text{ZF}} = \frac{E_x}{\|\tilde{\mathbf{w}}_{s,\text{ZF}}\|^2 4 \sigma_z^2} \quad (14)$$

where,  $\tilde{\mathbf{w}}_{s,\text{ZF}}$  is the  $s$ th row vector of  $(\tilde{\mathbf{H}}^H \tilde{\mathbf{H}})^{-1} \tilde{\mathbf{H}}^H$ . The detection ordering is well known to reduce the error propagation effect [16], thereby increasing the probability that true  $\mathbf{x}$  is contained in the set  $\mathcal{S}_{\text{QR-LRL}}$ . Considering the criterion of choosing the first symbol as well as the criterion for ordering the remaining symbols, we can determine the reliability of the virtual 4 subchannels. The subchannel corresponding to the symbol chosen in (11) is the least reliable, which is denoted as subch( $r=4$ ). The subchannel for the first detected symbol in (13) is the most reliable, which is denoted as subch( $r=1$ ), the subchannels for the secondly and thirdly detected symbol using (13) are the second and third most reliable subchannels denoted as subch( $r=2$ ) and subch( $r=3$ ), respectively.

Fig. 3 shows that the average PER performance of QR-LRL is slightly worse than that of ML. Simulation environment is the same as in the simulations for Fig. 2. Note that the subch( $r=2$ ) achieves the optimal PER performance and that subch( $r=1$ ) outperforms the optimal ML detection. It

can be also observed that the four subchannels of QR-LRL have differentiated transmission reliability.

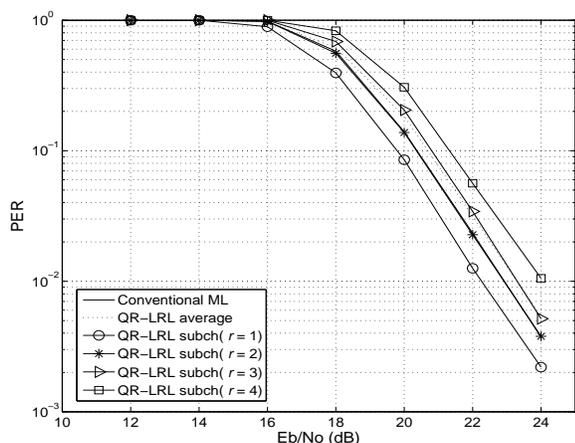


Figure 3. PER performance of the suboptimal QR-LRL and the conventional ML. The performance of conventional ML and QR-LRL subch( $r = 2$ ) are almost identical. Also, the average PER of QR-LRL is almost the same as the QR-LRL subch( $r = 3$ ).

#### IV. SIMULATION RESULTS

In this section, we perform a set of computer simulations to show the improved PSNR performance by the proposed UEP technique based on QR-LRL. Reference software JM16.0 was used for encoding and decoding of H.264/AVC video stream. The total number of frames  $F = 256$ , GOP size is 4, 30 frames per second (fps) Foreman video in CIF resolution was used, the number of slices in a frame  $S=4$ , and a slice is composed of 99 macroblocks, IDR period and I period are the same as 4. QP of I and P slices are set as 28, and no B slice was used. The search range for motion vector estimation is  $\pm 16$  pixels with resolution of  $1/4$  pixel, the number of reference frame for motion vector estimation was set to 1, entropy coding of CABAC was used. The transmission of the first GOP was also assumed error-free for a simple implementation of error concealment that copies a previous slice. In PHY layer, 16-QAM constellation was used, 31 bit CRC parity bits based on the polynomial  $X^{31} + X^{30} + X^{26} + X^{24} + X^{18} + X^{15} + X^{14} + X^{12} + X^{11} + X^{10} + X^8 + X^6 + X^5 + X^4 + X^3 + X + 1$  are used to check if a slice or a PHY packet was successfully received. If NACK is detected, the NACK information is passed to APP layer that does not try to decode but instead copy a recent slice.

Fig. 4 compares the PSNR performance of the two schemes at  $E_b/N_0 = 22$ (dB), the conventional ML and the suboptimal QR-LRL-based UEP. The PSNR in Fig. 4 is the average of 7 PSNRs for 7 different set of i.i.d. PHY channel gains and noises. It can be seen that the suboptimal QR-LRL-based UEP achieves a better performance than the

conventional ML detection. Thus it can be stated that a better average PER performance does not guarantee a better PSNR performance. We note that the conventional optimal ML detection requires  $|\mathcal{C}|^4$  times ML metric calculations, while the suboptimal QR-LRL requires only  $|\mathcal{C}|$  times ML metric calculations.

#### V. CONCLUSION

In this paper, we proposed a novel UEP technique for prioritized video streaming over SM MIMO systems with a suboptimal joint detection. OSTBC MIMO system and SM MIMO system with linear signal detection suffer from a low data transmission speed and a degraded PER performance, respectively. SM MIMO system with joint detection achieves both high transmission speed and high transmission reliability, however, UEP is not allowed in general. We showed the suboptimal joint detection method QR-LRL can be used to implement UEP. With the aid of computer simulations, we demonstrated that the PSNR by the suboptimal QR-LRL, requiring only  $|\mathcal{C}|$  times ML metric calculations, is slightly superior to that by the conventional optimal ML detection that requires  $|\mathcal{C}|^4$  times ML metric calculations, when the number of transmit antenna is 4.

#### REFERENCES

- [1] T. Wiegand, G. J. Sullivan, G. Bjontegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560-576, July 2003.
- [2] T. Stockhammer, M. M. Hannuksela, and T. Wiegand, "H.264/AVC in wireless environments," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 657-673, July 2003.
- [3] S. Kumar, L. Xu, M. K. Mandal, and S. Panchanathan, "Error resiliency schemes in H.264/AVC standard," *J. Vis. Commun. Image R.*, April 2006.
- [4] M. K. Jubran, M. Bansal, and L. P. Kondi, "Low-delay low-complexity bandwidth-constrained wireless video transmission using SVC over MIMO systems," *IEEE Trans. Multimedia*, vol. 10, no. 8, pp. 1698-1707, Dec. 2008.
- [5] M. K. Jubran, M. Bansal, and L. P. Kondi, "Accurate distortion estimation and optimal bandwidth allocation for scalable H.264 video transmission over MIMO systems," *IEEE Trans. Image Process.*, vol. 18, no. 1, pp. 106-116, Jan. 2009.
- [6] V. Tarokh, N. Seshardi, and A. Calderbank, "Space-time block codes from orthogonal designs," *IEEE Trans. Inform. Theory*, vol. 45, no. 6, pp. 1456-1467, July 1999.
- [7] S. Sanhdu and A. Paulraj, "Space-time block codes: a capacity perspective," *IEEE Commun. Letters.*, vol. 4, no. 12, pp. 384-386, Dec. 2000.
- [8] S. Song and C. W. Chen, "Scalable H.264/AVC video transmission over MIMO wireless systems with adaptive channel selection based on partial channel information," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1218-1226, Sept. 2007.

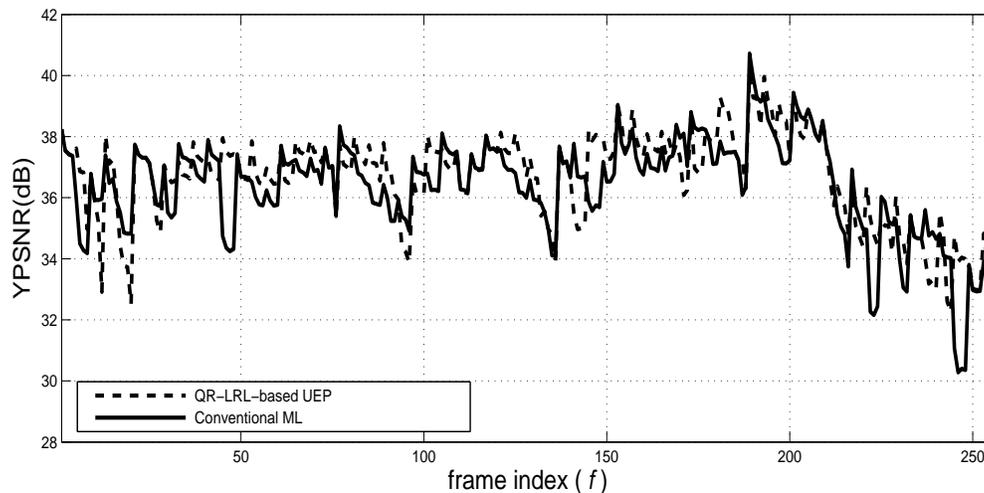


Figure 4. PSNR performance of the conventional ML and QR-LRL-based UEP.

- [9] D. Song and C. W. Chen, "Maximum-throughput delivery of SVC-based video over MIMO systems with time-varying channel capacity," *J. Vis. Commun. Image R.*, pp. 520-528, 2008.
- [10] E. Viterbo and J. Boutros, "A universal lattice code decoder for fading channels," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1639-1642, July 1999.
- [11] K. J. Kim and J. Yue, "Joint channel estimation and data detection algorithms for MIMO-OFDM systems," *Proc. 36th Asilomar Conf. Signals, Syst., Comput.*, pp. 295-300, 2002.
- [12] D. Wubben, R. Bohnke, V. Kuhn, and K.-D. Kammeyer, "Near-Maximum- Likelihood Detection of MIMO Systems using MMSE-Based Lattice- Reduction," *Proc. IEEE Int. Conf. Commun.*, vol. 2, pp. 798-802, June 2004.
- [13] Y. S. Cho, J. Kim, W. Y. Yang, and C. G. Kang, *MIMO-OFDM Wireless Communications with MATLAB*, Wiley and Sons, 2010.
- [14] S. Bahng, Y. Park, and J. Kim, "QR-LRL signal detection for spatially multiplexed MIMO systems," *IEICE Trans. Commun.*, vol. E91-B, no. 10, pp. 3383-3386, Oct. 2008.
- [15] G. J. Foschini, "Layered Space-Time Architecture for Wireless Communication in a Fading Environment When Using Multiple Antennas," *Bell Labs Tech. J.*, vol. 1, no. 2, pp. 41-59, 1996.
- [16] R. Narasimhan, "Error propagation analysis of VBLAST with channel-estimation errors," *IEEE Trans. Commun.*, vol. 53, no. 1, pp. 27-31, 2005.

# Prefilter Bandwidth Effects in Sequential Symbol Synchronizers based on Pulse Comparison by Positive Transitions at Half Rate

Antonio D. Reis<sup>1,2</sup> and José P. Carvalho<sup>1</sup>

Dep. Física / Unidade D. Remota

<sup>1</sup>Universidade da Beira Interior, 6200 Covilhã, Portugal  
adreis@ubi.pt, pacheco@ubi.pt

José F. Rocha<sup>2</sup> and Atílio S. Gameiro<sup>2</sup>

Dep. Electrónica e Telecom. / Instituto Telecom.

<sup>2</sup>Universidade de Aveiro, 3810 Aveiro, Portugal  
frocha@det.ua.pt, amg@det.ua.pt

**Abstract-** This work studies the effects of the prefilter bandwidth in the sequential symbol synchronizers based on pulse comparison at bit rate and at half bit rate. We consider three different prefilter bandwidth namely  $B1=\infty$ ,  $B2=2.tx$  and  $B3=1.tx$ , where  $tx$  is the bit rate. The synchronizer has two variants one operating by both transitions at bit rate and other operating by positive transitions at half rate. Each variant has two versions namely the manual and the automatic. The objective is to study the prefilter bandwidth with four synchronizers and to evaluate their output jitter UIRMS (Unit Interval Root Mean Square) versus input SNR (Signal Noise Ratio).

**Keywords -** Prefilter; Synchronizers; Communication systems.

## I. INTRODUCTION

This work studies the prefilter bandwidth effects on the jitter-SNR behavior of four sequential symbol synchronizers.

The prefilter, applied before the synchronizer, switches their bandwidth between three values namely first  $B1=\infty$ , after  $B2=2.tx$  and next  $B3=1.tx$ , where  $tx$  is the bit rate [1, 2].

The synchronizer has four types supported in two variants, one operating by both transitions at the rate with versions manual (b-m) and automatic (b-a) and other operating by positive transitions at half rate with versions manual (p-m/2) and automatic (p-a/2) [3, 4, 5, 6].

The difference between the four synchronizers is only in the phase comparator, since the other blocks are equal [7, 8].

The synchronizer VCO (Voltage Controlled Oscillator) is the clock whose performance determines, in good part, the system quality [9, 10, 11, 12].

Fig. 1 shows the prefilter followed of the synchronizer.

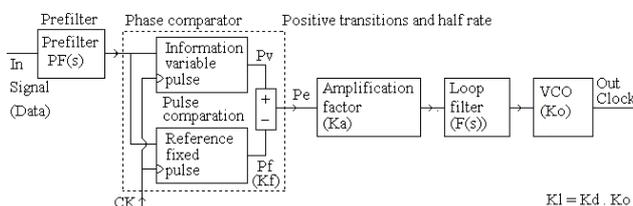


Fig. 1 Prefilter with the synchronizers based on pulse comparison

$PF(s)$  is the prefilter (low pass). The synchronizer has various blocks, namely  $K_f$  is the phase detector gain,  $F(s)$  is the loop filter,  $K_o$  is the VCO gain and  $K_a$  is the loop gain factor that controls the root locus and loop characteristics.

Next, we present the state of the art. Then, we present the prefilter with three bandwidths ( $B1=\infty$ ,  $B2=2.tx$ ,  $B3=1.tx$ ).

Following, we present the variant by both transitions at rate with their manual and automatic versions. Next, we present the variant by positive transitions at half rate with their manual and automatic versions.

After, we present the design and tests. Then, we present the results. Finally, we present the conclusions.

## II. STATE OF THE ART, PROBLEM AND SOLUTION

In priori and actual-art state, various synchronizers have been developed. The motivation is to create new synchronizers operating at half rate and evaluate their performance with the noise. This contribution increases the knowledge about the synchronizers [1, 2, 3, 4].

The problem is that the synchronizers' output jitter increases when the input SNR decreases. To solve or to minimize the problem, we propose a prefilter that attenuates the noise, but unfortunately distorts slightly the signal [5, 6].

## III. PREFILTER BANDWIDTH EFFECTS

The prefilter, applied before the synchronizer, filters the noise but disturbs slightly the signal. The prefilter bandwidth  $B$  switches between three values ( $B1=\infty$ ,  $B2=2.tx$ ,  $B3=1.tx$ ).

Fig. 2 shows the prefilter with their three bandwidths.

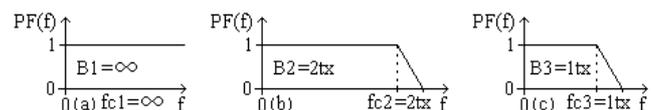


Fig. 2 Three prefilter bandwidths: a)  $B1=\infty$ ; b)  $B2=2.tx$ ; c)  $B3=1.tx$

a) First, as shown in Fig.2a, the prefilter has a bandwidth equal to infinite ( $B1=\infty$ ).

b) Second, as shown in Fig.2b, the prefilter has a bandwidth equal to times the bit rate ( $B2 = 2.tx$ ).

c) Third, as shown in Fig.2c, the prefilter has a bandwidth equal to the bit rate ( $B3 = 1.tx$ ).

We will evaluate the three bandwidth effects ( $B1$ ,  $B2$ ,  $B3$ ) on the jitter-SNR curves of the four symbol synchronizers.

## IV. SYNCHRONIZERS OPERATING AT THE RATE

The synchronizer with its phase comparator operates, here, by both transitions at the data transmission rate.

This variant has the manual and the automatic versions, the difference is only in the phase comparator. The variable pulse  $P_v$  supported in the first flip flop with exor is equal in the two versions, but the fixed pulse  $P_f$  is different [1, 2].

**A. Both transitions at the rate and manual**

The manual version has a phase comparator, where the fixed pulse Pf is produced by an exor with a delay  $\Delta t=T/2$ , that needs a previous manual adjustment (Fig. 3)

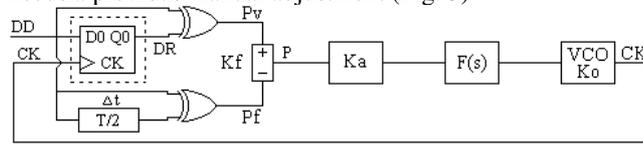


Fig. 3 Synchronizer both at the rate and manual (b-m)

The variable pulse Pv minus the fixed pulse Pf ( $Pv-Pf$ ) determines the error phase that controls the VCO.

**B. Both transitions at the rate and automatic**

The automatic version has a phase comparator where the fixed pulse Pf is produced automatically by the second flip flop with exor, without previous adjustment (Fig. 4).

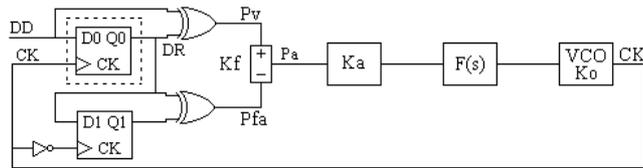


Fig. 4 Synchronizer both at the rate and automatic (b-a)

The variable pulse Pv minus the fixed pulse Pf ( $Pv-Pf$ ) determines the error phase that controls the VCO.

**V. SYNCHRONIZERS OPERATING AT HALF RATE**

The synchronizer with its phase comparator operates, here, by positive transitions at half data transmission rate.

This variant has the manual and the automatic versions, the difference is only in the phase comparator. The variable pulse Pv, based in the two first flip flops with multiplexer, is equal in the two versions, but the fixed pulse Pf is produced from a different way [3, 4].

**A. Positive transitions at half rate and manual**

The manual version has a phase comparator, where the fixed pulse Pf is produced by an exor with a delay  $\Delta t=T/2$ , that needs a previous manual adjustment (Fig. 5).

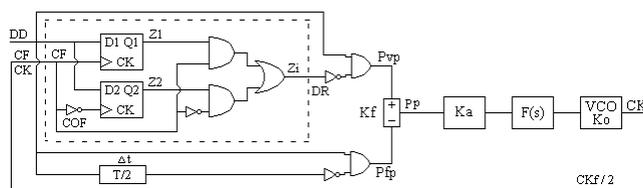


Fig. 5 Synchronizer positive at half rate and manual (p-m/2)

The variable pulse Pv minus the fixed pulse Pf ( $Pv-Pf$ ) determines the error phase that controls the VCO.

**B. Positive transitions at half rate and automatic**

The automatic version has a phase comparator, where the fixed pulse Pf is produced automatically by the seconds flip flops and multiplexer with exor, without previous adjustment (Fig. 6).

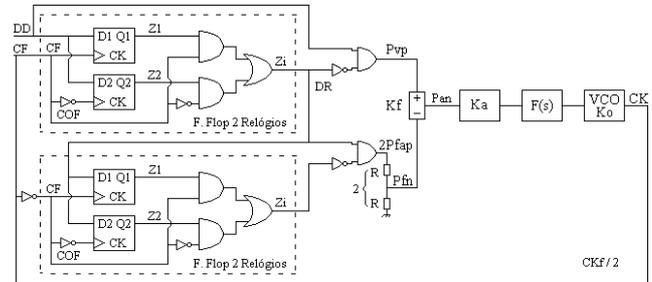


Fig. 6 Synchronizer positive at half rate and automatic (p-a/2)

The variable pulse Pv minus the fixed pulse Pf ( $Pv-Pf$ ) determines the error phase that controls the VCO.

**VI. DESIGN, TESTS AND RESULTS**

We will present the design, the tests and the results of the referred synchronizers [5].

**A. Design**

To get guaranteed results, it is necessary to dimension all the synchronizers with equal conditions. Then it is necessary to design all the loops with identical linearized transfer functions.

The general loop gain is  $Kl=Kd.Ko=Ka.Kf.Ko$  where Kf is the phase comparator gain, Ko is the VCO gain and Ka is the control amplification factor that permits the desired characteristics.

For analysis facilities, we use a normalized transmission rate  $tx=1$ baud, what implies also normalized values for the others dependent parameters. So, the normalized clock frequency is  $fCK=1$ Hz.

We choose a normalized external noise bandwidth  $Bn = 5$ Hz and a normalized loop noise bandwidth  $B1 = 0.02$ Hz. Later, we can disnormalize these values to the appropriated transmission rate tx.

Now, we will apply a signal with noise ratio SNR given by the signal amplitude Aef, noise spectral density No and external noise bandwidth Bn, so the  $SNR = A_{ef}^2 / (No \cdot Bn)$ . But, No can be related with the noise variance  $\sigma n$  and inverse sampling  $\Delta\tau=1/Samp$ , then  $No=2\sigma n^2 \cdot \Delta\tau$ , so  $SNR=A_{ef}^2 / (2\sigma n^2 \cdot \Delta\tau \cdot Bn) = 0.5^2 / (2\sigma n^2 * 10^{-3} * 5) = 25/\sigma n^2$ .

After, we observe the output jitter UI as function of the input signal with noise SNR. The dimension of the loops is

- 1<sup>st</sup> order loop:

The loop filter  $F(s)=1$  with cutoff frequency 0.5Hz ( $B_p=0.5$  Hz is 25 times bigger than  $B_l=0.02$ Hz) eliminates only the high frequency, but maintain the loop characteristics. The transfer function is

$$H(s) = \frac{G(s)}{1+G(s)} = \frac{KdKoF(s)}{s+KdKoF(s)} = \frac{KdKo}{s+KdKo} \quad (1)$$

the loop noise bandwidth is

$$B_l = \frac{KdKo}{4} = Ka \frac{KfKo}{4} = 0.02Hz \quad (2)$$

Then, for the analog synchronizers, the loop bandwidth is  $B_l=0.02=(Ka.Kf.Ko)/4$  with ( $K_m=1, A=1/2, B=1/2; K_o=2\pi$ )

$$(Ka.Km.A.B.Ko)/4 = 0.02 \rightarrow Ka=0.08*2/\pi \quad (3)$$

For the hybrid synchronizers, the loop bandwidth is  $B_l=0.02=(Ka.Kf.Ko)/4$  with ( $K_m=1, A=1/2, B=0.45; K_o=2\pi$ )

$$(Ka.Km.A.B.Ko)/4 = 0.02 \rightarrow Ka=0.08*2.2/\pi \quad (4)$$

For the combinational synchronizers, the loop bandwidth is  $B_l=0.02=(Ka.Kf.Ko)/4$  with ( $K_f=1/\pi; K_o=2\pi$ )

$$(Ka*1/\pi*2\pi)/4 = 0.02 \rightarrow Ka=0.04 \quad (5)$$

For the sequential synchronizers, the loop bandwidth is  $B_l=0.02=(Ka.Kf.Ko)/4$  with ( $K_f=1/2\pi; K_o=2\pi$ )

$$(Ka*1/2\pi*2\pi)/4 = 0.02 \rightarrow Ka=0.08 \quad (6)$$

The jitter depends on the RMS signal  $A_{ef}$ , on the power spectral density  $N_o$  and on the loop noise bandwidth  $B_l$ .

For analog PLL the jitter is

$$\sigma\phi^2 = B_l N_o / A_{ef}^2 = B_l \cdot 2 \cdot \sigma_n^2 \cdot \Delta\tau = 0.02 * 10^{-3} * 2 \sigma_n^2 / 0.5^2 = 16 * 10^{-5} \cdot \sigma_n^2$$

For the others PLLs the jitter formula is more complicated.

- 2<sup>nd</sup> order loop:

The second order loop is not shown here, but the results are identical to the ones obtained above for the first order loop.

### B. Tests

The following figure (Fig. 7) shows the setup that was used to test the various synchronizers.

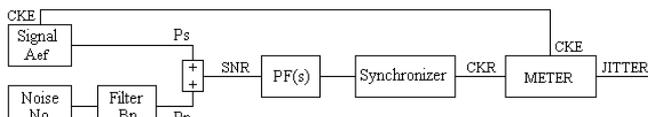


Fig. 7 Block diagram of the test setup

The receiver recovered clock with jitter is compared with the emitter original clock without jitter, the difference is the jitter of the received clock.

### C. Jitter measurer (Meter)

The jitter measurer (Meter) consists of a RS flip flop, which detects the random variable phase of the recovered clock (CKR), relatively to the fixed phase of the emitter clock (CKE). This relative random phase variation is the recovered clock jitter (Fig. 8).

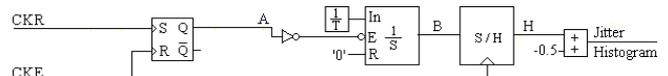


Fig. 8 The jitter measurer (Meter)

The other blocks convert this random phase variation into a random amplitude variation, which is the jitter histogram. Then, the jitter histogram is sampled and processed by an appropriate program, providing the RMS jitter and the peak to peak jitter.

### D. Results

We will present the results (jitter-noise graphics) for the prefilter with the four synchronizers.

Fig. 9 shows the jitter-SNR curves of the prefilter bandwidth  $B_1=\infty$  with the four synchronizers namely both transitions at the rate manual (b-m), both transitions at the rate automatic (b-a), positive transitions at half rate manual (p-m/2) and positive transitions at half rate automatic (p-a/2).

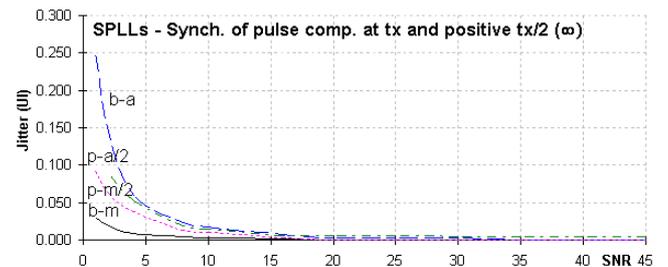


Fig. 9 Jitter-SNR curves of B1- 4 synchronizers(b-m,b-a,p-m/2,p-a/2)

We verify that, generally, the output jitter UIRMS decreases more or less exponentially with the input SNR increasing.

For prefilter B1, for high SNR, the four synchronizer jitter curves tend to be similar. However, for low SNR, the manual versions (b-m, b-m/2) are slightly better than the similar automatic versions (b-a, b-a/2). The both transitions at rate manual (b-m) is slightly the best.

Fig. 10 shows the jitter-SNR curves of the prefilter bandwidth  $B_2=2.tx$  with the four synchronizers namely both transitions at the rate manual (b-m), both transitions at the rate automatic (b-a), positive transitions at half rate manual (p-m/2) and positive transitions at half rate automatic (p-a/2).

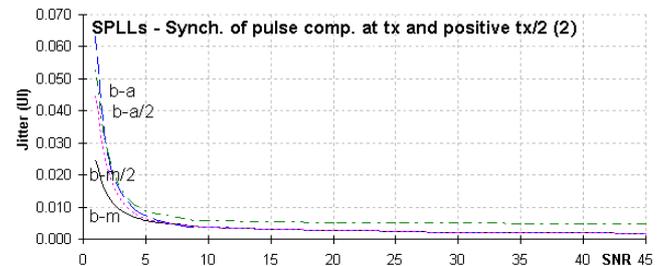


Fig. 10 Jitter-SNR curves of B2-4 synchronizers(b-m,b-a,p-m/2,p-a/2)

For prefilter B2, we verify that, it becomes the jitter-SNR curves more similar between themselves. For high SNR, it harms slightly the jitter-SNR curves. However, for low SNR, it benefits significantly the jitter - SNR curves.

Fig. 11 shows the jitter-SNR curves of the prefilter bandwidth B3=1.tx with the four synchronizers namely both transitions at the rate manual (b-m), both transitions at rate automatic (b-a), positive transitions at half rate manual (p-m/2) and positive transitions at half rate automatic (p-a/2).

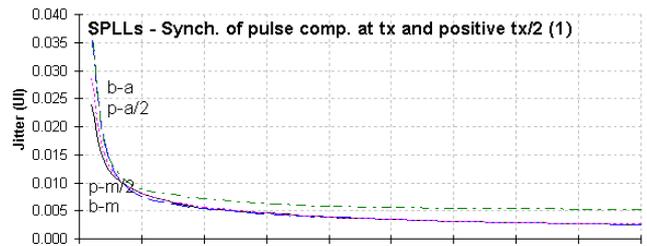


Fig. 11 Jitter-SNR curves of B3-4 synchronizers(b-m,b-a,p-m/2,p-a/2)

For prefilter B3, we verify that, it becomes the jitter-SNR curves still more similar between themselves. For high SNR, it harms more slightly the jitter-SNR curves. However, for low SNR, it benefits less significantly the jitter - SNR curves.

## VI. CONCLUSION AND FUTURE WORK

We studied three prefilter bandwidths ( $B1=\infty$ ,  $B2=2.tx$ ,  $B3=1.tx$ ) with four symbol synchronizers, namely both transitions at rate manual (b-m), both transitions at rate automatic (b-a), positive transitions at half rate manual (b-m/2) and positive transitions at half rate automatic (b-a/2). Then, we measured their jitter-SNR.

We observed that, in general, the output jitter decreases almost exponentially with the input SNR increasing.

For prefilter  $B1=\infty$ , we verified that, for high SNR, the four synchronizers jitter curves tend to be similar, this is comprehensible since all the four synchronizers are digital and have similar noise margin. However, for low SNR, the manual versions (b-m, b-m/2) are significantly better than the similar automatic versions (p-a, p-a/2), this is comprehensible since the automatic versions have more error states propagation that contributes to the jitter. The version both transitions at rate (b-m) is slightly the best because has less digital states.

For prefilter  $B2=2.tx$ , we verified that, it becomes the jitter curves more similar between themselves. For high SNR, it harms slightly the jitter-SNR curves. However, for low SNR, it benefits significantly the jitter-SNR curves.

For prefilter  $B3=1.tx$ , we verified that, it becomes the jitter curves still more similar between themselves. For high SNR, it harms more slightly the jitter-SNR curves. However, for low SNR, it benefits less significantly the jitter-SNR curves.

So, the prefilter, for high SNR, degrades slightly the jitter-SNR curves, but for low SNR, benefits significantly the jitter-SNR curves and becomes them more similar.

In the future, we are planning to extend the present study to other types of synchronizers.

## ACKNOWLEDGMENTS

The authors are grateful to the program FCT (Foundation for sScience and Technology) / POCI2010.

## REFERENCES

- [1] Jean C. Imbeaux, "performance of the delay-line multiplier circuit for clock and carrier synchronization in Digital Satellite Communications". IEEE Journal on Selected Areas in Communications pp.82-95 Jan. 1983.
- [2] Werner Rosenkranz, "Phase Locked Loops with limiter phase detectors in the presence of noise", IEEE Transactions on Communications com-30 N°10 pp.2297-2304. Oct 1982.
- [3] Hans H. Witte, "A Simple Clock Extraction Circuit Using a Self Sustaining Monostable Multivibrator Output Signal", Electronics Letters, Vol.19, Is.21, pp.897-898, Oct 1983.
- [4] Charles R. Hogge, "A Self Correcting Clock Recovery Circuit", IEEE Transactions on Electron Devices pp.2704-2706 Dec 1985.
- [5] Antonio D. Reis, Jose F. Rocha, Atilio S. Gameiro and Jose P. Carvalho "A New Technique to Measure the Jitter", Proc. III Conf. on Telecommunications pp.64-67 Ffoz-PT 23-24 Apr 2001.
- [6] Marvin K. Simon and William C. Lindsey, "Tracking Performance of Symbol Synchronizers for Manchester Coded Data", IEEE Transactions on Communications Vol. com-2.5 N°4, pp.393-408, April 1977.
- [7] Jeffrey B. Carruthers, D. D. Falconer, H. M. Sandler and L. Strawczynski, "Bit Synchronization in the Presence of Co-Channel Interference", Proc. Conf. on Electrical and Computer Engineering pp.4.1.1-4.1.7, Ottawa-CA 3-6 Sep. 1990.
- [8] Johannes Huber and Weilin Liu "Data-Aided Synchronization of Coherent CPM-Receiver" IEEE Transactions on Communications Vol.40 N°1, pp.178-189, Jan. 1992.
- [9] Antonio A. D'Amico, Aldo N. D'Andrea and Ruggero Reggianni, "Efficient Non-Data-Aided Carrier and Clock Recovery for Satellite DVB at Very Low SNR", IEEE Jou. on Satellite Areas in Comm. Vol.19 N°12 pp.2320-2330, Dec. 2001.
- [10] Rostislav Dobkin, Ran Ginosar and Christos P. Sotiriou "Data Synchronization Issues in GALS SoCs", Proc. 10th International Symposium on Asynchronous Circuits and Systems, pp.CD-Ed., Crete-Greece 19-23 Apr. 2004.
- [11] N. Noels, H. Steendam and M. Moeneclaey, "Effectiveness Study of Code-Aided and Non-Code-Aided ML-Based Feedback Phase Synchronizers", Proc. IEEE Intern. Conference on Communications (ICC'06) pp.2946-2951, Istanbul-TK, 11-15 Jun 2006.
- [12] Antonio D. Reis, Jose F. Rocha, Atilio S. Gameiro and Jose P. Carvalho "Effects of the Prefilter Type on Digital Symbol Synchronizers", Proc. VII Symposium on Enabling Optical Network and Sensors (SEONS 2009) pp.35-36, Lisbon-PT 26-26 June 2009.

# Sequential Symbol Synchronizers based on Pulse Comparison at Quarter Rate

Antonio D. Reis<sup>1,2</sup> and José P. Carvalho<sup>1</sup>

Dep. Física / Unidade D. Remota

<sup>1</sup>Universidade da Beira Interior, 6200 Covilhã, Portugal

adreis@ubi.pt, pacheco@ubi.pt

José F. Rocha<sup>2</sup> and Atílio S. Gameiro<sup>2</sup>

Dep. Electrónica e Telecom. / Instituto Telecom.

<sup>2</sup>Universidade de Aveiro, 3810 Aveiro, Portugal

frocha@det.ua.pt, amg@det.ua.pt

**Abstract-** This work presents the synchronizer based on pulse comparison, between variable and fixed pulses. This synchronizer has two variants, one operating by both transitions at the bit rate and other operating by both transitions at quarter rate. Each variant has two versions which are the manual and the automatic. The objective is to study the four synchronizers and evaluate their output jitter UIRMS (Unit Interval Root Mean Square) versus input SNR (Signal Noise Ratio).

**Keywords - Prefilter; Synchronizers; Communication systems.**

## I. INTRODUCTION

This work studies the sequential symbol synchronizer, with a phase comparator based on a pulse comparison, between a variable pulse  $P_v$  and a fixed reference pulse  $P_f$ .

The synchronizer has four types supported in two variants, one operating by both transitions at rate and other operating by both transitions at quarter rate. The variant at the rate has two versions namely the manual (b-m) and the automatic (b-a). The variant at quarter rate has also two versions, namely the manual (b-m/4) and the automatic (b-a/4) [1, 2, 3, 4, 5].

The difference between them is only in the phase comparator since the other blocks are equal [6, 7, 8].

The error pulse  $P_e$  ( $P_v - P_f$ ) controls the VCO (Voltage Controlled Oscillator) to synchronize with the input data.

The variable pulse  $P_v$  is common to manual and automatic versions but the fixed pulse  $P_f$  is different [9, 10].

The VCO output is the clock, with good quality, that samples the input data at the maximum opening eye diagram and retimes its bit duration [11, 12].

Fig. 1 shows the blocks diagram of the synchronizer.

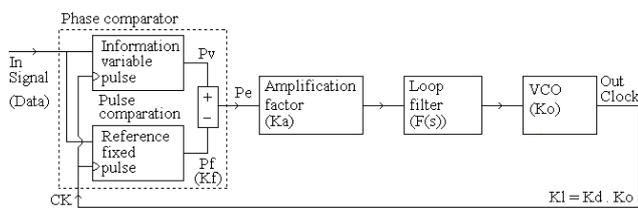


Fig. 1 Synchronizer based on pulse comparison

$K_f$  is the phase comparator gain,  $F(s)$  is the loop filter,  $K_o$  is the VCO gain and  $K_a$  is the loop amplification factor that controls the root locus and then the loop characteristics.

Following, we present the variant at bit rate with their manual and automatic versions. Next, we present the variant at the half bit rate with their manual and automatic versions.

After, we present the design and tests. Then, we present the results. Finally, we present the conclusions.

## II. STATE OF THE ART, PROBLEM AND SOLUTION

In priori and actual-art state, were developed various synchronizers operating, initially, at the bit rate. After, we developed synchronizers operating at half rate. Now, we present synchronizers operating at a quarter rate. This contribution increases the know-how about synchronizers.

Our motivation is to create new synchronizers and evaluate their performance. The problem is that the synchronizers have speed limitations. One solution proposes internal low frequency operation, but external high speed rate [1, 2, 3, 4].

## III. SYNCHRONIZERS OPERATING AT THE RATE

The synchronizer with its VCO operates, here, at the data transmission rate.

This variant has the manual and the automatic versions, the difference is only in the phase comparator. The variable pulse  $P_v$  consists of first flip flop with exor and is equal in the two versions, but the fixed pulse  $P_f$  is different [1, 2].

### A. Operation at the rate and manual version

The manual version has a phase comparator, where the fixed pulse  $P_f$  is produced by an exor with a delay  $\Delta t = T/2$ , that needs a previous manual adjustment (Fig. 2)

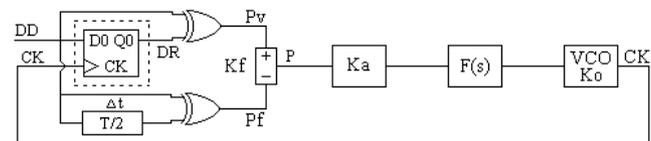


Fig. 2 Synchronizer at the rate and manual (b-m)

The variable pulse  $P_v$  minus the fixed pulse  $P_f$  ( $P_v - P_f$ ) determines the error phase that controls the VCO.

Fig. 3 shows the waveforms of the synchronizer operating at the rate and manual version.

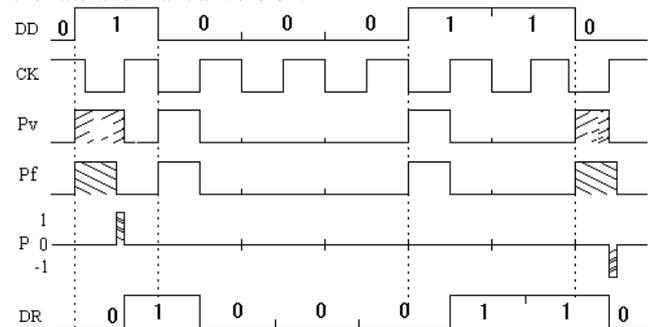


Fig. 3 Waveforms of the synchronizer at the rate and manual

The error pulse  $P_e$  diminishes and disappear at the equilibrium point.

**B. Operation at the rate and automatic version**

The automatic version has a phase comparator where the fixed pulse Pf is produced automatically by the second flip flop with exor, without previous adjustment (Fig. 4).

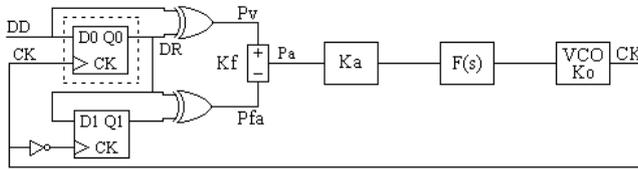


Fig. 4 Synchronizer at the rate and automatic (b-a)

The variable pulse Pv minus the fixed pulse Pf (Pv-Pf) determines the error phase that controls the VCO.

Fig. 5 shows the waveforms of the synchronizer operating at the rate and automatic version.

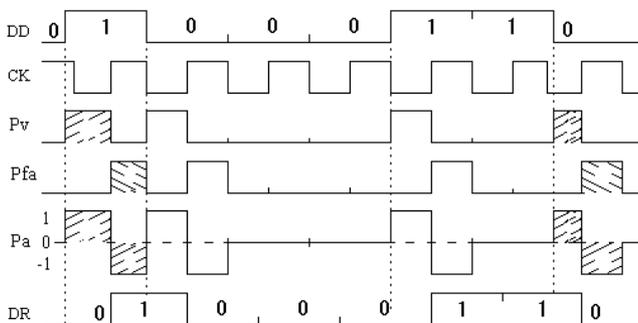


Fig. 5 Waveforms of the synchronizer at the rate and automatic

The error pulse Pe don't disappear, but the variable area Pv is equal to the fixed one Pf at the equilibrium point.

**IV. SYNCHRONIZERS OPERATING AT QUARTER RATE**

The synchronizer with its phase comparator operates, here, by both transitions at quarter data transmission rate.

This variant has the manual (b-m/4) and the automatic (b-a/4) versions, the difference is only in the phase comparator. The variable pulse Pv, based in the four first flip flops with multiplexer and exor, is equal in the two versions, but the fixed pulse Pf is produced from a different way [3, 4].

**A. Operation at quarter rate and manual version**

The manual version has a phase comparator, where the fixed pulse Pf is produced by an exor with a delay  $\Delta t = T/2$ , that needs a previous manual adjustment (Fig. 6).

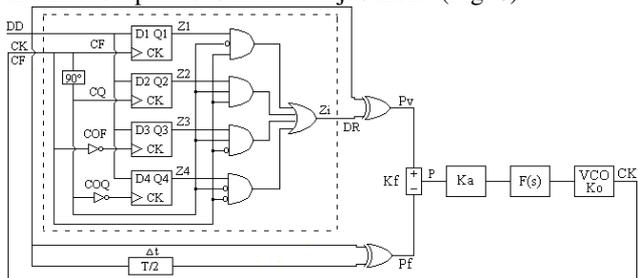


Fig. 6 Synchronizer at quarter rate and manual (b-m/4)

The variable pulse Pv minus the fixed pulse Pf (Pv-Pf) determines the error phase that controls the VCO.

Fig. 7 shows the waveforms of the synchronizer operating at quarter rate and manual version.

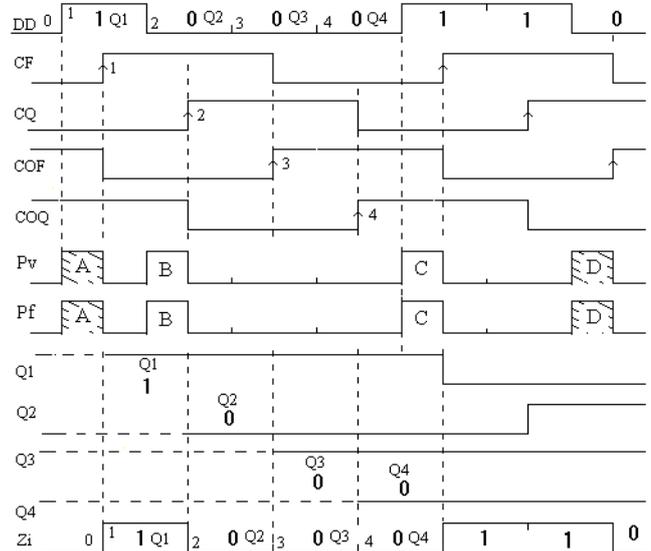


Fig. 7 Waveforms of the synchronizer at quarter rate and manual

The error pulse Pe diminishes and disappear at the equilibrium point

**B. Operation at quarter rate and automatic version**

The automatic version has a phase comparator, where the fixed pulse Pf is produced automatically by the seconds flip flops and multiplexer with exor, without previous adjustment (Fig. 8).

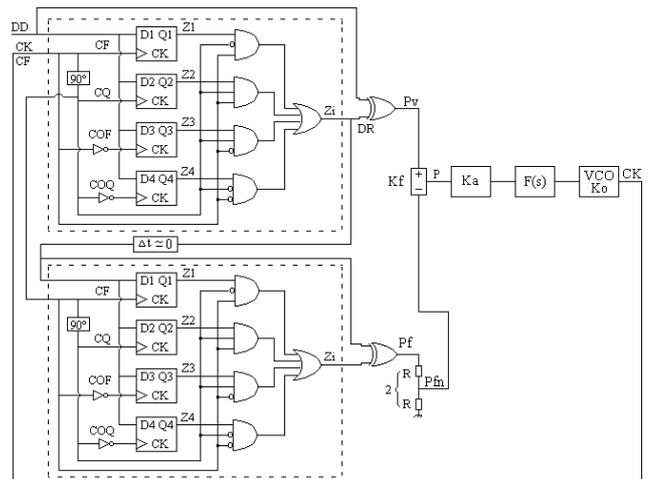


Fig. 8 Synchronizer at quarter rate and automatic (b-a/4)

The variable pulse Pv minus the fixed pulse Pf (Pv-Pf) determines the error phase that controls the VCO.

Fig. 9 shows the waveforms of the synchronizer at quarter rate and automatic version.

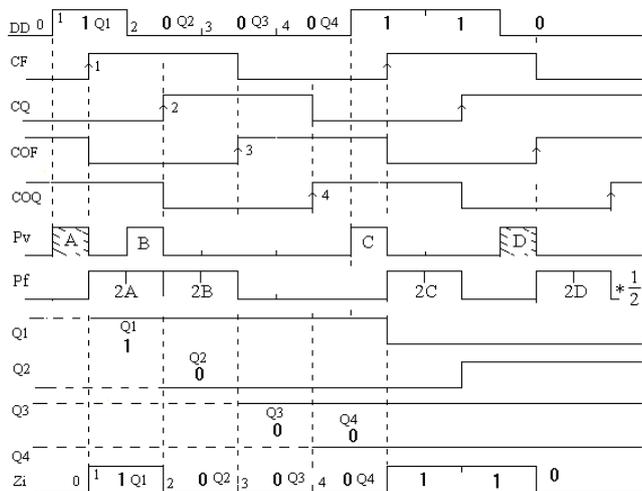


Fig. 9 Waveforms of the synchronizer at quarter rate and automatic

The error pulse  $P_e$  don't disappear but the positive area  $P_v$  is equal to the negative  $P_f$  at the equilibrium point.

### V. DESIGN, TESTS AND RESULTS

We will present the design, the tests and the results of the referred synchronizers [5].

#### A. Design

To get guaranteed results, it is necessary to dimension all the synchronizers with equal conditions. Then it is necessary to design all the loops with identical linearized transfer functions.

The general loop gain is  $Kl=Kd.Ko=Ka.Kf.Ko$  where  $Kf$  is the phase comparator gain,  $Ko$  is the VCO gain and  $Ka$  is the control amplifier factor that permits the desired characteristics.

For analysis facilities, we use a normalized transmission rate  $t_x=1$ baud, what implies also normalized values for the others dependent parameters. So, the normalized clock frequency is  $f_{CK}=1$ Hz.

We choose a normalized external noise bandwidth  $B_n = 5$ Hz and a normalized loop noise bandwidth  $B_l = 0.02$ Hz. Later, we can disnormalize these values to the appropriated transmission rate  $t_x$ .

Now, we will apply a signal with noise ratio SNR given by the signal amplitude  $A_{ef}$ , noise spectral density  $N_o$  and external noise bandwidth  $B_n$ , so the  $SNR = A_{ef}^2/(N_o.B_n)$ . But,  $N_o$  can be related with the noise variance  $\sigma_n$  and inverse sampling  $\Delta\tau=1/Samp$ , then  $N_o=2\sigma_n^2.\Delta\tau$ , so  $SNR=A_{ef}^2/(2\sigma_n^2.\Delta\tau.B_n) = 0.5^2/(2\sigma_n^2*10^{-3}*5) = 25/\sigma_n^2$ .

After, we observe the output jitter  $UI$  as function of the input signal with noise SNR. The dimension of the loops is

- 1<sup>st</sup> order loop:

The loop filter  $F(s)=1$  with cutoff frequency 0.5Hz ( $B_p=0.5$  Hz is 25 times bigger than  $B_l=0.02$ Hz) eliminates only the high frequency, but maintain the loop characteristics.

The transfer function is

$$H(s) = \frac{G(s)}{1+G(s)} = \frac{KdKoF(s)}{s+KdKoF(s)} = \frac{KdKo}{s+KdKo} \quad (1)$$

the loop noise bandwidth is

$$B_l = \frac{KdKo}{4} = Ka \frac{KfKo}{4} = 0.02Hz \quad (2)$$

Then, for the analog synchronizers, the loop bandwidth is

$$B_l=0.02=(Ka.Kf.Ko)/4 \text{ with } (K_m=1, A=1/2, B=1/2; K_o=2\pi) \\ (Ka.K_m.A.B.Ko)/4 = 0.02 \rightarrow Ka=0.08*2/\pi \quad (3)$$

For the hybrid synchronizers, the loop bandwidth is

$$B_l=0.02=(Ka.Kf.Ko)/4 \text{ with } (K_m=1, A=1/2, B=0.45; K_o=2\pi) \\ (Ka.K_m.A.B.Ko)/4 = 0.02 \rightarrow Ka=0.08*2.2/\pi \quad (4)$$

For the combinational synchronizers, the loop bandwidth is

$$B_l=0.02=(Ka.Kf.Ko)/4 \text{ with } (K_f=1/\pi; K_o=2\pi) \\ (Ka*1/\pi*2\pi)/4 = 0.02 \rightarrow Ka=0.04 \quad (5)$$

For the sequential synchronizers, the loop bandwidth is

$$B_l=0.02=(Ka.Kf.Ko)/4 \text{ with } (K_f=1/2\pi; K_o=2\pi) \\ (Ka*1/2\pi*2\pi)/4 = 0.02 \rightarrow Ka=0.08 \quad (6)$$

The jitter depends on the RMS signal  $A_{ef}$ , on the power spectral density  $N_o$  and on the loop noise bandwidth  $B_l$ .

For analog PLL the jitter is

$$\sigma_{\phi}^2 = B_l.N_o/A_{ef}^2 = B_l.2.\sigma_n^2.\Delta\tau = 0.02*10^{-3}*2\sigma_n^2/0.5^2 = 16*10^{-5}.\sigma_n^2$$

For the others PLLs the jitter formula is more complicated.

- 2<sup>nd</sup> order loop:

The second order loop is not shown here, but the results are identical to the ones obtained above for the first order loop.

#### B. Tests

The following figure (Fig. 10) shows the setup that was used to test the various synchronizers.

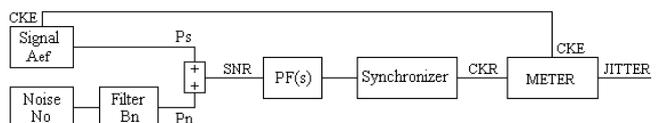


Fig. 10 Block diagram of the test setup

The receiver recovered clock with jitter is compared with the emitter original clock without jitter, the difference is the jitter of the received clock.

#### C. Jitter measurer (Meter)

The jitter measurer (Meter) consists of a RS flip flop, which detects the random variable phase of the recovered clock (CKR), relatively to the fixed phase of the emitter clock (CKE). This relative random phase variation is the recovered clock jitter (Fig. 11).

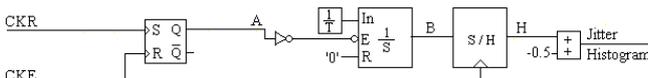


Fig. 11 The jitter measurer (Meter)

The other blocks convert this random phase variation into a random amplitude variation, which is the jitter histogram.

Then, the jitter histogram is sampled and processed by an appropriate program, providing the RMS jitter and the peak to peak jitter.

## D. Results

We will present the four synchronizer results in terms of output jitter UIRMS versus input SNR.

Fig. 12 shows the jitter-SNR curves of the four synchronizers with both transitions, at rate manual version (b-m), at rate automatic version (b-a), at quarter rate manual version (b-m/4) and at quarter rate automatic version (b-a/4).

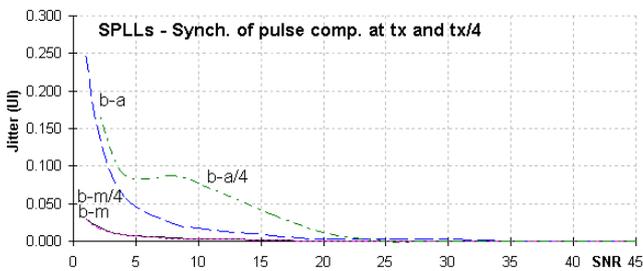


Fig. 12 Jitter-SNR curves of the 4 synchronizers(b-m,b-a,b-m/4,b-a/4)

We see that, in general, the output jitter UIRMS decreases gradually with the input SNR increasing. However, the both quarter rate automatic (b-a/4) has some irregularities.

For high SNR, the four synchronizer jitter curves tend to be similar. However, for low SNR, the manual versions (b-m, b-m/4) are significantly better than the automatic versions (b-a, b-a/4), the both transition at rate manual (b-m) is slightly the best. Also, for an intermediate SNR ( $SNR \approx 10$ ), the both transitions quarter rate automatic (b-a/4) has a significant jitter perturbation, due to some losses of synchronism.

## V. CONCLUSION AND FUTURE WORK

We studied four synchronizers using both transitions, with two variants, one operating at the rate that has two versions namely the manual (b-m) and the automatic (b-a) and other variant operating at quarter rate that has also two versions namely the manual (b-m/4) and the automatic (b-a/4). Then, we tested their jitter - noise curves.

We observed that, in general, the output jitter curves decreases gradually with the input SNR increasing. However, the quarter rate automatic (b-a/4) has some irregularities.

We verified that, for high SNR, the four synchronizers jitter curves tend to be similar, this is comprehensible since the synchronizers are digital and have similar noise margin.

However, for low SNR, the manual versions (b-m, b-m/4) are significantly better than the automatic versions (b-a, b-a/4), this is comprehensible since the automatic versions have more digital states than the manual versions, then the error state propagation effects is aggravated.

The version at rate manual (b-m) is slightly the best because has less digital states. On the other hand, the version at quarter rate automatic (b-a/4) has a significant jitter perturbation ( $SNR \approx 10$ ) due to some losses of synchronism.

In the future, we are planning to extend the present study to other types of synchronizers.

## ACKNOWLEDGMENT

The authors are grateful to the program FCT (Foundation for sScience and Technology) / POCI2010.

## REFERENCES

- [1] Jean C. Imbeaux, "performance of the delay-line multiplier circuit for clock and carrier synchronization in Digital Satellite Communications". IEEE Journal on Selected Areas in Communications pp.82-95 Jan. 1983.
- [2] Werner Rosenkranz, "Phase Locked Loops with limiter phase detectors in the presence of noise", IEEE Transactions on Communications com-30 N°10 pp.2297-2304. Oct 1982.
- [3] Hans H. Witte, "A Simple Clock Extraction Circuit Using a Self Sustaining Monostable Multivibrator Output Signal", Electronics Letters, Vol.19, Is.21, pp.897-898, Oct 1983.
- [4] Charles R. Hogge, "A Self Correcting Clock Recovery Circuit", IEEE Transactions on Electron Devices pp.2704-2706 Dec 1985.
- [5] Antonio D. Reis, Jose F. Rocha, Atilio S. Gameiro and Jose P. Carvalho "A New Technique to Measure the Jitter", Proc. III Conf. on Telecommunications pp.64-67 Foz-PT 23-24 Apr 2001.
- [6] Marvin K. Simon and William C. Lindsey, "Tracking Performance of Symbol Synchronizers for Manchester Coded Data", IEEE Transactions on Communications Vol. com-25 N°4, pp.393-408, April 1977.
- [7] Jeffrey B. Carruthers, D. D. Falconer, H. M. Sandler and L. Strawczynski, "Bit Synchronization in the Presence of Co-Channel Interference", Proc. Conf. on Electrical and Computer Engineering pp.4.1.1-4.1.7, Ottawa-CA 3-6 Sep. 1990.
- [8] Johannes Huber and Weilin Liu "Data-Aided Synchronization of Coherent CPM-Receiver" IEEE Transactions on Communications Vol.40 N°1, pp.178-189, Jan. 1992.
- [9] Antonio A. D'Amico, Aldo N. D'Andrea and Ruggero Reggianni, "Efficient Non-Data-Aided Carrier and Clock Recovery for Satellite DVB at Very Low SNR", IEEE Jou. on Satellite Areas in Comm. Vol.19 N°12 pp.2320-2330, Dec. 2001.
- [10] Rostislav Dobkin, Ran Ginosar and Christos P. Sotiriou "Data Synchronization Issues in GALS SoCs", Proc. 10th International Symposium on Asynchronous Circuits and Systems, pp.CD-Ed., Crete-Greece 19-23 Apr. 2004.
- [11] N. Noels, H. Steendam and M. Moeneclaey, "Effectiveness Study of Code-Aided and Non-Code-Aided ML-Based Feedback Phase Synchronizers", Proc. IEEE Intern. Conference on Communications (ICC'06) pp.2946-2951, Istanbul-TK, 11-15 Jun 2006.
- [12] Antonio D. Reis, Jose F. Rocha, Atilio S. Gameiro and Jose P. Carvalho "Effects of the Prefilter Type on Digital Symbol Synchronizers", Proc. VII Symposium on Enabling Optical Network and Sensors (SEONs 2009) pp.35-36, Lisbon-PT 26-26 June 2009.

# Prefilter Bandwidth Effects in Sequential Symbol Synchronizers based on Pulse Comparison at Quarter Rate

Antonio D. Reis<sup>1,2</sup> and José P. Carvalho<sup>1</sup>

Dep. Física / Unidade D. Remota

<sup>1</sup>Universidade da Beira Interior, 6200 Covilhã, Portugal  
adreis@ubi.pt, pacheco@ubi.pt

José F. Rocha<sup>2</sup> and Afílio S. Gameiro<sup>2</sup>

Dep. Electrónica e Telecom. / Instituto Telecom.

<sup>2</sup>Universidade de Aveiro, 3810 Aveiro, Portugal  
frocha@det.ua.pt, amg@det.ua.pt

**Abstract-** This work studies the effects of the prefilter bandwidth before the sequential symbol synchronizers based on pulse comparison at quarter rate. We consider three prefilter bandwidth namely  $B1=\infty$ ,  $B2=2.tx$  and  $B3=1.tx$ , where  $tx$  is the bit rate. We consider four synchronizers, one variant operates at the rate with versions manual (b-m) and automatic (b-a) and the other operates at quarter rate with versions manual (b-m/4) and automatic (b-a/4). The objective is to study the prefilter with the synchronizers and evaluate the output jitter UIRMS (Unit Interval Root Mean Square) versus input SNR (Signal Noise Ratio).

**Keywords - Prefilter; Synchronizers; Communication systems.**

## I. INTRODUCTION

This work studies three prefilter bandwidth effects on the jitter-SNR behavior of four sequential symbol synchronizers.

The Butterworth prefilter, applied before the synchronizer, switches their bandwidth between three values, first  $B1=\infty$ , after  $B2=2.tx$  and next  $B3=1.tx$ , where  $tx$  is the bit rate [1, 2].

The sequential symbol synchronizer has four types that can operate at rate with versions manual (b-m) and automatic (b-a) and can operate at quarter rate with versions manual (b-m/4) and automatic (b-a/4) [3, 4, 5, 6].

The difference between the four synchronizers is only in the phase comparator, since the other blocks are equal [7, 8].

The VCO (Voltage Controlled Oscillator) is the clock that samples the data and retimes its bits duration [9, 10, 11, 12].

Fig. 1 shows the prefilter followed of the synchronizer.

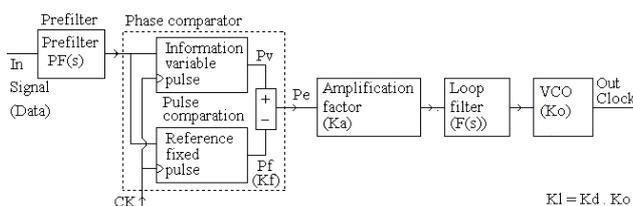


Fig. 1 Prefilter with the synchronizer based on pulse comparison

PF(s) is the prefilter. The synchronizer has various blocks, namely  $Kf$  is the phase comparator gain,  $F(s)$  is the loop filter,  $Ko$  is the VCO gain and  $Ka$  is the loop amplification factor that controls the root locus and loop characteristics.

Next, we present the state of the art. Then, we present the prefilter with three bandwidths ( $B1=\infty$ ,  $B2=2.tx$ ,  $B3=1.tx$ ).

After, we present the variant at bit rate with their manual and automatic versions. Next, we present the variant at quarter bit rate with their manual and automatic versions.

Following, we present the design and tests. Then, we present the results. Finally, we present the conclusions.

## II. STATE OF THE ART, PROBLEM AND SOLUTION

In priori and actual-art state various synchronizers were developed, some ones operate at the rate, others at half rate and others at quarter rate. We study their performance. This contribution increases the knowledge about synchronizers.

The motivation is to create new synchronizers operating at sub bit rate and evaluate their performance with noise [1, 2].

The problem is that the synchronizers' output jitter increases with the input SNR decreases. To solve or to minimize this problem, we propose a prefilter that attenuates the noise but unfortunately distorts slightly the signal [3, 4, 5, 6, 7, 12].

## II. PREFILTER BANDWIDTH EFFECTS

The prefilter, applied before the synchronizer, filters the noise but distorts slightly the signal. The prefilter bandwidth  $B$  switches between three values ( $B1=\infty$ ,  $B2=2.tx$ ,  $B3=1.tx$ ).

Fig. 2 shows the prefilter with their three bandwidths.

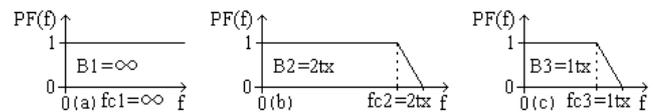


Fig. 2 Three prefilter bandwidths: a)  $B1=\infty$ ; b)  $B2=2.tx$ ; c)  $B3=1.tx$

a) First, as Fig.2a, the prefilter has a bandwidth equal to infinite ( $B1= \infty$ ).

b) Second, as Fig.2b, the prefilter has a bandwidth equal to times the bit rate ( $B2 = 2.tx$ ).

c) Third, as Fig.2c, the prefilter has a bandwidth equal to the bit rate ( $B3 = 1.tx$ ).

We will evaluate the three bandwidth effects ( $B1, B2, B3$ ) on the jitter-SNR curves of the four symbol synchronizers.

## III. SYNCHRONIZERS OPERATING AT THE RATE

The synchronizer with its VCO operates, here, at the data transmission rate.

This variant has the manual and the automatic versions, the difference is only in the phase comparator. The variable pulse  $Pv$  consists of first flip flop with exor and is equal in the two versions, but the fixed pulse  $Pf$  is different [1, 2].

### A. Operation at the rate and manual version

The manual version has a phase comparator, where the fixed pulse  $Pf$  is produced by an exor with a delay  $\Delta t=T/2$ , that needs a previous manual adjustment (Fig. 3)

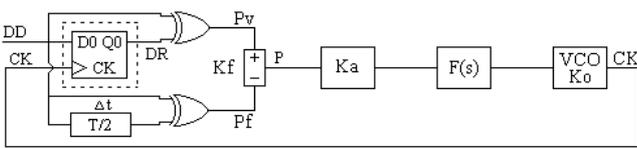


Fig. 3 Synchronizer at the rate and manual (b-m)

The variable pulse Pv minus the fixed pulse Pf (Pv-Pf) determines the error phase that controls the VCO.

**B. Operation at the rate and automatic version**

The automatic version has a phase comparator where the fixed pulse Pf is produced automatically by the second flip flop with exor, without previous adjustment (Fig. 4).

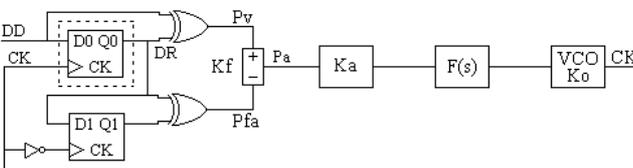


Fig. 4 Synchronizer at the rate and automatic (b-a)

The variable pulse Pv minus the fixed pulse Pf (Pv-Pf) determines the error phase that controls the VCO.

**IV. SYNCHRONIZERS OPERATING AT QUARTER RATE**

The synchronizer with its phase comparator operates, here, by both transitions at quarter data transmission rate.

This variant has the manual (b-m/4) and the automatic (b-a/4) versions, the difference is only in the phase comparator. The variable pulse Pv, based in the four first flip flops with multiplexer and exor, is equal in the two versions, but the fixed pulse Pf is produced from a different way [3, 4].

**A. Operation at quarter rate and manual version**

The manual version has a phase comparator, where the fixed pulse Pf is produced by an exor with a delay Δt=T/2, that needs a previous manual adjustment (Fig. 5).

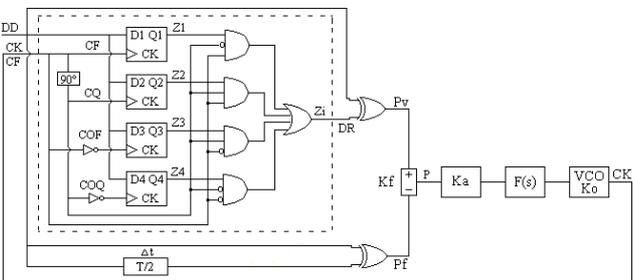


Fig. 5 Synchronizer at quarter rate and manual (b-m/4)

The variable pulse Pv minus the fixed pulse Pf (Pv-Pf) determines the error phase that controls the VCO.

**B. Operation at quarter rate and automatic version**

The automatic version has a phase detector, where the fixed pulse Pf is produced automatically by the seconds flip flops and multiplexer - exor, without previous adjustment (Fig. 6).

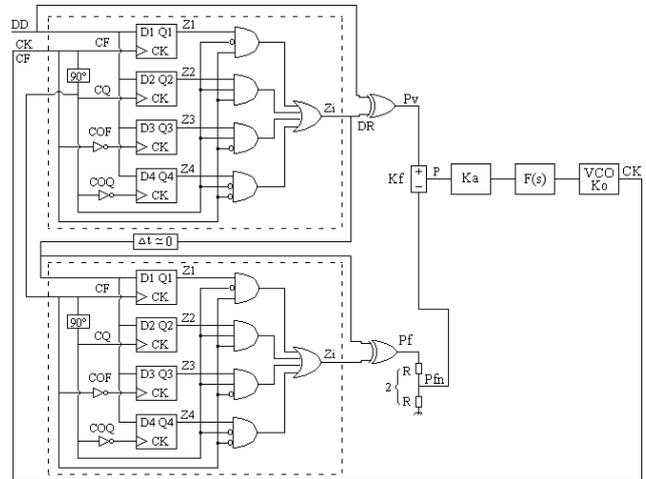


Fig. 6 Synchronizer at quarter rate and automatic (b-a/4)

The variable pulse Pv minus the fixed pulse Pf (Pv-Pf) determines the error phase that controls the VCO.

**V. DESIGN, TESTS AND RESULTS**

We will present the design, the tests and the results of the referred synchronizers [5].

**A. Design**

To get guaranteed results, it is necessary to dimension all the synchronizers with equal conditions. Then it is necessary to design all the loops with identical linearized transfer functions.

The general loop gain is  $Kl=Kd.Ko=Ka.Kf.Ko$  where Kf is the phase comparator gain, Ko is the VCO gain and Ka is the control amplification factor that permits the desired characteristics.

For analysis facilities, we use a normalized transmission rate  $t_x=1\text{baud}$ , what implies also normalized values for the others dependent parameters. So, the normalized clock frequency is  $f_{CK}=1\text{Hz}$ .

We choose a normalized external noise bandwidth  $B_n = 5\text{Hz}$  and a normalized loop noise bandwidth  $B_l = 0.02\text{Hz}$ . Later, we can disnormalize these values to the appropriated transmission rate  $t_x$ .

Now, we will apply a signal with noise ratio SNR given by the signal amplitude  $A_{ef}$ , noise spectral density  $N_o$  and external noise bandwidth  $B_n$ , so the  $SNR = A_{ef}^2/(N_o.B_n)$ . But,  $N_o$  can be related with the noise variance  $\sigma_n$  and inverse sampling  $\Delta\tau=1/\text{Samp}$ , then  $N_o=2\sigma_n^2.\Delta\tau$ , so  $SNR=A_{ef}^2/(2\sigma_n^2.\Delta\tau.B_n) = 0.5^2/(2\sigma_n^2*10^{-3}*5) = 25/\sigma_n^2$ . After, we observe the output jitter UI as function of the input signal with noise SNR. The dimension of the loops is

- 1<sup>st</sup> order loop:

The loop filter  $F(s)=1$  with cutoff frequency 0.5Hz ( $Bp=0.5$  Hz is 25 times bigger than  $B1=0.02$ Hz) eliminates only the high frequency, but maintain the loop characteristics.

The transfer function is

$$H(s) = \frac{G(s)}{1+G(s)} = \frac{KdKoF(s)}{s + KdKoF(s)} = \frac{KdKo}{s + KdKo} \quad (1)$$

the loop noise bandwidth is

$$B1 = \frac{KdKo}{4} = Ka \frac{KfKo}{4} = 0.02\text{Hz} \quad (2)$$

Then, for the analog synchronizers, the loop bandwidth is

$$B1=0.02=(Ka.Kf.Ko)/4 \text{ with } (Km=1, A=1/2, B=1/2; Ko=2\pi) \\ (Ka.Km.A.B.Ko)/4 = 0.02 \rightarrow Ka=0.08*2/\pi \quad (3)$$

For the hybrid synchronizers, the loop bandwidth is

$$B1=0.02=(Ka.Kf.Ko)/4 \text{ with } (Km=1, A=1/2, B=0.45; Ko=2\pi) \\ (Ka.Km.A.B.Ko)/4 = 0.02 \rightarrow Ka=0.08*2.2/\pi \quad (4)$$

For the combinational synchronizers, the loop bandwidth is

$$B1=0.02=(Ka.Kf.Ko)/4 \text{ with } (Kf=1/\pi; Ko=2\pi) \\ (Ka*1/\pi*2\pi)/4 = 0.02 \rightarrow Ka=0.04 \quad (5)$$

For the sequential synchronizers, the loop bandwidth is

$$B1=0.02=(Ka.Kf.Ko)/4 \text{ with } (Kf=1/2\pi; Ko=2\pi) \\ (Ka*1/2\pi*2\pi)/4 = 0.02 \rightarrow Ka=0.08 \quad (6)$$

The jitter depends on the RMS signal  $Aef$ , on the power spectral density  $No$  and on the loop noise bandwidth  $B1$ .

For analog PLL the jitter is

$$\sigma\phi^2 = B1.No/Aef^2 = B1.2.\sigma_n^2.\Delta\tau = 0.02*10^{-3}*2\sigma_n^2/0.5^2 = 16*10^{-5}.\sigma_n^2$$

For the others PLLs the jitter formula is more complicated.

- 2<sup>nd</sup> order loop:

The second order loop is not shown here, but the results are identical to the ones obtained above for the first order loop.

**B. Tests**

The following figure (Fig. 7) shows the setup that was used to test the various synchronizers.

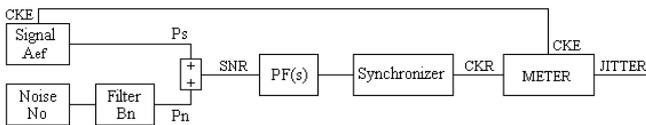


Fig. 7 Block diagram of the test setup

The receiver recovered clock with jitter is compared with the emitter original clock without jitter, the difference is the jitter of the received clock.

**C. Jitter measurer (Meter)**

The jitter measurer (Meter) consists of a RS flip flop, which detects the random variable phase of the recovered clock (CKR), relatively to the fixed phase of the emitter clock (CKE). This relative random phase variation is the recovered clock jitter (Fig. 8).

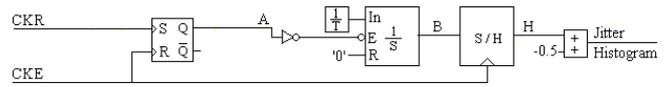


Fig. 8 The jitter measurer (Meter)

The other blocks convert this random phase variation into a random amplitude variation, which is the jitter histogram.

Then, the jitter histogram is sampled and processed by an appropriate program, providing the RMS jitter and the peak to peak jitter.

**D. Results**

We will present the results (jitter - noise graphics) for the prefilter with the four synchronizers.

Fig. 9 shows the jitter-SNR curves of the prefilter bandwidth  $B1=\infty$  with the four synchronizers at the rate manual (b-m), at the rate automatic (b-a), at quarter rate manual (b-m/4) and at quarter rate automatic (b-a/4).

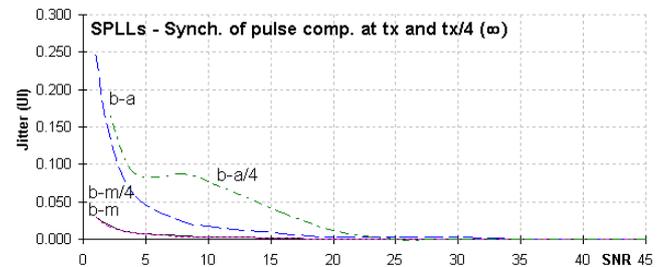


Fig. 9 Jitter-SNR curves of B1- 4 synchronizers(b-m,b-a,b-m/2,b-a/2)

We see that, in general, the output jitter  $UI_{RMS}$  decreases gradually with the input SNR increasing. However, the type at quarter rate (b-m/4) has some irregularities.

For prefilter  $B1=\infty$ , for high SNR, the four synchronizer jitter curves tend to be similar. However, for low SNR, the manual versions (b-m, b-m/4) are similar but slightly better than the automatic versions (b-a, b-a/4), the type at rate manual (b-m) is slightly the best. Also, for an intermediate SNR ( $SNR \approx 8$ ), the type at quarter rate automatic (b-a/4) has a very significant jitter perturbation, due to various losses of synchronism.

Fig. 10 shows the jitter-SNR curves of the prefilter bandwidth  $B2=2.tx$  with the four synchronizers at the rate manual (b-m), at the rate automatic (b-a), at quarter rate manual (b-m/4) and at quarter rate automatic (b-a/4).

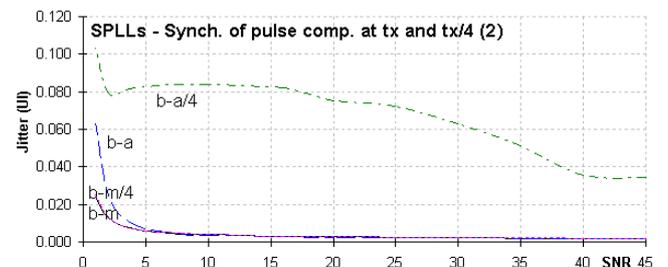


Fig. 10 Jitter-SNR curves of B2-4 synchronizers(b-m,b-a,b-m/2,b-a/2)

For prefilter B2=2.tx, we verify that, it becomes the jitter-SNR curves more similar between themselves. For high SNR, it harms the jitter-SNR curves. However, for low SNR, it benefits significantly the jitter-SNR curves. Also, for an intermediate SNR ( $SNR \approx 10$ ), the type both transitions at quarter rate automatic (b-a/4) has a very great jitter perturbation due to various losses of synchronism.

Fig. 11 shows the jitter-SNR curves of the prefilter bandwidth B3=1.tx with the four synchronizers at the rate manual (b-m), at the rate automatic (b-a), at quarter rate manual (b-m/4) and at quarter rate automatic (b-a/4).

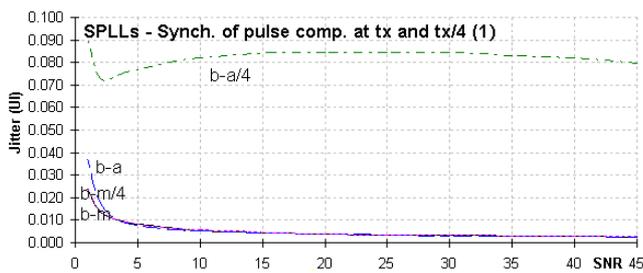


Fig. 11 Jitter-SNR curves of B3-4 synchronizers(b-m,b-a,b-m/2,b-a/2)

For prefilter B3=1.tx, we verify that, it becomes the jitter-SNR curves still more similar between themselves. For high SNR, it harms more the jitter-SNR curves. However, for low SNR, it benefits less the jitter-SNR curves. Also, for an intermediate SNR ( $SNR \approx 12$ ), the type both transitions at quarter rate automatic (b-a/4) has a very great jitter perturbation due to various losses of synchronism.

## VI. CONCLUSION AND FUTURE WORK

We studied the prefilter with three bandwidths (B1= $\infty$ , B2=2.tx, B3=1.tx) followed of four synchronizers, at rate manual (b-m), at rate automatic (b-a), at quarter rate manual (b-m/4) and quarter rate automatic (b-a/4). We saw that, in general, the output jitter curves decreases gradually with the input SNR increasing. However, the type at quarter rate automatic (b-a/4) has some undesired irregularities.

For prefilter B1= $\infty$ , we verified that, for high SNR, the four synchronizers jitter curves tend to be similar, this is comprehensible since all the synchronizers are digital and have similar noise margin. However, for low SNR, the manual versions (b-m, p-m/4) are significantly better than the automatic versions (b-a, b-a/4), this is comprehensible since the automatic versions have more digital states, what aggravates the error state propagation. The type at rate manual (b-m) is slightly the best, because has less digital states. Also, for an intermediate SNR ( $SNR \approx 8$ ) the quarter rate automatic (p-a/4) has a very significant jitter perturbation due to various losses of synchronism.

For prefilter B2=2.tx, we verify that, it becomes the jitter-SNR curves more similar between themselves. For high SNR, it harms slightly the jitter-SNR curves. However, for low

SNR, it benefits slightly the jitter-SNR curves. Also, for an intermediate SNR ( $SNR \approx 10$ ), the type (b-a/4) has a great jitter perturbation, due to various losses of synchronism.

For prefilter B3=1.tx, we verify that, it becomes the jitter-SNR curves still more similar between themselves. For high SNR, it harms more the jitter-SNR curves. However, for low SNR, it benefits less the jitter-SNR curves. Also, for an intermediate SNR ( $SNR \approx 12$ ), the type (b-a/4) has a great jitter perturbation, due to various losses of synchronism.

In the future, we are planning to extend the present study to other types of synchronizers.

## ACKNOWLEDGMENT

The authors are grateful to the program FCT (Foundation for sScience and Technology) / POCI2010.

## REFERENCES

- [1] Jean C. Imbeaux, "performance of the delay-line multiplier circuit for clock and carrier synchronization in Digital Satellite Communications". IEEE Journal on Selected Areas in Communications pp.82-95 Jan. 1983.
- [2] Werner Rosenkranz, "Phase Locked Loops with limiter phase detectors in the presence of noise", IEEE Transactions on Communications com-30 N°10 pp.2297-2304. Oct 1982.
- [3] Hans H. Witte, "A Simple Clock Extraction Circuit Using a Self Sustaining Monostable Multivibrator Output Signal", Electronics Letters, Vol.19, Is.21, pp.897-898, Oct 1983.
- [4] Charles R. Hogge, "A Self Correcting Clock Recovery Circuit", IEEE Transactions on Electron Devices pp.2704-2706 Dec 1985.
- [5] Antonio D. Reis, Jose F. Rocha, Atilio S. Gameiro and Jose P. Carvalho "A New Technique to Measure the Jitter", Proc. III Conf. on Telecommunications pp.64-67 Foz-PT 23-24 Apr 2001.
- [6] Marvin K. Simon and William C. Lindsey, "Tracking Performance of Symbol Synchronizers for Manchester Coded Data", IEEE Transactions on Communications Vol. com-25 N°4, pp.393-408, April 1977.
- [7] Jeffrey B. Carruthers, D. D. Falconer, H. M. Sandler and L. Strawczynski, "Bit Synchronization in the Presence of Co-Channel Interference", Proc. Conf. on Electrical and Computer Engineering pp.4.1.1-4.1.7, Ottawa-CA 3-6 Sep. 1990.
- [8] Johannes Huber and Weilin Liu "Data-Aided Synchronization of Coherent CPM-Receiver" IEEE Transactions on Communications Vol.40 N°1, pp.178-189, Jan. 1992.
- [9] Antonio A. D'Amico, Aldo N. D'Andrea and Ruggiero Reggianni, "Efficient Non-Data-Aided Carrier and Clock Recovery for Satellite DVB at Very Low SNR", IEEE Jou. on Satellite Areas in Comm. Vol.19 N°12 pp.2320-2330, Dec. 2001.
- [10] Rostislav Dobkin, Ran Ginosar and Christos P. Sotiriou "Data Synchronization Issues in GALS SoCs", Proc. 10th International Symposium on Asynchronous Circuits and Systems, pp.CD-Ed., Crete-Greece 19-23 Apr. 2004.
- [11] N. Noels, H. Steendam and M. Moeneclaey, "Effectiveness Study of Code-Aided and Non-Code-Aided ML-Based Feedback Phase Synchronizers", Proc. IEEE Intern. Conference on Communications (ICC'06) pp.2946-2951, Istanbul-TK, 11-15 Jun 2006.
- [12] Antonio D. Reis, Jose F. Rocha, Atilio S. Gameiro and Jose P. Carvalho "Effects of the Prefilter Type on Digital Symbol Synchronizers", Proc. VII Symposium on Enabling Optical Network and Sensors (SEONs 2009) pp.35-36, Lisbon-PT 26-26 June 2009.

# Particle Swarm Optimization with Time-Varying Acceleration Coefficients Based on Cellular Neural Network for Color Image Noise Cancellation

Te-Jen Su<sup>1</sup>

<sup>1</sup>College of Information Technology  
Kun Shan University  
Tainan 710, Taiwan, R.O.C.

tejensu@mail.ksu.edu.tw

Jui-Chuan Cheng<sup>2</sup>

<sup>2,3</sup>Department of Electronic Engineering  
National Kaohsiung University of Applied Sciences  
Kaohsiung 807, Taiwan, R.O.C.

eagle@cc.kuas.edu.tw

Yang-De Sun<sup>3</sup>

1097305128@cc.kuas.edu.tw

**Abstract**—This paper proposes a novel method for designing templates of Cellular Neural Network (CNN) for color image noise removal. The control of CNN systems is achieved via Particle Swarm Optimization (PSO) with Time-Varying Acceleration Coefficients (PSO-TVAC). Based on PSO-TVAC method, the proposed approach can automatically update the parameters of the templates of CNN to optimize them for diminishing noise interference in polluted image. The demonstrated examples are compared favorably with other available methods, which illustrate the better performance of the proposed PSO-TVAC-CNN methodology.

**Keywords**- Cellular Neural Network; Color Image Noise Removal; Particle Swarm Optimization with Time-Varying Acceleration Coefficients

## I. INTRODUCTION

Particle swarm optimization is a population based stochastic optimization technique developed by Dr. Eberhart and Dr. Kennedy in 1995 [1, 2, 9–11], inspired by social behavior of bird flocking or fish schooling. It is easily implemented in most programming languages and has proven both very effective and quick for a diverse set of optimization problems. However, local convergence problem and slow later convergence problem are the two critical shortcomings of PSO that limit its applications [3]. A Particle Swarm Optimization with Time-Varying Acceleration Coefficients (PSO-TVAC) is presented in this paper, which allows to effectively overcome the two mentioned problems [12].

A novel class of information processing system called cellular neural networks was proposed by L.O. Chua and L. Yang in 1988 [6, 7]. CNN is characterized by the parallel computing of simple processing cells locally interconnected. It has been widely used for image processing, pattern recognition, signal processing, etc. In recent years, the problems of CNN templates design for image processing have received considerable attention [4, 5].

A new method that combines the discrete-time cellular neural network template learning method with an adaptive particle swarm optimization, and applies to gray image noise cancellation was developed [8]. The approach is able to find the template values easily without complex mathematic

computing processes but also to obtain the balance of convergence speed and convergence accuracy. This work is extended from the previous study [8]; we attempt to apply the technique of gray image noise cancellation to color image noise cancellation by separating the color image into three Gray-Scale RGB elements.

The rest of this paper is organized as follows: in Section 2, the Particle Swarm Optimization techniques, while in Section 3, the Cellular Neural Network is discussed. In Section 4, the CNN based on PSO-TVAC template learning for images noise cancellation is presented. Examples are given in Section 5 to demonstrate the proposed methodology. Finally, conclusion is drawn in Section 6.

## II. PARTICLE SWARM OPTIMIZATION WITH TIME-VARYING ACCELERATION COEFFICIENTS (PSO-TVAC)

In PSO, suppose that the search space is D-dimensional, and then the  $i$ th particle is represented as  $X_i = (x_{i1}, x_{i2}, \dots, x_{iD})$ . The velocity (rate of the position change) of this particle is denoted as  $V_i = (v_{i1}, v_{i2}, \dots, v_{iD})$ . The best previous position of the  $i$ th particle is represented as  $P_i = (p_{i1}, p_{i2}, \dots, p_{iD})$ . In other words,  $P_i$  involves the best previous position, which  $X_i$  has visited (the best position called  $pbest$ ). The index of the best particle among all the particles in the swarm is defined as the symbol  $g$  (called  $gbest$ ). The particles are manipulated according to the equations 1 and 2. In its canonical form, Particle Swarm Optimization is modeled as follows [6-8]:

$$v_{id}(t+1) = wv_{id}(t) + c_1 rand()_1(p_{id} - x_{id}) + c_2 rand()_2(p_{gd} - x_{id}) \quad (1)$$

$$x_{id}(t+1) = x_{id}(t) + v_{id}(t+1) \quad (2)$$

where

$v_{id}(t+1)$	: velocity of particle $i$ at iteration $t+1$
$v_{id}(t)$	: velocity of particle $i$ at iteration $t$
$x_{id}(t+1)$	: position of particle $i$ at iteration $t+1$
$x_{id}(t)$	: position of particle $i$ at iteration $t$
$c_1$	: cognitive parameter
$c_2$	: social parameter
$rand()_1$	: random number uniform distribution $U(0,1)$
$rand()_2$	: random number uniform distribution $U(0,1)$
$p_{id}$	: $pbest$ position of particle $i$

$p_{gd}$  : *g*best position of swarm  
 $w$  : inertia weight

The objective of PSO-TVAC is to enhance the global search in the early part of the optimization and to encourage the particles to converge toward the global optimum at the end of the search. With a large cognitive parameter and small social parameter at the beginning, particles are allowed to move around the search space, instead of moving toward the population best. However, a small cognitive parameter and a large social parameter allow the particles to converge to the global optimum in the latter part of the optimization. Under this development, the cognitive parameter  $c_1$  starts with a high value  $c_{1max}$  and linearly decreases to  $c_{1min}$ . Whereas the social parameter  $c_2$  starts with a low value  $c_{2min}$  and linearly increases to  $c_{2max}$ . This modification can be mathematically represented as follows:

$$c_1(t) = (c_{1max} - c_{1min})\left(\frac{T_{max} - t}{T_{max}}\right) + c_{1min} \quad (3)$$

$$c_2(t) = (c_{2min} - c_{2max})\left(\frac{T_{max} - t}{T_{max}}\right) + c_{2max} \quad (4)$$

where  $T_{max}$  is the maximal number of iterations and  $t$  is the current number of iterations.

### III. CELLULAR NEURAL NETWORK

A two-dimensional CNN array is considered in which the cell dynamics are described by the following nonlinear ordinary differential equation with linear and nonlinear terms [13–16]:

$$\dot{x}_{ij}(t) = -x_{ij}(t) + \sum_{C(k,l) \in N_r(i,j)} A(i,j;k,l)y_{kl}(t) + \sum_{C(k,l) \in N_r(i,j)} B(i,j;k,l)u_{kl}(t) + \sum_{C(k,l) \in N_r(i,j)} D_{ij;kl}(\Delta v) + I_{ij} \quad (5)$$

$$y_{ij} = f(x_{ij}(t)) = \frac{1}{2}(|x_{ij}(t) + 1| - |x_{ij}(t) - 1|) \quad (6)$$

where

$$\Delta v = u, x, y_{kl}(t) - u, x, y_{ij}(t), |x_{ij}(t)| \leq 1, |u_{ij}(t)| \leq 1, 1 \leq i \leq M, 1 \leq j \leq N$$

$x_{ij}, u_{ij}, y_{ij}, I_{ij}$  are the state, input, output and threshold voltage of the specified CNN cell, respectively.  $A(i,j;k,l)$  is called the feedback cloning template,  $B(i,j;k,l)$  is called the feedforward or input control template,  $D_{ij;kl}$  are nonlinear terms applied for  $\Delta v$  ( $\Delta v$  is a generalized difference). The state and output vary in time, the input is static (time independent), and the CNN is single-layer with nearest neighbor linear.

In this paper,  $D_{ij;kl}$  is the generalized nonlinear term applied to  $\Delta v = u_{kl}(t) - x_{ij}(t)$ , the voltage difference of either the input or state values. The nonlinear template D is as follows:

$$D = \begin{bmatrix} d_2 & d_1 & d_2 \\ d_1 & d_0 & d_1 \\ d_2 & d_1 & d_2 \end{bmatrix} d(\Delta v) \quad (7)$$

$$d(\Delta v) = (\Delta v) e^{-\left(\frac{\Delta v}{K}\right)^2}, \Delta v = u_{kl}(t) - x_{ij}(t)$$

where  $d_0, d_1, d_2$  are the parameters in the nonlinear template D.  $K$  should be set very close to this value in an attempt to separate the noise effects and the image structure.

### IV. CNN BASED ON PSO-TVAC TEMPLATE LEARNING

A digital image is composed of pixels which can be thought of as small dots on the screen and may be represented as  $m$ -by- $n$  matrices. MATLAB is a matrix processing language for a wide range of applications. The color image in MATLAB is described as a two dimensional matrix in uint8 format. In order to apply the gray image noise cancellation from the precious study [8], we separate the color image into RGB elements. These three Gray-Scale images will be blend into a colorful image after finishing noise cancellation processes proposed. The process followed to perform noise cancellation is shown in Fig. 1

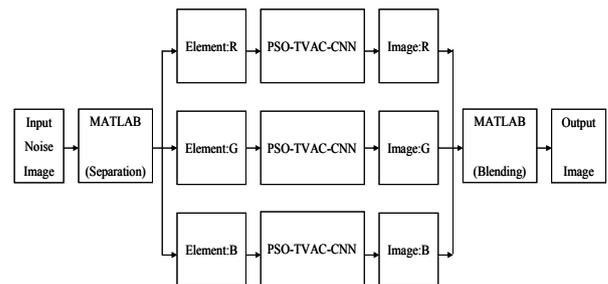


Figure 1. Block diagram of the image separation and blending

In this case, PSO-TVAC is employed to design templates of CNN for canceling the noise interference in Gray-Scale images. The templates are designed as following pattern structures, respectively:

$$A = \begin{bmatrix} a_2 & a_1 & a_2 \\ a_1 & a_0 & a_1 \\ a_2 & a_1 & a_2 \end{bmatrix}, B = \begin{bmatrix} b_2 & b_1 & b_2 \\ b_1 & b_0 & b_1 \\ b_2 & b_1 & b_2 \end{bmatrix}, D = \begin{bmatrix} d_2 & d_1 & d_2 \\ d_1 & d_0 & d_1 \\ d_2 & d_1 & d_2 \end{bmatrix}, I, K \quad (8)$$

where  $a_0, a_1, a_2, b_0, b_1, b_2, d_0, d_1, d_2, I, K$  are elements of the swarm, in order to satisfy output saturation effectively, we set  $a_2 = 0, b_0 = 0, b_1 = 0, b_2 = 0, I = 0, x_{ij}(0) = u_{ij}(t)$ . The training image is corrupted by the salt and pepper noise shown in Fig. 2.



Figure 2. The training images (a) Input image to the CNN (b) Desired output image of CNN.

The following equation is used as an objective function (error function); the block diagram is shown in Fig. 3.

$$Error = \sum_{i=1}^k (P_c(i) - P_d(i))^2 \quad (9)$$

where  $k$  denotes the total pixel of the picture,  $P_c(i)$  is the value of the  $i$ th pixel of the input image and  $P_d(i)$  stands for the pixel of the desired image. Each resulting image is compared with the desired image which should be obtained in the end. The comparison allows to compute the value of the error function, and consequently obtain the best template.

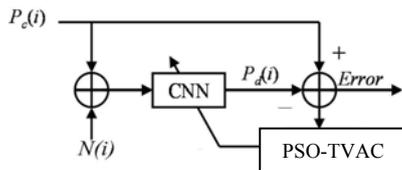


Figure 3. Block diagram of the objective function.

The process for implementing the PSO-TVAC based on CNN is shown as Fig. 4.

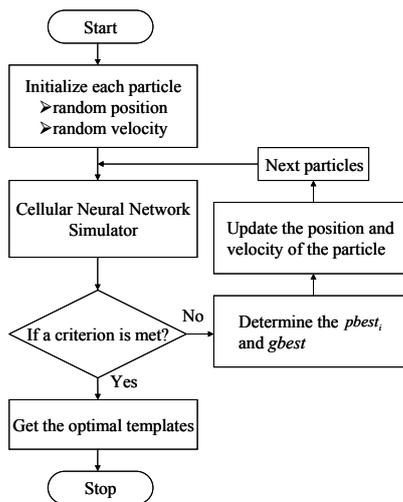


Figure 4. Flow chart of PSO-TVAC-CNN.

## V. EXAMPLES AND RESULTS

In this section, we present the examples polluted by different percentage of noise density interference and using our proposed method PSO-TVAC-CNN compares with PSO-CNN [5] for gray and color image noise cancellation respectively.

### A. Examples 1

Consider a  $256 \times 256$  LENA Gray-Scale image Fig. 5(a) which is polluted by the salt and pepper noise 10%, 20%, 30% in Fig. 5(b) - Fig. 5(d), respectively. The parameters of the proposed method PSO-TVAC-CNN and PSO-CNN are set as indicated in Table 1 and Table 2, respectively. The self-adapting inertia weight  $w$  is defined in [17].

TABLE I. PSO-TVAC PARAMETERS SETTING

The number of swarm size	12
The maximum position $X_{max}$	10
The maximum velocity $V_{max}$	1
Acceleration coefficient $C_{1max}$	2.5
Acceleration coefficient $C_{1min}$	0.5
Acceleration coefficient $C_{2max}$	2.5
Acceleration coefficient $C_{2min}$	0.5
Inertia weight $W$	0.8
Iterations	500

TABLE II. PSO PARAMETERS SETTING

The number of swarm size	12
The maximum position $X_{max}$	10
The maximum velocity $V_{max}$	1
Acceleration coefficient $c_1$	2.05
Acceleration coefficient $c_2$	2.05
Inertia weight $W$	0.8
Iterations	500

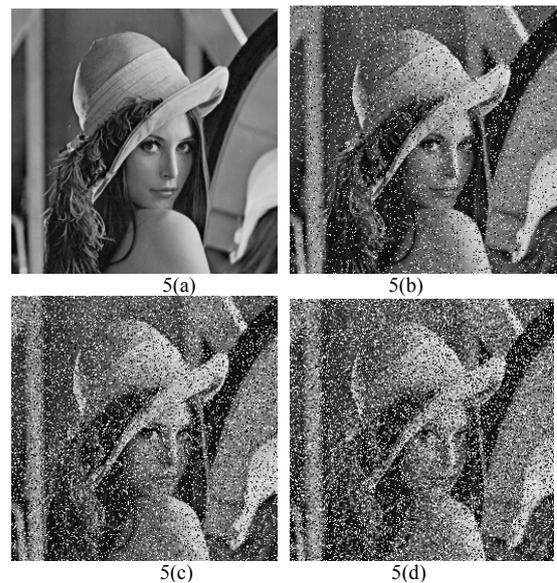


Figure 5. (a) Original LENA gray-Scale image (b) The contaminated image with 10% noise (c) The contaminated image with 20% noise (d) The contaminated image with 30% noise.

According to these parameters, the consequences of approximated optimal templates  $A$ ,  $D$ , and threshold  $K$  were found by the PSO-TVAC after a few iterations.

The contaminated image with 10% noise:

$$A = \begin{bmatrix} 0 & 0.3698 & 0 \\ 0.3698 & -0.6665 & 0.3698 \\ 0 & 0.3698 & 0 \end{bmatrix}, D = \begin{bmatrix} 0.2810 & 0.6154 & 0.2810 \\ 0.6154 & 5.5066 & 0.6154 \\ 0.2810 & 0.6154 & 0.2810 \end{bmatrix}, K = 0.2495$$

The contaminated image with 20% noise:

$$A = \begin{bmatrix} 0 & 0.4256 & 0 \\ 0.4256 & -0.8792 & 0.4256 \\ 0 & 0.4256 & 0 \end{bmatrix}, D = \begin{bmatrix} 0.2881 & 0.7562 & 0.2881 \\ 0.7562 & 4.9813 & 0.7562 \\ 0.2881 & 0.7562 & 0.2881 \end{bmatrix}, K = 0.2488$$

The contaminated image with 30% noise:

$$A = \begin{bmatrix} 0 & 0.4144 & 0 \\ 0.4144 & -0.8347 & 0.4144 \\ 0 & 0.4144 & 0 \end{bmatrix}, D = \begin{bmatrix} 0.2475 & 0.6196 & 0.2475 \\ 0.6196 & 3.6130 & 0.6196 \\ 0.2475 & 0.6196 & 0.2475 \end{bmatrix}, K = -0.2721$$

Similarly according to above parameters setting, the PSO found the consequences of approximated optimal templates  $A$ ,  $D$  and threshold  $K$  as following:

The contaminated image with 10% noise:

$$A = \begin{bmatrix} 0 & 0.2732 & 0 \\ 0.2732 & -0.1733 & 0.2732 \\ 0 & 0.2732 & 0 \end{bmatrix}, D = \begin{bmatrix} 0.2824 & -0.0459 & 0.2824 \\ -0.0459 & 3.4591 & -0.0459 \\ 0.2824 & -0.0459 & 0.2824 \end{bmatrix}, K = -0.2460$$

The contaminated image with 20% noise:

$$A = \begin{bmatrix} 0 & 0.6230 & 0 \\ 0.6230 & -1.5564 & 0.6230 \\ 0 & 0.6230 & 0 \end{bmatrix}, D = \begin{bmatrix} 0.0928 & 1.1943 & 0.0928 \\ 1.1943 & 4.4690 & 1.1943 \\ 0.0928 & 1.1943 & 0.0928 \end{bmatrix}, K = 0.1440$$

The contaminated image with 30% noise:

$$A = \begin{bmatrix} 0 & 0.3224 & 0 \\ 0.3224 & -1.0504 & 0.3224 \\ 0 & 0.3224 & 0 \end{bmatrix}, D = \begin{bmatrix} 0.1076 & 0.6847 & 0.1076 \\ 0.6847 & 3.9886 & 0.6847 \\ 0.1076 & 0.6847 & 0.1076 \end{bmatrix}, K = 0.4617$$

The error after a few iterations for PSO-TVAC-CNN and PSO-CNN are shown in Fig. 6-8. Table 3 shows the PSNR [18] of the image noise cancellation with both cases.

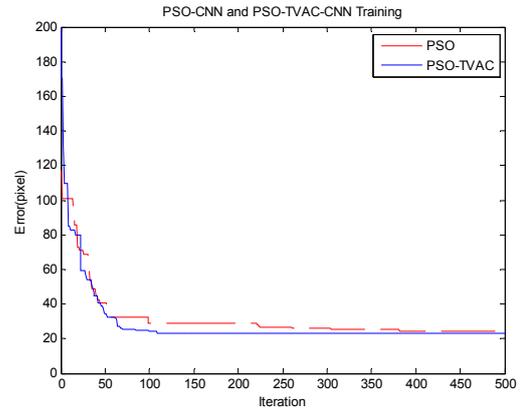


Figure 6. PSO-TVAC-CNN and PSO-CNN Training for Gray image with 10% noise

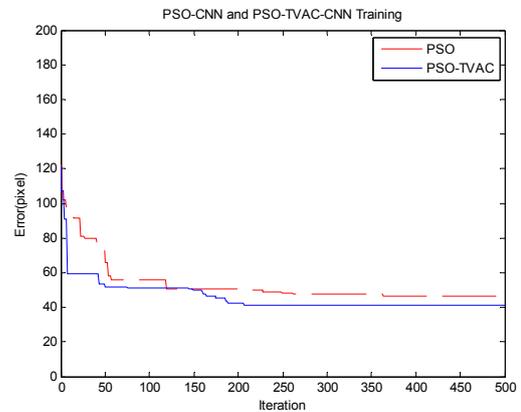


Figure 7. PSO-TVAC-CNN and PSO-CNN Training for Gray image with 20% noise

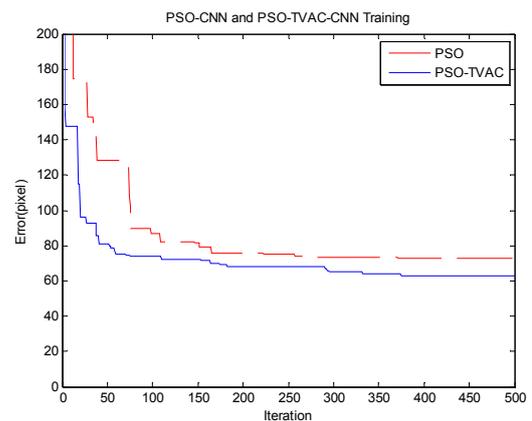


Figure 8. PSO-TVAC-CNN and PSO-CNN Training for Gray image with 30% noise

TABLE III. PSNR OF THE GRAY IMAGE (256×256 LENA) NOISE CANCELLATION

Salt and Pepper	Contaminated Image	PSO-TVAC-CNN	PSO-CNN
10% Noise	15.1659 dB	34.3663 dB	33.3802 dB
20% Noise	12.0989 dB	31.9834 dB	27.3987 dB
30% Noise	10.3124 dB	29.8721 dB	26.9026 dB



Figure 9. Results by using PSO-TVAC-CNN algorithm for the Gary-Scale image with noise of (a) 10%, (b) 20%, (c) 30%



Figure 10. Results by using PSO-CNN algorithm for the Gary-Scale image with noise of (a) 10%, (b) 20%, (c) 30%

Using the above templates, the output images processing by PSO-TVAC-CNN and PSO-CNN are shown in Fig. 9(a) - 9(c) and Fig. 10(a) - 10(c) respectively. By comparing Fig. 6 - 8, Table 3 and Fig. 9(a) - 9(c) with Fig. 10(a) - 10(c) respectively, our proposed method PSO-TVAC-CNN could restrain from noise of the polluted image more effectively than PSO-CNN.

Next we apply the same optimal templates found by PSO-TVAC-CNN to the Color images to prove the better performance.

B. Examples 2

In order to demonstrate that the optimal template has the same performance when processing color images, we have performed similar tests with the color version of the Lean image (Fig. 11(a)). The 256×256 LENA color image in Fig. 11(b) - 11(d) which is polluted by the salt and pepper noise 10%, 20%, 30%, respectively.

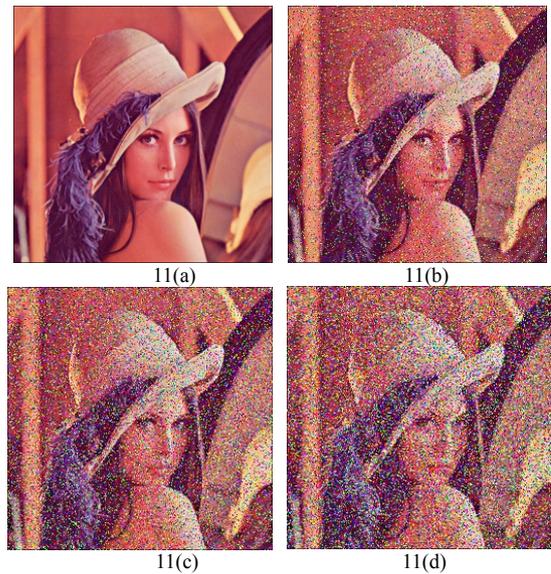


Figure 11. (a) Original LENA color image (b) The contaminated image with 10% noise (c) The contaminated image with 20% noise (d) The contaminated image with 30% noise.

By using the proposed method PSO-TVAC-CNN and median filter, the results for the output images obtained in Fig. 12(a) - 12(c) and 13(a) - 13(c). Table 4 show the PSNR of the color image noise cancellation with both cases.

TABLE IV. PSNR OF THE COLOR IMAGE (256×256 LENA) NOISE CANCELLATION

Salt and Pepper	Contaminated Image	PSO-TVAC-CNN	Median filter
10% Noise	15.1767 dB	34.0704 dB	30.6477 dB
20% Noise	12.1143 dB	30.9633 dB	27.4666 dB

30% Noise	10.3164 dB	28.3820 dB	22.9037 dB
-----------	------------	------------	------------



Figure 12. Results by using PSO-TVAC-CNN algorithm for the color image with noise of (a) 10%, (b)20%, (c)30%.

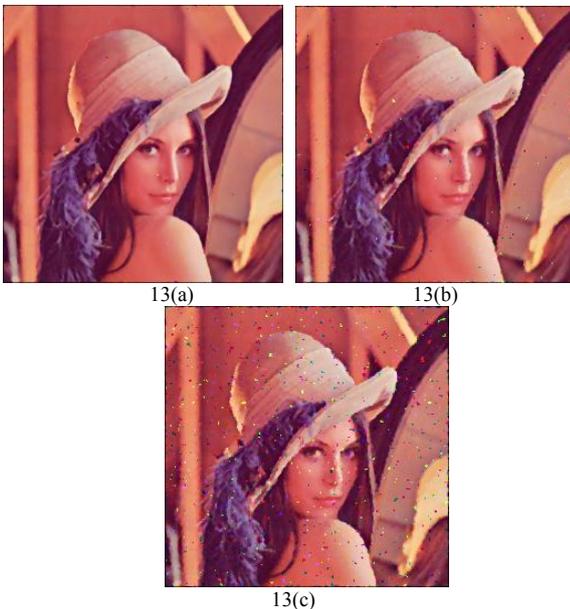


Figure 13. Results by using median filter algorithm for the color image with noise of (a) 10%, (b)20%, (c)30%.

By comparing Fig. 12(a) - 12(c) with Fig. 13(a) - 13(c), and Table 4, our proposed method (PSO-TVAC-CNN)

could restrain from noise of the polluted image more effectively than PSO-CNN.

## VI. CONCLUSION

In this paper, we have presented a Cellular Neural Network templates learning method that combined Particle Swarm Optimization with Time-Varying Acceleration Coefficients, applied to color image noise cancellation. Template learning is a crucial step in cellular neural network technology. The implementation of PSO-TVAC-CNN is a contribution to the modern heuristics research in the image processing area. From the demonstrated examples, the proposed algorithm shows the better performance of the noise cancellation for color image than PSO-CNN.

In the future research, we hope to improve the adaptive templates training to repair the color images that are polluted by the higher density of miscellaneous noises. For real applications, the proposed method may be implemented and fabricated on FPGA or VLSI technology.

## REFERENCES

- [1] J. Kennedy and R. C. Eberhart, "Particle Swarm Optimization", *Proceedings of IEEE International Conference on Neural Networks*, Perth, Australia, pp. 1942–1948, 1995.
- [2] J. Kennedy and R. C. Eberhart, "A New Optimizer Using Particle Swarm Theory", *Proceedings of the Sixth International Symposium on Micromachine and Human Science*, Nagoya, Japan, pp. 39–43, 1995.
- [3] D.X. Zhang, Z. Hong Guan, X.Z. Liu, "An adaptive particle swarm optimization algorithm and simulation", *Proceedings of the IEEE International Conference on Automation and Logistics*, pp. 2399–2402, 2007.
- [4] M. Nakagawa, T. Inoue and Y. Nishio, "CNN Template Learning To Obtain Desired Images Using Back Propagation Algorithm", *IEEE Workshop on Nonlinear Circuit Networks*, pp. 93-95, 2009.
- [5] T. J. Su, T. H. Lin, J. and, W. Liu "Particle Swarm Optimization for Gray-Scale Image Noise Cancellation", *International Conference on Intelligent Information Hiding and Multimedia Signal Processing*, pp.1459-1462, 2008.
- [6] L. O. Chua and L. Yang, "Cellular Neural Networks: Theory", *IEEE Trans. Circuits Syst.*, vol. 35, pp.1257–1272, Oct. 1988.
- [7] L. O. Chua and L. Yang, "Cellular Neural Networks: Applications", *IEEE Trans. Circuits Syst.*, vol. 35, pp.1273–1290, Oct. 1988.
- [8] T. J. Su, J. C. Cheng, M. Y. Huang, T. H. Lin, and, C. W. Chen "Applications of Cellular Neural Networks to Noise Cancellation in Gray Images Based on Adaptive Particle-swarm Optimization", *Circ. Syst. Signal*, Jan. 2011, doi: 10.1007/s00034-011-9269-x.
- [9] F. Heppner and U. Grenander: "A Stochastic Nonlinear Model for Coordinated Bird Flocks", *The Ubiquity of Chaos, AAAS Publications*, Washington DC, 1990.
- [10] Y. Shi and R. Eberhart, "A Modified Particle Optimizer", *Proceedings of the 1998 IEEE World Congress on Computational Intelligence*, pp. 69–73, May 1998.
- [11] Y. Shi and R. Eberhart, "Parameter Selection in Particle Swarm Optimization", *Proceedings of the 7th International Conference on Evolutionary Programming VII*, Lecture Notes In Computer Science, pp. 591–600, 1998.
- [12] A. Ratnaweera, Saman K. Halgamuge and Harry C. Watson, "Self-Organizing Hierarchical Particle Swarm Optimizer with Time-

- Varying Acceleration Coefficients”, *IEEE Transactions on Evolutionary Computation*, Vol. 8, No. 3, pp. 240–255, 2004.
- [13] P. L. Venetianer, T. Roska and L. O. Chua, “Analogic CNN Algorithms for Some Image Compression, Decompression and Restoration Tasks”, *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, Vol. 42, No. 5, pp. 278–284, 1995.
- [14] A. Zarandy, F. Werblin, T. Roska and L. O. Chua, “Novel Types of Analogic CNN Algorithms for Recognizing Bank-Notes”, *CNNA-94 Third IEEE International Workshop on Cellular Neural Networks and their Applications*, pp. 273–278, Dec. 8–21, 1994.
- [15] A. Zarandy, A. Stoffels, T. Roska and L. O. Chua, “Implementation of Binary and Gray-Scale Mathematical on the CNN Universal Machine”, *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, Vol. 45, No. 2, pp. 163–168, 1998.
- [16] C. Rekeczky, T. Roska and A. Ushida, “CNN-Based Difference-Controlled Adaptive Nonlinear Image Filters”, *International Journal of Circuit Theory and Applications*, Vol. 26, pp. 375–423, 1998.
- [17] M. Zhang, C. J Li, X.H. Yuan, Y.C. Zhang, “An improved PSO and its application in research on reservoir operation function of long-term”, *Third International Conference on Natural Computation*, vol. 4, pp. 118–122, 2007.
- [18] R. Lukac and K. N. Plataniotis, “A Taxonomy of Color Image Filtering and Enhancement Solutions”, *Advances in Imaging and Electron Physics*, Vol. 140, pp. 187–264, 2006.

## Peer Clustering System for Different Start Video Broadcasting

Hathairat Ketmaneechairat  
King Mongkut's University of  
Technology North Bangkok  
Bangkok, Thailand  
e-mail: hathairatk@kmutnb.ac.th

Phoemphun Oothongsap  
King Mongkut's University of  
Technology North Bangkok  
Bangkok, Thailand  
e-mail: phoempn@kmutnb.ac.th

Anirach Mingkhwan  
King Mongkut's University of  
Technology North Bangkok  
Bangkok, Thailand  
e-mail: anirach@ieee.org

**Abstract**—The different start video broadcasting is a new approach to manage video streaming applications. This approach allows unpunctual users to view broadcast programs from the beginning. This paper proposes a cluster model for different start video broadcasting. The model is composed of five processes: peer joining, super node selection, backup-node selection, download paths and leaving node process. The model is simulated and verified by NS-2. The results show that (i) the video server load is diminished (ii) the tracker load is reduced due to the existing of super nodes or clusters, (iii) the play out delay is only to 10 seconds and (iv) the bandwidth utilization is improved as the consequence of the reduced number of control messages. Moreover, the authors have compared the performance of the cluster model with the one of the non-cluster model.

**Keywords**—Peer-to-Peer (P2P); live video streaming; video on demand (VoD); different start video broadcasting; cluster;

### I. INTRODUCTION

Peer-to-Peer applications (P2P) have attracted a large number of Internet users. P2P technologies offer obvious advantages over content delivery network or content distribution network (CDN). P2P technologies improve system scalability with low implementation cost. P2P content delivery is an important technique for commercial systems such as IPTV. There are a lot of popular P2P file-sharing systems that support downloading such as Gnutella [1], Napster [2], Kazaa [3], BitTorrent [4], and eDonkey [5]. The main area of usage is P2P-based file sharing systems, like BitTorrent. Unlike traditional client-server architectures, peers in networks act as clients (leeches) and servers (seeds). A peer is not only downloads file from the network, but also uploads the downloaded file to other users in the network. Parts of the files are exchanged over direct connections between the peers. To enhance the system scalability and reduce the cost, several P2P video streaming applications have been published to use P2P technologies for the streaming of video and audio content. P2P technologies are provided content distribution service for both live video streaming and video-on-demand (VoD). PPLive [6], CoolStreaming [7], UUSee [8], Sopcast [9] and PPStream [10] are demonstrated by the huge popularity of P2P video streaming applications.

These works [6][7][8][9][10] cause unpleasant problems. The first problem is that a far away connection increases network traffic and thus decreases network resource utilization. The second problem is a heavy tracker load. These problems can be eliminated by using a hierarchical architec-

ture as explained in [17]. In [15][16][17][18], the cluster concepts for P2P systems are introduced.

For the live video streaming, a live video content is disseminated to all users in real-time. Hence, all users in the system can watch the same part of the video stream at the same time. If users join the program later on, they will miss the beginning of the stream. For the video-on-demand, the users can watch the video stream anywhere at any time. Multiple users may watch the same movie at the different playback times. Besides these two separated categories, there is a different start video broadcasting [11][19][20] that allows unpunctual viewers to view the stream from the beginning during server broadcast time. By mixing a Peer-to-Peer download concept with a live broadcasting one, a new node can find users who have the needed parts of the stream, and use them as sources for download.

For example, suppose that there is a game of FIFA World Cup which starts at 3 PM and the game will end at 5 PM. A big amount of viewers will connect to the network and select the channel of the game. When the game starts at 3 PM, the viewers can load and view the game in real-time. After the game has started for 15 minutes, a new viewer decides to join the stream. The new viewer will have two choices: (i) view the game as the server broadcasts or (ii) view the game from the beginning. With the first choice, this broadcast will feature live video streaming while the second choice will employ both live streaming and video-on-demand features.

In this paper, a peer clustering system for different start video broadcasting is proposed. This model is created by five steps: peer joining, super node selection, backup-node selection, download paths and leaving node. There are three peer types: super node (SN), backup-node (BN) and normal node (NN). The buffer is used to store existing chunks of each peer in order to offer them to others. This is shown in section III.

The cluster model will be compared with non-cluster model for different start video broadcasting [11][19][20] to check the improved performance of the overall system. The cluster model for different start video broadcasting can be applied for live video streaming and video on demand. In contrast, the cluster model for live video streaming and video on demand will not comply with this approach because the unpunctual viewer has no chance to view the first chunk.

The remainder of this paper is organized as follows: Section II describes related work, Section III explains the system architecture, cluster-based system design and pseudo-code of cluster creation. A comparison of cluster and non-cluster

model system is proposed in Section IV. Section V indicates experimental results. Section VI presents the conclusion and future work.

## II. RELATED WORK

This section will briefly explain P2P streaming applications development [6][7][8][12][13][14] and P2P content delivery network (CDN) [15][16][17][18].

Coolstreaming or DOnet [7] is a P2P live video streaming application for only one channel. Every node in the network can be a video-source which produces the content for neighbor nodes. Coolstreaming does not use any tree, mesh or any structure. Additional infrastructure and implementations are necessary to provide more channels via Coolstreaming. Furthermore, there is only one node act as the origin node that stores all of video segments. The departure of an origin node can cause a single point of failure.

Sopcast, UUSEE and PPLive are channel-based systems, which provide a lot of different video streams on different channels. So each of the application networks need at least one media encoding server, where the video streams are created and stored, and a well known channel server where the clients can get information about available programs [6][8][12][13][14].

Sopcast [12] has a set of root servers, which maintains the information what peer is available for what channel. Sometimes also peer lists are exchanged between the peers. The most important difference of Sopcast is the usage of UDP as transport protocol [13]. This leads to fast packet transmission but also causes a lot of overhead for control. The usage of an external media player and a second buffer are very inefficient and lead to a huge start-up delay.

UUSEE provides the videos by several dedicated streaming servers, so that there is no single point of failure and the video streaming quality especially the playback continuity is improved. The TCP protocol is used to communicate with all peers, exchange the buffer map, measure the round trip time (RTT) and estimate the throughput [8]. If a huge number of peers try to join the same channel in the network at short time duration, a noticeable influence on the network performance has been recognized.

PPLive [14] uses different methods to exchange information about the availability of channels or movies, chunks and pieces. A distributed hash table (DHT) is used to assign dedicated movies to dedicated trackers and to achieve a load balancing [6]. On the other side, PPLive tries to improve its playback quality at the expensive of the network architecture. A locality mechanism, which prefers physically near peers (e.g., of the same ISP) is implemented, but peers with high bandwidth are preferred. This may lead to a bad network performance also for other participants.

Most of these works [6][7][8][12][13][14] have drawbacks related with low bandwidth utilization, high delay and a single point of failure. Thus to improve the performance of content distribution, the peers can be grouped in clusters. Many peers clustering approaches are proposed as the following:

The hierarchical architecture to group peers into clusters called CBT is proposed in [15]. The CBT has two novel al-

gorithms: a peer joining algorithm and a super-peer selection algorithm. The proximity measurements of the RTT value and the TTL value between a pair of peer and super-peer are used. The CBT system improves the performance and scalability, and can be used to build a large-scale BitTorrent-like P2P overlay network.

A novel super node overlay based on information exchange called SOBIE is proposed in [16]. The main contributions are to select the super nodes by considering the aggregation of not only the delay and distance, but also the information exchange frequency, exchange time and query similarity. The SOBIE is guarantees the matching between the physical network and logical network. Moreover, the SOBIE has small-world characteristic to improve the efficiency and robustness.

The super node selection problem for Peer-to-Peer applications is presented in [17]. Three super nodes selection protocols for overlay P2P networks are proposed: SOLE, PoP-Corn and H2O. An integrated approach to the super node selection problem built on strong graph theoretic foundations and guided by realistic applications, can benefit the Peer-to-Peer community through cross-fertilization of ideas and sharing of protocols.

An effective real-time Peer-to-Peer streaming system for the mobile environment is proposed in [18]. The peers are grouped into clusters according to their proximity using RTT values between peers as criteria for the cluster selection. The cluster leaders are using to help a service discovery server. The partial streams help to utilizing the upload capacity with finer granularity than just per one original stream. This is beneficial in mobile environments where bandwidth is scarce.

The non-cluster model for different start video broadcasting is proposed in [11][19][20]. The users can join the real time video stream at different times and view from the beginning or the same part as the server. The proposed model is based on an application layer MESH network. The advantages of the proposed model are a proper buffer size will smooth video quality, the starting delay can be bounded, the video server load is reduced drastically and users with different joining time can view the first frame or beginning of the video stream. The details of this system model are shown in Figure 1.

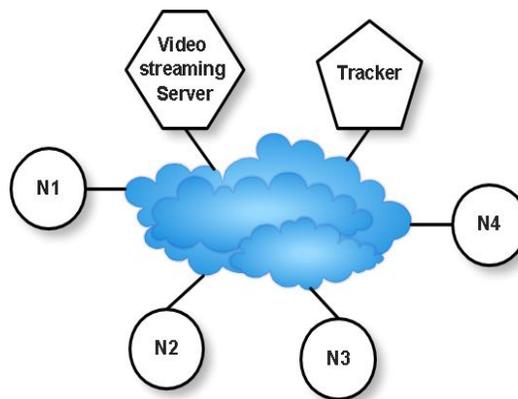


Figure 1. The non-cluster model for different start video broadcasting.

This paper presents a peer clustering system for different start video broadcasting to improve the content distribution for all peers in the network. The peers are grouped into clusters according to the availability of chunks and the join time of each node. A peer can download and upload video content with all peers in the same cluster. The global tracker and local tracker are created. The super node acts as a local tracker. This model can reduce server and tracker load. The control packet will be decreased.

### III. CLUSTERING SYSTEM DESIGN

This section introduces the concept of system architecture, cluster-based system design and pseudo-code of cluster creation. When a new node joins and wants to download chunks from the peers in the cluster model for different start video broadcasting, the global tracker has to decide which cluster and super node will be joined.

#### A. System Architecture

The peer clustering system architecture groups nodes into clusters according to the join time of each node or the available chunks. This peer clustering system is composed of a server, a global tracker (GT), super node (SN) or local tracker (LT), backup-node (BN) and normal nodes (NN, seed and leech). The server is a node that shares all chunks of a live video stream. The global tracker is known by all nodes and maintains the list of all super nodes. The super node acts as a local tracker keeping the list of all nodes in the cluster. All super nodes are connected with global tracker to synchronize the lists of all nodes in the cluster. The super node, normal node and backup-node are all downloader (leech) and uploaders (seed) of chunks.

#### B. Cluster-Based System Design

For the peer clustering for the different start video broadcasting model, the clustering means the grouping of node partnerships according to their network proximity. The proximity is measured by the joining time of each node and the availability of chunks. In this work, we assume that each peer has enough bandwidth in order to upload and download chunks. Several users join in a short period of time. For the initial work, the number of users is defined as a constant. In the future work, the optimal number of users in the cluster will be calculated as a function of user bandwidth and round trip delay. The clustering is used to control the traffic streams within a P2P system and additionally helps to decrease the load of the server and global tracker. Based on the non-cluster model for the different start video broadcasting presented in [11][19][20], the algorithms of peer exchange information, peer selection, buffer organization, segment scheduling are not changed but extended by a logical clustering mechanism. The clustering is realized by the separation of nodes, super nodes, backup-nodes and the introduction of local trackers.

The cluster based system model consists of a global part and clusters. The global part composes of the server that provides the video stream and the global tracker that is known by all nodes in the network. The clusters consist of a super node (SN), a local tracker (LT), a backup-node (BN)

and normal nodes (NN). The local tracker functionality is hosted by the super node itself. Both together are the coordinators for all nodes in their cluster. Whereas the local tracker acts as the source of neighbor lists, the super node is the only node in the cluster that can download from the server and other super nodes of the previous clusters. The normal nodes in a cluster know only other normal nodes within their own cluster and hence can only exchange chunks within the cluster. When the super node leaves from cluster, the backup-node will act as a new super node. The BN maintains a list of all nodes in the cluster. All normal nodes are candidates for the SN and BN. SN and BN selection process is explained in the next section. There are two major advantages of the cluster based system model. First, the clusters are only logical entities that are controlled by different known neighbor lists and the resulting chunk exchange paths. Second, the introduction of 3 levels of chunk sources (server, super nodes, nodes) leads to a controlled distribution of traffic and hence to a decreasing load for dedicated components mainly the server and tracker. An overview of the cluster-based system model is shown in Figure 2. The different components of the cluster based system can be divided into five processes as follows: (i) peer joining process, (ii) super node selection process, (iii) backup-node selection process, (iv) download process and (v) leaving node process.

#### 1) Peer Joining Process

When a new node joins to use a video streaming and wants to participate with neighbor peers in the cluster, it will contact with the global tracker to ask for the cluster and super node. Therefore, the peer joining algorithm has 2 important phases: connect to global tracker and connect to local tracker. For the first phase, all nodes know the address of the global tracker. When a new node contacts with the global tracker and asks for the first chunk. The global tracker will contact SN of each cluster to search for nodes having the first available chunk. The global tracker then selects the cluster which has the maximum number of nodes containing the first chunk. The new node gets the address of the local tracker (SN) and registers there. For the second phase, the new node contacts with the local tracker. The local tracker returns a random list of neighbor peers in the same cluster to the new node. The new node receives a random list of neighbor peers and sends the message to exchange buffer maps with neigh-

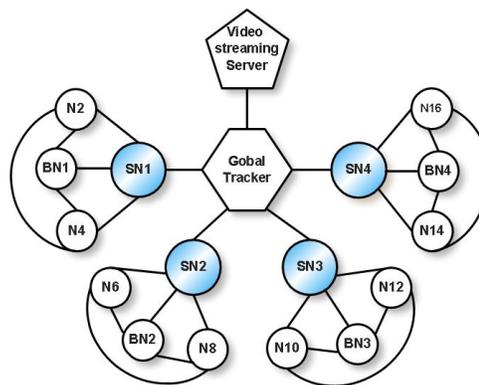


Figure 2. The cluster model for different start video broadcasting.

bor peers. The new node selects neighbor peers to download chunks.

2) Super Node Selection Process

When the first node joins, the global tracker will set the first node to be a super node and local tracker of the first cluster. The cluster size is limit to  $m$  nodes. After the global tracker received joining message from a new node, it will check the cluster size to select appropriate cluster. The global tracker will verify the member size of the selected cluster. If the size of the selected cluster is less than  $m$ , the address of SN in that selected cluster will be sent to the joining node. If the size of the selected cluster is full (equals  $m$ ), a new cluster is created. The global tracker will split a node that has first chunk available in the old cluster to be a SN for a new cluster. If there is no cluster in the system, the first cluster is created and the first joining node will be a SN of the first cluster.

3) Backup-Node Selection Process

If the size of cluster is full (equals  $m$ ), the backup-node will be selected. All normal nodes can be selected as backup-node. The backup-node selection is used to keep a list of all peers in the cluster by contact with SN. When the super node leaves from the cluster, a backup-node will be a new SN. The BN will receive the list of all peers in the cluster from SN. The backup node can be selected by three different methods as follow.

- Select the node joining the cluster after the first node. ( the second joining node, 2<sup>nd</sup>)
- Select the node joining the middle of the group. (the  $\frac{m}{2}$  node)
- Select the lasted node that joining the group. (the  $n^{\text{th}}$  node)

For the first method, all nodes in the cluster will have an equal chance to be a SN and BN. The drawback of this approach is a frequent SN and BN selection. The second method selects a new SN and BN not often and works well. The third method selects a new super node not often but may cause packet losses. In this paper, the second method is implemented in the simulation.

4) Download Process

When the new node joins in the cluster and contact with the SN, the local tracker (SN) prepares a random list of peers for a new node. The node knows only peers from the same cluster and can only download from neighbor peers. The buffer of each node is organized into three parts: data buffer, buffer map and sliding window. The data buffer is used to store video frames. The buffer map is a bit vector representing the information of available segments on a node. Each node exchanges its buffer map with its partners periodically. From buffer map information, the peer will decide which partner nodes are used to fetch required segments. If there is more than one partner having the same segments, the peer node will randomly select the partners or select the partners with minimum delay or maximum bandwidth. Besides buffer map, each node needs to have a sliding window which is used to store a number of displaying segments. From this buffer organization, the video segments

will be displayed continuously, and the starting delay of each node will be bounded. In this work, the circular buffer is used as buffer management as shown in Figure 3. The buffer data is divided into three parts: playback buffer, displaying buffer, and next available buffer as follows [11][19][20].

- The playback buffer is used to buffer data stream for a certain period of time before playing the stream. The number of frames in playback buffer is calculated from the delay called playback delay between the sending and receiving peer. The playback delay between each peer is random since the mesh-based architecture is used. For simplicity, the playback delay is defined as a maximum delay bound in this group of users in this particular network. Thus, every peer will have the same playback delay.

- The displaying buffer is used to store data that will be viewed by users. This buffer is designed by using a sliding window. The frame in the beginning for buffer is the displaying frame and the next frame is the next frame in the window will be viewed in the next minutes.

- The next available buffer is used to receive new frames. The new frames are received from other peers or partners by using sequential or rarest-first scheduling.

The node will exchange buffer map with the neighbor peers and will select peers to download chunks. The most important concept to download chunks has two ways: sequential download and rarest-first download. The sequential download is used when more chunks are available on the maximum number of neighbor peers. The rarest-first download is used when less chunks is available on the minimum number of neighbor peers. The SNs can download chunks from the server and all other SNs.

5) Leaving-Node Process

If a node leaves from cluster, the local tracker will delete it from list of peers. If the leaving node is a local tracker (SN), the backup-node will be a new local tracker (SN). If the last node leaves the cluster, the cluster is deleted. The local tracker always tells the global tracker about leaving nodes to synchronize the list of SN in the global tracker. The leaving-node process can be divided into three cases: the leaving of super node, back-up node and normal nodes. For the first case, when the super node is leaving from the cluster, it sends flooding message to all nodes. The all nodes

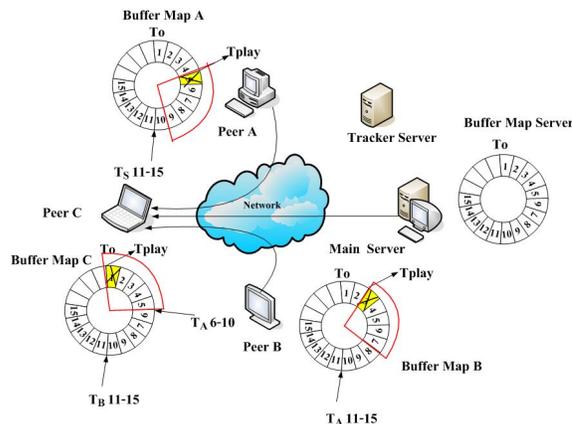


Figure 3. Filling buffer of peer C.

in the cluster will send keep-alive message to their super node. The super node sends keep-alive message to the global tracker. For the second case, the back-up node exchanges information periodically with the super node. If the back-up node leaves, it sends the message to inform the super node. For the last case, the normal node can leave the network at anytime.

C. Pseudo Code of Cluster Creation

The following pseudo code program demonstrates the cluster selection of global tracker, the cluster create and super node selection as shown in Table 1.

TABLE I. PSEUDO CODE OF CLUSTER CREATION

```

Function GetCluster(Node newNode, int neededChunk)
Begin
  Read N_P and Max_Node_In_Cluster
  For each cluster N determine occurrence of neededChunk
  Find cluster X with max occurrence
  If X is not found
    Generate new cluster C
    Set newNode as super node and local tracker of C
  Else
    If number of nodes in X < Max_Node_in_Cluster
      Add newNode to X
      Return X
    Else
      Find best node N for new cluster
      Cut N from cluster X
      Create new cluster C
      Set N to superNode and local tracker of C
      Add newNode to cluster X
      Return X
    End if
  End if
End
    
```

IV. A COMPARISON OF CLUSTER AND NON-CLUSTER MODEL SYSTEM

In this section, the cluster model will be compared with the non-cluster model for different start video broadcasting. Both system behaviors can be explain as shown in Table 2.

TABLE II. COMPARE STEPS OF NON-CLUSTER AND CLUSTER MODEL FOR DIFFERENT START VIDEO BROADCASTING

Non-cluster Model	Cluster model
<ul style="list-style-type: none"> <li>The requesting node searches for Tracker (T).</li> <li>The requesting node connects with T.</li> <li>T search for lists of all nodes in its table.</li> <li>T replies random lists of peers to the requesting node.</li> <li>The requesting node exchanges buffer map with neighbor peer lists.</li> <li>The requesting node selects peers to download.</li> </ul>	<ul style="list-style-type: none"> <li>The requesting node searches for Global Tracker (GT)</li> <li>The requesting node connects with GT.</li> <li>GT search for lists of all super nodes in its table. (Local Tracker: LT).</li> <li>GT replies selected super node and cluster to the requesting node.</li> <li>The requesting node connects with own super node (LT) and cluster.</li> <li>LT replied random lists of peers to the requesting node.</li> <li>The requesting node exchanges buffer map with neighbor peer lists.</li> <li>The requesting node selects peers to download.</li> </ul>

The peer clustering system for different start video broadcasting design can be reduced searching time and server load as the following:

A. Searching Time Reduction

For the non-cluster model, the tracker uses the sequential search to find the list of peers in the peer list table. With the non-cluster model, the searching time of tracker is  $O(n)$  where  $n$  is the total number of nodes.

For cluster model, the global tracker employs the binary search to seek the proper super node in the peer list table. The local tracker employs the sequential search to seek the list of peers in the peer list table. The global tracker keeps a list of all peers and arrival time of each node. The cluster model reduces the searching time in the global tracker and local tracker. It groups peers into cluster according to joining time (arrival time) of each node. Let server starts broadcast at the time,  $T = 0$  and ends at the time,  $T = t$ . The arrival time of each node will be referenced with the server broadcast time and sorted from minimum to maximum value. Each node will be grouped to each cluster according to its arrival time, as shown in Figure 4. Then, the number of clusters is in order of  $2^x$ . The tracker will check the arrival time of each node and then assign the proper cluster to that particular node. Thus, the number of nodes in each cluster is a random

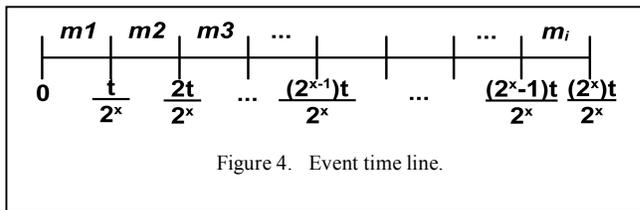


Figure 4. Event time line.

number. Let  $m_i$  is the number of nodes in each cluster,  $n$  is the total number of nodes in the system and  $2^x$  is the number of clusters. With this structure, the binary search is used to find the proper super node in GT. The searching time of GT is  $O(\log 2^x)$ . The sequential search is used to find the list of peers in LT. The searching time of LT is  $O(m)$ . Thus, the total searching time of cluster model is equal to is  $O(\log 2^x + m)$ .

$$\text{Non-cluster model} = O(n) \tag{1}$$

$$\text{Cluster Model} = O(\log 2^x + m) \tag{2}$$

B. Server Load Reduction

The peer clustering system for the different start video broadcasting is proposed to reduce the performance of server load. In the non-cluster model for the different start video broadcasting [11][19][20], the server supports all peers. For the cluster model the server supports only a super node in each cluster. Then, the server load can be calculated in Eq. (3), Eq. (4) and Eq. (5).

- Non-cluster model:

$$SL_{NC} = N_P \times N_C \tag{3}$$

- Cluster model with more than one nodes contacting server (Worst Case):

$$SL_C = N_{SP} \times N_C \quad (4)$$

- Cluster model with more only one node contacting server (Best Case):

$$SL_C = 1 \times N_C \quad (5)$$

Note:  $SL_{NC}$  denote the server load of non-cluster model.  
 $SL_C$  denote the server load of cluster model.  
 $N_P$  is the number of peers.  
 $N_C$  is the number of chunks.  
 $N_{SP}$  is the number of super nodes.

From Equations (3), (4) and (5),  $SL_C$  is less than or equal to  $SL_{NC}$ . The ratio of server load of non-cluster model and cluster model can be calculated as in Eq. (6) and Eq. (7).

$$\frac{SL_C}{SL_{NC}} = \frac{N_{SP} \times N_C}{N_P \times N_C} \quad (6)$$

$$SL_C = \frac{N_{SP}}{N_P} SL_{NC} \quad (7)$$

Since,  $N_{SP} \leq N_P$ , then  $\frac{N_{SP}}{N_P} \leq 1$ . Thus,  $SL_C \leq SL_{NC}$

### V. EXPERIMENTAL RESULTS

This section deals with the experimental performance evaluation of the peer clustering system for different start video broadcasting. The simulation experiments are conducted by the discrete event simulator NS-2 [21].

#### A. Simulation Setup

The experimental setup will create one video media server. The number of clusters is varied as 1, 2, 4, 6, 8 and 10, respectively. The non-cluster model is denoted by 1. The total number of nodes equals to 300 (limitation of simulation). For simplicity, the number of nodes in each cluster is equal. The bandwidth of all links is set to 2 Mbps, the delay is 10 ms and the joining time varies from 1 sec to 400 sec. The video stream bit rate is 512 Kbps. The video stream length is 32 MB, the size of each chunk is 64 Kb and the number of chunk is 512 chunks. The video play rate is 1 chunk/1 sec. The playback buffer and release buffer size are set to 10 sec and 15 sec, respectively. The buffer size of each node equals 54 sec.

#### B. Simulation Results

The simulation results are illustrated in Figures 5, 6, and 7. Figure 5 shows the relationship of server load and peer load of non-cluster and cluster model. Figure 5 plots the number of chunks download from server and neighbor peers in the network. The x-axis represents the number of clusters and the y-axis represents the number of downloaded chunks. The number of node is 300 nodes, each node downloads

only 512 chunks. Hence, the total number of chunks downloaded equals to 153,600 chunks. The result shows that the number of chunks downloaded from server in the cluster model system is less than in the non-cluster one, and it is a constant. It means that the server serves only one peer. The other peers can download chunk from neighbor peers in the cluster.

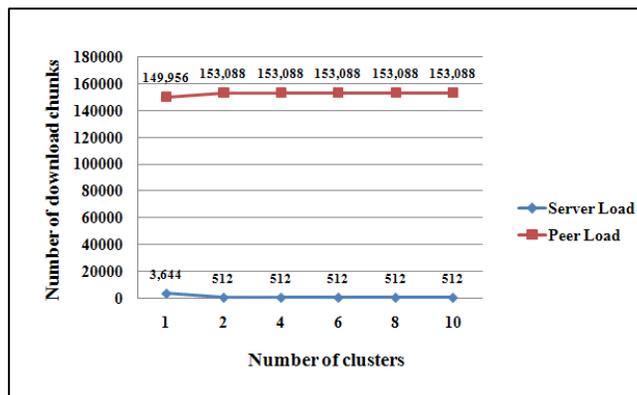


Figure 5. The relationship of server and peer load.

Figure 6 shows the relationship of global tracker connection with peers (tracker load). The x-axis represents the number of clusters (1, 2, 4, 6, 8, and 10) and non-cluster model (1). The y-axis represents the number of peers that has a connection with the tracker. The result shows that the tracker load of the cluster model system is less than the one of the non-cluster model. The tracker of the non-cluster model (1) serves for all peers in the system (300 nodes). The tracker of the cluster model serves for all super nodes or equals to number of clusters (which is equal to 2, 4, 6, 8, and 10 respectively).

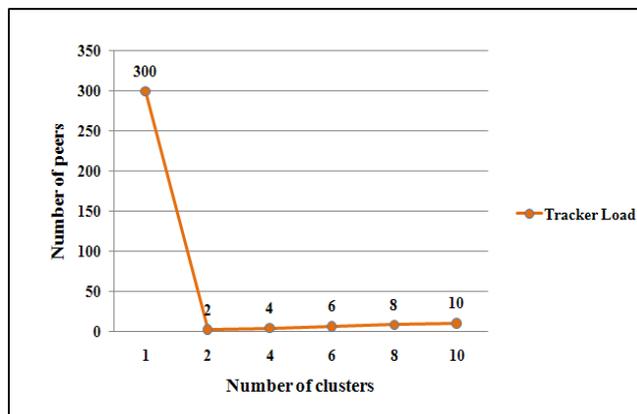


Figure 6. The relationship of tracker connect with peers.

Figure 7 shows the number of TCP control messages of each cluster used to exchange information between peers. The x-axis represents the number of clusters and non-cluster. The y-axis represents the number of control messages that used to connection between the peers. The result

shows that the control messages of the cluster model system are less than that of the non-cluster model.

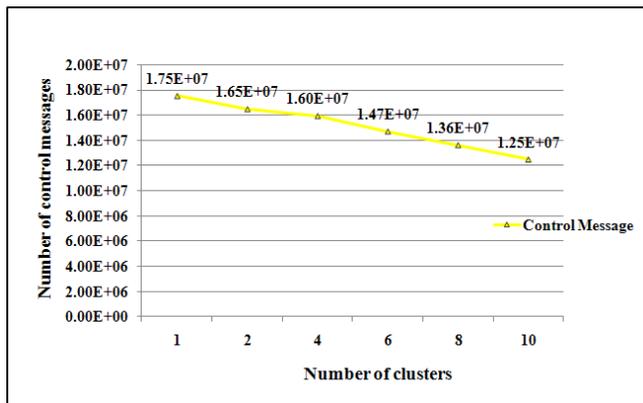


Figure 7. The control message of each cluster.

## VI. CONCLUSION AND FUTURE WORK

In this paper, a peer clustering system for different start video broadcasting is proposed. The peers are grouped into cluster according to joining time of each node or availability of chunks. The cluster model is used to control the streaming traffic. The proposed model is created algorithms for peer joining process, super node selection process, backup-node process, download process and leaving node to create the cluster model. The cluster model for the different start video broadcasting will be compared with the non-cluster model. As a result, the proposed model can reduce server load and tracker load. The number of control packet is decreased as the number of cluster or super node increase. The play out delay is only 10 sec. The performance of the cluster model is better than non-cluster model. The unpunctual viewers can join the broadcast video streaming at any point of time and can view any part of the video stream. For further performance evaluation of the different start video broadcasting system, the tracker traffic, and node dynamics are simulated. The backup-node selection method will be reconsidered.

## REFERENCES

- [1] Gnutella [online], available at <http://www.gnutella.com> [Accessed: Aug. 10, 2010]
- [2] Napster [online], available at <http://www.napster.com> [Accessed: Aug. 10, 2010]
- [3] KaZaA [online], available at <http://www.kazaa.com> [Accessed: Aug. 10, 2010]
- [4] BitTorrent [online], available at <http://www.bittorrent.com> [Accessed: Apr. 26, 2009]
- [5] eDonkey [online], available at [http://en.wikipedia.org/wiki/EDonkey\\_2000](http://en.wikipedia.org/wiki/EDonkey_2000) [Accessed: JAN. 14, 2010]
- [6] Y. Huang, T. Z. J. Fu, John C. S. Lui, and C. Huang, "Challenges, Design and Analysis of a Large-scale P2P-VoD System," in Proc. ACM SIGCOMM'08 conference on Data communication, pp. 375-388, Aug. 17-22, 2008.
- [7] X. Zhang, J. Liu, B. Li, and T.-S. P. Yum, "CoolStreaming/DONet: A data-driven Overlay Network for Efficient Live media streaming," in Proc. IEEE INFOCOM'05, vol. 3, pp. 2102-2111, Mar. 2005.
- [8] C. WU and B. LI, "Exploring Large-Scale Peer-to-Peer Live Streaming Topologie," in Proc. ACM Transaction on Multimedia Computing, Communications, and Applications, vol. 4, no. 3, pp. 19-23, pp. 1-25, Aug. 2008.
- [9] S. Tang, Y. Lu, J. Martín Hernández, F.A. Kuipers, and P. Van Mieghem, "Topology dynamics in a P2PTV network," in Proc. of the 8<sup>th</sup> International IFIP-TC 6 Networking conference'09, pp. 326-337, May 11-15, 2009.
- [10] X. Su and L.Chang, "A Measurement Study of PPStream," in Proc. of the 3<sup>th</sup> International conference on Communications and Networking IEEE ChinaCom'08, pp. 1162-1166, 2008.
- [11] H. Ketmaneechairat, P. Oothongsap, and A. Mingkhwan, "Smart Buffer Management for Different Start Video Broadcasting," in Proc. ACM ICIS'09, pp. 615-619, Nov. 24-26, 2009.
- [12] B. Fallica, Y. Lu, F. Kuipers, R. Kooij, and P. V. Mieghem, "On the Quality of Experience of SopCast," in Proc. of the 1st IEEE International Workshop on Future Multimedia Networking FMN'08, pp. 501-506, Sep. 17-18, 2008.
- [13] A. Sentinelli, G. Marfia, M. Gerla, and L. Kleinrock, "Will IPTV ride the peer-to-peer stream?," in Proc. IEEE Communications Magazine, vol. 45, no. 6, pp. 86-92, Jun. 2007.
- [14] X. Hei, C. Liang, J. Liang, Y. Liu, and K. W. Ross, "Insights into PPLive: A measurement study of a large-scale P2P IPTV system," in IPTV workshop, International World Wide Web www'06 Conference, May 2006.
- [15] J. Yu and M. Li, "CBT: A Proximity-Aware Peer Clustering System in Large-Scale BitTorrent-like Peer-to-Peer Networks," Computer Communication's Special issue on Foundation of Peer-to-Peer Computing, Vol 31/3: pp. 591-602, 2007
- [16] Z. Chen, J. Liu, D. Li, H. Liu, and A. Vasilakos, "SOBIE: A novel super-node P2P overlay based on information exchange," Journal of Computers, vol. 4, no. 9, pp. 853-861, Sep. 2009.
- [17] V. Lo, D. Zhou, Y.Liu, C. Gauthierdickey, and J. Li, "Scalable supernode selection in peer-to-peer overlay networks," in Proc.of the 2nd International Workshop on Hot Topics in Peer-to-Peer Systems, pp. 18-27, 2005.
- [18] J. Peltotalo, J. Harju, L. Väättämoinen, I. Bouazizi, and I. Curcio, "RTSP-based Mobile Peer-to-Peer Streaming System," published in International Journal of Digital Multimedia Broadcasting. Vol. 2010.
- [19] H. Ketmaneechairat, P. Oothongsap, and A. Mingkhwan, "Buffer Size Estimation for Different Start Video Broadcasting," in Proc. of the Electrical Engineering/Electronics Computer Telecommunications and Information Technology (IEEE ECTI-CON'10) International Conference, pp. 924-928, May 19-21, 2010.
- [20] H. Ketmaneechairat, P. Oothongsap, and A. Mingkhwan, "Communication Size and Load Performance for Different Start Video Broadcasting," in Proc. of the 11th Annual Conference on the Convergence of Telecommunications, Networking & Broadcasting PNET'10, Jun. 21-22, 2010.
- [21] The Network Simulator (NS-2) [online], <http://www.isi.edu/nsam/ns/> [Accessed: Feb. 15, 2009]

# Mining Web Usage and Content structure Data to Improve Web Cache Performance in Content Aggregation Systems

Carlos Guerrero  
*Department of Mathematics and  
 Computer Science  
 University of Balearic Island  
 Palma, E-07122, Spain  
 Email: carlos.guerrero@uib.es*

Carlos Juiz  
*Department of Mathematics and  
 Computer Science  
 University of Balearic Island  
 Palma, E-07122, Spain  
 Email: cjuiz@uib.es*

Ramon Puigjaner  
*Department of Mathematics and  
 Computer Science  
 University of Balearic Island  
 Palma, E-07122, Spain  
 Email: putxi@uib.cat*

**Abstract**—Web cache performance has been reduced in Web 2.0 applications due to the increase of the content update rates and the higher number of personalized web pages. This problem can be minimized by the caching of content fragments instead of complete web pages. We propose a classification algorithm to define the fragment design that experiences the best performance. To create the algorithm, we have mined data of content characterization, user behaviour and performance. We have obtained two classification tree as result of this process. These classification trees are used to determine the fragment design. We have optimized the model of a real web site using both classification trees and we have evaluated the user observed response time. We have obtained significant results which prove that the optimization of the fragment designs can achieve high speedups in the user perceived response time.

**Keywords**—Web caching; classification trees; web performance engineering; web content aggregation.

## I. INTRODUCTION

Web Caching is a widely used technique to save bandwidth, to reduce server workload and to improve user response time, i.e., Web Caching improves the performance of web architectures. This improvement is based on the *reusability* of web responses between different users and requests. This happens when several users request the same page or when a user requests the same web pages at least twice before its content changes.

In current web architectures, especially in Web 2.0 systems, content changes are more usual and the personalization of web pages is allowed. The behaviour of user generated content and pages created by collaboration are more unpredictable [1]. As a result of this, the web responses that are stored in the web cache become not *reusable*. It is widely accepted that this problem can be solved by reducing the minimum cacheable unit: content fragments instead of whole web pages [2], [3]. Nevertheless, from a performance point of view, there is a dilemma [4]: on one hand, a high level of fragmentation (a big number of content fragments) improves hit ratio, but response time could be increased due to overhead connection and fragment joining times; on the

other hand, a low level of fragmentation (a small number of fragments) minimizes overhead times but it makes hit ratio worse. So that, the problem results on determining when it is better to serve two fragments together (joined) and when it is better to do it separately (split).

To deal with this problem, we use a tree-based classification algorithm which optimizes the performance of the system. This algorithm uses the characteristics of the contents of the page (content fragments sizes, update rates, request rates, ...) to obtain a design of the web pages (which elements are served split and which ones are served joined) for an optimal performance.

We have compared the performance of using *joined* and *split* states in a high number of synthetic pages. We have applied data mining algorithms to the data obtained in these exploration phase and we have produced different classification trees. These classification trees have been tested in a web page model extracted from a real system (The New York Times web site, <http://www.nytimes.com>). The performance results obtained show high speedups of them with the two basic design alternatives: either all the fragments are joined or all the fragments are split.

Our main contribution is the developing of a *classification algorithm* which improves the performance of the systems in where web pages are created by the aggregation of content atomic fragments. This solution can be applied to systems in where these fragments can be served joined or split. We represent these two ways of delivering the content fragments by a state of the aggregation relationship. The performance of the system changes depending on the state of each aggregation relationship. The inputs of the algorithm are the characteristics of the fragments.

In Section III, we give the details about how the content aggregation application works and about the model we use to represent the contents fragments, the web pages, the characterization parameters of the fragments and the states of the aggregation relationships. In Section IV, we explain in which type of applications we can use our propose and how it fits in these applications. We explain the process

and the results of obtaining the classification trees in Section V. Finally, we analyse the performance of a system where we have implemented our solution to optimise the page fragment design (Section VI). Last section is used to explain the conclusion of our work and possible future works (Section VII).

## II. RELATED WORK

The application of techniques based on decision trees and data mining to improve the performance of the system, and more concretely, the performance of web architectures is not either new.

Pallis et al. [5] present a clustering-based scheme in order to improve the performance of a prefetching and web caching system. The input of the data mining process is the users' access patterns. Their solution identifies clusters of correlated Web pages. The idea of mining users' access sequence is also present in studies of Huang et al. [6]. They propose a system to mine popular surfing sequences to obtain rules. The rules will be used by a prediction-based manager to achieve better buffer utilization. Kumar et al. [7], use data mining processes to detect dynamic deviations from normal usage patterns of the users. But the problem of improving the fragmentation design of the web pages have not been addressed before by using classification trees. Our work differs from the previous in two main aspects: the scope and the inputs parameters of the data mining process. Our propose improved the performance of the system by determining the fragmentation design of the content in web pages. The other researches improve it by doing more accurately the selection of the pages to be evicted from the cache or the next web pages to be prefetched. Finally, they only use user access patterns as inputs of the data mining process. In the other hand, we also take into account the structure of the web pages.

Other approaches have deal with the same scope than us: to determine the fragmentation design of the content web pages [8]. But they have based the solution on cost-benefit functions. Their studies are in preliminary phases and they have not validated it in real web systems. We are not able to contrast our results with them.

Similar methods have been used in the study presented by Li et al. [9]. Their objective is to predict the performance for services instead of content delivery. Their study also differs in the data analysed. They use object timing information from packet traces. We use data extracted from higher layers (application layer).

## III. CONTENT AGGREGATION MODEL AND PERFORMANCE

Web systems have evolved from static and general content pages to dynamic generated, personalized and updated content pages. These characteristics are specially present in a group of web applications, for example: content aggregation

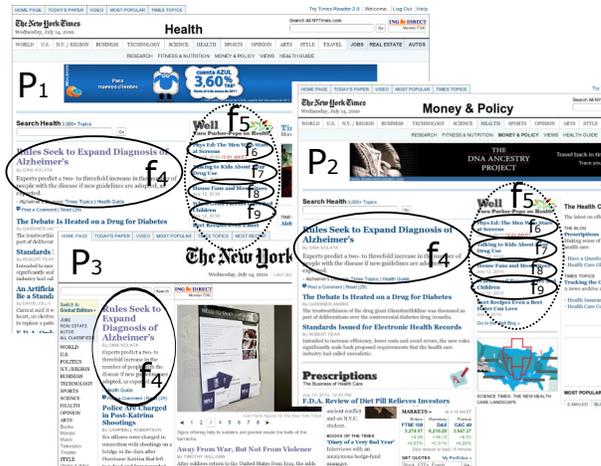


Figure 1: Example of pure content fragments, aggregated fragments and web pages from a real web site (The New York Times, <http://www.nytimes.com>)

systems, Web 2.0 applications, weblogs, newspaper sites, etc.

Content aggregation technologies are systems that combine content from different sources to create new content elements. These contents could be extracted from server-local databases or from remote systems using RSS, XML, or similar technologies. There are two main types of aggregation systems: mashups and portlets. The differences between them are related to the side where the aggregation takes place: client-side or server-side, respectively. For example, newspaper sites and weblogs are content aggregation systems where contents are stored in server-local database and the contents are aggregated on the server (this aggregation can be done by the application server, the web server or the proxy cache).

So, in these content aggregation systems, pages are created by the aggregation of independent contents (for example, news in a newspaper site or entries in weblogs). These contents are usually included in more than one web page (for example, news included in political and economy section or entries in a blog classified by tags and by sections). But the structure of a web page in these systems can be more complex than aggregations of only one degree of depth. The aggregation of a set of contents can be considered as a new content fragment which could be aggregated by other pages of fragments (for example, the group of highlighted news or the group of the most recent entries in a weblog).

Therefore, we could distinguish between two content element types: pure content elements (which are created/stored by a user/system) and aggregated content elements (which are created from the combination of others content elements). Also we consider another type of elements, the aggregation relationship which defines that a fragment includes the content of other fragment (pure or aggregated). The num-

ber of aggregation relationships is unlimited. In Figure 1, we can observe an example of three web pages extracted from the The New York Times web site ( $P_1, P_2, P_3$ ). Three of them share the aggregation of a pure content fragment ( $f_4$ ) and two of them share an aggregated fragment ( $f_5$ ). This aggregated fragment is created by the aggregation of several fragments ( $f_6, f_7, f_8, \dots$ ).

Authors of [10] define an Object Dependence Graph (ODG) to model web pages that are created by aggregation of contents. An ODG is a Directed Acyclic Graph (DAG) where pure and aggregated content elements are represented by vertices (nodes). ODGs have two vertex types:  $P_i$  which represents parent vertices (user web pages) and  $f_i$  which represents pure or aggregated fragments. The sink vertices correspond to pure elements and source vertices to user web pages. The rest of the vertices correspond to aggregated elements. Finally, the edges of the graph represent aggregation relationships between a pair of elements.

We use ODGs in our system to represent the design of the web pages. Nevertheless, ODGs are not able to model all the information we need. Thus, we have extended the ODG model to include information about the characteristics of the fragments and information about how fragments are delivered by the server: joined or split (Figure 2).

Our main goal is to determine the fragment design of a web pages which experiences a better performance. The fragment design is the information that refers to how the fragments are delivered by the server and stored by the web cache: if a set of contents are delivered and stored independently (split) or together (joined). This information can be represented in a ODG using labelled edges which indicates if a pair of content fragments are joined or split.

In the results presented in [11], we studied the relation between characterization parameters of the content fragments and the performance experienced by the web cache for aggregated content systems. We concluded that, in these type of web sites, there is a strong correlation between compared performance speedups of aggregation states (joined or split) and the characterization parameters for a given pair of related content fragments. The characterization parameters that we took into account in the study were: for both fragments, update rates, request rates and sizes; for the *father* fragment, the number of *children*; and for the *child* fragment, the number of *fathers*. We conclude also that other parameters as fragment service times or ODG structure characteristics (as fragment depth degree in the ODG) are not able to predict the difference of the performance between joining or splitting two fragments.

Our classification algorithm uses these characterization parameters, for a given pair of fragments, to determined if it is better to deliver them joined or split. Some of this information is in the ODG (number of fathers and children of a vertex) but other is not (fragment update rate, content size, ...). We have added some attributes to each vertex to

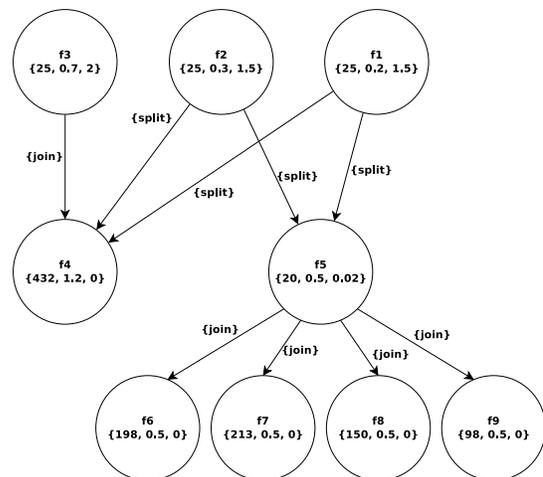


Figure 2: ODG of the example of Figure 1. The text in vertices corresponds to the identifiers of the fragments and the characterization attributes vector {size in bytes, request rate in seconds<sup>-1</sup>, update rate in seconds<sup>-1</sup>}

represent these characteristics.

To sum up, our web page model is represented by a DAG where the edges represent the aggregation relationships and they are labelled with the state of the aggregation (the contents are joined or not in the server). The vertices have a direct correspondence with the content fragments of the web page. Sink vertices are pure content fragment (user or third-party generated), source vertices are the user web pages and the rest of vertices are groups of aggregated contents. Each vertex has additional information about its request rate, update rate and size (Figure 2).

In Figure 2 we can observe a ODG example. The labels of the edges and the values of the characterization attributes vector have been chosen randomly. We can observed, in one hand, that the web server, using this model, has all the information to determine how to serve the web pages (labels of the edges). In the other hand, our design optimization algorithm also has all the information required to determine the states of the aggregation relationships (characterization vector and number of incoming and outgoing edges).

#### IV. APPLICABILITY OF THE SOLUTION

Our propose is applicable to web architectures that are able to manage fragments of contents. In these types of architecture, web server is able to serve fragments of the web pages instead of only complete HTML document. These fragments must be related using special tags inside the HTML code. Edge Side Includes (ESI) is the *de facto* standard used for that.

If a server responds using fragments of the pages, they must be gathered together in other tier of the system in order to to create the complete web page. This task could be done by client browsers, CDN networks, web proxies, web

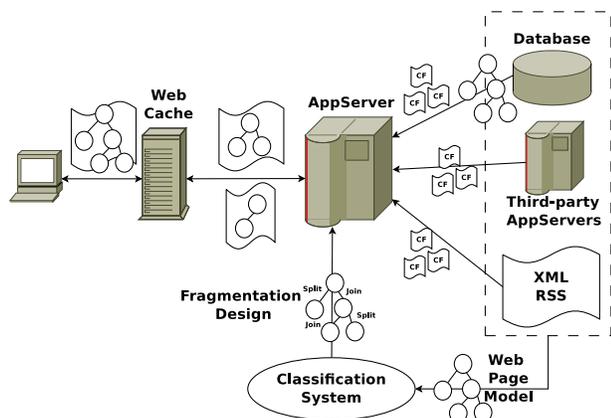


Figure 3: Architecture of content aggregation systems and fragments Web Caching

caches, etc. Nevertheless, if the task is done by elements placed between the user and the web cache, the web cache is able to manage fragments of web pages (it is able to store fragments independently and request them to the web server). As a result of this, the minimum cacheable units are content fragments instead of whole page. Therefore, the performance of the web cache will depend on the groups of fragments, i.e., the content fragments that the web server will join or not.

In Figure 3 we can observe the general architecture of a content aggregation web system. In this model, the cache is the element in charge of create the HTML documents using the fragments responded by the web server. The applications, which are based on this model, are able to manage the contents independently. The contents can be extracted from local databases, third-party applications or by the use of feeds. The application server requests all the contents and gets together them according the page design. The main goal of our research work is that the web server will create the design which makes the web cache experience the best performance. In this paper we present an optimization algorithm to solve the problem of obtaining the best web page design.

This application layout is not applicable to all kind of web applications. The applicability is restricted to web systems where the pages can be easily fragmented. Aggregation content applications satisfy this feature. In these types of applications, the atomic fragments of the page can be associated with the independent contents extracted from data sources. These correspond to pure content elements in our model. There are an important number of applications that are based on the aggregation of contents. Next elements can be considered as pure content elements which are aggregated to create complete web pages: news in a newspaper website; entries in a weblog; user update advices in social networks; users update advices in social networks; gadgets in per-

sonalized home page systems; bookmarks in social tagging systems; feeds in online feed reader systems; etc. Therefore, the solution we propose can improve the performance of a wide range of traditional web applications and in the most of Web 2.0 applications. These types of applications usually are developed allowing to the web server to manage fragments of the pages. If not, the additional work to develop this feature is not very difficult because the pure content fragments are managed independently. In both cases, the web application has to be upgraded to manage groups of fragments. It has to join the contents with *joined* state. But this feature is also simple to be implemented.

## V. OPTIMIZATION ALGORITHM

In previous sections, we have presented the problem we have to deal with: adapt the design of the content fragments of a web page to improve the performance of a cache system. To adapt the design, we limit the parameters to use to a subset of the ones available before that the web page is requested. As we have justified in Section III, these parameters are: (a) request rates; (b) update rates; (c) content sizes; (d) number of aggregations of and over a fragment. We name them as characterization parameters. Therefore, these parameters are the inputs of the classification algorithm we want to create.

The output of our algorithm is the fragment design of a web page. The algorithm decides which fragments must be requested grouped (or joined) and which ones must be requested independently (split). So, the outputs of the algorithm are the states of each aggregation relationship.

Once we have decided the inputs and outputs of our algorithm, we need to determine the scope of action of the algorithm. The set of characterization parameters to be used by the algorithm can be done taking into account: all the vertices of the ODG; a subset of vertices; only one vertex; the pair of vertices related by an aggregation. In [11] we study how the characterization parameters of the related pair of vertices are enough to decide the state of relationship (if both vertex are delivered joined or split).

In order to develop the classification algorithm, we have decided to mining the data monitored on an emulation application. The model for the pages and the behaviour of the users is synthetic. We randomly produced the pages of an content aggregation system: we produced enough content fragments to cover a wide range of values for every characterization parameter; these contents are randomly aggregated to create several web pages. We assigned update rates randomly to each content fragment (during the emulation phase, a daemon uses these values to change the content). Finally, we developed a user-emulator that uses random values to request the web pages. All the random values of all the parameters are created using uniform distributions. There are a wide number of studies that determine that these parameters follows specific statistical distributions in real

systems [12], [13]. But in this first phase, we are not only interested in samples from usual scenarios. We also want to cover a wide range of values in order to study a big search space. This search space will give us information either of usual scenarios and extreme ones. This is the reason because we use uniform distributions to create the samples.

Data mining algorithms need a set of data that is used to extract knowledge from it. This is usually called training data set. It usually also needs a second set of data to evaluate the credibility of the knowledge extracted from the first set. This is called test data set. This second set is used to test the outputs of the knowledge model created. This is the reason because we created two different page models. Each of the models had over 5000 pure content fragments and 12000 aggregation relationships creating over 1500 user web pages.

To study the influence of the aggregation states in the performance of the web cache, we randomly assigned *joined* and *split* states to the aggregation relationships and we changed periodically the state of one of the aggregation of each web page. By this way, we could monitor the user observed response time of the web pages for both states of a given aggregation relationship. Both web page models are emulated during the enough time to get confiable meanings. At this point, we were able to create a vector with the characterization parameters of the father fragment and the child fragment of the aggregation and performance information for the aggregation states (joined and split). This performance information is represented by an attribute which indicates if the performance is better for *split* or *joined* state. The vector pattern is:: ((CharacterizationParameters<sub>father</sub>), (CharacterizationParameters<sub>child</sub>), ResponseTime<sub>joined</sub>, ResponseTime<sub>split</sub>) where (CharacterizationParameters<sub>i</sub>) is (RequestRatio<sub>i</sub>, UpdateRatio<sub>i</sub>, Size<sub>i</sub>, FathersNumber<sub>i</sub>, ChildrenNumber<sub>i</sub>).

We preprocessed the data before we applied data mining on it. We transformed both response time attributes in a class one (PerformanceClass = {JoinedClass, Split-Class}). This (ResponseTime<sub>joined</sub> > ResponseTime<sub>split</sub>) (ResponseTime<sub>joined</sub> < ResponseTime<sub>split</sub>). Finally, the attributes ChildrenNumber<sub>child</sub> and FatherNumber<sub>father</sub> does not influence on the performance; so, they were eliminated from the data set.

At this point we had two data set represented by instances of the vector of characterization parameters which was used as input of the data mining algorithm and a class attribute which is the attributed to infer with the knowledge extracted. In [14], we presented a rule-based system that infers the class attribute. The rules were created manually observing the classes created after apply a clustering process to the input data set. The clustering process was done using WEKA [15]. WEKA is a collection of machine learning algorithms for data mining tasks which could be used in a GUI application or integrated in your own JAVA programs. With the obtained rules we assigned the states of the aggregation to a ODG

Table I: Summary of the complexity of the classification trees

	Number of leaves	Size of the tree
$C_{Tree}_{abs}$	46	91
$C_{Tree}_{rat}$	32	63

of a real web site and we emulate the behaviour of the user to obtain performance results (in Section VI we explain the details of the validation process). The results obtained were positive but not very significant: the optimized page model had an average speedup of 1.0488 with the page model where all the aggregation are split (all the fragments are served independently); and it had an average speedup of 1.1970 with the page model where all the aggregation are joined (all the fragments are served together). The reason of these modest results is the low credibility of the data mining process. We obtained a coverage of 40% after testing the rules over the test data set, This low coverage is explained by the way the process to extract the rules was made: manually by the observation of the obtained clusters.

This previous result helped us to conclude we have to improve the data mining process. This is the reason because we changed the algorithm to assign the states of the aggregations. Instead of a group of rules extracted from a clustering process, we have tried to develop an algorithm based on classification trees. The machine learning scheme used to obtain the classification tree is C4.5 algorithm [16]. We have used WEKA to obtain the classification tree. The implementation of the C4.5 algorithm in Weka is called J48. We used the default set up parameters of the C4.5 algorithm.

We have used two different input data sets. The data was the same in both sets but not the way to express it. In one of them we represented the absolute values of all the parameters. In the second one, we used the ratios of the characterization parameter shared by both fragments: RequestRate<sub>father</sub>/RequestRate<sub>child</sub>, UpdateRate<sub>father</sub>/UpdateRate<sub>child</sub> and Size<sub>father</sub>/Size<sub>child</sub>; while we used the absolute values of the rest of the parameters: ChildrenNumber<sub>father</sub> and FatherNumber<sub>child</sub>. We have respectively called  $C_{Tree}_{abs}$  and  $C_{Tree}_{rat}$  the trees obtained from these data sets.

In this section, we have explained how has been created the classification trees. These models are used to optimize the design of the web pages in order to improve the performance. In the next section (Section VI), we analyse how the web design models obtained with both classification trees affect this performance.

## VI. PERFORMANCE ANALYSIS

In order to validate the method we have proposed to improve the performance of web cache systems we need to monitor a real application and compare the results using the traditional fragment designs (all the fragments joined

and all the fragments split) with the results using the fragment design determined by our classification algorithm. To achieve this goal we have to deal with three problems: (a) how to create a model of a real system (b) how to monitor and execute the emulator; (c) and how to compare the results. We explain them in following sections.

#### A. Modelling a real system

The training data set used to obtain the classification trees was randomly created using uniform distributions because our goal was to cover a wide range of values. Otherwise, we are interested in using results from a real system in order to validate the use of our design optimization algorithm. Due to we are not able to use our solution in a representative web system, we need to emulate it. The models used by the emulator can be created by the use of suitable statistical distributions or mining data from a real system. We have used a combination of the two alternatives.

The model created for our emulator is based on the The New York Times web site. We have mined all the data which a final user can do. This data is related to structure of the pages. Thus, we have mined data about the structure of the aggregation relationships (news and pages created as aggregation of news) and about the size of the contents of each news. This process took place during one week on March 2010, and we got 3082 fragments aggregated in 482 pages using 13979 relationships.

We are not able to obtain information about the behaviour of the users and publishers of the The New York Times web site. This is the reason because we have to use statistical distributions in order to create the parameters of the model referred to update rates and request rates. The number of studies which characterize the behaviour of the users of current web system is very large. But it is commonly accepted that the popularity of web objects follows a power law distribution [13], [1]. These studies also detail the more usual values for the parameters of these statistical distributions. In our case, the usual accepted solution is a power law with  $\alpha = 0.83$  for request rates and  $\alpha = 0.54$  for update rates, both with  $R^2 = 0.99$  [17].

#### B. Monitoring emulator performance

The emulator has been executed several periods of time to obtain performance results for the four fragment design scenarios. The four emulation scenarios are: (a) *SplitScenario*, fragments are served independently, so the almost 14000 relationships of our model are labelled as *split* (0 joined and 13979 splitted); (b) *JoinedScenario*, fragments are served joined, relationships are labelled as *joined* (13979 joined and 0 splitted); (c) *Ctree<sub>abs</sub>OptimizedScenario*, relationships are labelled using the classification tree created with data set of absolute values, *Ctree<sub>abs</sub>* (12815 joined and 1164 splitted); (d) *Ctree<sub>rat</sub>OptimizedScenario*, in this last scenario we use the classification tree created with data set of ratios values,

Table II: Summary of the states for the four scenarios

	Number of joined states	Number of split states
<i>JoinedScenario</i>	13979	0
<i>SplitScenario</i>	0	13979
<i>Ctree<sub>abs</sub>OptimizedScenario</i>	12815	1164
<i>Ctree<sub>rat</sub>OptimizedScenario</i>	8544	5435

Table III: Hit ratios and byte hit ratios of the emulation scenarios

	Hit ratio (%)	Byte hit ratio (%)
<i>JoinedScenario</i>	8.4184	10.4679
<i>SplitScenario</i>	91.7699	80.7882
<i>Ctree<sub>abs</sub>OptimizedScenario</i>	60.0031	18.2495
<i>Ctree<sub>rat</sub>OptimizedScenario</i>	78.9625	18.2986

*Ctree<sub>rat</sub>* (8544 joined and 5435 splitted). Table II is a summary of the states values for each of the scenarios.

For each of the scenarios, we have run the emulator the enough time to obtain reliable mean values of the response times. After deleting all the samples of the transient period, we get more than 300,000 requests during 20 hours of execution for each of the scenarios, obtaining a concurrence level of 4 requests/second. The transient period is the initial period of the execution where the elements in the architecture are not in a stationary state. We consider the stationary state is reached when the hit ratio of the cache is stabilized.

#### C. Results analysis

To show the strength of our new scenarios (*Ctree<sub>abs</sub>OptimizedScenario* and *Ctree<sub>rat</sub>OptimizedScenario*), we need to compare with other scenarios. We have considered that the traditional scenarios for aggregation applications are good comparison elements. These scenarios are the ones in which : the fragments are all served independently (*SplitScenario*) and the fragments are served as a whole web page (*JoinedScenario*). We have focused in the user response time metric to study the performance of the system, because we are more interested in improve the user performance metrics than in the server performance metrics.

In general, *Ctree<sub>abs</sub>OptimizedScenario* shows the best response times (Figure 4). This is explained because the classification tree used to create the fragment design of this scenario is more complex than the one used to create the design of the *Ctree<sub>rat</sub>OptimizedScenario*. Usually, when you have a more complex classification tree, the result of the classification is more accurate. We need to study if the complexity of the tree influences the computational requirements. In this case, we will have to balance the overhead of the execution of the algorithm and the performance improvement obtained by its page model. This work is out of the scope of the presented work, but it is considered as future work.

If we compare the performance of the optimized scenarios with the traditional ones (Figure 4), we can observe the speedup is bigger when we compared with the *JoinedScenario* than with the *SplitScenario*. This is because, as we mentioned in Section I, one of the first improvements for caching in content aggregation applications is to serve the fragments of content independently [18].

Finally, we analyse the hit ratios for the scenarios. In current web systems, where the updates are more usual and the pages are more personalized, invalidation times are shorter and, as a result of this, the system reduces its hit ratios. We can see that in Table III for the *JoinedScenario*, which has lower ratios in comparison with the other scenarios. If we compare the number of joined elements (Table II) and the hit ratios (Table III), we can observed a clear correlation between them. The bigger the number of joined elements is, the worse the hit ratio is.

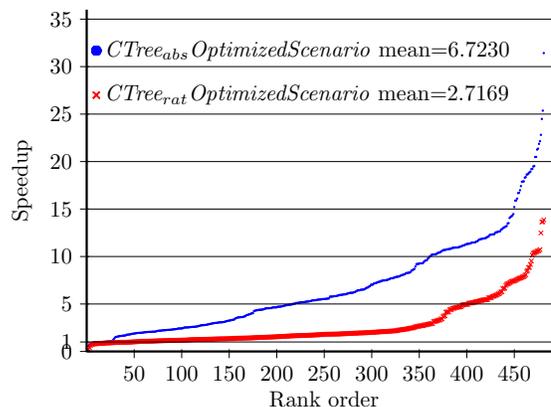
The scenarios with best response times are not the ones which has the best hit ratio. As we explained at the beginning of the document, it is best to sacrifice the hit ratio (by joining fragments) in order to reduce the overhead times of having too many fragments. This is the reason because the *Ctree<sub>abs</sub>OptimizedScenario* shows the best response time and the worse hit ratio. The *Ctree<sub>abs</sub>OptimizedScenario* takes profit of reducing overhead times in spite of it gets worse the hit ratio.

It can be very surprising that, between the optimized scenarios, the byte hit ratio is almost equal, but the hit ratios are considerably different. It is also to emphasize the high difference between the hit ratios and the byte hit ratios of both optimized scenarios. This can be explained considering that the mining process was able to inference that it is more important to obtain cache hits in small fragments than in big ones [19], [20].

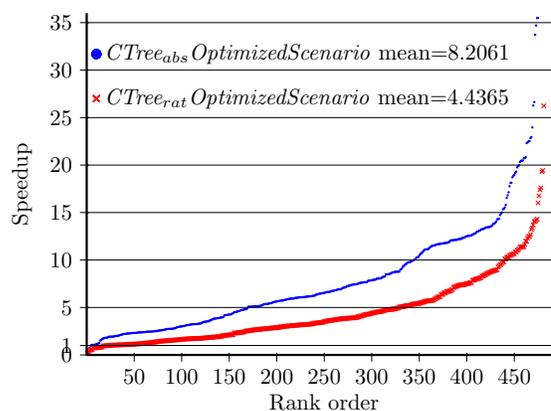
After the analysis of the results, we can conclude that the optimized scenarios show best response times in spite of their worse hit ratios. So it is interesting to adapt this improvement if we focus on improving user performance metrics. To get conclusions about server performance metrics, we would need to analyse how affects these worse hit ratios to the performance of the servers. This study is out of the scope of this work. It is consider as future work.

## VII. CONCLUSION

We have presented an improvement for web caches used in content aggregation systems. These systems have worse cache performance than others because content aggregation systems have high content update rates and high page personalization level. To solve this problem, we have presented a algorithm based on classification trees, which tries to find the best fragment design of the web page in order to minimize the overload times of communication and joining process and to maximize the hit ratios of the web cache.



(a) *SplitScenario* ( $Speedup = \frac{RT_{Split}}{RT_{CtreeOpt}}$ )



(b) *JoinedScenario* ( $Speedup = \frac{RT_{Joined}}{RT_{CtreeOpt}}$ )

Figure 4: Rank order chart of the response time speedups

To create the classification trees we have used data mining techniques, more concretely, we have used the J48 algorithm (the Weka's implementation of the C4.5 algorithm). The inputs of the algorithm are a subset of the characterization parameters of the content fragments of the web pages. The output of the algorithm is the state for an aggregation relationship between two content fragments. These states (*joined* and *split*) indicate if the server has to serve the fragments together or independently. We have extended the fragment web pages representation presented in [10] in order to cover all our modelling requirements.

The training and test data sets have been created with the results monitored in an emulation application. The page models and user behaviour models have been created randomly. They cover a wide range of values for the characterization parameters. Depending on the way we express the characterization parameters we can obtain different data sets and different classification trees. We have studied the coverage level of the most usual ways to express the parameters

and we have finally chosen two classification trees: one tree using the absolute values of the parameters (CTree<sub>abs</sub>) and the other using the ratios of the parameters shared between the two related fragments (CTree<sub>rat</sub>).

In order to validate the use of the classification trees, we have tested them in a emulation application which uses a page model extracted from a real web site, and a user behaviour model created using the most suitable statistical distributions. We have monitored the user response times in the execution of the emulation scenarios (the two corresponding to the classification trees, the scenario with all the fragments joined and the scenario with all the fragments split).

As future work, we need to contrast the results obtained in these experiments with others web pages to validate that the strength of our solution in general terms. Finally, we need to study how influence alternative classification trees.

#### ACKNOWLEDGMENT

This work is partially financed by the Spanish Ministry of Education and Science through TIN2007-60440 project and TIN2009-11711 project.

#### REFERENCES

- [1] M. Cha, H. Kwak, P. Rodriguez, Y.-Y. Ahn, and S. Moon, "I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system," in *IMC '07: Proceedings of the 7th ACM SIGCOMM conference on Internet measurement*. New York, NY, USA: ACM, 2007, pp. 1–14.
- [2] L. Ramaswamy, A. Iyengar, L. Liu, and F. Douglis, "Automatic fragment detection in dynamic web pages and its impact on caching," *IEEE Trans. on Knowl. and Data Eng.*, vol. 17, no. 6, pp. 859–874, 2005.
- [3] C. Yuan, Y. Chen, and Z. Zhang, "Evaluation of edge caching/offloading for dynamic content delivery," *IEEE Trans. on Knowl. and Data Eng.*, vol. 16, no. 11, pp. 1411–1423, 2004.
- [4] C. Guerrero, C. Juiz, and R. Puigjaner, "The applicability of balanced esi for web caching," in *Proceedings of the 3rd International Conference on Web Information Systems and Technologies*, March 2007.
- [5] G. Pallis, A. Vakali, and J. Pokorny, "A clustering-based prefetching scheme on a web cache environment," *Comput. Electr. Eng.*, vol. 34, no. 4, pp. 309–323, 2008.
- [6] Y.-F. Huang and J.-M. Hsu, "Mining web logs to improve hit ratios of prefetching and caching," *Know.-Based Syst.*, vol. 21, no. 1, pp. 62–69, 2008.
- [7] C. Kumar and J. B. Norris, "A new approach for a proxy-level web caching mechanism," *Decis. Support Syst.*, vol. 46, no. 1, pp. 52–60, 2008.
- [8] O. A.-H. Hassan, L. Ramaswamy, and J. A. Miller, "Mace: A dynamic caching framework for mashups," in *ICWS '09: Proceedings of the 2009 IEEE International Conference on Web Services*. Washington, DC, USA: IEEE Computer Society, 2009, pp. 75–82.
- [9] Z. Li, M. Zhang, Z. Zhu, Y. Chen, A. Greenberg, and Y.-M. Wang, "Webprophet: automating performance prediction for web services," in *Proceedings of the 7th USENIX conference on Networked systems design and implementation*, ser. NSDI'10. Berkeley, CA, USA: USENIX Association, 2010, pp. 10–10. [Online]. Available: <http://portal.acm.org/citation.cfm?id=1855711.1855721>
- [10] J. Challenger, P. Dantzig, A. Iyengar, and K. Witting, "A fragment-based approach for efficiently creating dynamic web content," *ACM Trans. Internet Technol.*, vol. 5, no. 2, pp. 359–389, 2005.
- [11] C. Guerrero, C. Juiz, and R. Puigjaner, "Web cache performance correlation with content characterization parameters in content aggregation systems," in *Proceedings of the XXXVIIth Latin American Informatics Conference*, 2010.
- [12] M. A. Goncalves, J. Almeida, L. Santos, A. H. Laender, and V. Almeida, "On popularity in the blogosphere," *IEEE Internet Computing*, vol. 99, no. PrePrints, 2010.
- [13] L. Guo, E. Tan, S. Chen, X. Zhang, and Y. E. Zhao, "Analyzing patterns of user content generation in online social networks," in *KDD '09: Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*. New York, NY, USA: ACM, 2009, pp. 369–378.
- [14] C. Guerrero, C. Juiz, and R. Puigjaner, "Rule-based system to improve performance on mash-up web applications," in *Distributed and Artificial Intelligence: 7th International Symposium*, 2010.
- [15] Weka Machine Learning Project, "Weka," URL <http://www.cs.waikato.ac.nz/ml/weka>, October 2010.
- [16] I. H. Witten and E. Frank, *Data Mining: Practical Machine Learning Tools and Techniques*, 2nd ed. San Francisco, CA: Morgan Kaufmann, 2005.
- [17] F. Duarte, B. Mattos, A. Bestavros, V. Almeida, and J. Almeida, "Traffic characteristics and communication patterns in blogosphere," in *Proceedings of the International AAAI Conference on Weblogs and Social Media*, 2007.
- [18] C. Yuan, Y. Chen, and Z. Zhang, "Evaluation of edge caching/offloading for dynamic content delivery," in *WWW '03: Proceedings of the 12th international conference on World Wide Web*. New York, NY, USA: ACM, 2003, pp. 461–471.
- [19] L. Cherkasova, "Improving www proxies performance with greedy-dual-size-frequency caching policy," HP Technical Report HPL-98-69, Tech. Rep., 1998. [Online]. Available: <http://www.hpl.hp.com/techreports/98/HPL-98-69R1.html>
- [20] L. Cherkasova and G. Ciardo, "Role of aging, frequency, and size in web cache replacement policies," in *In Proceedings of HPCN Europe*. Springer-Verlag, 2001, pp. 114–123.

## Enhancing Situational Awareness by Means of Combat-ID to Minimize Fratricide and Collateral Damage in the Theater

Tapio Saarelainen

Department of Military Technology  
National Defence University, Finland  
Helsinki, Finland  
tapio.saarelainen@mil.fi

**Abstract**—The Future Battlefield has expanded to a battlespace where its commanders rely on SA-tools to perform optimally in their given tasks. Operations may include combat settings as well as counter insurgency actions, peace-keeping operations and disaster relief activities. In multi-national operations taking place in versatile and hostile environments, it is essential to detect, classify and identify the encountered objects in the battlespace early enough. The concept of war has changed in the direction of multi-symmetric warfare involving enemy troops, own forces and impartial entities. This paper describes existing technical solutions and oncoming tools applicable in enhancing Situational Awareness (SA) and minimizing fratricide and collateral damage in the battlefield. This paper adopts a meta-analytic approach and examines the current capability of utilized Combat Identification (CID) systems to minimize fratricide and avoid collateral damage in the theatre as described in relevant existing studies and military exercises. This, furthermore, involves introducing means to enhance the overall SA in the battlespace.

**Keywords** - Situational Awareness (SA), Common Operational Picture (COP), fratricide, collateral damage.

### I. INTRODUCTION AND DEFINITIONS

This paper begins by first taking a look in Combat Identification (CID) issues. Secondly, a comprehensive approach to Target Combat Identification (TCID) is discussed. Thirdly, the paper examines how an ongoing training program, the Bold Quest 2011, is testing these presented methods. Lastly, the concept of Shared Situational Awareness (SSA) is introduced with the help of utilizing Unmanned Aerial Vehicles (UAVs) and Unmanned Ground Vehicles (UGVs). Combining a swarm of UAVs with Free Space Optics (FSO) offers a reliable platform for aerial ad-hoc networking [6]. The capabilities of these platforms are implemented by enhancing SA and data distribution to enable near real-time Common Operational Picture (COP) to be implemented into SSA. Once the location data of different entities can be reliably forwarded to respective command posts, the number of fratricide incidents and collateral damage can be significantly minimized. The significance of Target Combat Identification (TCID) in minimizing fratricide is introduced.

Once the information and its distribution in the battlespace is defined as the key in Network Centric Operations (NCO), every effort to ensure the information flow between own warriors and sensors needs to be analyzed

[5]. Contemporary weapon systems require greater amounts of intelligence data at a higher fidelity than ever before [19]. Since operations tend to be multi-national, different sensors and systems are required to communicate understandably between each entity to minimize fratricide and collateral damage by maximizing the distribution of the near real-time Common Operational Picture (COP). One solution is to utilize Battle Management Language (BML) [5].

This paper tackles the following three questions: What are the means to locate the soldier by employing the existing CID and SA technology? How to increase SA with the available technical solutions? And, furthermore, how to test these technologies in peace-time?

As for defining terminology, a new network structure called the Wireless Polling Sensor Network (WPSN) is explained in [1]. Since nodes do not form a network per se but rather are polled by a selected node of the mobile network, they remain undetected due to their passive nature. The network structure offers a new and ubiquitous way to share and forward all kinds of data, including data collected by various sensors. Moreover, the outdated Identification Friend or Foe (IFF) systems are replaced and supplemented with effective and accurate means to identify the prevailing objects.

Examining the means to minimize fratricide and collateral damage presupposes applying the model presented in Figure 1 below. This terminologically updated model emphasizes how Tactics Techniques and Procedures (TTP), CID, COP, and SA play a central role in minimizing incidents of fratricide and collateral damage.

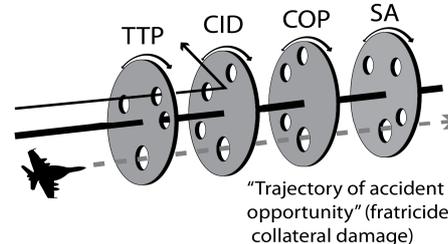


Figure 1. The Reason's Swiss Cheese Model updated by applicable terminology as a tool to explain the mechanism of avoiding fratricide and collateral damage, Blue-on-Blue (BoB).

An applicable definition for SA is given in Army Field Manual 1-02 (September 2004): “Knowledge and understanding of the current situation which promotes timely, relevant and accurate assessment of friendly, competitive and other operations within the battle space in order to facilitate decision making. An informational perspective and skill that fosters an ability to determine quickly the context and relevance of events that is unfolding.”

The process of determining the affiliation of detected objects in the battlefield equals Target Identification (TI) [4]. When using this categorization, blue denotes the friendly force, red the enemy, and white refers to neutral (impartial) entities. The traditional method of TI is based on visual signature of the object of interest. In contemporary warfare TI is also based on utilizing the electromagnetic spectrum of the target. Properly applied data and sensor fusion can be seen as a means to prevent collateral damage and fratricide. As a matter of fact, TI can be divided into two categories: Cooperative Target Identification (CTI) and Non-Cooperative Target Identification (NCTI). CTI allows a human shooter or sensor to interrogate a potential target and thereby forces the potential target to respond to the interrogation in a timely manner as described in Figure 2 [3].

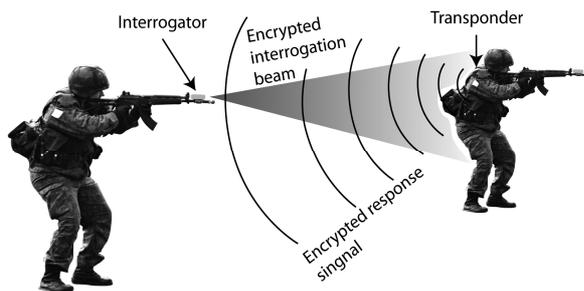


Figure 2. The process of Cooperative Target Identification (CTI).

NCTI in turn does not require a cooperative response from the target. NCTI involves systems or methods which exploit physical characteristics of entities in the battle space to help identify and determine affiliation. NCTI systems include optics such as Thermal Weapon Sights (TWS), night Vision Goggles (NVG), Forward Looking Infrared Radar (FLIR), as well as vehicle and personnel markings such as Joint Combat Identification Marking Systems (JCIMS). JCIMS are used in conjunction with TWS, NVG and FLIR assisting in friendly identification at the point of engagement [4].

CID can be defined as a process of attaining an accurate and timely characterization of detected objects in the joint battle space to the extent that high confidence, timely application of military options and weapons resources can occur [4][25]. An extension of this can be understood as a process of accurately characterizing the detected objects via the operational environment sufficiently to support engagement decisions [4]. The purpose of CID is to enhance unit combat effectiveness and simultaneously minimizing

fratricide. In the form of an equation CID reads as:  $SA + TI = CID$  [4].

The core capability in SA is COP that fosters effective decision making, rapid staff actions, and appropriate mission execution [4][26]. COP is employed to collect, share and display multi-dimensional information to facilitate collaborative planning and response to security incidents. Each organization involved in applying COP typically comprises three types of modules: 1) information gathering sources that observe events and report information to the command and control module, 2) a command and control module that makes decisions based on both information received directly from its information gathering sources and information reported by other peers, and 3) display units at the emergency location that receive instructions from the command and control module [4].

The acronym MOUT (Military Operations on Urbanized Terrain) denotes military actions planned and conducted on a terrain complex where manmade constructions impact the tactical options available to commanders. Urban combat operations may be conducted in order to capitalize on the strategic or tactical advantages gained by the possession or control of a particular urban area or to deny these advantages from the enemy [1]. The characteristics of MOUT include complex situations brought about by urban environments (ambushes, civilians). The maze-like boxed surroundings hamper command and control leading to combat engagement taking place at squad level with low coordination with higher echelons.

Combat Effectiveness (CE) can be defined as the ability of a (friendly) unit to rapidly and accurately sort and characterize detected objects into categories (blue, white, red) and make a decision as to whether or not to employ deadly force against the identified object/target. Effectively applying the CE guarantees a minimum level of collateral damage and fratricide. Now, to exemplify the previously defined terms, the following briefly examines Rules of Engagement (ROE) together with tactics, techniques and procedures (TTP). ROE defines the situations and guidelines which then support an individual in a situation when a decision is made about whether or not to open fire. TTP supports the decision making process regarding force implementation in the Area of Operations (AOR). Depending on the ROE formulations, the orders concerning using force may vary as indicated in Figure 3.

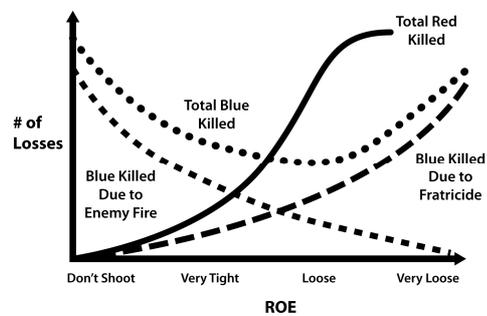


Figure 3. ROE in a relation to the number of troops killed (blue, white, red) and the number of losses and fratricide.

All warriors depend on SA [2] which can be provided also by using WPSN-systems introduced in [1]. The Blue Force Tracking-systems (BFT) along with the White Force Tracking (WFT) presented in [3] provide vital information for improving commanders' decision-making and avoiding fratricide and collateral damage. Blue Forces, allies and White Forces need to be constantly precisely located. It is crucial to improve the efficiency of dismounted operations with smaller and more capable units. These units require a great degree of flexibility and reliability to obtain their set goals to ensure the desired end state.

This paper discusses recent research in the Finnish Defence Forces and elsewhere on Combat Identification. Since the author's research has targeted the networking of the future warfighter in particular, this paper examines components relevant in this regard. The remaining of the paper is arranged as follows: Section II introduces the related work, Section III describes challenges in combat identification, Section IV explains the process of the comprehensive targeting process, Section V deals with the problems causing fratricide, whereas Section VI explains the challenges in distributing the SA, Section VII focuses on problems designated in Military Operations in built-up areas, Section VIII introduces the Bold Quest exercise, and finally, Section IX concludes the paper.

## II. RELATED WORK

In military operations in the Persian Gulf, BFT was used during Operation Iraqi Freedom (OIF) for coordinating operations among the Joint Services and with allies and resulted in reduced casualties due to enhanced SA [21]. Obviously, Soldier Modernization Programmes (SMPs) are significantly important in enhancing the performance of the militaries. These SMPs concentrate on improving and updating dismounted soldiers' equipment. These Future Soldier Programs are currently underway along with a series of demonstrations and exercises in which collateral damage and fratricide are to be minimized, whereas the means to increase SA via improved BFT and WFT are also in progress. One example of these is a series of Bold Quest (BQ) exercises carried out since 2007. The incoming Bold Quest 2011 (BQ11) features a primary emphasis on enhancing the capability of CID as related to the needs of dismounted coalition warfighters and those providing them with timely and effective supporting fires.

The principal contributing efforts, technical and procedural, involve the following [21]. First, CTI, automated query or response systems for dismounted personnel and light vehicles need to be addressed. Secondly, a means to share SA systems for employment at the platoon, squad, team, and individual levels must be applied. Thirdly, digitally-aided supporting fires' coordination and control must be defined. Fourth, Digitally-aided Close Air Support (DCAS) coordination and control has to be applied. In addition, challenges with Combat Identification Server (CIS) interoperability and Personnel Recovery command and control (C2) need to be solved. Lastly, marking and beacon systems for dismounted personnel, light vehicles, and

friendly locations need to be applied. In fact, the US Army is fielding its new SA system known as Force XXI Battle Command and Brigade and Below (FBCB2) [2][21]. There is also a European Defence Agency (EDA) proposal, named the SAFE, a cancelled project, which was to increase SA and efficiency in crisis management operations in the urban environment ranging from mid-intensity battle to peace keeping operations. One of the keys into the success is careful mission analysis and thorough evaluation of Courses of Actions (COAs). Both processes can save time and minimize collateral damage. The use of available Blue and friendly Forces and resources can be optimized. This increases efficiency and along with minimum casualties, leads to minimum recovery times.

Another objective of the BQ11 for the Finnish Defence Forces (FDF) is currently developing new joint fires and joint C2 capabilities both for national defence and coalition combined operations [22]. The recent progress involves starting to develop tactics, TTP for Air-to-Ground (AG) operations and Joint Fires (JF). One of the Finnish goals involves purchasing an interoperable, net-enabled and digitally-aided material solution for close air support. Engaging the target accurately and cost-effectively is vital especially for a nation with limited resources. At the moment, fourteen nations participate in the BQ11 exercise to improve the outlined objectives [21].

Obviously, all nations want to enhance the SA, minimize fratricide and collateral damage in the battlespace. This paper provides a meta-analysis and describes the results gathered in existing studies and military exercises thereby highlighting the significance of technology in maximizing the effectiveness of own forces. This is applicable until robotic armies replace human warriors in the battlespace. Until then, the continuous developing of the SA and CID systems is instrumental for saving lives in the ongoing and ensuing conflicts.

## III. CHALLENGES IN COMBAT IDENTIFICATION

In military operations everything is done to prevent BoB. Currently, identifying a warrior regardless of the visibility conditions is essential. As evident in Figure 2 above, both an interrogation unit and a responder unit are necessary, presupposing, first of all, that the systems are fully operational, and, secondly, that the distance between the warriors is appropriate. In case the identification system doesn't reply, a human is making the decision to open fire based on the TTP. The Identification to whether or not to open fire is based on the visual signature of the uniform, weapon and gear [23].

However, one needs to keep in mind that there is always the possibility that the location device gets stolen or misused by a third party, for example, an insurgent tries to function as a member of the White Force (WF) [3]. In order to increase the reliability of the system, the tracking devices have to be pre-coded and tied in pairs in advance before entering the battlefield to prevent the stealing of the tracking device. Once paired devices are torn apart, they stop functioning as planned – and devices become dysfunctional [3]. After the

separation process, the devices must be re-paired and re-coded by the operator. During this process, the operator re-identifies the person.

IV. COMPREHENSIVE TARGETING PROCESS

When returning to ROE/TTP (whether or not to open fire), a link behind the targeting process deserves a closer look. The process is known as Detect, Identify, Decide; Engage and Assess (DIDEA) [4]. The DIDEA provides an iterative, standardised and systematic approach supporting targeting and decision making, being generic enough to be used as a systematic process for C2 node targeting and decision making. Separate actions inside DIDEA area as follows:

**Detect:** The process of acquiring and locating an object in the battle space by analysing the phenomena in the electromagnetic spectrum.

**Identify:** The process of classifying an object into the category of blue, white (neutral) or enemy. This represents a primary step where specified CID tasks are accomplished.

**Decide:** The decision making process that follows the detection and identification phases. This is the most generic step within the process and represents the primary step where a specific ROE application occurs. In the decision-making phase the executive officer/warrior has to decide and define what type of weaponry is appropriate for to the mission. In cases of opting for the use of deadly force, the following questions need to be addressed: 1. Can I engage (ROE application)? 2. If there are several targets, what is the order to engage the selected targets? 3. Which one is the most appropriate weapon system (most cost-effective, least collateral damage and fratricide causing).

**Engage:** The execution of selected weapons in a selected order starting from the most dangerous target moving on according the panned sequence.

**Assess:** Monitoring the gained effects with the use of destruction power. Employing the force of various weapon systems available is repeatedly executed until the required level of destruction is achieved.

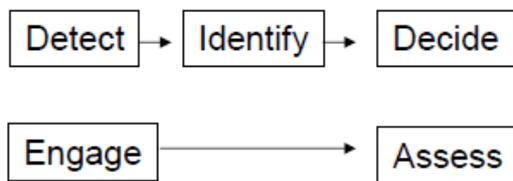


Figure 4. The simplified DIDEA process.

V. WHAT IS THE PROBLEM?

Self-evidently, cases of BoB and collateral damage are bound to surface to some extent. Militaries are interested in locating own troops and increasingly also in the neutral entities of Non-Governmental Organizations (NGOs) and

Governmental Organizations (GOs), the WF, the members of which can be tracked by using WFT described in [5].

Briefly put, the problem relies in relating the TTP, CID, COP and SA to the rules of ROE. This involves dealing with the balance described in Figure 2. If ROE formulations are too strict – for example, the commander’s intent is to avoid the use of deadly force unless it is absolutely certain that the targeted object is positively identified to be an enemy – the Blue Force will suffer on the basis of actions caused by the enemy. And, if ROE formulations leave too much room for interpretation, various types of casualties (red, blue and white) are bound to occur. Thereby the transmission of combat-critical location and identification data plays a crucial role in the battle space.

A warrior can be equipped with the appropriate locating devices as presented in Figure 5, but this not, however, apply in the case of representatives of the White Force.

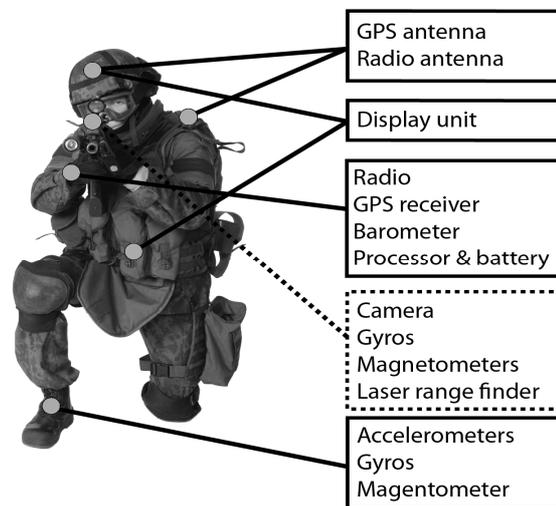


Figure 5. An example of a fully integrated Warrior for location purposes outdoors, indoors and in MOUT for contemporary warfare [18].

The destruction power of a chosen weapon system has to be optimized according to the enemy location (forest, open area, Urban Territory) state of movement on-the Move (OTM) or at-the-halt (ATH) and the protection-level (mounted, dismounted, dig). Aside, the commanding officer has to keep in mind that operations are executed with improper SA, COP and with lack of precise real- time CID.

The cruel reality remains that an executive commander is always aware of fact the there is always the possibility of fratricide and collateral damage. The commander who is not employing the firepower in his possession will have his troops killed by the actions of the enemy as presented in Figure 2 earlier.

To enhance improved SA and COP, Geographical Based Situational Awareness (GBSA) can be utilized [9]. The system utilizes the VHF-frequency operated Combat Net Radios (CNRs). When the CNRs are on the connectivity range, they recognize and identify radios in the system. Once

the radios are at the same channel and the clock (hopping sequency) of CNRs are in a correct time, a reliable SA-tool [9]. The main problems related to this system have to do with the clock and hopping sequence. This is one possibility to avoid fratricide and collateral damage. At the moment, the main benefit of this concept is in preventing from being fired at by own weapon systems, BoB, by means of improved SA-information.

Figure 6 below emphasizes the importance of both CID and BFT in a battlefield environment. The end-user of the weapon system, the one opening fire, has to be aware of the locations and status of both own troops and the enemy. Before employing the deadly force, the impact of the given weapon system needs to be in proportion to the chosen target and its perimeters in order to avoid unnecessary collateral damage caused by possibly choosing too effective weapon systems to destroy a particular target.

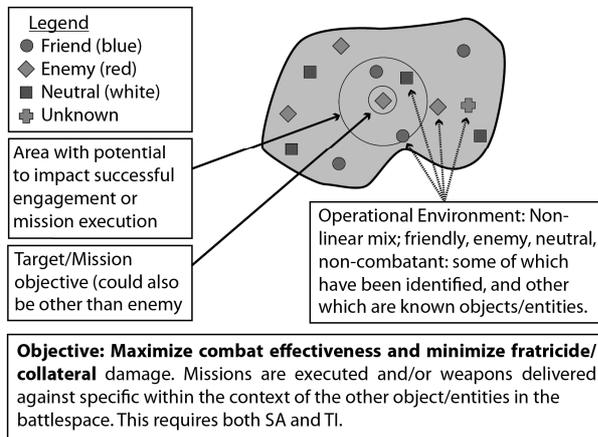


Figure 6. Military problems of CID-BFT identified from the perspective of collateral damage and fratricide assets.

### VI. CHALLENGES INVOLVED IN DISTRIBUTING SITUATIONAL AWARENESS DATA

The amount of gathered data via sensors and tracking systems is vast. To distribute the location information filtered and fused through various systems remains a challenge. A warrior has to fight, not to monitor his palm, wrist computer or lap-top. Besides, the disturbances in electromagnetic spectrum, quality of service (QoS) and transmitting power along with the limited bandwidth set limitations to the communication systems. As indicated in Figure 7, the possibilities of communication are vast, since almost all the sensors are somehow linked together to maximize BFT, CID, COP and SA.

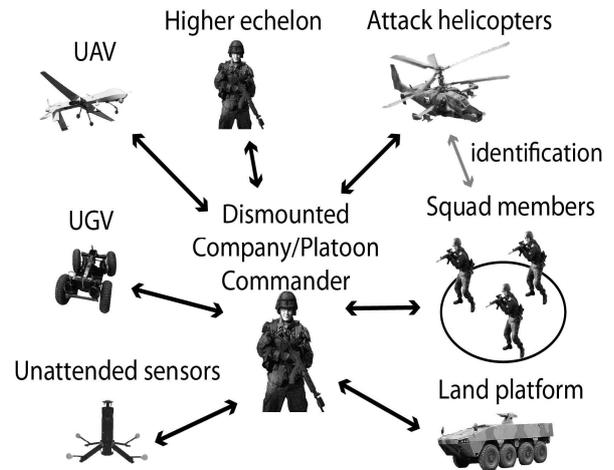


Figure 7. The types of possible platforms serving as sensors and network nodes.

The problems in data distribution are linked to various devices and data in interfaces. BML can be seen as a common language enabler between machines and interfaces [5] along with almost ubiquitous swarms of UAVs described in [6]. Limitations in energy and bandwidth play an important role. Locating instruments of various types consumes reasonable amounts of energy, not to mention the increase in weight and number of devices in warrior gear and required maintenance. Due to the lack of accessible wire line infrastructures, unmanned systems have to be powered through a combination of batteries, solar power, and power scavenging [15]. When FSO-technology is adopted in backbone networks and between selected ground stations, an intelligent, dynamic and secure data transmission with high data rates can be offered to mobile end-user [7]. FSO-technology offers high-speed, reliable and cost-effective connectivity for heterogeneous wireless services provision in both urban and rural deployments when Dense Wavelength Division Multiplexing (DWDM) is utilized in Radio-on-FSO (RoFSO) system [8]. It has been demonstrated in tests that the advanced DWDM RoFSO offers a viable solution to provide broadband wireless connectivity. Radio over Fiber (RoF) technology will most likely offer a reliable data transmission rate of 10 Gbps in the next generation FSO-systems [8].

Furthermore, older existing systems are available for distributing data gathered by various types of sensors in various types of military and humanitarian crises environments. These technologies are based on WPSNs described in [1] and Wireless Sensor Networks (WSNs) described in [10][16]. The former is passive, so it will remain hidden once the latter is active, representing an easier detected system. Both systems are applicable to be used in transmitting constant data from a sensor to a node, for example, to a vehicle or an Unmanned Vehicle (UV).

To maximize the possibility of devices communicating in a proper and planned manner, the topology of network

systems has to be correctly coordinated (manage spectrum usage with group mobility patterns) [24]. Also the hierarchy of a network has to support and enable this. Both the goals can be achieved by hierarchical design where devices are only to interact with their peers from the same group [11]. In addition, the transmit antenna selection is a practical technique for achieving significant power gain, even with commodity hardware and without changes to different waveform protocols [12].

#### VII. LOCATION POSSIBILITIES IN URBAN AREAS

An Army tactical warfighter needs network services both OTM and ATH [3]. One of the lessons learned from Iraq and Afghanistan was the need for a more robust Beyond-Line-Of-Sight (BLOS) communication capacity between the lower Army echelon Land Warriors, from Squad Leaders to Battalion Commanders [3].

The proposed and described solutions have to be based on novel, generic and robust battlefield-proven solutions in order to meet the given needs, and this in turn involves addressing the topology of the network system carefully. In MOUT transmitting and receiving signals of different waveforms simultaneously is challenging due to the nature of the combat environment [17].

Since the power production and power consumption will remain as a challenge, certain issues need to be addressed. Thus when defining the network design, it has to be emphasized that network coding enables a more efficient, scalable and reliable wireless network [14].

The MOUT environment features no GNSS indoors and indoors propagation poses a serious problem. The placement of an antenna platform is challenging. One solution can be the installing of a high-bandwidth conformal antenna in the soldier's helmet with the coverage of over 750 MHz through a 2,7 GHz frequency band [13]. The combat-critical solutions involve improving communicating, SA and transmitting C2 information among highly dispersed battlefield units in dynamic environments, such as MOUT [14] [17].

Next, let us assume that there is a WPSN-system available for positioning and location services. If the capability of GPS-Pseudolite, better known as the Self-Calibrating Pseudolite Array (SCPA), is attached into the satellite-based Carrier-phase Differential GPS-type (CDGPS), it is possible to determine positioning in locations without access to the GPS satellite constellation [16] [1] [20]. This in turn will improve locating own troops inside buildings dramatically, thereby significantly improving CID, TID and SA. The particular challenges set by the MOUT environments are presented in Figure 8.

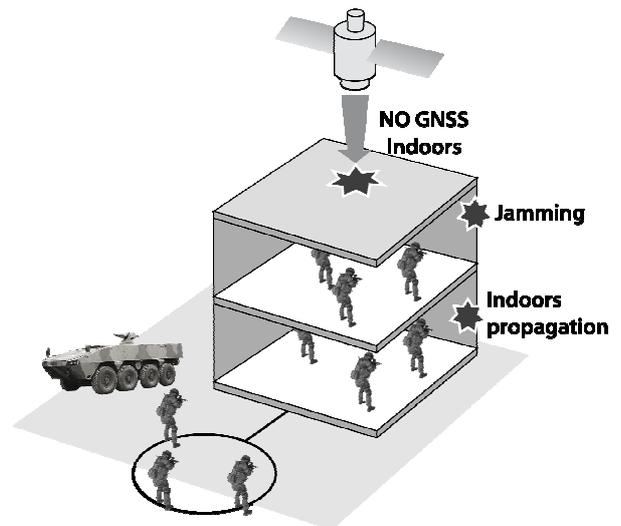


Figure 8. The challenging environment of MOUT.

#### VIII. BOLD QUEST EXERCISE

The Bold Quest is a name of exercises which have been executed as a series of exercises since 2007. One purpose of these exercises is to enhance SA, TID and minimize collateral damage and fratricide. The exercises are planned to be executed biannually. Between the exercises there is an evaluation period, where collected data are analyzed. Damage control in operations can be minimized by adopting advanced combat identification technologies via executing series of experiments. In Bold Quest 2009 (BQ09), organized by the US Joint Forces Command at Camp Lejeune and Marine Corps Air Station Cherry Point, North Carolina, Coalition Combat Identification – Advanced Concept Technology Demonstration (CCID – ACTD), various systems were tested. Ten partner nations participated, and the objective of the exercise was to demonstrate and assess air-to ground CID technologies in a stressful and challenging operational environment [18].

In BQ09, technical initiatives included prototype-level systems to enable aircrew and controllers to exchange position information digitally among friendly ground elements relative to their proximity to potential ground targets [4][25]. The results improve the capabilities in target acquisition and minimize the risk and level of fratricide.

One of the most promising technologies demonstrated during BQ09 was the Smart-Pull Warfighter Information for Targeting (SWIFT) Combat Identification Server (CIDS) [17]. SWIFT provides the fighter pilot of a combat aircraft with on-demand request capability for friendly forces location information. Information is forwarded via a secure transmission from the ground located CIDS as request based service. This is the process for the fighter pilot to verify the known friendly locations according to the ground commanders close air-support request. Figure 9 portrays the CCID-ACTD system.

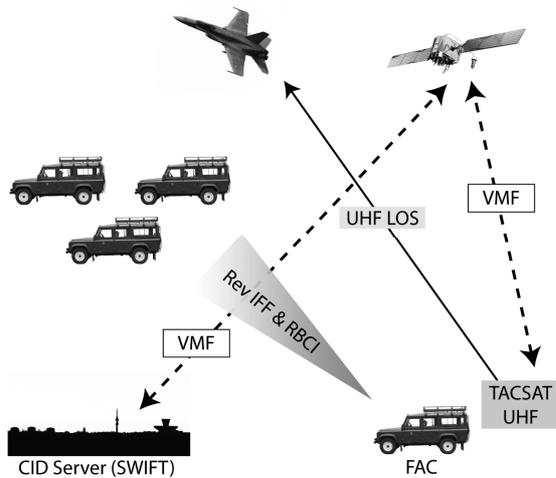


Figure 9. The CCID-ACTD system of collateral damage and fratricide assets.

When it comes to CIDS, the requirement is to contribute to improving Tactical Situational Awareness (TSA) for allied forces involved in the delivery and control of indirect and direct fires to land operations. The CIDS is about to correlate BFT information from 15 different sources, including LINK-16, BOWMAN [25]. Besides, the CIDS will enable the joint fires assets and Close Air Support (CAS)/Close Combat Air (CCA) aircraft on request. The CIDS will utilize Link-16, Variable Message Format (VMF) and other tactical networks to redistribute BFT information [17].

Once the next Bold Quest series exercise is arranged in June 2011 (BQ11), the focus will be on Fires on Dismounts. The CIDS technical demonstrator will be ready for the exercise for the extensive testing of the planned concept [17].

The air-to-air and ground-to-air systems utilize the IFF as a method to define entities in their battle space. In the near future, ground-to-ground systems may start using Battlefield Target Identification Device (BTID), CTI systems. As demonstrated in Figure 10 below, Radio-Based Combat Identification (RBCI), along with different types of IFF-systems available, can be utilized.

A key factor is the efficiency of a warrior, which can be gained via improved SA, BFT and Command, Control, Computers, Communication, Information, Intelligence, Surveillance and Reconnaissance (C<sup>4</sup>I<sup>2</sup>SR). However, since supplementary gear can never fully substitute human intelligence, a warrior must remain active and alert maintain in the battlefield.

The constant objective of military forces is to gain the initiative and turn this into success – to maximize performance in military operations with minimized casualties in minimal recovery times. Operational time spent in the battlefield can be minimized by careful mission analysis and thorough evaluation of Courses of Actions (COAs).

These presented technologies expand the possibilities to carry out the set missions in the future with ever increasing

performance capabilities. On a practical level, this involves, for instance, minimizing fratricide and collateral damage by detecting roadside bombs early and precisely enough. One of the most promising Combat Identification systems is introduced in Figure 10 below. The collected CID data can be forwarded via different means of communication to everyone requiring these CID data.

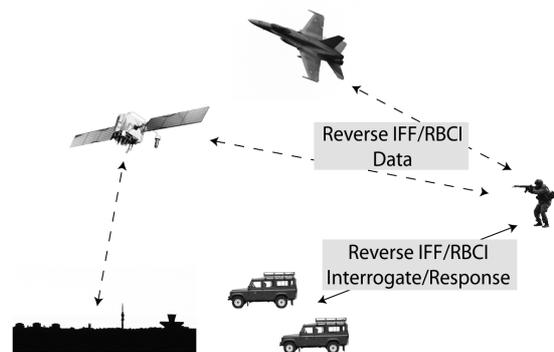


Figure 10. Radio-Based Combat Identification (RBCI).

To facilitate capability development, the Finnish Defence Forces plans to participate in the BQ11 in to increase combat effectiveness in national defence and coalition combat operations. This constant and continuing process enhances the capability to avoid BoB along with the capability to avoid collateral damage by tracking the WF.

The battle can be won only by careful mission planning and comparing different COAs, whereas winning wars presupposes winning the hearts and minds as well. Avoiding collateral damage, minimizing fratricide and increasing SA among the coalition forces enables performing maximally in minimum time, and enhances the probability to succeed in this complex challenge. The adoption of existing technologies and their viable solutions offers a key to constant success when appropriately applied.

## IX. CONCLUSIONS

This paper is a synthesis of recent research work in the Finnish Defence Forces and elsewhere on Combat Identification, and many of the proposed components in this synthesis are based on the author's research on the networking of the future warfighter. Since accurate and timely identification in the battlefield is life-preserving for each warrior, a careful analysis of the performance and capabilities of chosen systems needs to be executed before introducing any of these systems in the battlespace. CID equals the process that warriors and sensors go through in order to identify battlefield objects prior to deciding whether or not to open fire. Warfighters are trained to employ all available means at their disposal to define and assess potential targets in the battlespace prior to employing combat power. CID can be viewed as a complex series of networked systems, procedures and doctrine as presented in Figures 1 –

4. These systems also include the definitions of TTP, COP, SA, ROE and DIDEA.

More specifically, problems can arise in particular in commanding and being commanded, as indicated in Figures 5 – 7. Moreover, the functioning capacities of a chosen network are relevant, as demonstrated in Figures 8 and 9. CIDS offers military commanders and fighter pilots access to accurate and near real-time BFT and WFT systems. Besides this, CIDS offers commanders a tool which can foster improved mission planning resulting in increased accuracy and tempo of missions. To sum up, CIDS aids commanders to reduce the number of unexpected incidents and minimize collateral damage.

Once the TTP, CID, COP, and SA systems discussed in this paper (cf. Figure 1) are designed, tested and become fully implemented as part of the combat gear, some progress may be discernible in minimizing fratricide and collateral damage. The reality is that for as long as human actors remain part of any decision-making processes, incidents of fratricide and collateral damage are bound to occur. All efforts to minimize the human error factor by improving existing technologies, TTP, CID, COP and SA together with defining explicitly the formulations in ROE, are to be saluted. The efforts to minimize unwanted phenomena are to be applied, for example, in the ongoing series of Bold Quest exercises.

So far, all the decision-making processes in battlespace settings have culminated in a human being making the final decision to apply combat power. In the future, this decision maker's position may be manned by Artificial Intelligence (AI). And, needless to say, detailed planning, testing and implementation are necessary prerequisites for all future warrior systems to be successfully deployed by any robotic militaries of the future.

#### REFERENCES

- [1] T. Saarelainen and J. Jormakka, C4I2-Tools for Future Battlefield Warrior, IEEE Conference on Digital Communications (ICDT), 13 – 19 June 2009, Athens, Greece, pp. 38 – 43.
- [2] K. Chevli, P.Kim, A.Kagel, D. Moy, R. Pattay, R. Nichols, and A. Goldfinger, Blue Force Tracking Network Modeling and Simulation, IEEE Military Communications Conference, MILCOM 2006, pp. 23-25 Oct. 2006, Washington, U.S.A., pp. 1 – 7.
- [3] T. Saarelainen, White Force Tracking (WFT), ECIW2009, Portugal, pp. 216 – 223.
- [4] E.Ospital and A. Wojack, A Holistic Approach to Combat Identification, Military Technology, vol XXXIV, issue 9, 2010, ISSN-3226, pp. 68 –73.
- [5] K. Rein, U. Schade and M. Hieb, Battle Management language (BML) as an enabler, IEEE Conference on Communications, ICC'09, 14 – 18 June, Wachtberg, Germany, 2009, pp. 1 – 5.
- [6] C. Chlestil, E. Leitgeb, N. Schmitt, S. Sheikh Muhammad, K. Zetl, and W. Rehm, Reliable Optical Wireless Links within UAV Swarms, Transparent Optical Networks, 18 – 22 June, 2006, Nottingham, Great-Britain, pp. 39 – 42.
- [7] A. Desai and S. Milner, Autonomous Reconfiguration in Free-Space Optical Sensor Networks, *IEEE Journal on Selected Areas in Communication*, vol 23, issue 8, 2005, pp. 1556 – 1563.
- [8] K. Kazaura, K. Wakamori, M. Matsumoto, T. Higashino, K. Tsukamoto, and S. Komaki, A Proposal for a Broadband Wireless Access Technology based on Radio-onFSO Links, IEEE GLOBECOM Workshops, Nov. 30 – Dec 4, 2008, New Orleans, U.S.A., pp. 1 – 6.
- [9] T. Sierksma, J. Hoekstra, and B. Jansen, Geographical Based Situational Awareness in Military Mobile Domain, IEEE Military Communications Conference, MILCOM2007, 29 – 30 Oct, 2007, Orlando, U.S.A., pp. 1 – 7.
- [10] M. Bahrepour, N. Meratnia, and P. Havinga, Sensor fusion-based Event Detection in Wireless Sensor Networks, Mobile and Ubiquitous Systems, Networking Services, MobiQuitous, 2009, 13 – 16 July, Toronto, Canada, 2009, pp. 1 – 8.
- [11] H. Zheng, J. Shi, and L. Cao, Group-Mobility-Aware Spectrum Management for Future Digital Battlefields, MILCOM 2006, Washington, DC, U.S.A., pp. 1 – 7.
- [12] C-M. Cheng, P-H. Hsiao, H. T. Kung, and D. Vlah, Transmit Antenna Selection Based on Link-layer Channel Probing, World of Wireless, Mobile and Multimedia Networks, 2007, Cambridge, MA, U.S.A.
- [13] D. Herold, L. Griffiths, and T.Y. Fung, Lightweight, High-Bandwidth Conformal Antenna System for Ballistic Helmets, MILCOM 2007, Orlando, FL, U.S.A., pp. 1 – 6.
- [14] G. Jakobson, J. Buford, and L. Lewis, A Framework of cognitive situation modelling and recognition, MILCOM 2006, 23 – 26 Oct. 2006, Washington, DC, pp. 1 – 7.
- [15] O. Kosut, A. Turovsky, J. Sun, M. Ezovski, L. Tong, and G. Whipps, Integrated Mobile and Static Sensing for Target Tracking, MILCOM 2007, 29 – 31 Oct. 2007, Orlando, FL, U.S.A., pp. 1 – 7.
- [16] D. Al-Abri and J. McNair, Improving Localization Accuracy in Wireless Sensor Networks using Location Verification Feedback, MILCOM 2007, 29 – 31, Oct. 2007, Orlando, FL, U.S.A., pp. 1 – 7.
- [17] J. Miller, Coalition Combat Identification Capability Assessments “Bold Quest”, A presentation in the Combat-ID and Situational Awareness Conference, IQPC, London 16.11.2010, [www.sa-combatid.co.uk/presentations](http://www.sa-combatid.co.uk/presentations), accessed on 1.1.2011.
- [18] J. Rantakokko, Towards a Robust Communications System for the Dismounted Soldier through an European Consortium, A presentation in the Combat-ID and Situational Awareness Conference, IQPC, London 16.11.2010, [www.sa-combatid.co.uk/presentations](http://www.sa-combatid.co.uk/presentations), accessed on 2.1.2011.
- [19] R. Buchanan, Intelligence as Fuel Data Requirements for Modern CID Systems, A presentation in Combat-ID and Situational Awareness Conference, IQPC, London 16.11.2010, [www.sa-combatid.co.uk/presentations](http://www.sa-combatid.co.uk/presentations), accessed on 31.1.2011.
- [20] M. Matsuoka, S.M. Rock, and M. G. Bualat, Autonomous Deployment of a Self-Calibrating Pseudolite Array for Mars Rover Navigation, Position Location and Navigation Symposium, PLANS 2004, 26 – 29 April 2004, pp. 733 – 739.
- [21] Bold Quest 2011 Overview, a presentation attached to the Finnish Defence Forces' Letter of Intent concerning the Bold Quest 2011.
- [22] The Finnish Defence Forces' Letter of Intent concerning the Bold Quest 2011.
- [23] T. Saarelainen, Emerging Technologies for Improving Situational Awareness and Preventing Fratricide and Collateral Damage, A presentation in the Combat-ID and Situational Awareness Conference, IQPC, London 16.11.2010, [www.sa-combatid.co.uk/presentations](http://www.sa-combatid.co.uk/presentations), accessed on 1.1.2011.
- [24] T. Saarelainen and J. Jormakka, Computer-Aided Warrior for Future Battlefields, Proceedings of the 8th European Conference on Information Warfare and Security, ECIW2009, Braga, Portugal, 6-7 July 2009, pp. 224 – 233.
- [25] J. Donner, Combat Identification with German Air Force, A presentation in the Combat-ID and Situational Awareness Conference, IQPC, London 16.11.2010, [www.sa-combatid.co.uk/presentations](http://www.sa-combatid.co.uk/presentations), accessed on 1.1.2011.
- [26] A. Bhattacharyya1, C. Mazumdar, and V. Saraswat, A Real-Time Robust Decision Support System for Network Centric Warfare,

Computer-Aided Warrior for Future Battlefields, Proceedings of the  
8th European Conference on Information Warfare and Security,

ECIW2009, Braga, Portugal, 6-7 July 2009, pp. 9 – 19.

# The MediaSense Framework: Ranking Sensors in a Distributed Architecture

Jamie Walters, Patrik Österberg, Theo Kanter  
 Department of Information Technology and Media  
 Mid Sweden University  
 Sundsvall, Sweden  
 {jamie.walters, patrik.osterberg, theo.kanter}@miun.se

**Abstract**—The evolution of mobile applications and services, largely enhanced by their ability to respond to changes in a user's situation is key driver towards an *Internet of Things*. With a size expected to exceed the current Internet, solutions are required for self organisation based on relevance and importance. This, underpinning new applications and services exposed to reliable and evolving context information around a presentity. One such element is the ability to rank heterogeneous and distributed sensors in response to real time interaction between users and the digital ecosystem. In this paper, we consider one approach to calculating sensor ranking based on their general usage patterns. We present it relevant to our work in progress MediaSense framework, showing the ability for it to be implemented without the need for any centralized coordination.

**Keywords**-Sensor Ranking; MediaSense; Context-Awareness; Presence

## I. INTRODUCTION

As with any typical day within an urban environment, people are constantly on the move for business or pleasure. Within such a future cityscape, there exists a digital ecosystem capable of providing enough information in order to derive support for services wishing to affect changes or deliver experiences to a user based on some context. This includes audio-visual devices, internet connections, or a range of sensors such temperature, humidity or even traffic and air quality.

William, a 10 year old child with a mobile phone would be able to connect to, and derive representations of, context from these points in order to support his applications. His mother has a application on her mobile phone which reports on William's current context situation; i.e., is it too warm, too cold, raining, etc. He should therefore be able to connect to the most accurate and reliable sensors reporting the current outdoor temperature in order for his mum whether he is safe or not.

Current solutions implementing context awareness services rely on the availability of fixed information points from which to derive indications of context. William could be connected to the SenSei architecture [1], which enables the deployment of applications and services in response to his context. He would, however be given the nearest sensor, or the sensor connected to his infrastructure in order to derive a temperature value. While the addition of new sensors

would be made available to William, they would not be recommended based on any metric outside of being a part of his domain or infrastructure. If he is in a room, he would like to be assigned the temperature sensor attached to the room.

In a wide and heterogeneous *Internet of Things* [2], there can and will exist multiple sensors in the same locality being offered as sources of context information. William must be able to derive enough information in order to, whether manually or autonomously, select the most suitable sensor for reporting his temperature.

One such element of information is the current ranking, the implied reputation the sensors available to William. An indication of reputation could be derived from the behaviour of the users; continually connecting to and using the sensor as a source of context information. An accurate and reputable sensor, by any measure, would more likely be chosen and used over a sensor that is considered to be inaccurate and usually unavailable. For William, the temperature sensor attached to the building might not be the best sensor, as it might situated close to the heating radiator and reporting a reading that is several degrees higher than the actual temperature in the room. However a sensor temporarily placed in the room would not be made available by [1] which would first require it to be added to profiles and made available, which is not be feasible in dynamic environments. Therefore, there exists, a need to be able to identify changing patterns in user behaviour, such as most users not connecting to the existing sensor but to the alternative, i.e., a sensor ranking approach complementing the sensor proximity approaches mandated by [3].

Current ranking approaches such as Internet search engines consider the theory of connected things, however relative to static document content on the Internet. A document's connectivity determines its relevance. This concept of ranking has been explored and used both in a centralized solutions [4] as well as distributed solutions [5]. However, centralized solutions such as Google index only a tiny portion, less than 10 billion of the estimated 550 billion pages, on the relatively static Internet [6], [5]. Any attempt to apply such a centralized solution to the ranking of sensors in an *Internet of Things* would be undermined by their ability to scale well. Distributed solutions such as [5] which are



Figure 1. The MediaSense Framework

based on the PageRank [7] concept would not scale well to accommodate highly dynamic document sets. Current *real-time* searches are realized by targeting known content providers, an approach that could not scale to accommodate the vast and mostly ad-hoc nature of a connect things infrastructure.

Latencies with respect to scalability could not guarantee William's mother freshness or accuracy with regards to the information being used in her judgement. The MediaSense project in its ongoing work, realises the need to create solutions capable of a real-time distributed ranking algorithm.

## II. THE MEDIASSENSE FRAMEWORK

The MediaSense framework seeks to create an enabling platform for the *Internet of Things*. At its center are presentities [8] and their interaction with the *things* within a digital ecosystem. Supported by an overlay, this solution enables the distributed provisioning of context information reflecting such real world interactions. Figure 1 illustrates the composite layers; each contributing to this realisation. Enabling an interactive environment with respects to a user's state, interactions and preferences. The framework is comprised of the following layers:

1) *Overlay*: The MediaSense Framework is underpinned by a distributed overlay network used to maintain backbone communication as well as providing an indexing mechanism for information that must be persisted among nodes. As with typical peer-to-peer protocol implementations, the nodes participating within the overlay act as entry points for application and services wishing to execute a query over all or a subset of the participating nodes. Early implementations employed a Chord [9] based overlay. This had the inherent drawbacks of distributed hash table (DHT) based implementations with respect to supporting range queries. Citing this and the advantages of P-Grid [10] over DHTs, we have migrated the overlay to a P-Grid based implementation.

2) *DCXP*: Residing immediately on top of the overlay is the Distributed Content Exchange Protocol (DCXP). As

Table I  
THE PRIMITIVES OF THE DISTRIBUTED EXCHANGE PROTOCOL

REGISTER_UCI	Registers a UCI with the overlay making it available for use.
RESOLVE_UCI	Resolves a UCI to the node which is responsible for it.
GET	Retrieves the current context value from the node responsible for a UCI. The reply is sent using a <i>NOTIFY</i> .
SUBSCRIBE	Submits a subscription request to the node responsible for a UCI. The node in turn sends a <i>NOTIFY</i> message containing the current context value, either at regular intervals or when the value changes.
NOTIFY	Dispatches the current context value associated with a specified UCI to an interested node..
TRANSFER	Reassigns the management responsibility of a resource to another node. This might be full or partial responsibility, where the requester re-creates a local copy of the resource permitting improved real time performance.

with the early implementation of our architecture, this layer implements the core protocols employed in the provisioning of context information. However one key departure is that the protocol is no longer utilized in the maintenance of the overlay; this being completely managed by the P-Grid overlay itself. As a result of this, network composition and state is abstracted from the protocol layer. With this key modification, the DCXP primitives have been adjusted, with the new set of primitives listed in Table I.

We introduced a single new primitive, the *TRANSFER* primitive. This provides the ability to relocate context resources to nodes closer to where their demand is greater. This is in an effort to reduce network messaging overhead, and the considerable demands that can be placed on nodes responsible for a context resource. Such an action could be achieved autonomously for load balancing or in response to application requirements.

3) *Persistence*: Persistence is offered at each node in the form of object-oriented databases. The purpose of this is two fold: firstly, it provides a persistence mechanism for context information generated by sensors local to the node. These values are made available in response to DCXP *GET* and *SUBSCRIBE* requests, providing a source of sensor information to the nodes in the overlay. Secondly, it provides a persistence mechanism for the objects created in the *Object Layer* and enable the searching and browsing of context objects over the framework. This enables a single view across all the entire collection of information in the overlay, similarly to a distributed database.

4) *Object Layer*: The object layer exposes all the underlying information as a collection of objects that are accessible and can be used to realise the provisioning of an application or service in response to context information. Objects as made available through the API which permits

the definition of the *object-predicate-object* relationships constituting the Context Information Integration Model (CII) [11].

Objects are accessible across the entire framework by virtue of the underpinning overlay and its ability to *TRANSFER* resources between end-points. Objects may be comprised into schemata relative to presentities, permitting an easy reference to the resources relative to a presentity and thus available to an application or service.

5) *Schema*: We introduced the concept of a *Context Schema* [12], defined as:

*The collection of information points associated with and contributing to a presentity's current context*

where an *Information Point* is defined as:

*Any source providing information about the context of an entity or any sink capable of accepting an input effecting changes to an entity's context*

Within the schema layer, such a schema is attached to a presentity and encapsulates all the information points and the relationships related to a presentity. An application or service with a requirement to deliver some user-context centric experience subscribes to the current schema description; it realises a collection of information objects underpinned by a *publish/subscribe* interface to the end points described by the schema. As a presentity traverses a connected things infrastructure it discovers new entities and consequently updates its schema to reflect this. As a result, all subscribing end points receive an updated schema and can adjust their services to accommodate this. For simplicity, we refer to information points as *sensors* within the remainder of this paper.

6) *API*: The API layer presents itself as a facade to be utilized by application developers in accessing the framework's functionalities. It masks the complexity of lower layers and their interactions, enabling users to focus on developing context objects, applications or services; having them transparently shared across the network with relative ease.

### III. RANKING SENSORS IN AN INTERNET OF THINGS

One problem being addressed by the MediaSense project is the ranking of sensors within a heterogeneous and distributed landscape. While traversing an *Internet of Things*, users will be expected to encounter masses of information sources such as sensors. Such sensors, whether physical or virtual are positioned to the user as points from which to derive dimensions of context which can be used to support a multitude of available applications and services. A supporting solution should therefore be able to provide application developers and users with as much information as possible in order to select the most relevant and recommended sources available. We therefore look at

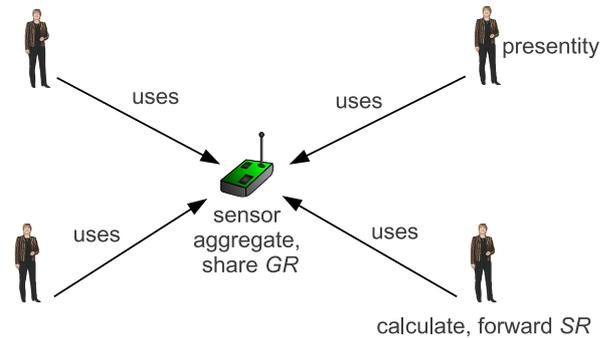


Figure 2. Calculating Sensor Ranking

approaches to adding useful metrics to sensors enabling applications and service providers to rank and select suitable sensors from any groups of sensors encountered.

#### A. Approach

The schema objects described in Section II-5 permit presentities to create collection of sensors, contributing to and expressing the context over a presentity. Using this approach, we are locally aware of all the instances where a sensor  $s$  has been utilized by a presentity  $P$  in schema construction. We also know all the schemata that have been used by this presentity  $P$  and can therefore derive some representation of the importance of  $s$  relative to  $P$ . We argue that such a value, represents the localized ranking of  $s$  and as such the node where  $s$  resides is therefore able to collect and aggregate these values, indicating an overall, near global, ranking value of  $s$ .

#### B. Localized Ranking

Our ranking algorithm consists of two main components, illustrated in Figure 2. Firstly we need to determine the local ranking value for  $s$  with respects to  $P_i$ . We then need to aggregate the global ranking value for  $s$ . In approaching this problem, we adapt a modified version of the *Inverse Document Frequency* algorithm [13]. This algorithm, initially used to calculate the importance of a query term with respect to a document corpus, provided a simple but representative metric for ranking documents with respects to a search query.

The algorithm shown in Equation (1) is modified with respects to a sensor  $s$ , a schema  $r$  and a presentity  $P$ . A sensor in a schema is considered to be analogous to the query term in a document and is expressed as follows:

$$SR_{si}^{(P)} = \log \frac{|R|}{\{r : s_i \in r\}} \quad (1)$$

Where  $SR$  is the sensor ranking of the sensor  $s$ ,  $R$  is the corpus, the total collection of schemata relative to presentity  $P$  with  $r$  being all the schemata relevant to  $P$  containing a reference to sensor  $s$ . This provides us with a representative

metric as to the importance of sensor  $s$  relative to  $P$ . We consider further, that there exists scenarios where some presentities will be less dynamic or mobile with respect to  $s$ . Such an example might be a sensor located in a store; the employees working in the store will by default almost always utilize the sensors that are local to the store accounting for a disproportionately higher value for:

$$\frac{|R|}{\{r : s_i \in r\}} \quad (2)$$

In such scenarios, all stores within a shopping area would have high values, granted solely by the employees themselves. Therefore, by taking:

$$\log \frac{|R|}{\{r : s_i \in r\}} \quad (3)$$

we consider more dynamic presentities traversing an *Internet of Things*. A person that travels around the city interacts with more sensors subsequently creates more context schemas in fulfilment of service delivery. This is represented by larger ratio of  $R$  to  $\{r : s_i \in r\}$ . Such presentities we argue, indicate a more accurate representation of the ranking that should be associated with  $s$ , relative to the wider sensor ecosystem.

Another scenario being that we penalize malicious nodes may that might attempt to collude or independently attempt to inflate their rankings by creating a disproportionate number of schemas. In such scenarios, the result of Equation 2 would move closer to a value of 1. By taking its  $\log$ , instead, we adjust this such that for  $|R| = \{r : s_i \in r\} = 1$ ,  $SR = 0$ . Thereby having a very small effect on the sensor ranking.

### C. Time Limited Localization

While Equation (1) permits the calculation of ranking over the entire interactions of  $P$ , the need will arise to be able to calculate sensor rankings at some given point  $t$  in time. Such a scenario would be useful when trying to rank sensors that are in use at an event or at a situation occurring in a localized area. Here, we could calculate ranking limited by some time duration of interest,  $t$ .

$$SR_{s_i}^{(P)} = \log \frac{|R_t|}{\{r : s_i \in r_t\}} \quad (4)$$

Each presentity now has a value for its ranking of  $s$ , both historically or relative to some interesting duration. This metric permits the presentity to evaluate the usefulness of  $s$ . This being available should it again encounter  $s$  either isolated or as a potential sensor amongst several other sensors.

---

### Algorithm 1 Ranking Sensors and Information Points

---

```

loop
  {at the local node,  $i$ }

  determine the size of the local corpus  $R_i$  of schemata
  for all information points  $s$  attached to  $P_i$  do
    determine the ranking value with respect to  $R_i$ 
    assign this as the local ranking value  $SR_i$ 
    forward this to the global domain owner  $D$ 
  end for
  {at each domain owner,  $j$ }

  for all information points  $s_i$  residing at  $D$  do
    aggregate the all values received for  $SR_{s_k}$  to find  $GR_{s_k}$ 
    calculate the new value for the domain rank  $DR_j$ 
    from all  $GR_s$ 
  end for
end loop

```

---

### D. Global Aggregation

The second component of our approach is a global aggregation of all the local ranking values  $SR$  assigned to  $s$ . To achieve this, we calculate the Global Ranking  $GR$  by finding the sum of all  $SR$  of  $s$  such that:

$$GR_s = \sum_{k=0}^n SR_k^{(P)} \quad (5)$$

This value is continually calculated as new schemas referencing  $s$  are created. We however take into consideration the owner of  $s$ , the presentity or domain where it resides or to which it belongs. This we regard as the domain  $D$  and assign it a value equal to the average ranking of all the sensors belonging to  $D$ . This we call  $DR$ , the ranking of  $D$ . This value is important to us as it would permit us to identify more connected and important spaces such as domains, buildings or just a collection of deployed sensors. We calculate  $DR$  as:

$$DR = \frac{\sum_{k=0}^n SR_k^{(P)}}{k} \quad (6)$$

The resulting values derived above, can be used as indicators or relevancy or importance of sensors in an *Internet of Things*.

### IV. APPROACH ON A DISTRIBUTED ARCHITECTURE

Within a distributed architecture, the implementation of such a sensor ranking approach gains the best implementation with respect to performance and its ability to scale. A presentity would likely not reside on the same node as the sensor that it is trying to use.

Our distributed solution, could be implemented on the MediaSense Framework making use of the schemas and the underlying overlay for messaging support. A node using a sensor in a schema in order to represent the context information for a presentity, calculates the localized ranking value and forwards it to the node responsible for the sensor. The node owning the sensor, aggregates the sensor ranking and makes this available on the overlay to any other interested nodes. This is summarised in Algorithm 1. The benefit of this approach is that sensors calculate values locally with no centralization needed, deriving its scalability properties from the underlying infrastructure.

## V. CONCLUSION AND FUTURE WORK

In an *Internet of Things*, users will require useful metrics in order to understand the digital ecosystem in which they are embedded. Apart from the sensor information themselves, we require metrics that construct an overview of a sensor's ranking. We have shown above that we can adapt existing document ranking algorithms towards solving this problem, providing algorithms that can derive these metrics representative of a sensor's importance in a context centric infrastructure.

This is achievable through additional extension of MediaSense to include a distributed context model and the introduction of context schemas attached to presentities.

We further elaborated on how such a set algorithms can be implemented relative to the distributed approach of the MediaSense framework.

Future work on this area would involve the inclusion of sensor proximity factor with sensor ranking to improve the filtering of undesirable sensors. The reputation of the presentities awarding the ranking values could be taken into consideration as well as the reputation on the nodes calculating the global ranking value.

## ACKNOWLEDGMENT

This research is partially funded by the regional EU target 2 funds, regional public sector, and industry. The authors would like to thank Victor Kardeby, Stefan Forsström, Enrico Saviolli for their feedback and contribution.

## REFERENCES

- [1] M. Presser, P. Barnaghi, M. Eurich, and C. Villalonga, "The SENSEI project: integrating the physical world with the digital world of the network of the future," *Communications Magazine, IEEE*, vol. 47, no. 4, pp. 1–4, Apr. 2009. [Online]. Available: [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=4907403](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=4907403)
- [2] H. Sundmaeker, P. Guillemin, P. Friess, and S. Woelfflé, "Vision and Challenges for Realising the Internet of Things," *Cluster of European Research Projects on the Internet of Things (CERP-IoT)*, no. March, 2010. [Online]. Available: [http://docbox.etsi.org/tispan/Open/IoT/20100601\\_IoT\\_Conference/ALL\\_SLIDES/EU\\_IOT\\_Clusterbook\\_2009.pdf](http://docbox.etsi.org/tispan/Open/IoT/20100601_IoT_Conference/ALL_SLIDES/EU_IOT_Clusterbook_2009.pdf)
- [3] J. Hong and J. Landay, "An infrastructure approach to context-aware computing," *Human-Computer Interaction*, vol. 16, no. 2, pp. 287–303, Dec. 2001. [Online]. Available: <http://portal.acm.org/citation.cfm?id=1463119>
- [4] "Google <http://www.google.com>," p. <http://www.google.com>, 2010. [Online]. Available: <http://www.google.com/>
- [5] Y. Zhu, S. Ye, and X. Li, "Distributed PageRank computation based on iterative aggregation-disaggregation methods," in *Proceedings of the 14th ACM international conference on Information and knowledge management*. New York, New York, USA: ACM, 2005, pp. 578–585. [Online]. Available: <http://portal.acm.org/citation.cfm?id=1099554.1099705>
- [6] J. Li, B. Loo, J. Hellerstein, M. Kaashoek, D. Karger, and R. Morris, "On the feasibility of peer-to-peer web indexing and search," *Peer-to-Peer Systems II*, pp. 207–215, 2003.
- [7] K. Sankaralingam, S. Sethumadhavan, and J. Browne, "Distributed pagerank for p2p systems," in *High Performance Distributed Computing, 2003. Proceedings. 12th IEEE International Symposium on*. IEEE, 2003, pp. 58–68. [Online]. Available: [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=1210016](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1210016)
- [8] H. Christein, *Distributed Communities on the Web*, ser. Lecture Notes in Computer Science, J. Plaice, P. G. Kropf, P. Schulthess, and J. Slonim, Eds. Berlin, Heidelberg: Springer Berlin Heidelberg, Dec. 2002, vol. 2468. [Online]. Available: <http://www.springerlink.com/index/10.1007/3-540-36261-4>
- [9] I. Stoica, R. Morris, D. Karger, M. Kaashoek, and H. Balakrishnan, "Chord: A scalable peer-to-peer lookup service for internet applications," in *Proceedings of the 2001 conference on Applications, technologies, architectures, and protocols for computer communications*, vol. 31, no. 4. ACM, 2001, pp. 149–160. [Online]. Available: <http://portal.acm.org/citation.cfm?id=964723.383071>
- [10] K. Aberer, P. Cudré-Mauroux, A. Datta, Z. Despotovic, M. Hauswirth, M. Puceva, and R. Schmidt, "P-Grid," *ACM SIGMOD Record*, vol. 32, no. 3, p. 29, Sep. 2003. [Online]. Available: <http://portal.acm.org/citation.cfm?doid=945721.945729>
- [11] F. Dobsław, A. Larsson, T. Kanter, and J. Walters, "An Object-Oriented Model in Support of Context-Aware Mobile Applications," *Mobile Wireless Middleware, Operating Systems, and Applications*, pp. 205–220, 2010. [Online]. Available: <http://www.springerlink.com/index/UH0745GR616N7387.pdf>
- [12] J. Walters, T. Kanter, and R. Norling, "Distributed Context Models in Support of Ubiquitous Mobile Awareness Services," in *The 2nd International ICST Conference on Sensor Systems and Software*. Miami: Springerlink, 2010.
- [13] S. Robertson, "Understanding inverse document frequency: on theoretical arguments for IDF," *Journal of Documentation*, vol. 60, no. 5, pp. 503–520, 2004.

# Evaluation of End-to-End Quality of Service over VPN Networks through Various Priority Mechanisms

Nasser-Eddine Rikli  
 Department of Computer Engineering  
 King Saud University  
 Riyadh, Saudi Arabia  
 e-mail: rikli@ksu.edu.sa

Saad Almogari  
 Department of Computer Engineering  
 King Saud University  
 Riyadh, Saudi Arabia  
 e-mail: smogri@nic.org.sa

**Abstract**— VPN networks running over MPLS have found widespread acceptance as both an efficient and cost effective means to provide connectivity for large organizations and companies. However, providing QoS is still a major challenge that needs to be addressed. Using realistic input traffic, a simulation model is built for a large network where various queueing policies are implemented and evaluated for the provision of certain QoS requirements. After a thorough analysis the merits and shortcomings of each policy are determined, and recommendations are given along with future research directions.

**Index Terms**—Virtual private networks; quality of service; multimedia; MPLS; queueing mechanisms.

## I. INTRODUCTION

Quality-of-Service (QoS) over Virtual Private Networks (VPN) is prone to many challenges, among which setting policies for a flexible and scalable support of QoS is of primordial importance [1][2]. Any provider of VPN service should be able to offer customers various Classes of Service (CoS) per VPN [3]. Furthermore, depending on the customer choice and selection, the CoS that a particular application would get within one VPN could be different from the CoS that exactly the same application would get within another VPN. Thus, the set of policies to support QoS should allow the decision to be made on a per-VPN basis.

VPN has used two models in providing QoS, namely the *pipe* model and the *hose* model [4]. In the former, a customer is supplied with certain QoS guarantees for the traffic from one Customer Edge (CE) router to another. While in the latter, a customer is supplied with certain guarantees for the traffic that the customer's CE router sends to and receives from other CE routers over the same VPN.

In [5], a programmable framework for CoS Based Resource Allocation (CBRA) in Multi Protocol Label Switching (MPLS) tunneled VPNs is proposed. The resources are partitioned in a way that facilitates the creation of multiple VPNs on a demand basis.

In [6], the QoS over a VPN IP network is presented from

a service provider point of view. The study includes the provision of QoS guarantees both at the network level and at the node level.

In [7], a CoS classification with associated QoS parameter set for VPNs over an IP WAN is presented. Various scenarios were studied, and it was determined that by policing the aggregate arrival rates of each class from each VPN access interface into the IP network, the appropriate QoS can be guaranteed for each CoS.

The main purpose of this paper is to propose a simulation model and to study the behavior of a VPN network under various queueing mechanisms and for various types of traffic. A thorough network performance analysis will be carried out for various traffic types with different QoS requirements. A special emphasis will be given to the effects of the bandwidth of last mile link at the main site.

The rest of the paper is organized as follows. In Section II, the architecture of the network to be studied will be presented. Then, in Section III the traffic models and traces to be used in the simulation will be described. In Section IV, the queueing models to be used in the various routers will be introduced. The results will be presented in Section V, along with some network specific data. Finally, in Section VI, conclusions will be summarized.

## II. NETWORK ARCHITECTURE MODEL

Based on an existing network, a simulation model for a customer with four sites connected through a VPN service provider (VPN-SP) network was built. The general network architecture is shown in Fig. 1. The network topology of the VPN-SP consists of:

1. Three Provider (P) routers, located at the customer headquarter.
2. One P router and one Provider Edge (PE) router, located at each one of the three satellite locations.
3. Four CE routers: one at site 1, one at site 2, one at site 3, and one at the main site.

The VPN services are assumed to be provided through a

hose model, and most traffic is assumed to pass through the router at the main site (whether it is coming from other sites or passing through towards them).

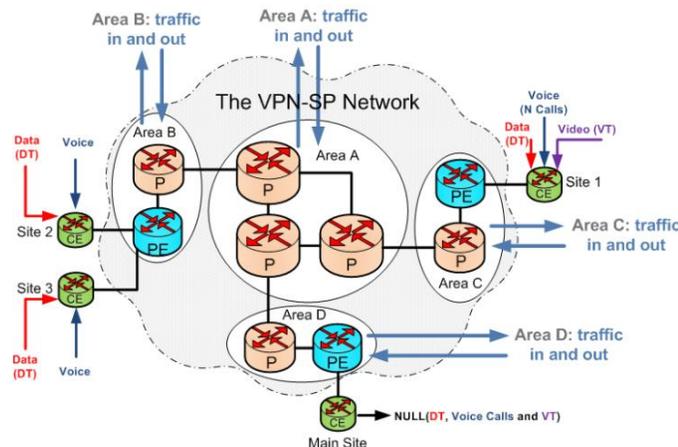


Fig. 1. Network architecture and traffic input locations and types.

The routing protocol used between a CE router and a PE router is the Border Gateway Protocol (BGP). At the PE router, each site connects its customers through an interface that marks all outgoing traffic with a unique VPN label to mark its traffic between PE routers.

Routing table information is exchanged between PE routers using Multiprotocol BGP (MP-BGP). The VPN-SP uses Multiprotocol Label Switching (MPLS) over Open Short Path First (OSPF) network.

### III. TRAFFIC MODELS

#### A. Types

The VPN-SP network carries various types of traffic generated by the different customers. We divided the aggregate traffic into three kinds: voice traffic, video traffic, and data traffic.

Voice traffic is assumed to be generated using a G.729 coder. The aggregate traffic model for VoIP was modeled by an ON-OFF source with Exponential durations. During the ON period, packets of fixed size are generated at fixed time intervals [9].

The two other types of traffic, i.e. MPEG-4 video and data, were captured into trace files from the real traffic flows at the various locations of the actual VPN-SP network using a sniffer tool. These files were used as input at their corresponding locations to simulate real traffic from site-to-site of the chosen customer (or inside the VPN-SP network when coming from other customers).

#### B. Load Distribution

The diagram in Fig. 1 illustrates the distribution of the three types of traffic over the various sites. Voice, video,

and data traffic were sent from site 1 to the main site, while only voice and data traffic were sent from site 2 to main site, and the same thing from site 3 to the main site. Furthermore, each one of the four Areas (A, B, C, and D) has both external input traffic and output traffic leaving the network. It is assumed that all flows include the three types of traffic.

#### C. Requirements

The QoS traffic requirements are shown in Table I. They were chosen to satisfy both generic requirements of the types of application carried over the network, and the specific requirements of the equipment existing on the premises.

TABLE I. TRAFFIC REQUIREMENTS

Criteria	Voice	Video	Data
packet delay (msecs)	< 200	< 250	-
Jitter (msecs)	< 40	< 40	-
packet loss <sup>1</sup> (%)	< 5	< 10	-
packets resent (%)	-	-	< 10

### IV. QUEUEING MODELS

#### A. Description

Various queueing policies may be implemented at the different routers of the considered network. In this study, four types will be considered:

1. *Fair queueing (FQ)*: where the traffic is divided into three flows (video, voice, and data) with separate FIFO queues, and served through a round-robin scheduling (each queue sends one byte in every round).
2. *Priority queueing (PQ)*: similarly packets are classified into three queues but served with priority one for voice traffic, priority two for video traffic, and priority three for data traffic. Within each queue packets are served in FIFO. If a newly arriving packet finds the queue full, then it will be dropped.
3. *Custom queueing (CQ)*: it is similar to PQ in that it also supports a certain classification option. The scheduling, however, is completely different. It uses a round-robin service, in which each queue is allowed to forward a certain number of bytes (not packets). The queues are served in a weighted round-robin scheme. Depending of the weight (% of share) the available bandwidth is distributed among queues. Tail dropping is still used with each individual queue. We study two cases of the custom queueing which are commonly used in real networks: (1) 10% voice, 20% video, and 70% data, and (2) 20% voice, 30% video, and 50% data.
4. *low-latency queueing (LLQ)*: it is a combination of PQ

<sup>1</sup> In here, packet loss includes both the number of dropped packets and delayed packets.

and CQ policies. The first queue has the highest priority, and is still served first. If the first queue is empty then the second and third queues will be served based on a partition of 40% for the second queue and 60% for the third queue. In this study, the first queue was assigned to voice flow, the second to video flow, and the third to data flow.

**B. Placement**

Fig. 2 shows the location of the ports of each router where the proposed queueing mechanisms will be implemented. So, each P router in Area A has three ports, while each P router in the remaining areas (B, C, and D) has only two ports. Also, the PE routers in areas C and D have two ports each, while the PE router in area B has three (since it is connected to two sites). Finally, all CE routers have a single port.

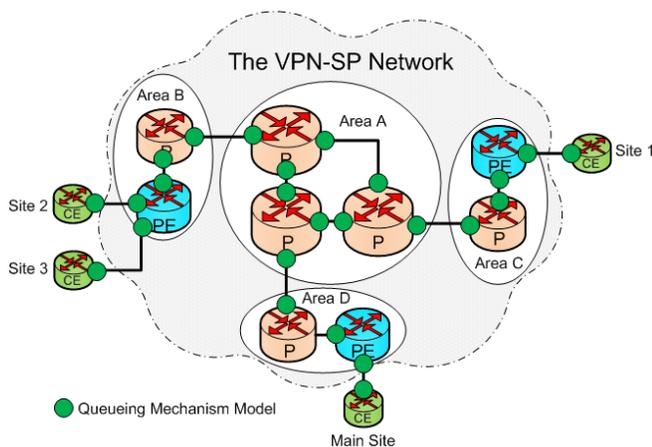


Fig. 2. Points of implementation of the queueing mechanisms.

**V. RESULTS**

To investigate various aspects of the effects of the queueing policy on the performance of our network, two sets of experiments have been designed. Using different queueing mechanisms, five experiment variations were undertaken in each set.

The simulation experiments were built using NS2, and run for one hour of simulation time. All router queues were assumed of finite buffer sizes and had a total size of 512 KBytes (KB) with 128 KB for the first queue, 128 KB for the second, and 256 KB for the third.

The router capacities were 1 Gbps for the core P routers, 10 Mbps for the area P satellite routers, 1 Gbps for the PE routers, and 1 Mbps for all CE routers except the one at the main site which had a 2 Mbps.

**A. Effects of the Number of Channel Calls**

In the first set, the effects of the voice traffic on the VPN-

SP's network was studied by increasing the number of voice calls, initiated from site 1 and going to the main site, from 1 to 7 channel calls. The same experiment was repeated using five different queueing mechanisms. Our focus will be on the traffic flowing from site 1 to the main site, including voice, video, and data.

**1) Effects on Voice Traffic**

Fig. 3 shows the percentage of voice packets dropped due to an excess delay of 200 msec. The best results were obtained when using the PQ and LLQ mechanisms, which have very similar results. This is due to the fact that voice has the highest priority in both schemes.

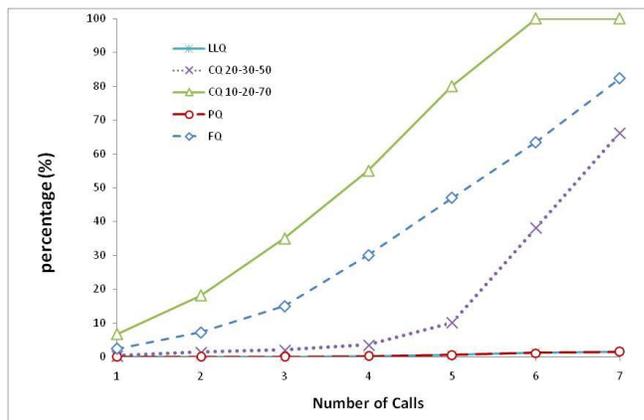


Fig. 3. Percentage of Voice Packets with Delay over 200ms.

The CQ 20-30-50 mechanism was able to handle up to four voice calls dropping rate less than 5%), while the CQ 10-20-70 mechanism barely handled one call. However, in both cases, the results were worst than the ones achieved with PQ and LLQ. This is because not all voice traffic has the highest priority, with an advantage of the 20% scheme over the 10% scheme since a higher share of its traffic was privileged.

Lastly, the FQ mechanism was not able to handle even one call, since there is no priority mechanism implemented. Also, we notice that the performance trend is almost constant with PQ and LLQ mechanisms, while with all other mechanisms it deteriorates rapidly after a certain number of calls.

**2) Effects on Video Traffic**

Fig. 4 shows the dropping rate for video traffic exceeding 250 msec as the voice traffic is increased. The best performance was achieved through the CQ 20-30-50 and FQ mechanisms, with a slight advantage of the latter. As the voice traffic increased, the video performance was kept very close to the required bound.

In the case of the LLQ and PQ mechanisms, the video traffic performance was kept acceptable up to four calls, and

then it deteriorated very quickly.

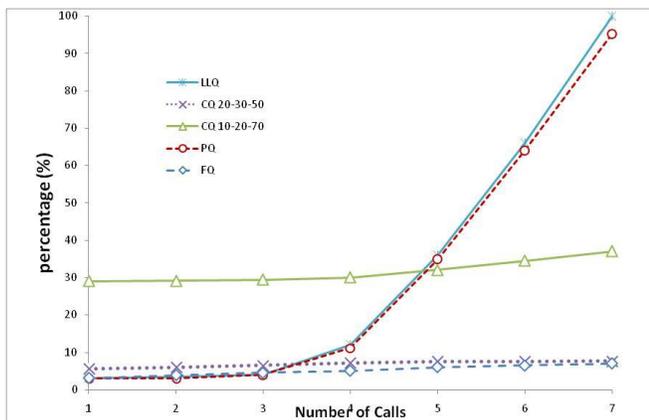


Fig. 4. Percentage of Video Packets with Delay over 250ms.

Lastly, for the CQ 10-20-70 mechanism, although the performance was kept almost constant, it was very far from the required limit.

These results are in concordance with the fact that video traffic has the second priority in the LLQ and PQ mechanisms, where the performance was the best when the first priority traffic (i.e., voice) was comparatively low (< 5 sources). As the first priority traffic was increased, all lower priority traffic suffered. In the case of the other mechanisms, the share of the video traffic was not affected by the increase in voice traffic.

Here also, the CQ mechanisms have better performance than the FQ mechanism, since they use some sort of priority for video. Furthermore, the 30% CQ case performed better than the 20% one, although the voice share also was decreased from 20% to 10%.

### 3) Effects on Data Traffic

Fig. 5 shows the retransmission rate of data traffic as the voice traffic was increased. The best performance was achieved through the CQ 10-20-70, CQ 20-30-50, and FQ, with the former being the best and the latter the worst. In the three cases, the results were kept almost constant, in accordance with non-prioritized mechanism or partially prioritized ones. Here also, the mechanism that allowed 70% of the data traffic to be served as a third priority performed better than the one allowing only 50%.

In the case of LLQ and PQ mechanisms, the performance was kept constant up to three calls, and then increased rapidly. However, while the LLQ performance was acceptable before the three calls knee, the PQ's was unacceptable in all cases. This is similar to the video traffic results, but with a much larger gap in favor of LLQ.

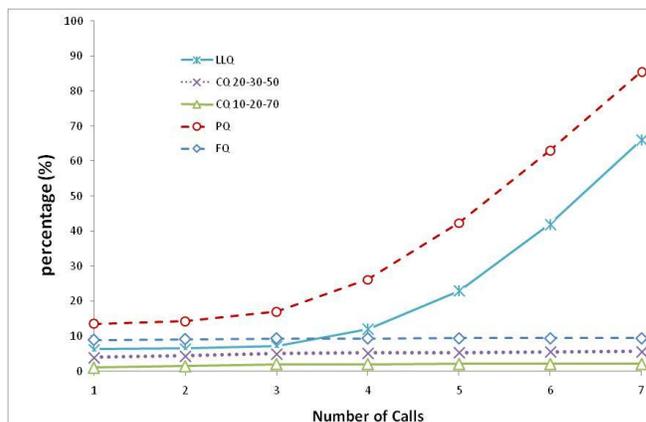


Fig. 5. Percentage of data packets being resent.

## B. Effects of Last-Mile Bandwidth

In the second part of experiments we want to study the effects of the last-mile bandwidth. It is the channel capacity of link connecting the CE router to the PE router at the main site, and it is expected to be the bottleneck for the customer's traffic behavior in the VPN-SP's network.

Its effects will be studied by increasing the capacity of the link from 128 Kbps to 8 Mbps. Here also, the five different queuing mechanisms will be tested, and the performance of the voice, video, and data traffic from site 1 to the main site will be monitored.

### 1) Effects on Voice Traffic

Fig. 6 shows the dropping rate for voice traffic that exceeds a 200 msec delay as a function of the last-mile bandwidth and for the various queuing mechanisms. In all cases the dropping rate decreases as more bandwidth is made available at the bottleneck link. The same relative performances were obtained as in Fig. 3.

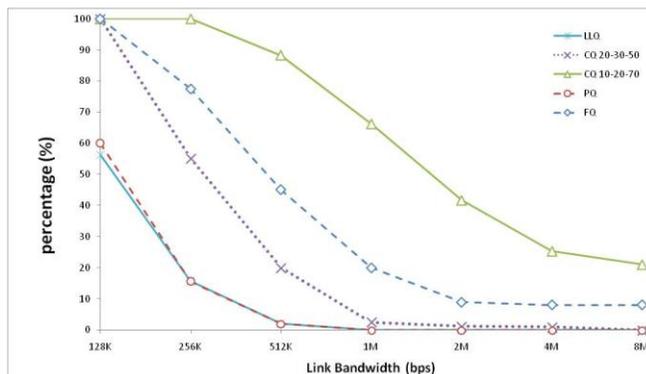


Fig. 6. Percentage of voice packets with delay over 200ms.

The LLQ and PQ mechanisms achieved acceptable performance for bandwidths larger than 512 Kbps, the CQ 20-30-50 mechanism required at least 1 Mbps, while FQ and CQ 10-20-70 failed for all bandwidths.

2) Effects on Video Traffic

Fig. 7 shows the dropping rate of video traffic as a function of the last-mile bandwidth. The CQ 10-20-70 mechanism had a poor performance for all bandwidth values, while the remaining mechanisms had very close performance, with a bandwidth requirement of at least 2 Mbps. The PQ mechanism achieved the best performance for all bandwidths.

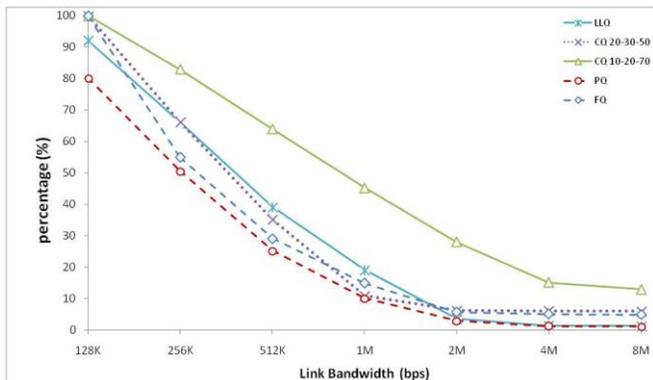


Fig. 7. Percentage of video packets with delay over 250ms.

3) Effects on Data Traffic

Fig. 8 shows the retransmission rate of data traffic as a function of the last-mile bandwidth. The minimum required bandwidth for acceptable data traffic performance were summarized in Table 2. The two CQ mechanisms achieved the best performance, with a noticeable advantage of CQ 10-20-70, which had a larger fraction reserved for data (70%), and this was true for all bandwidth values.

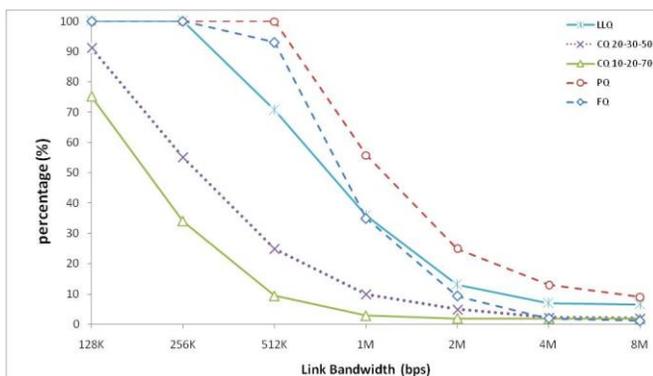


Fig. 8. Percentage of data packets being resent.

TABLE II. MINIMUM BANDWIDTH FOR ACCEPTABLE DATA PACKETS RESENT.

Mechanism	PQ	FQ	LLQ	CQ20-30-50	CQ10-20-70
$BW_{min}$ (Mbps)	8	4	4	1	0.512

The PQ mechanism, which gives data traffic the least priority, achieved the worst performance. With high bandwidths, the FQ mechanism reaches the same level of performance as the CQ mechanisms.

VI. CONCLUSION

In this paper, we have considered a large VPN-SP network providing service to a customer with four remote sites. A simulation model was built with real traffic input, and run under various service policies with the QoS performance being observed.

Four queuing mechanisms were considered, namely: FQ, PQ, CQ (two versions), and LLQ. Criteria for acceptable performance was set for each carried traffic type which was assumed to be carried over the network.

As a result, an estimation of the impact of a new voice call on the performance of the other traffic types being carried over the network was quantified. Consequently, it was possible to determine the limitation on the number of calls in each customer's sites.

Finally, we varied the bandwidth of the last-mile link located at the customer's main site, given that it was considered as the main bottleneck to the traffic being carried. Consequently, it was possible to advise the service provider whether to increase the bandwidth of the last-mile link at the main site if the need for accepting more customers of certain type may arise.

REFERENCES

- [1] M. Rahimi, H. Hashim, and R.A. Rahman , "Implementation of Quality of Service (QoS) in Multi Protocol Label Switching (MPLS) networks," 5th International Colloquium on Signal Processing and Its Applications, pp. 98-103, March 2009.
- [2] M. El Hachimi, M.-A Breton, and M. Bennani, "Efficient QoS Implementation for MPLS VPN," International Conference on Advanced Information Networking and Applications, pp. 259-263, March 2008.
- [3] F. Luyuan, N. Bitu, J.-L. Le Roux, and J. Miles, "Interprovider IP-MPLS services: requirements, implementations, and challenges," IEEE Communications Magazine, vol.43, no.6, pp. 119-128, June 2005.
- [4] N.G. Duffield et al., "Resource management with hoses: point-to-cloud services for virtual private networks," IEEE/ACM Transactions on Networking, vol.10, no.5, pp. 679- 692, Oct 2002.
- [5] P. Kumar, N. Dhanakoti, S. Gopalan, and V. Sridhar, "CoS Based Resource Allocation (CBRA) in VPNs over MPLS", IEEE Workshop on IP Operations and Management, pp. 140-145, 2004.
- [6] J. Zeng and N. Ansari, "Toward IP Virtual Private Network Quality of Service: A Service Provider Perspective", IEEE Communications Magazine, vol.41, no.4, pp. 113-119, April 2003.
- [7] M. Girish, J. Yu, and T. Soon, "A QoS Specification Proposal for IP Virtual Private Networks", IEEE Workshop on IP Operations and Management, pp. 85-90, Oct. 2003.
- [8] C. Metz, "The latest in VPNs: part II," IEEE Internet Computing, vol.8, no.3, pp. 60- 65, May-Jun 2004.
- [9] H. Hassan, J.M. Garcia, and C. Bockstal, "Aggregate Traffic Models for VoIP Applications," IEEE international conference on Digital Telecommunications, pp. 70, Aug. 2006.