INNOV 2013

The Second International Conference on Communications, Computation, Networks and Technologies

ISBN: 978-1-61208-315-5

November 17 - 22, 2013

Lisbon, Portugal

INNOV 2013 Editors

Petre Dini, Concordia University, Canada || China Space Agency Center - Beijing, China

Pascal Lorenz, University of Haute Alsace, France

INNOV 2013

Foreword

The Second International Conference on Communications, Computation, Networks and Technologies (INNOV 2013), held between November 17-22, 2013 in Lisbon, Portugal, continued a series of events addressing recent research results and forecasting challenges on selected topics.

Considering the importance of innovative topics in today's technology-driven society, there is a paradigm shift in classical-by-now approaches, such as networking, communications, resource sharing, collaboration and telecommunications. Recent achievements demand rethinking available technologies and considering the emerging ones.

We take here the opportunity to warmly thank all the members of the INNOV 2013 Technical Program Committee, as well as the numerous reviewers. The creation of such a high quality conference program would not have been possible without their involvement. We also kindly thank all the authors who dedicated much of their time and efforts to contribute to INNOV 2013. 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 INNOV 2013 organizing committee for their help in handling the logistics and for their work to make this professional meeting a success.

We hope that INNOV 2013 was a successful international forum for the exchange of ideas and results between academia and industry and for the promotion of progress in the area of communications, computations, networks and technologies.

We are convinced that the participants found the event useful and communications very open. We hope that Lisbon, Portugal, provided a pleasant environment during the conference and everyone saved some time to enjoy the charm of the city.

INNOV 2013 Chairs:

Petre Dini, Concordia University, Canada || China Space Agency Center - Beijing, China
Pascal Lorenz, University of Haute Alsace, France

INNOV 2013

Committee

INNOV 2013 Technical Program Committee

Omar Alhazmi, Taibah University, Saudi Arabia
Alargam Elrayah Elsayed Ali, University of Khartoum, Sudan
Wan D. Bae, University of Wisconsin-Stout, USA
Henri Basson, University of Lille North of France, France
Michael Bauer, The University of Western Ontario, Canada
Khalid Benali, LORIA - Université de Lorraine, France
Li-Der Chou, National Central University, Taiwan
Morshed U. Chowdhury, Deakin University-Melbourne Campus, Australia
Jacques Demongeot, IMAG/University of Grenoble, France
Uma Maheswari Devi, IBM Research, India
Tarek El-Ghazawi, George Washington University, USA
Mohamed Y. Eltabakh, Computer Science Department - Worcester Polytechnic Institute, USA
Agata Filipowska, Poznan University of Economics, Poland
David A. Gustafson, Kansas State University, USA
Fred Harris, University of Nevada - Reno, USA
Shih-Chang Huang, National Formosa University, Taiwan
Sajid Hussain, Fisk University, Nashville, USA
Tazar Hussain, King Saud University (KSU) - Riyadh, Kingdom of Saudi Arabia
Sergio Ilarri, University of Zaragoza, Spain
Abdessamad Imine, LORIA-INRIA, France
Wassim Jaziri, Taibah University, Saudi Arabia
Miao Jin, University of Louisiana - Lafayette, USA
Khaled Khankan, Taibah University, Saudi Arabia
Igor Kotenko, St. Petersburg Institute for Informatics and Automation, Russia
Raquel Trillo Lado, University of Zaragoza, Spain
Marcela Castro León, Universitat Autònoma de Barcelona, Spain
Jonathan C.L. Liu, University of Florida - Gainesville, USA
Emilio Luque, University Autonoma of Barcelona (UAB), Spain
Xun Luo, Qualcomm Research Center, USA
Leslie Miller, Iowa State University, USA
Maria Mirto, University of Salento - Lecce, Italy
Mena Badieh Habib Morgan, University of Twente, Netherlands
Zahraa Muhsin, Al-Isra University, Jordan
Kok-Leong Ong, Deakin University, Australia
Satish Penmatsa, University of Maryland - Eastern Shore, USA
Ilia Petrov, Reutlingen University, Germany
Bharat Rawal, Loyola University Maryland, USA
Dolores I. Rexachs, University Autonoma of Barcelona (UAB), Spain
Daniel Riesco, National University of San Luis, Argentina

Table of Contents

A Novel Compact Preamble Structure for Timing Synchronization in MIMO-OFDM Systems Using CAZAC Sequences

Ali S. Rachini and Fabienne Nouvel
Institut National des Sciences Appliquées
INSA-IETR
Rennes, France
e-mail: {Ali.Rachini, fabienne.nouvel}@insa-rennes.fr

Ali Beydoun and Bilal Beydoun
PRASE-EDST - Lebanese University
Hadath, Lebanon
e-mail: beydoual@yahoo.fr, bilbey@ul.edu.lb

***Abstract*— To increase the throughput of transmission systems, MIMO-OFDM technology allows for a better reach and rate of transmission and improves the reception. The synchronization between the transmitter and the receiver has become a big challenge. A bad timing synchronization causes the loss of a lot of information in a MIMO-OFDM system. In this paper, we propose a novel compact preamble structure based on orthogonal CAZAC sequences. Simulations results show that the proposed solution presents a good performance at a low SNR in AWGN and multipath fading Rayleigh channels.**

Keywords- MIMO-OFDM system; Timing synchronization; CAZAC sequences; Compact preamble.

I. INTRODUCTION

The world of wireless communications and mobile communication is currently at a very important crossroads in its evolution. This crossroads introduces a variety of challenges such as multi-path signal reflections, and interference. These can reduce the performance of a receiver. To address these challenges, Orthogonal Frequency Division Multiplexing (OFDM) modulation combined with Multiple Input-Multiple Output (MIMO) is proposed in 802.11n [1].

The OFDM [2]-[5] is a transmission technique of distributing symbols on multiple orthogonal carrier frequencies. To transmit a signal, OFDM divides a frequency range into several closely spaced sub-carrier frequencies to carry data thanks to the Fast Fourier Transform's algorithm (FFT). The Inverse Fast Fourier Transformer (IFFT) algorithm is also used to demodulate the message at the receiver.

The main advantage of OFDM system is the orthogonality between subcarriers. The orthogonality of frequencies is primordial to eliminate the Inter Carrier Interference (ICI). The ICI occurs, especially, when the duration of a symbol is short compared to the delay spread involved by the channel. On the other hand, it should be noted that the most important disadvantage of the OFDM is the sensitivity to Doppler shift and frequency synchronization problems.

MIMO technique can be divided into two main categories: Space Time Code (STC) and Space Division Multiplexing (SDM). SDM transmits independent data streams over different transmit antennas. SDM can increase throughput rate of wireless communication links between transmitter and receiver. Foshini et al. [6] and Telatar [7] have shown that the theoretical capacity of the MIMO channel, with N_t transmit antennas and N_r receive antennas, increases linearly with $\min(N_t, N_r)$. MIMO systems are mainly used to increase the flow of wireless communications. The STC increases the performance by sending redundant data over different transmit antennas [8].

In this paper, we will focus on Space-Time Block Code (STBC) [9] [10]. Several applications, based on MIMO technology, are already considered in wireless and local area networks, and 3rd and 4th generation of mobile network. The combination of MIMO-OFDM systems is proposed in the standard of wireless LAN IEEE 802.11n [1], where the objective is to achieve 100 Mbps for video applications.

One of the main challenges for MIMO-OFDM system is the synchronization between transmitter and receiver. Two types of synchronization are necessary, the timing and the frequency synchronization. The frequency synchronization is based on the detection of the frequency offset between the transmitter and the local oscillator at the receiver [11], while the coarse timing synchronization detects the arrival of the OFDM frame and the fine timing synchronization needed to detect the beginning of OFDM symbols.

In the literature, several synchronization approaches have been proposed for OFDM and MIMO-OFDM systems [12]-[16]. The main idea is the use of good synchronization preamble, at the transmitter, in order to detect the packet arrival, at the receiver.

In this paper, we propose a novel compact preamble structure for timing synchronization in MIMO-OFDM systems using Constant Amplitude Zero Auto-Correlation (CAZAC) sequences. The CAZAC sequences [17] are a class of complex-valued sequences with cyclic autocorrelation equal to zero. The main characteristics of CAZAC sequences are their Zero Auto-Correlation; it means that a CAZAC code is always orthogonal with its cyclic shifted versions. Furthermore, they have constant amplitude. The main benefits of CAZAC sequences are the reduction of

 ISBN: 978-1-61208-315-5

Inter-Symbol Interference (ISI), they avoid interferences between multiple antennas, and have a lower Peak-to-Average Power Ratio (PAPR). As a result, CAZAC sequences are regarded as preamble for timing synchronization in MIMO-OFDM systems.

Based on this approach, a compact preamble design for synchronization in distributed MIMO-OFDM Systems has been proposed in [18]. Wang et al. [18] have designed training symbols based on exclusive subband. The main drawback of this approach is that this approach is limited when the number of transmit antennas increases. With a Signal to Noise Ratio (SNR) of 15dB, [18] shows a good timing synchronization. On the other hand, our proposed method at low SNR shows a perfect synchronization against the proposed method in [18].

The aim of this paper is to present a novel timing synchronization approach based on CAZAC sequences. Section II briefly describes the MIMO-OFDM system structure based on STBC code. The proposed method and preamble structure are presented in Section III. Simulation results and conclusion are given in Sections IV and V, respectively.

II. SYSTEM MODEL

Like any telecommunications system, MIMO-OFDM system consists of a transmitter, a channel, and a receiver. Figure 1 presents a MIMO-OFDM system with N_t transmit antennas and N_r receive antennas using M-ary Modulation and N subcarriers per transmit antenna. A parallel data stream is generated by data generator's block. The parallel data is encoded according to a QAM constellation and distributed over different sub carriers with OFDM modulator. The output of the OFDM modulator is fed into STBC encoder in order to encode the data stream with Alamouti encoder [9].

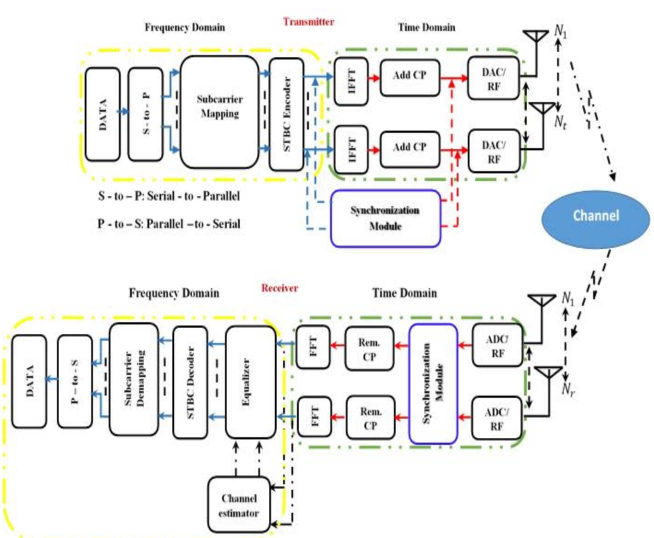

Figure 1 . Transmission system MIMO-OFDM-STBC

Then, we apply the IFFT on these symbols to convert the signal into time domain. A cyclic prefix (CP) is appended to the beginning of each OFDM data symbol, containing a copy of the latest samples.

A synchronization preamble is added at the beginning of each OFDM frame. The preamble may be generated in the time domain [19] or in the frequency domain. In our case, the preamble is generated in the frequency domain. At the receiver, the synchronization is performed in the time domain.

Let $C(l,k)$ denote the transmit data, corresponding to symbol number *l* and carrier number *k*. The OFDM transmit signal $x_i(t)$ from i^{th} transmit antenna, can be expressed as follows:

$$x_i(t) = e^{j(2\pi f_c t + \phi)} \sum_{l=-\infty}^{+\infty} \sum_{k=0}^{K-1} C(l,k)\Psi(l,k,t) \qquad (1)$$

Here:

$$\Psi(l,k,t) = \begin{cases} e^{j2\pi \frac{k-K_c}{T_u}\left(t-T_g-lT_s\right)} & lT_s < t < \left(l+1\right)T_s \\ 0 & t < lT_s, \left(l+1\right)T_s < t \end{cases} \qquad (2)$$

where *k* is the current carrier number, *K* is the total number of sub-carriers, *l* is symbols numbers, T_s is the length of symbol period ($T_s = T_g + T_u$), T_g is the length of guard-interval period, T_u is the length of effective symbol period, f_c is center frequency of RF signal, K_c is the carrier number corresponding to center frequency of RF signal, and ϕ is the carrier phase.

The multipaths fading channel between transmit antenna T_i and receive antenna R_j is given by:

$$h_{i,j}(t) = \sum_{p_{ij}=1}^{P_{ij}} \left[\alpha_{p_{ij}} . \delta_{p_{ij}}\left(t - \tau_{p_{ij}}\right)\right] \qquad (3)$$

where $i = 1, 2, \cdots, N_T$, $j = 1, 2, \cdots, N_R$, P_{ij} is the number of multipaths between the transmit antenna T_i and the receive antenna R_j , and $\alpha_{p_{i,j}}, \delta_{p_{i,j}}$ are respectively the gain and the propagation delay of the path p_{ij} .

At the receiver, a synchronization module is presented in time domain, in order to detect the beginning of frame and then the beginning of useful OFDM symbols. After good timing frame synchronization, a CP removal block is presented to remove the cyclic prefix, because it has no value as information. The FFT is applied to work in the frequency domain. The output of the FFT is fed into a STBC-Alamouti decoder. In order to retrieve the digital data, the received signal is demodulated, using OFDM demodulator.

The transmitted signal from i^{th} transmit antenna undergoes fading by the channel before reaching the j^{th} receive antenna. Let the received signal r_j be:

$$r_j(t) = \sum_{i=1}^{N_t} h_{ij}(t,\tau) * x_i(t)\, e^{-j(2\pi f_c t + \phi)} + n_{ij}(t) \quad (4)$$

$$r_j(t) = \sum_{i=1}^{N_t} \sum_{p_{ij}=1}^{P_{ij}} x_i(t) * \alpha_{p_{ij}} . \delta_{p_{ij}} \left(t - \tau_{p_{ij}}\right) e^{-j(2\pi f_c t + \phi)} + n_{ij}(t) \quad (5)$$

where h_{ij} is the channel between the transmit antenna T_i and the receive antenna R_j, τ is the propagation delay for the different channels paths, x_i is the OFDM transmitted signal, $*$ is the convolution, and n_{ij} is the Additive White Gaussian Noise (AWGN) noise between T_i and R_j.

III. Proposed Method

The proposed frame timing synchronization method aims to estimate the beginning of the OFDM received frame and to detect the beginning of OFDM symbol. Our timing synchronization method relies on sending a preamble structure performed in the frequency domain, as shown in Figure 2.

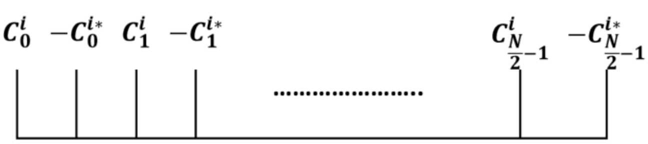

Figure 2. Preamble Structure in Frequency Domain

The combination of a CAZAC sequence with its minus conjugate (-Conj) gives a time-domain complex envelope form that have a good autocorrelation and cross-correlation functions. This combination does not destroy the orthogonality between subcarriers, and it retains the orthogonality between different transmit preambles over different transmit antennas. The preamble structure in time domain is presented in Figure 3.

Different orthogonal preambles are transmitted on each transmit antenna T_i, as shown in Figure 4. Each preamble contains a CAZAC (C) sequence mapped on the odd subcarrier, and the "-Conj" of C is mapped on the even subcarrier. The length of each sequence L_C is:

$$L_C = L_{FFT}/2 \quad (6)$$

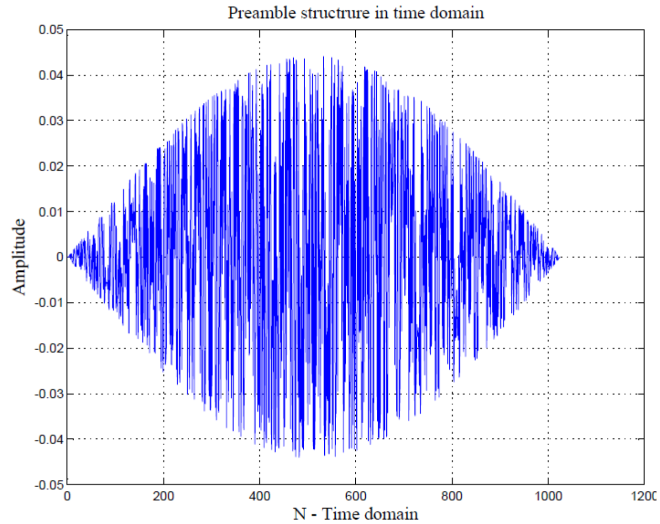

Figure 3. Preamble structure in time domain

Figure 4 represents the different preamble structure over different transmit antennas. The proposed method can be applied regardless of the number of transmit or receive antennas for different Quadrature Amplitude Modulation (M-QAM).

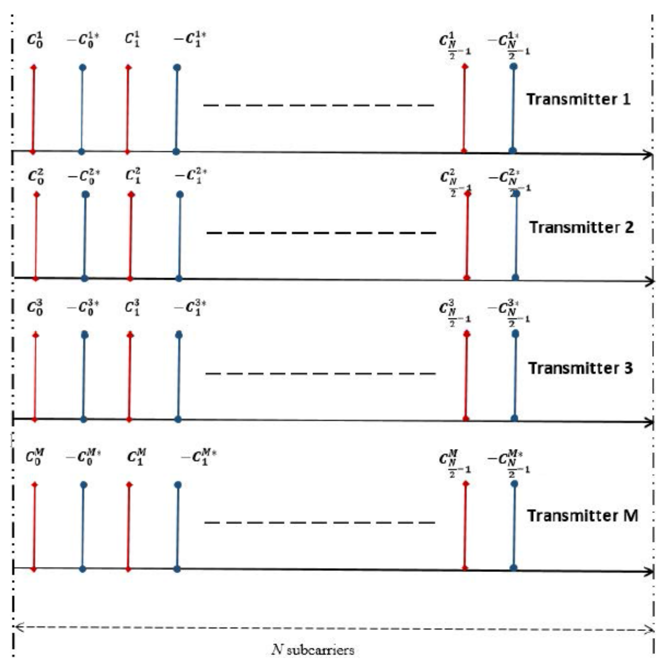

Figure 4. Frame structure in frequency domain

In order to detect the timing synchronization peak, the correlation function $\Re_{r_j seq_j}$ between the received signal r_j

 ISBN: 978-1-61208-315-5

and the local sequence seq_j at the receive antenna R_j is calculated as:

$$\Re_{r_j seq_j}(n) = \sum_{n=1}^{L} \left[r_j(n) * seq_j(n-\tau) \right] \quad (7)$$

where n is the index of the sample, equivalent to the subcarrier index.

IV. SIMULATION RESULTS

The simulations have been done, in both AWGN channel and multipaths fading channel, to evaluate the performance of our proposed preamble against existing preamble. A Single-Input/Single-Output (SISO) and MIMO-OFDM systems up to 8x8 transmit and receive antennas were simulated. An OFDM system with 512 and 1024 subcarriers (L_{FFT}=512, L_{FFT}=1024 resp.) was considered in Rayleigh multipath fading channel with 6 paths sample-spaced with T_S (Sampling Time), suggested by the IEEE 802.11 Working Group [20]. The simulations parameters are summarized in Tables I and II.

TABLE I. SIMULATION PARAMETERS

Simulations parameters	Values
MIMO	up to 8x8
FFT/IFFT Length	1024 & 512
Cyclic Prefix Length (N_C) in samples	$L_{FFT}/4$
Channel Type	Multi-path Rayleigh and AWGN channel
Synchronization sequences	Orthogonal CAZAC
Length of each code L_C	$L_{FFT}/2$
Number of synchronization symbol	1
Number of channel taps between different antennas	6
Signal to Noise Ratio (SNR)	from -5dB to 25dB

TABLE II. THE AVERAGE POWER PROFILE OF THE MULTIPATH RAYLEIGH CHANNEL MODEL

Simulations parameters	Values
Propagation delay between different multipath	[0.Ts, 1.Ts, 2.Ts, 3.Ts, 4.Ts, 5.Ts]
The power of each multipath	[0.8111, 0.1532, 0.0289, 0.0055, 0.0010,0.0002]

At each j^{th} receive antenna a correlation function $\Re_{r_j seq_j}$ (7), in time domain, is applied in order to detect the synchronization point. Due to the distribution of CAZAC sequence C and $-C^*$ in each preamble, the correlation between received signal and the local sequence may give a high peak's value.

By correlating the received signal r_j with a local sequence seq_j at each receive antenna R_j, a timing synchronization estimate ($\hat{ind}_n$) is presented and given by:

$$\hat{ind}_n = \arg\max_n \left\{ \left\| \Re_{r_j seq_j}(n) \right\| \right\} \quad (8)$$

The $\hat{ind}_n$ is the timing estimate where n is considered as the timing synchronization point.

The probability of successful timing frame synchronization P_{SYNC} is evaluated in Figures 5 and 6. The figures showed the acquisition probability for different OFDM systems (SISO-OFDM 1x1, MIMO-OFDM up to 8x8), using CAZAC sequences. The lengths of preamble are 1024 and 512, respectively (L_{FFT}=1024, L_{FFT}=512 resp.).

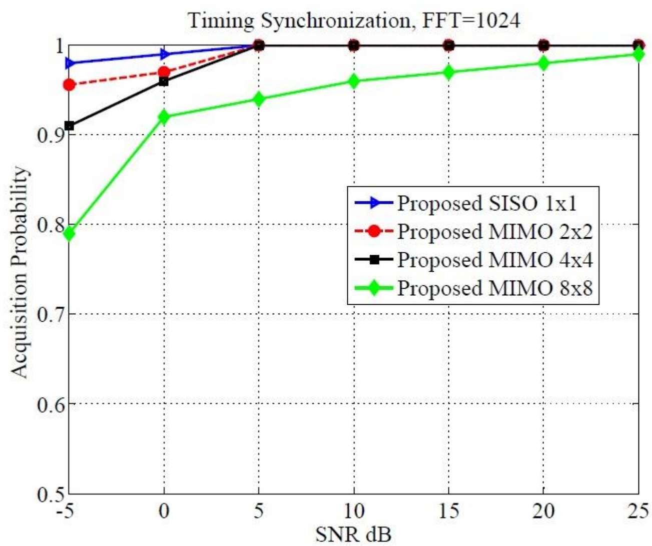

Figure 5. Timing synchronization performance of the proposed method (L_{FFT}=1024)

Figure 5 shows that the system has achieved a synchronization probability over ***95%*** at an SNR=-5dB for both SISO-OFDM and MIMO-OFDM 2x2 systems. For a MIMO-OFDM 4x4 system the P_{SYNC} > 95% at an SNR=0dB. On the other hand, for MIMO-OFDM 8x8 system the acquisition probability P_{SYNC} reaches 90% at an SNR=0dB.

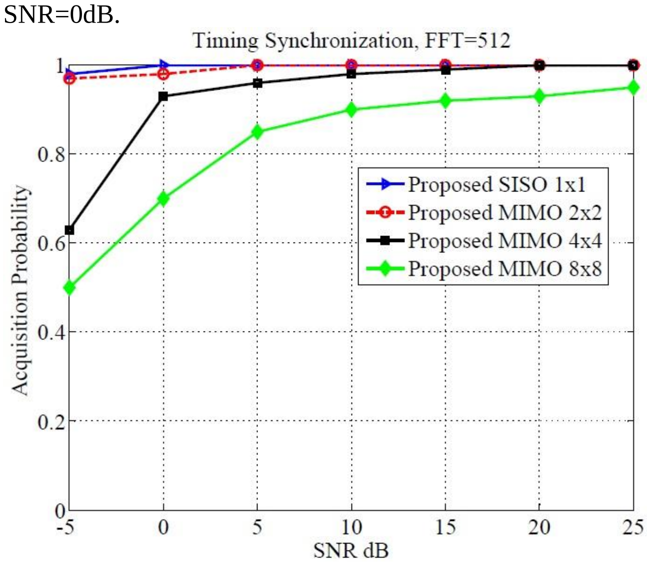

Figure 6. Timing synchronization performance of the proposed method (L_{FFT}=512)

 ISBN: 978-1-61208-315-5

In Figure 6 and for L_{FFT} = 512, it should be mentioned that the acquisition probability P_{SYNC} is greater than 99% for both SISO-OFDM and MIMO-OFDM 2x2 systems at an SNR=-5dB, and P_{SYNC}>90% for MIMO-OFDM 4x4 system at an SNR=0dB. For MIMO-OFDM 8x8 system the P_{SYNC} reaches 80% at an SNR=3dB.

For MIMO-OFDM systems, the limitations of all timing synchronization methods, are the large number of transmit antennas and the length of the synchronization preamble.

The simulation results of our proposed method (Figure 5) show a good performance in terms of timing synchronization acquisition probability for MIMO-OFDM system up to 8x8, where the length of preamble is equal to 1024 (samples). A degradation of performances in term of P_{SYNC} (Figure 6) is observed when the length of preamble is smaller than the length used in Figure 5. The degradations of P_{SYNC}, for a large number of transmit antennas, are due to the length of the synchronization preamble. Otherwise, at the receiver, as the length of the preamble is longer, the value of the correlation peak is higher.

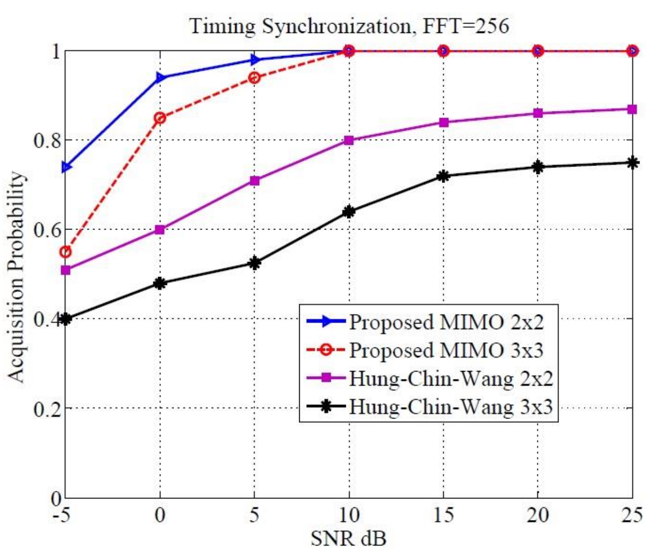

Figure 7. Comparisons between the proposed approach and subband-based preamble [18]

In Figure 7, we compared the performance between our proposed approach and the synchronization scheme of [18], based on a compact preamble design for synchronization, using Zadoff-Chu sequences [17] to generate the training preamble in distributed MIMO-OFDM systems. At a low SNR and by using the same simulation parameters presented in Tables I and II, our proposed method has a good performance against those obtained by [18]. The acquisition probability P_{SYNC} of our method is greater than 90% at an SNR > 3dB for both MIMO-OFDM 2x2 and 3x3 system, while the method proposed by [18] shows that the acquisition probability is between 50% and 75% at the same SNR value.

V. Conclusion

One of the main challenges in a MIMO-OFDM system is the synchronization. We proposed a novel compact preamble structure, in order to detect the timing frame synchronization. At the receiver, and due to the combination of CAZAC sequence C and $-C^*$, the correlation function between received signal and local sequence shows a good frame detection with a large number of transmit antennas. In comparison to the subband-based preamble timing synchronization method proposed by [18], our approach presents a better timing frame synchronization at a low SNR. This approach can be implemented in MIMO-OFDM systems, with a large number of transmit antennas.

Acknowledgment

This work was supported by the GET of Lebanese University and IETR de Rennes-France.

References

[1] "Ieee draft standard for information technology – telecommunications and information exchange between systems - local and metropolitan area networks - specific requirements - part 11: Wireless lan medium access control (mac) and physical layer (phy) specifications," IEEE Draft P802.11-REVmb/D3.0, March 2010 (Revision of IEEE Std 802.11- 2007, as amended by IEEE Std 802.11k-2008, IEEE Std 802.11r- 2008, IEEE Std 802.11y-2008, IEEE Std 802.11w-2009 and IEEE Std 802.11n-2009), 11 2010, pp. 1 –2228.

[2] R. Prasad, OFDM for Wireless Communications Systems. Norwood, MA, USA: Artech House, Inc., 2004.

[3] S. Weinstein and P. Ebert, "Data transmission by frequency-division multiplexing using the discrete fourier transform," Communication Technology, IEEE Transactions on, vol. 19, no. 5, 1971, pp. 628–634.

[4] J. Heiskala and J. Terry, Ph.D., OFDM Wireless LANs: A Theoretical and Practical Guide. Indianapolis, IN, USA: Sams, 2001.

[5] R. v. Nee and R. Prasad, OFDM for Wireless Multimedia Communications, 1st ed. Norwood, MA, USA: Artech House, Inc., 2000.

[6] G. J. Foschini and M. J. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," Wireless Personal Communications, vol. 6, 1998, pp. 311–335.

[7] I. E. Telatar, "Capacity of multi-antenna gaussian channels," European transactions on telecommunications, vol. 10, 1999, pp. 585–595.

[8] P. Wolniansky, G. Foschini, G. Golden, and R. Valenzuela, "V-blast: an architecture for realizing very high data rates over the rich-scattering wireless channel," in Signals, Systems, and Electronics, 1998. ISSSE 98. 1998 URSI International Symposium on, 1998, pp. 295–300.

[9] S. Alamouti, "A simple transmit diversity technique for wireless communications," vol. 16, no. 8, Oct. 1998, pp. 1451 –1458.

[10] V. Tarokh, H. Jafarkhani, and A. Calderbank, "Space-time block codes from orthogonal designs," Information Theory, IEEE Transactions on, vol. 45, no. 5, 1999, pp. 1456–1467.

[11] V. Tarokh, A. Naguib, N. Seshadri, and A. Calderbank, "Space-time codes for high data rate wireless communication: performance criteria in the presence of channel estimation errors, mobility, and multiple paths," Communications, IEEE Transactions on, vol. 47, no. 2, Feb 1999, pp. 199–207.

[12] T. Schmidl and D. Cox, "Robust frequency and timing synchronization for ofdm," Communications, IEEE Transactions on, vol. 45, no. 12, Dec 1997, pp. 1613 –1621.

[13] U. Rohrs and L. Linde, “Some unique properties and applications of perfect squares minimum phase cazac sequences,” in Communications and Signal Processing, 1992. COMSIG ’92., Proceedings of the 1992 South African Symposium on, Sep 1992, pp. 155 –160.

[14] B. Yang, K. Letaief, R. Cheng, and Z. Cao, “Burst frame synchronization for ofdm transmission in multipath fading links,” in Vehicular Technology Conference, 1999. VTC 1999 - Fall. IEEE VTS 50th, vol. 1, 1999, pp. 300 –304 vol.1.

[15] J.-J. van de Beek, M. Sandell, M. Isaksson, and P. Ola Borjesson, “Low- complex frame synchronization in ofdm systems,” in Universal Personal Communications. 1995. Record., 1995 Fourth IEEE International Conference on, Nov 1995, pp. 982 –986.

[16] L. Li and P. Zhou, “Synchronization for b3g mimo ofdm in dl initial acquisition by cazac sequence,” in Communications, Circuits and Systems Proceedings, 2006 International Conference on, vol. 2, June 2006, pp. 1035 –1039.

[17] R. Frank, S. Zadoff, and R. Heimiller, “Phase shift pulse codes with good periodic correlation properties (corresp.),” Information Theory, IRE Transactions on, vol. 8, no. 6, Oct 1962, pp. 381 –382.

[18] H.-C. Wang and C.-L. Wang, “A compact preamble design for synchronization in distributed mimo ofdm systems,” in Vehicular Technology Conference (VTC Fall), 2011 IEEE, Sep 2011, pp. 1 –4.

[19] A. Rachini, A. Beydoun, F. Nouvel, and B. Beydoun. Timing synchronization method for mimo-ofdm system using orthogonal preamble. In Telecommunications (ICT), 2012 19th International Conference on, April 2012, pages 1–5.

[20] B. O’Hara and A. Petrick, The IEEE 802.11 Handbook: A Designer’s Companion. Standards Information Network IEEE Press, 1999. [Online]. Available: http://portal.acm.org/citation.cfm?id=553211

 ISBN: 978-1-61208-315-5

Highlights on a Study of Bicriteria Routing Methods with Protection for WDM Networks with Dynamic Traffic

Carlos Simões
Polytechnic Institute of Viseu/
INESC Coimbra
Viseu, Portugal
e-mail: csimoes@ipv.pt

José Craveirinha and Teresa Gomes
Dept. of Electrical and Computer Engineering
University of Coimbra / INESC Coimbra
Coimbra, Portugal
e-mail: jcrav@deec.uc.pt; teresa@deec.uc.pt

***Abstract*—Dedicated path protection is a robust and efficient way to increase the resiliency of a telecommunication network. Two variants of a bi-criteria dedicated path protection model for wavelength division multiplexing (WDM) networks in a dynamic traffic scenario, are presented. The proposed routing approach is based on a bi-criteria optimization model for the calculation of disjoint lightpath pairs in WDM networks with dedicated protection. A performance analysis study of the proposed routing method considering two different metric pairs as objective functions and using an exact resolution approach, in a reference test network, is highlighted. The presented simulation results show the potential advantages of this approach when compared with single criteria optimization approaches.**

Keywords-optical networks; multicriteria optimization; routing; protection.

I. INTRODUCTION

Given the enormous transfer rate available in the links of a WDM network, a single failure potentially results in huge data losses and service disruptions to many customers. To circumscribe this potential problem and ensure a high degree of availability, network operators adopt protection and/or restoration schemes [1]. In protection mechanisms, an optical connection (or *lightpath*) can be protected against failures by a precomputed backup route. Alternatively, in restoration schemes, the necessary resources to recover a failed connection are only determined and reserved after a failure occurs. Usually, restoration mechanisms are more resource efficient since they do not reserve resources in advance, and provide recovery against different types of failures (including multiple failures), but they need more time for resource determination and reservation, and the recovery will only be successful if there are sufficient resources in the network. Conversely, protection schemes have faster recovery times and can insure resource availability in the failure scenarios for which they have been designed, at the expense of more resources.

In path protection against any single failure, it is necessary to calculate a pair of disjoint paths - the *Active Path* (AP) to be used in normal conditions and the *Backup Path* (BP) to be used if a failure occurs. With dedicated path protection, both the resources of the AP and the BP are assigned to the connection and will only be released when connection ends.

The problem of finding two link disjoint paths can be solved in polynomial time using *Suurballe*'s algorithm [2].

Optical WDM networks consist of optical fiber links, interconnected by optical nodes (routers). Each fiber link comprises several channels, with different wavelengths, and each optical node supports wavelength based switching or routing. The configuration of these optical nodes allows the establishment of point-to-point optical paths (lightpaths), between node pairs. A lightpath may span several links and consist of a wavelength channel in each of these links, interconnected at the nodes by means of optical routing. In order to set up a lightpath, the network needs to decide on the topological path and the wavelength(s) assigned to the lightpath - the so-called Routing and Wavelength Assignment problem (RWA). In the absence of wavelength converters at the optical nodes, a lightpath must use the same wavelength on all the links of its route (the wavelength continuity constraint), but wavelengths can be reused by different lightpaths in the network, as long as they do not share any fiber link.

In a dynamic traffic scenario (Dynamic Lightpath Establishment problem – DLE), a lightpath is set up for each connection request as it arrives, and is released after some finite amount of time. The objective is usually to choose routes and assign wavelengths in a manner that minimizes the overall blocking. To overcome the high complexity of this problem, the routing sub-problem and the wavelength assignment sub-problem are solved separately.

Concerning dedicated path protection in WDM networks (Survivable Routing and Wavelength Assignment – SRWA), the problem of finding a pair of disjoint paths under the wavelength continuity constraint is *NP*-complete [3].

Even though, in general, the routing methods try to optimise only one metric, typically using some variant of a shortest path algorithm, optical networks are characterised in terms of performance by multiple metrics. Furthermore, the design of real networks involves multiple, often conflicting objectives and constraints.

In this context, it seems potentially advantageous to develop multicriteria models that explicitly represent the different performance objectives to be optimised, enabling to treat in a consistent manner the trade offs among these objectives. In this type of models, the concept of optimal solution is replaced by the concept of non-dominated solution. A feasible solution is said to be non-dominated if it is not possible to find any other

 ISBN: 978-1-61208-315-5

feasible solution which enables one criteria to be improved without sacrificing, at least, one of the remaining criteria.

Reference [4] presents a state-of-art review on multicriteria approaches in communication networks and includes a section dedicated to routing models. A survey of multicriteria routing models can be seen in [5].

In this paper, we focus on the problem of dedicated path protection against link and node failures with dynamic traffic, and present a performance analysis study of two variants of a bicriteria model for the route calculation problem using an exact resolution approach. The first variant of the model was initially proposed in [6], but in a context of incremental traffic.

The two major contributions of this work in progress will be: *i*) the consideration of specific variants of a bi-criteria model for dynamic traffic instead of incremental traffic (which is a fundamental change in the underlying assumptions of the model) using an exact resolution method; *ii*) an extensive performance comparison of the bi-criteria model variants with the associated single-criterion models using reference test networks. The analysis in this paper considers relevant network performance measures, namely, the global blocking probability, the carried traffic and percentage of used bandwidth. A dynamic traffic model in a benchmark network will be considered.

The remaining of this paper is organised as follows. In Section II, the bicriteria model is presented, the metrics are introduced and two versions of the bicriteria model are considered. Section III identifies the relevant parameters for performance assessment of the model, presents the simulation results and a comparative study of the models. Finally, some conclusions are drawn in Section IV.

II. Model Description

In this section, we describe the features of the proposed bicriteria routing model associated with the Dynamic Lightpath Establishment problem (DLE) with dynamic traffic in WDM networks. The model was developed for application in large WDM networks, with multiple wavelengths per fiber and multifibers per link. Network reliability is improved through dedicated path protection, so a node disjoint path pair must be obtained for each connection request. In order to cover a wide variety of networks, different types of nodes are considered (with complete wavelength conversion capability, limited range conversion or no wavelength conversion capability) in the model. Due to the real-time nature of the intended application, solutions should be obtained in a short time. This requirement lead to the separation of the routing and wavelength assignment problems, also having in mind an automatic selection of the solution (among the non-dominated solutions, previously identified). The wavelength assignment problem is solved separately, after the bicriteria routing problem.

Let the WDM network be represented by $R = \{N, L\}$ where N is the set of nodes and L is the set of links. A *topological path*, p, in the network is defined by: a source node s, a destination node t and an ordered sequence of nodes and links from s to t. In addition to the ordered sequence of nodes and links, a *lightpath* p^λ also comprises the fiber used in each link and the wavelength in each of the fibers.

A. Bicriteria Model

In a first variant of the model (*BiC_A Model*), we consider two additive metrics for the active and the protection paths as objective functions (*o.f.*). The first *o.f.* is the sum of the inverse of the free bandwidth in the links of each path:

$$\min_{z\in D}\left\{c_1(z) = \sum_{l\in z}\frac{1}{b_l^T} = \sum_{l\in p}\frac{1}{b_l^T} + \sum_{l\in q}\frac{1}{b_l^T}\right\} \quad (1)$$

where D represents the set of topological path pairs $z \equiv (p, q)$ for the origin–destination node pair (s, t) such that p and q are node disjoint and correspond to viable optical lightpaths (taking into account the availability of wavelengths in the links and the wavelength conversion capabilities of the nodes), and b_l^T is the total available capacity in link l, in terms of free wavelengths. This criterion aims to avoid more congested links, promoting a balanced distribution of the traffic throughout the network, hence decreasing the blocking probability and increasing the expected revenue.

The second objective consists of minimizing the number of links of the two paths (hop count), and is intended to avoid bandwidth waste, hence favouring global efficiency in the use of network resources:

$$\min_{z\in D}\{c_2(z) = h(p) + h(q)\} \quad (2)$$

In many cases, whenever the objective functions are conflicting, there is no feasible solution which optimises the two *o.f.* simultaneously. The algorithmic approach described in section II-B is used to obtain a good compromise solution among the non-dominated solutions of the bicriteria problem:

$$(\mathcal{P}_A) \quad \begin{cases} \min_{z\in D}\left\{c_1(z) = \sum_{l\in p}\frac{1}{b_l^T} + \sum_{l\in q}\frac{1}{b_l^T}\right\} \\ \min_{z\in D}\{c_2(z) = h(p) + h(q)\} \end{cases} \quad (3)$$

Another version of the bicriteria model (*BiC_B Model*) was also implemented. As in the previous model, the second *o.f.* is the hop count, but the first one is a load cost function suggested in [7] for IP/MPLS networks. This load cost function is an additive metric where the cost of a link is modeled through a piecewise linear increasing and convex function, the value of which is inversely proportional to the free capacity. The main goal of this metric is to avoid congestion, i.e., overloading of links, and obtain a balanced distribution of the traffic throughout the network.

The *load cost*, LC_l, of a link l is defined by [7]:

$$LC_l = \begin{cases} \rho_l & \text{if} \quad 0 < \rho_l < \frac{1}{3} \\ 3\rho_l - \frac{2}{3} & \text{if} \quad \frac{1}{3} \le \rho_l < \frac{2}{3} \\ 10\rho_l - \frac{16}{3} & \text{if} \quad \frac{2}{3} \le \rho_l < \frac{9}{10} \\ 70\rho_l - \frac{178}{3} & \text{if} \quad \frac{9}{10} \le \rho_l < 1 \end{cases} \quad (4)$$

where ρ_l denotes the relative occupancy of the link (the ratio between the occupied bandwidth and the total bandwidth on

 ISBN: 978-1-61208-315-5

the link). The second bicriteria problem, considering dedicated path protection, is now:

$$(\mathcal{P}_B) \quad \begin{cases} \min_{z\in D} \left\{c_1(z) = \sum_{l\in p} LC_l + \sum_{l\in q} LC_l\right\} \\ \min_{z\in D} \left\{c_2(z) = h(p) + h(q)\right\} \end{cases} \tag{5}$$

B. Resolution Approach

Reference [8] proposes an exact algorithm for finding non-dominated shortest pairs of disjoint paths that uses the *k*-shortest path algorithm MPS [9]. Given an origin-destination node pair (s,t), the algorithm starts by making a network topology modification, where all nodes and links of the graph, (N,L), representing the network topology are duplicated and a new link, with null cost, is added between node t and node s' (the duplicate of s). In this new modified graph, (N',L'), each path from s to t' (the duplicate of the destination node t) will correspond to a pair of paths from s to t in the original topology [8]. Finally, the adapted version of MPS is used for ranking by non-decreasing order of cost the pairs of paths $z \equiv (p,q)$, such that p and q are node disjoint. In order to select the best compromise solution, priority regions are defined in the objective function space [10]. The final solution is chosen among the non dominated solutions in the highest priority region, by using a weighted Chebyshev distance to a reference point (bottom left corner of the region). Further details of the topological path pair selection can be seen in [6].

C. Wavelength selection

Having selected a node disjoint path pair, we use the maximization of the *wavelength bottleneck bandwidth*, which corresponds to the choice of the Least Loaded wavelength (LL) along the links of each of paths in the pair, to make the wavelength and fiber selection. Further details and an illustrative example of this selection heuristic can be seen in [11]. Note that, if all the nodes of the network enable full wavelength conversion, once a viable topological path is chosen, the choice of the wavelength(s) to be used is irrelevant in terms of network performance. If the nodes have no conversion capability, the proposed criterion of wavelength selection is known in the literature (see, e.g., [12]) to give good results. In any case, it is also known that in these cases the critical factor in terms of network performance is the selection of topological paths, the choice of wavelength having a minor impact [12].

III. Performance Analysis Study

A. Network Performance parameters

Highlights of the results obtained using the two variants of the bicriteria model, *BiC_A* corresponding to $\mathcal{P}_A$ (3) and *BiC_B* corresponding to $\mathcal{P}_B$ (5), will be analysed and compared with the results obtained using single objective formulations, namely, the shortest path in terms of either the number of hops *SP_hop* (2), the inverse of the free bandwidth *SP_BW* (1), or the load cost *SP_LC* ($c_1(z)$ in (5)), by recurring to relevant network global performance metrics. The considered network metrics are: the global blocking probability (fraction of connection requests that are blocked); the carried traffic; the percentage of used bandwidth in the network and the point-to-point blocking probabilities.

B. Comparative Study

This section presents the simulation results in the NSFNET network [13], with 14 optical nodes and 21 multifiber links (with 16 wavelengths per fiber). The results were obtained from discrete events stochastic simulations, with Poisson traffic generated with intensities proportional to the traffic matrix presented in [13], which takes into account the distance and population of the nodes. Concerning the wavelength conversion capabilities, simulations were conducted considering three different scenarios: all nodes without conversion capability, five nodes with total conversion capability, (central nodes were chosen with this capability), and total conversion capability in all the nodes.

Figure 1 shows the global blocking probability and the percentage of used bandwidth for several values of the offered traffic, with no conversion capability in the network. Clearly, the single criterion routing model based on hops (*SP_hop*) has the worst performance, which confirms that choosing the path pair based only on the hop count is a poor strategy, in this type of networks. Model *BiC_A* presents the lower blocking, and hence more carried traffic and expected revenue. Thus, *BiC_A* outperforms SP_hop and SP_BW, the corresponding single objective formulations. Concerning *BiC_B*, its blocking is quite similar to that presented by *SP_LC*. However, *BiC_B* uses slightly less bandwidth (see Figure 1).

Although not shown here, the results with total wavelength conversion in the network or with 5 nodes with conversion exhibit the same behavior pattern: *BiC_B* and *SP_LC* have similar performance, which is better than *SP_BW*, but worse than *BIC_A*. Again, the single objective *SP_hop* exhibits the worst performance. As expected, it turns out that the presence of wavelength conversion enables a slight improvement in the global blocking probability. Furthermore, much more important than this improvement in global blocking, an additional advantage provided by the conversion capability is a substantial improvement in QoS fairness. Indeed, comparing the point-to-point blocking probabilities in both cases, it is noted that, for 1000 *Erlangs* of offered traffic, *BIC_A* has mean blocking probability of 0.63% without wavelength conversion and 0.53% with total conversion, but the maximum point-to-point blocking probability (usually observed between farther away nodes) is 2.44% in the former case and 1.42% with total conversion. For higher values of the offered traffic, without wavelength conversion, the QoS among node pairs is even more unbalanced. For example, with *BIC_A* and 1050 *Erlangs* the global blocking probability is 3.16% without conversion and 2.67% with total conversion, but there are pairs of nodes for which blocking is 11.8% without conversion and 6.21%

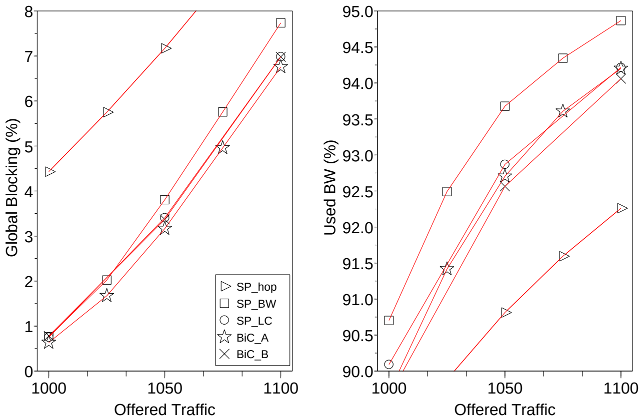

Figure 1. Global Blocking and Used Bandwidth.

with total conversion. The maximal point-to-point blocking is even worse in single criterion models.

IV. Conclusion and Further Study

A multicriteria approach of the type proposed can explicitly represent the different performance objectives, enabling to address, in a mathematically consistent manner, the trade offs among the various criteria. Two variants of a bicriteria model for obtaining a topological path pair for every lightpath request in a WDM network with dynamic traffic were presented. The proposed resolution approach uses an exact bicriteria shortest path pair algorithm and an automated solution selection procedure. Having obtained a non-dominated topological path pair, a heuristic procedure was then used to assign wavelengths to the links.

The performance of two variants of the bicriteria model was analysed in comparison with the corresponding single objective approaches. The model *BiC_B* did not show a significant advantage in comparison with *SP_LC*. But the *BiC_A* exhibits better performance than any of the single objective models. Furthermore, in conjunction with wavelength conversion, *BiC_A* also allows a higher fairness regarding the point-to-point blocking probabilities.

Further experimental studies of similar type, with other reference test networks, will be carried out.

It is important to note that, from a methodological point of view, the used modeling approach, as indeed any other shortest path based approach, is a flow oriented optimization approach. That is, the path pairs are calculated for each node-to-node connection request at a time, so that the objective functions do not integrate explicitly the combined effect of all possible connections offered to the network. This is an inherent limitation of this type of approaches, as analysed in depth in [14] in the context of multicriteria routing models for MPLS networks. This is an additional reason why it is so important to make extensive performance analysis studies with flow oriented models, using relevant network performance global metrics as in the present work. A more in depth analysis of these issues in the context of WDM networks, both at methodological and experimental level (considering different objective functions and other networks), deserves further research.

Acknowledgment

This work has been partially supported by programme COMPETE of the EC Community Support Framework III and cosponsored by the EC fund FEDER, by Fundação para a Ciência e a Tecnologia (FCT) under project grants PTDC/EEA-TEL/101884/2008 and PEst-C/EEI/UI0308/2011, by Project CENTRO-07-0224-FEDER-002003, by FLAD 2013/CON29/CAN9 - PROJ 277/12 and by the PhD scholarship SFRH/BD/49739/2009, granted by Instituto Politécnico de Viseu.

References

[1] J. Zhang and B. Mukherjee, "A review of fault management in WDM mesh networks: Basic concepts and research challenges," IEEE Network, vol. 18, no. 2, March/April 2004, pp. 41–48.
[2] J. W. Suurballe and R. E. Tarjan, "A quick method for finding shortest pairs of disjoint paths," Networks, vol. 14, no. 2, 1984, pp. 325–336.
[3] S. Yuan and J. P. Jue, "Dynamic lightpath protection in WDM mesh networks under wavelength-continuity and risk-disjoint constraints," Computer Networks, vol. 48, no. 2, Jun. 2005, pp. 91–112.
[4] J. Clímaco and J. Craveirinha, *Multiple Criteria Decision Analysis: State of the Art Surveys*, ser. International Series in Operations Research & Management Science. New York: Springer Science, 2005, vol. 78, ch. Multicriteria Analysis in Telecommunication Planning and Design - Problems and Issues, pp. 899–951.
[5] J. C. N. Clímaco, J. M. F. Craveirinha, and M. M. B. Pascoal, *Advances in Multiple Criteria Decision Making and Human Systems Management*. IOS Press, 2007, ch. Multicriteria Routing Models in Telecommunication Networks - Overview and a Case Study, pp. 17–46.
[6] C. Simões, T. Gomes, J. Craveirinha, and J. Clímaco, "Performance analysis of a bi-objective model for routing with protection in WDM networks," Journal of Telecommunications and Information Technology, vol. 2010/3, Oct. 2010, pp. 25–35.
[7] B. Fortz and M. Thorup, "Optimizing OSPF/IS-IS weights in a changing world," IEEE Journal on Selected Areas in Communications, vol. 20, no. 4, May 2002, pp. 756–767.
[8] J. C. N. Clímaco and M. M. B. Pascoal, "Finding non-dominated bicriteria shortest pairs of disjoint simple paths," Computers & Operations Research, vol. 36, no. 11, Nov. 2009, pp. 2892–2898.
[9] E. Martins, M. Pascoal, and J. Santos, "Deviation algorithms for ranking shortest paths," International Journal of Foundations of Computer Science, vol. 10, no. 3, 1999, pp. 247–263.
[10] L. Martins, J. Craveirinha, J. Clímaco, and T. Gomes, "On a bi-dimensional dynamic alternative routing method," European Journal of Operational Research, Advances in Complex Systems Modeling, vol. 166, no. 3, November 2005, pp. 828–842.
[11] T. Gomes, J. Craveirinha, J. Clímaco, and C. Simões, "A bicriteria routing model for multi-fibre WDM networks," Photonic Network Communications, vol. 18, no. 3, Dec. 2009, pp. 287–299.
[12] H. Zang, J. P. Jue, and B. Mukherjee, "A review of routing and wavelength assignment approaches for wavelength-routed optical WDM networks," Optical Networks Magazine, vol. 1, no. 1, Jan. 2000, pp. 47–60.
[13] A. Betker et al., "Reference transport network scenarios," MultiTeraNet Report, Tech. Rep., Jul. 2003.
[14] J. Craveirinha, R. Girão-Silva, and J. Clímaco, "A meta-model for multiobjective routing in MPLS networks," Central European Journal of Operations Research, vol. 16, no. 1, Mar. 2008, pp. 79–105.

Cognitive Control Based on Genetic Algorithm for Routing and Wavelength Assignment in Optical OBS/WDM Networks.

Tatiana Peña Valencia, Adriana M. Hincapié Moncayo, José Giovanny López Perafán
GNTT
University of Cauca
Popayán, Colombia
{tatianap, adrihin, glopez}@unicauca.edu.co

Abstract - **The exponential services demand concerning the Information and Communications Technology (ICT) has led to increasingly efficient implementation techniques to satisfy the information flow required by users. Optical Burst-Switched Wavelength Division Multiplexing networks (OBS/WDM) are currently set as a technology capable of supporting wide bandwidth, allowing high information's transmission and different types of traffic. OBS also combines the advantages of Optical Circuit Switching (OCS) and Optical Packet Switching (OPS). Furthermore, recent advances in optical networking add cognitive mechanisms to any network plane, providing it with adaptive features. From this point of view, it is possible to carry out methods that address the problem of Dynamic Routing Wavelength Assignment (DRWA) in order to improve the network blocking probability. In this paper, we propose a cognitive control method, based on genetic algorithms for DRWA problems in OBS/WDM networks under wavelength continuity constraint, analyzing its performance in terms of blocking probability and execution time by making a comparison with the same network that uses genetic algorithms, but without cognitive control.**

***Keywords*- optical burst switching; dynamic routing and wavelength assignment; genetic algorithms; cognition.**

I. INTRODUCTION

The changes made during the last years in telecommunication networks are mainly based on the need to supply the traffic demand due to the exponential growth of the Internet and the Web evolution [1]. This implies having networks capable of supporting high bandwidth. For this reason, the optical fiber is becoming one of the most important means of transmission today. Moreover, among the advantages of optical networks is the ability to carry data simultaneously by the same fiber using different wavelengths; this technique is called Wavelength Division Multiplexing (WDM).

It is also expected that the current optical networks receive all traffic more efficiently. For this, switching techniques like Optical Burst Switching (OBS) are implemented due to its bandwidth capacity, low latency, high adaptability, and low processing overhead [2].

Before sending a burst through the network, it is divided into the Burst Control Packet (BCP) and data packet. The BCP is sent across a dedicated channel, and it is responsible for the burst's routing. On the other hand, the data packet will be sent only when a timeout has elapsed after sending the BCP. For this reason, the signaling used is out of band, allowing space-time separation between header and data, thus providing manageability and network resilience.

Dynamic Routing Wavelength Assignment (DRWA) in OBS/WDM (Optical Burst-Switched Wavelength Division Multiplexing networks) networks, implements methods for routing and wavelength assignment between links. Such process can be done between two nodes not necessarily adjacent, and requires an available wavelength and route between the optical paths. This connection is called lightpath, and due to the dynamic features, the requests to set it are made at real time.

However, as the DRWA is an Nondeterministic Polynomial Time Complete (NP-C) problem [3], is not possible to use deterministic algorithms for its solution, instead, different methods are proposed to obtain near optimal solutions. This kind of problem allows the use of Genetic Algorithms (GA) as heuristic methods for DRWA in OBS/WDM networks.

The GA resembles the evolution stages and natural selection mechanisms postulated by Charles Darwin. They work with individuals from a population, each representing a tentative solution to the problem. The main GA stages are: generation of individuals, selection, crossover, mutation, and reduction operators. The genetic algorithm will perform a number of iterations of the above stages, until one of the stopping criteria is met. An equally important factor when talking about GA corresponds to fitness function, whose objective is to assign a value to each individual, which measures the relevance of such solution to the problem being tackled. An appropriate fitness function guides the GA towards its goal, converging to optimal solutions.

Moreover, the search for increasingly flexible networks capable of supporting all types of traffic, transmission, and switching technologies, lead to intelligent network management through techniques that allow awareness and adaptation mechanisms at any optical architecture level. Cognitive control methods are a promising factor to meet the challenges the next-generation optical networks have [4][5].

Taking into account the information provided, this article uses two approaches based on GA to the DRWA problem in OBS/WDM networks, one of them using cognitive control methods.

To facilitate understanding, the article is organized as follows. In Section 2, we present the related work to the different topics involving the current research. The next section presents the model and the simulation tool used for

 ISBN: 978-1-61208-315-5

OBS/WDM networks. Sections 4 and 5 will focus on the DRWA problem from the perspective of GA. Then, in Section 6, the inclusion of a cognitive information base supported on GA is presented. Subsequently, we analyze the networks performance taking into account the blocking probability and processing time. Finally, conclusions will be shown in Section 8.

II. Related Work

The aim of this paper encloses a context that unifies several study subjects. Related work about the RWA problem in OBS/WDM networks using cognitive control methods based on GA has not been widely developed. However, certain topics are treated in some research.

For example, Gond and Goel [6] and Gonzales et al. [7] solve the RWA in DWDM networks taking into account physical constraints and including devices called optical converters enabling the switching of wavelengths to increase efficiency. However, they don't use GA, OBS, or cognitive control methods in their strategies.

Kavian et al. [8] and Kavian [9], apply genetic algorithms to the RWA in WDM networks, including in cases like Nakkeeran, et al. [10], strategies of fairness among connections and fault tolerance capability. This papers present disadvantages because there is no evidence of the used switching method.

Moreover, Kozlovski [11], although the authors reflect routing and wavelength assignment strategies specifically in OBS/WDM networks, they don't use GA or cognitive control methods to solve de DRWA problem.

The approach of Tomkos et al. [12] focuses on cognitive optical networks towards a perspective that provides autonomy and counteracts the increasing complexity in the use and network performance. Finally, (and arguably one of the most similar to the one developed in this paper) Durán [13] proposes a strategy based on cognition to solve Impairment-Aware Virtual Topology Design Problems. Although the author does not focus on the switching technique used, he develops a GA including a memory-based cognitive control method to optimize the virtual topology design.

At the moment, there are no evidences of researches related to the DRWA problem in OBS/WDM networks using GA and cognitive control to analyze the network performance. Considering the above, this work aims to apply a cognition control method to an OBS/WDM network that implements GA for RWA processes and analyze the performance, comparing it with the same network that only uses GA for DRWA.

III. OBS/WDM Network Simulation

In order to simulate an OBS/WDM network that enables the GA implementation, OMNeT++ [14] is chosen as the used tool since it has flexible, organized, and suitable software for the simulation of interconnection networks for high performance. To develop the network design, it must be noted that OBS works with two kinds of nodes. The first one is called Edge Node, and presents the following stages:

Assembler Module (EdgeNodeAssembler): it is responsible for assembling IP packets into bursts. To do this, it makes use of three submodules. The first one is the Classifier; the packets go across this module only if their destination and class of service are valid values to send through the network. The Burstifier is another submodule and performs package grouping tasks to form bursts; three assembly criteria are set: time, number, and packet length. Once the bursts are formed, the last submodule called Sender, is responsible for dividing them in BCP and data packet.

Disassembler Module (EdgeNodedisassembler): when the burst reaches its destination, is disassembled returning to the original packets. Figure 1 shows the Edge Node modules.

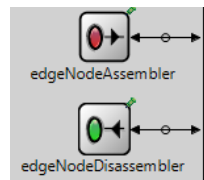

Figure 1. Edge Node modules.

The responsible for routing tasks and network planning is the Core Node (CoreNode). Inside of it are also several submodules:

The Input submodule (CoreInput) can differentiate between data packets and BCP to determine where to send them. If a burst arrives, then, is sent to the Core Control Unit (CoreControlUnit) by a dedicated cannel; moreover, if is a data burst, it will be sent to the Optical Switching Module (OXC) through which it pass in all optical way. Figure 2 shows the Core module with its corresponding submodules.

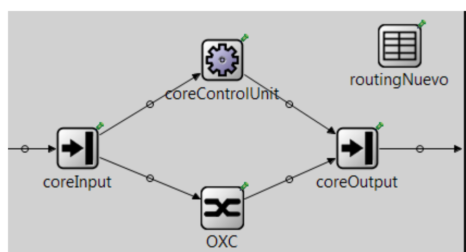

Figure 2. Core Node Submodules.

The Control Unit (CoreControlUnit) is responsible for the BCP treatment. Inside of it, the Logic Control submodule (ControlLogic) is set, which is in charge of routing, updating BCP information and generate wavelength requests. Optical–electrical (OE) and electrical-optical (EO) conversion submodules add a delay to the burst control packet.

The GatesHorizon Module contains information about the available wavelengths in order to assign one of them into the burst.

Figure 3 shows the components of the CoreControlUnit Module.

 ISBN: 978-1-61208-315-5

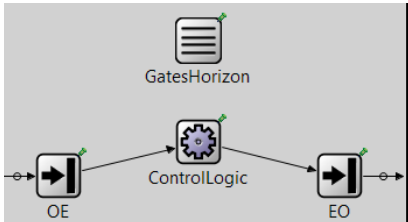

Figure 3. CoreControlUnit module.

Central and Edge nodes join together, creating the resulting OBS/WDM node as shown in Figure 4. This node is copied several times to form the evaluated network.

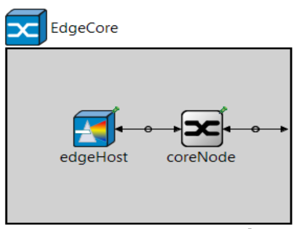

Figure 4. OBS/WDM Node.

To apply the OBS/WDM characteristics, we use The National Science Foundation Network (NSFNet), composed of fourteen nodes and twenty bidirectional connections as shown in Figure 5.

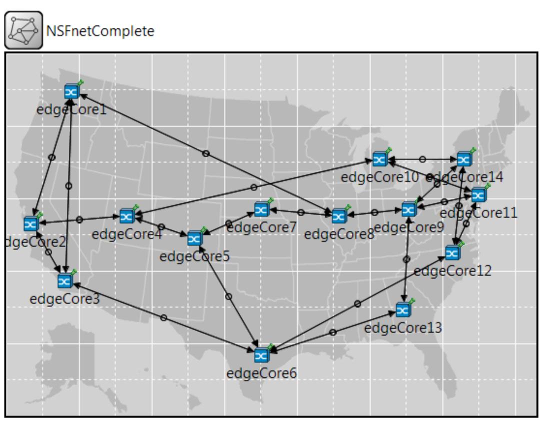

Figure 5. NSFNet.

NSFNet is a Wide Area Network (WAN) covering the U.S. territory and presents the information needed for implementation in the simulation tool OMNeT++.

IV. DWRA From the Perspective of Genetic Algorithms

When talking about routing and wavelength assignment in OBS/WDM networks, the genetic algorithm process begins with a particular number of individuals, each of them being presented as a tentative solution to the DRWA problem [15][16]. In this case, individuals refer to possible paths between the source and destination of a burst.

On the other hand, the analogy between networks and genetic processes is given by the information path, which is represented by a chromosome [17]. The latter is composed of genes, allusive to the nodes of the optical path. To provide greater clarity, the optical path is encoded using a string of integers, where each number refers to a path node, as shown in Figure 6.

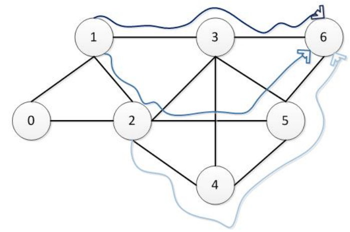

Figure 6. Coded representation of three routes. From top to bottom: {1, 3, 6}, {1, 2, 5, 6}, {2, 4, 5, 6}.

As can be seen in Figure 6, there are three chromosomes, ({1, 3, 6}, {1, 2, 5, 6}, {2, 4, 5, 6}) each one of them composed of three, four, and four genes, respectively.

V. Genetic Operators

A. *Generation of Random routes*

A set of random paths are generated with the source and destination values specified by the burst, which makes the request of a path in real time.

The function begins reaching the node from which the information is transmitted, then, it randomly chooses a node i whose position is adjacent to the source node; once is determined, it is marked as visited. Being in position i, the function performs the search and selection again to set the next node. The process is repeated several times and ends when the destination node is reached, or when all nodes adjacent to the current position have been visited, in the latter case, the function starts again.

B. *Fitness function*

The fitness function integrates information in order to reduce the blocking probability in the network. In this case, the information refers to the number of hops across the path, the distance between them and the number of fibers connected to each network node [15][18]. Now, before we continue talking about fitness function, we need to introduce an important term, i.e., the cost, which allows analyzing what network proportion is affected given a particular factor.

To assess individuals with the aforementioned characteristics, two cost functions are considered:

$$C1 = s + \alpha \sum_{i=1}^{s} (Li - Lf\,i) \qquad (1)$$

C1 cost function allows to analyze the impact created by the number of hops in a route and free wavelengths on it, where s is the number of hops, L is the total number of wavelengths in the path, Lf is the number of free wavelengths, and α is a design parameter which varies between 0 and 1. This function was used in the project of Vinh Tron Le et al. [19], which gives a maximum value for α that depends on the number of wavelengths across the path.

$$\alpha < \frac{1}{W-1} \qquad (2)$$

The second function represents the node congestion generated by the number of links connected to it; the more connections with other nodes present, the more the congestion link will increase. Cost C2 represents the sum of all links of the nodes across the route as shown below.

$$C2 = \sum_{i=1}^{n} Ei \qquad (3)$$

where E is the number of connections of each node across the route.

The fitness function is inversely proportional to the cost, thus, the resulting fitness functions are:

$$F1 = \frac{1}{s + \alpha \sum_{i=1}^{s}(Li - Lf\ i)} \qquad (4)$$

$$F2 = \frac{1}{\sum_{i=1}^{n}(Ei)} \qquad (5)$$

The second fitness function (F2) must have a restriction because if there are no intermediate nodes in the route, the sum of links is zero, which generates an infinite value for this function, i.e., an invalid individual in the search space; as the purpose is to find an optimal global solution, we appeal to a restraint method and reconstruction of that value. In the current project, if such feature generates an infinite result, it is replaced by 1, which represents the maximum value taken by the function.

Finally, two functions are employed as, in some cases, the development of n approximate fitness functions may be better than a single evaluation function [15].

C. Selection

This operator uses roulette wheel selection [20]. All individuals with its respective fitness value are stored in a vector, then; this method assigns to each individual a circular sector, proportional to its fitness function, as shown in Figure 7.

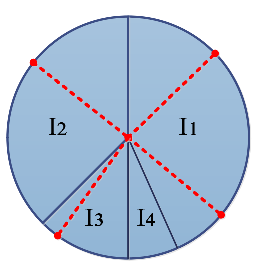

Figure 7. Roulette wheel selection.

The routes with higher fitness value will have higher probability of being selected to apply crossover operator.

D. Crossover

Once the selection process is finished, the routes are stored in vectors which are analyzed to discard those unfit (e.g., those having only two nodes), for crossover operator. Subsequently, the operator searches between the suitable paths, one common element (except source and destination nodes) to combine the routes and get the offspring, as shown in Figure 8.

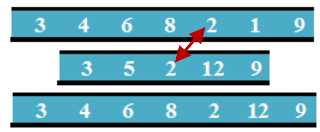

Figure 8. Crossover between two routes at node 2.

The obtained path cannot contain two equal nodes or be equal to a stored route (parents routes), if so, it is discarded, otherwise, it is saved and the fitness function is applied to each one of the children.

E. Reduction

In this process, from all resulting paths, only those with higher fitness function will be chosen. If it meets the stopping criterion, the route with the highest fitness is chosen as the solution, otherwise, the algorithm must perform more generational process until the maximum number of iterations is reached, or if a route is greater than or equal to the maximum fitness criterion.

VI. Cognitive Control Method Based on Genetic Algorithms for DRWA.

Cognitive optical networks have an architecture [12] that describes the functions performed by each layer of the network. In response to this, the control cognitive method is implemented to the Control Plane (CP) [21], since it treats directly the issues of routing, signaling, and network topology. The cognitive process is performed after the genetic algorithm had implemented in the optical network.

In order to explain the implemented cognitive method, the cognitive cycle is defined [4]. The first state corresponds to the Observation, and this refers to the network discovery. The implementation of this state allows the genetic algorithm to implement the methods and stages of its process, providing important information of the cognitive information base. Subsequently, Orientation emerges, where the information from the genetic algorithm takes significance. In this case, the factor of interest to the cognitive process is the resultant GA route, each time a burst is formed.

The third stage corresponds to the Planning cycle; this is where the information related to the best route becomes a target for cognitive control method. In other words, after obtaining the solution vector (path) in a particular network node, this is stored in a table of vectors that contain a maximum number of paths with the same origin but variable destination, which will be part of the cognitive strategy

 ISBN: 978-1-61208-315-5

implemented within the network, this cycle corresponds to Learning. The vector table size is limited and is set to design parameter. Then, and thanks to the Decision cycle, the final route vector corresponding to a node, interacts with the first stage of the genetic algorithm, "Generation of random routes."

Once inside this method, it must validate if one or more of the routes contain the same destination, (for which the request of the generation of the GA is made). If the result is positive and one or more vectors are selected, they will be part of the set of generated random routes (entering in the cognitive cycle called acting).

This method has an additional factor that adds dynamism and wide the solution space: if a new route is generated, and the vector of vectors is full, the latter discards the route corresponding to its last position and allows the entry of the new solution into the first position, displacing the other answers one box. Thus, the cognitive routes will not always be the same.

VII. Performance Evaluation of OBS/WDM Networks Based on GA with and Without Cognitive Control Methods

This section explores the parameters that affect the performance of OBS/WDM networks based on GA with and without cognitive control methods, evaluating the blocking probability and processing time.

Simulations are performed considering a dynamic traffic model where packets arrive according to independent Poisson model with arrival rate λ. The source and destination addresses of each packet are chosen randomly using a uniform distribution. The blocking probability is the first performance indicator to evaluate.

The execution times shown in the graphs are those corresponding to the OMNeT++ graphical user interface. It takes three time phases, each one of them corresponding to one, two, and three hours of real-time simulation. The tests correspond to the average value of the blocking probability and processing time, following the Maximum Likelihood Estimator method [22] with a confidence range of 85%.

Figures 9, 10, and 11 show the blocking probability for the two networks analyzed, taking maximum offset times of 72, 42, and 32 microseconds, respectively. The design parameters are:

- Mean time between packets sent: exponential (32 microseconds).
- Transmission speed: 1Gbps.
- Number of generations the algorithm: 1.
- Maximum fitness: 1.
- Size cognitive vector routes: 4
- Four wavelengths between central nodes.

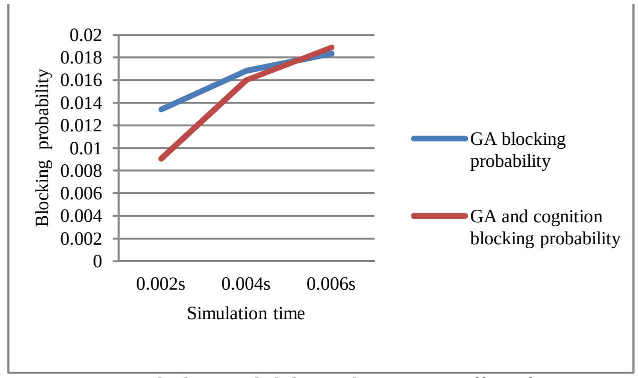

Figure 9. Blocking probability with maximum offset of 72 microseconds.

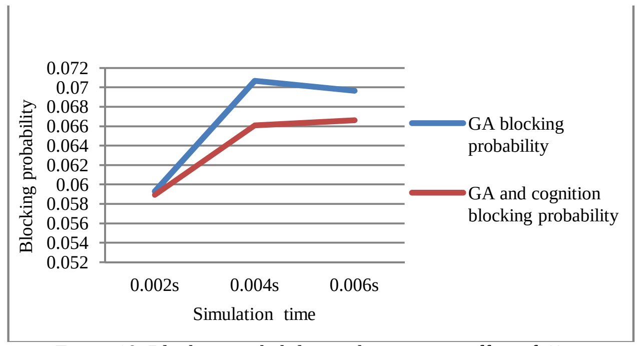

Figure 10. Blocking probability with maximum offset of 42 microseconds.

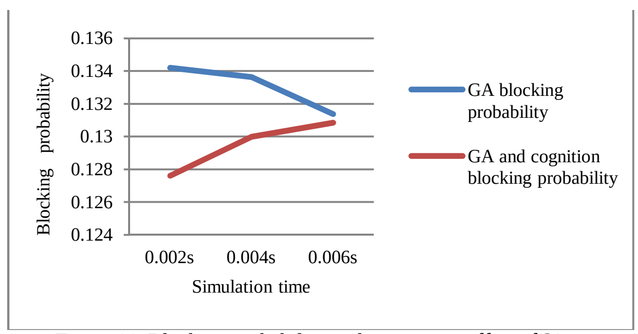

Figure 11. Blocking probability with maximum offset of 32 microseconds.

As shown in Figures 9, 10, and 11, for nine of the ten data taken, the genetic algorithm based on cognitive control performs better blocking probability values.

In Figure 12, we can see the performance of the algorithms by applying different data rates in the optical paths.

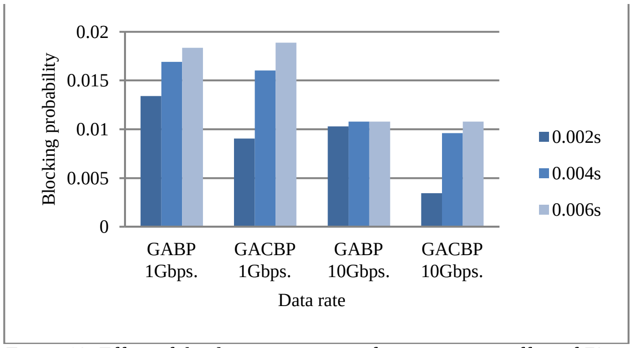

Figure 12. Effect of the data rate variation for a maximum offset of 72 microseconds.

Figure 12 shows that the genetic algorithm based on cognitive control has better response in terms of probability using 10Gbsp date rate.

Below are the results of the elapsed time for establishing the routes of a specific number of bursts, as shown in Figures 13, 14, and 15.

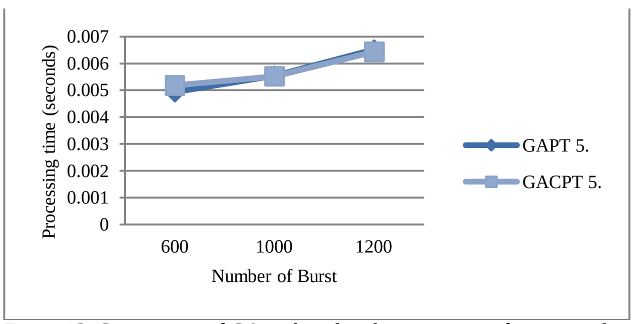

Figure 13. Comparison of GA with and without cognition for an initial population of 5 routes in the processing time estimation.

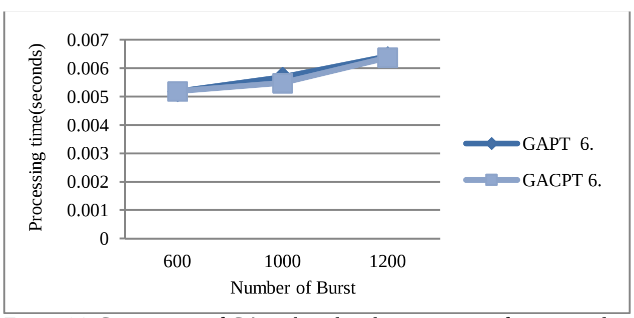

Figure 14. Comparison of GA with and without cognition for an initial population of 6 routes in the processing time estimation.

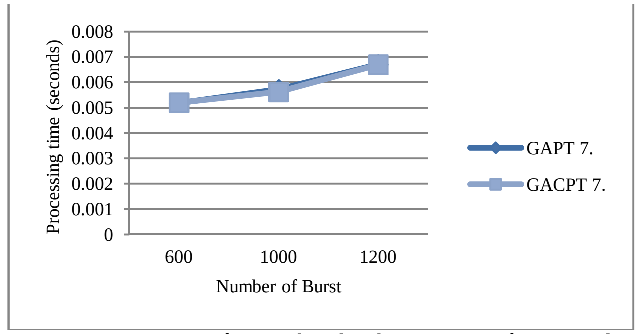

Figure 15. Comparison of GA with and without cognition for an initial population of 7 routes in the processing time estimation.

In the previous graphs, we analyzed the effect of the initial number of routes in the processing time. Although the difference does not differ significantly, we obtained better results using the genetic algorithm with the cognitive control method.

VIII. CONCLUSIONS

Among all analyzed time values and varying offset ranges, from a total of nine results, only in one case the blocking probability is lower for the network implementing genetic algorithms without cognitive control.

Using a transmission speed of 10Gbps and maximum offset 72US, the blocking probability values decrease for the genetic algorithm with cognitive control in two of the three results. Therefore, it can be inferred that in about 66% ± 15% of the cases analyzed, (using the above parameters) this heuristic method can obtain better network performance.

Conducting the evaluation of the processing time for the two designed networks it follows that in 6 of 9 results, the GA with cognitive control method has better processing time (minor), inferring that it will be in the 66.66% ± 15% of cases.

REFERENCES

[1] R. Ramaswami, K. Sivarajan, and G. Sasaki, “Optical Networks: A practical Perspective, 3rd ed., Elsevier, 30 Corporate Drive, Suite 400, Burlington, MA 01803, USA, 2010, pp. 1-30.

[2] R. Lamba and A. K. Garg, “Performance Analysis of Scheduling Algorithms In Optical Burst Switching Networks”, International Journal of Advanced Research in Computer Science and Software Engineering, vol. 1, Issue 1, Jan. 2012, pp. 62-66.

[3] X. Yu and M. Gen, Introduction to Evolutionary Algorithms, New York, Springer, 2010, pp. 3-37.

[4] CHRON (2010). “Cognitive Heterogeneous Reconfigurable Optical Network” Retrieved June, 21, 2013, from, www.ict-chron.eu.

[5] Q. Mahmoud, “Cognitive Networks, towards self aware networks,” University of Guelph: Canada, Wiley, 2007, pp. 1-19.

[6] J. Gond and A. Goel, “Performance Evaluation of Wavelength Routed Optical Network with Wavelength Conversion”, Journal of telecommunications, vol. 2, Issue 1, Apr. 2010, pp. 110-114.

[7] F. Gonzales, et al., “Lightpath routing and wavelength assignment by means of ant colony optimization”, University of Valladolid, Valladolid, Spain, February 2003.

[8] Y. S. Kavian, et al., “Genetic Algorithm for Designing DWDM Optical Networks under Demand Uncertainty”, Electrical Engineering Department, Iran University of Science and Technology, Tehran, Iran, 2009.

[9] Y. S. Kavian, “Robust DWDM Optical Networks Design using Genetic Algorithms,” Proc. First International Conference on Communications Engineering, University of Sistan and Baluchestan, Zahedan, Iran, Dec. 2010, pp. 143-146.

[10] R. Nakkeeran, A. V. Vivek, and P. Kota “Genetic algorithm based approach for routing and wavelength assignment,” Department of Electronics and Communications Engineering, Pondicherry, Engineering College Pondicherry, India, 2004.

[11] E. Kozlovski, “Survivability of Wavelength-Routed Optical Burst-Switched Networks with Guaranteed IP Services,” Department of Electronic and Electrical Engineering, University College London, Torrington Place, London, United Kingdom, 2003.

[12] L. Tomkos, et al., “Next Generation Flexible and Cognitive Heterogeneous Optical Networks - Supporting the Evolution to the Future Internet”, Future Internet Assembly, Springer, Nov. 2012, pp. 225-236.

[13] R. J. Durán, ”Advantages of Using Cognition when Solving Impairment- Aware Virtual Topology Design Problems”, University of Valladolid, Athens Information Technology, April 2011.

[14] OMNeT++ (2013). “User Manual, OMNeT++ versión 4.3”, Retrieved, June, 2, 2013, from, http://www.omnetpp.org/doc/omnetpp/manual/usman.html.

[15] U. Bhanja, S. Mahapatra, and R. Roy. “A Novel Solution to the Dynamic Routing and Wavelength Assignment Problem in Transparent Optical Networks,” International journal of Computer Networks & Comunications (IJCNC), vol.2, No.2, Mar. 2010, pp. 119-130.

[16] D. Bisbal, et al., “Dynamic Establishment of All-Optical Connections in Wavelength-Routed Optical Networks Using Genetic Algorithms”, Next Generation Optical Network Design and Modelling, Kluwer Academic Publishers, Sep. 2003, pp. 377-392.

[17] I de Miguel, R. Vallejos, A. Beghelli, and R. J. Durán, “Genetic Algorithm for Joint Routing and Dimensioning of Dynamic WDM Networks”, Journal of Optical Communications and Networking, vol. 1, Issue 7, Dec. 2009, pp. 608-621.

[18] B. Li and X. Chu, “Routing and Wavelength Assignment vs. Wavelength Converter Placement in All-Optical Networks”, University of Science and Technology Kazem Sohraby, University of Arkansas, IEEE Optical Communications, August 2003.

[19] V. T. Le, X. Jiang, S. Horiguchi, and Y. Inoguchi, “A New Fitness Function for GA-based Dynamic RWA Algorithms in Optical WDM Networks,” in Proc. of Networks, jointly held with the IEEE 7th Malaysia International Conference on Communication, 13th IEEE International Conference on, Malaysia, vol. 2, Nov. 2005, pp. 667-671.

[20] University of the Basque Country (2004), "Genetic Algorithms”, Retrieved May, 18, 2013, from http://www.sc.ehu.es/ccwbayes/docencia/mmcc/docs/temageneticos.pdf.

[21] G.S. Zervas and D. Simeonidou, “Cognitive Optical Networks: Need, Requirements and Architecture”, High-Performance Networks Group, School of CSEE, University of Essex, Colchester, United Kingdom, Jun. 2010, pp. 1-4.

[22] E. Box, J. S. Hunter, and W. G. Hunter, Statistics for Experimenters. Design, innovation and discovery, 2nd ed., Reverté S.A., 2008, pp. 17-94.

Using Visual Attention in Video Quality Assessment

Cristina C. Oprea, Radu O. Preda
Dept. of Telecommunications
Politehnica University of Bucharest
Bucharest, Romania
cristina@comm.pub.ro, radu@comm.pub.ro

***Abstract*—This paper presents the analysis of a modified structural similarity index metric. First, we describe briefly the adaptations we brought. The main idea is to integrate a visual attention model into a metric for perceptual video quality assessment. In order to be able to evaluate our proposed modifications, we brought two other quality metrics: the reduced reference Video Quality Metric (VQM) and the full reference Perceptual Distortion Metric (PDM). Results and conclusions are mentioned in the end, following scatter-plots and correlation coefficients computation.**

Keywords - visual attention; quality evaluation; perceptual assessment.

I. Introduction

Deciding which areas in a given frame have a perceptual significance has several applications. some of the most important applications include: the optimization of the compression scheme in the video coding stage and more effective information hiding in images and video signals in watermarking schemes.. In such applications, the main characteristics and also the thresholds of the visual system can be used to obtain the best performance with respect to visual quality of the output. This paper presents a comparison between three widely used video quality evaluation metrics, all of them trying to integrate a computational model of the human visual system into a tool for quality assessment. Only one of these metrics makes use of a visual attention model by performing a detection of the perceptual important areas.

The paper is structured in five sections. Section II introduces the latest achievements in the area of perceptual region detection and human visual system modelling. Section III contains a detailed presentation of the proposed method, while in the forth section of this paper we briefly introduce the metrics used for comparison. The final section focuses on results and conclusions.

II. Previous Work

The visual assessment task may seem simple, but it actually involves a set of very complex mechanisms that are not completely understood. The visual attention process can be reduced to two physiological mechanisms that combined together result in a usual selection of perceptual significant areas from a natural or artificial scene. Those mechanisms are bottom-up attentional selection and top-down attentional selection. The first mechanism is an automated selection performed very fast, being driven by the visual stimulus itself. The second one is started in the higher cognitive areas of the brain and it is driven by the individual preferences and interests. A complete simulation of both mechanisms can result in a tremendously complex and time-consuming algorithm.

The process of finding the focus of attention in a scene is usually done by building feature maps for that scene, following the feature integration theory developed by Treisman [1]. This theory states that distinct features in a scene are automatically registered by the visual system and coded in parallel channels, before the items in the image are actually identified by the observer. Independent features like orientation, color, spatial frequency, brightness, and motion direction are brought together and analysed in order to find that single object or area being in the focus of attention. Pixel-based, spatial frequency, and region-based models of visual attention are different methods of building feature maps and extracting saliency.

The pixel-based category is represented by Laurent Itti's work concerning the emulation of bottom-up and top-down attentional mechanisms [2]. Another possibility of building feature maps is by applying different filtering operations in the frequency domain. The most common type of such filtering is done using Gabor filters and Difference of Gaussians filters. Meur et al. [3] apply the opponent color theory and use contrast sensitivity functions for high contrast detection. The last category of visual attention models are the region-based algorithms. In this case, a clustering operation is usually performed and then feature maps are computed using these clusters [4].

The main works in human visual system modelling have been studied and evaluated by the Video Quality Experts Group (VQEG), resulting in several video quality metrics that are given as standards at this moment. Such metric is VQM, described by American National Standards Institute (ANSI) in [5].

III. Proposed Algorithm With Attention Modelling

A. Attention model

The first SSIM algorithm developed for video sequences did not take into account in any way the fact that different spatio-temporal regions have specific levels of perceptual importance. Mean-SSIM simply determined an objective score for each frame and at the end of the distorted video sequence computes the mean of all scores. This mean value represented the objective score for the entire video.

Instead of that simple approach, we propose a combination between the visual attention model described in [4] that detects the salient regions in each video frame and the SSIM metric, which is used only for the previously extracted salient areas. This approach saves some computation effort, but not enough to be a major improvement. The positive outcome results come from the determination of the structural distortions only in the regions where the observer's attention is focused. The rest of the video slightly escapes from the natural quality assessment operation performed by the visual system, since it is known that our vision does not process the entire visual information.

The idea is to extract all spatio-temporal regions that get the most attention from the viewer. This operation is performed frame by frame and similar regions that belong to consecutive frames in a video scene will form a spatio-temporal region. Some areas come foreward in terms of attention because they contain some highly saturated colour, powerfull contrast or a particular object shape. In order to achieve less computational effort, we used only the chromatic contrast saliency map and the dimension saliency map presented in [4]. Both maps are combined using equal weights and the result is a general saliency map.

Extracting chromatic saliency begins with a segmentation procedure applied for all frames in the video sequence. The aim at this point is to obtain regions having one colour or similar colors. These regions do not necessarily need to occupy the same place in space as the objects present in the current frame. One object may correspond to several regions. At this point, one region R_i in a frame has the position parameters and the colour parameter: $R_i(p_{x,y},\ c_j)$. $p_{x,y}$ represent the vector coordinates for the main diagonal of the bounding rectangle and c_j is the *j* index in the colour palette. Regions corresponding in size, location, and colour from consecutive frames will form a spatio-temporal region, R_i^f, the maximum length for the temporal dimension being $f_{max} = 5$ frames. The number of colors in the color map depends on the chromatic dynamic in the video. Each frame region R_i will also have a set of pointers to the immediate *N* neighbors and a region perimeter factor, P_i, used to eliminate small unnoticeable patches or large background areas.

The chromatic contrast saliency map is computed considering several scenarios that determine the observer's attention to focus on a specific area. An important aspect of a scene that generates saliency is the color contrast which is differently defined from intensity contrast. Two colors that are situated on opposite sides on the hue color wheel are generating contrast when situated next to one another. There are five color contrast situations evaluated. For a current region R_i^f, each region having an opposite hue brings a score for R_i^f, as well as regions having a sufficiently distant hue. Third situation is the contrast due to the warm and cool colors. Fourth, the saturation contrast, comming from regions having colors with completly different saturations and fifth, the usual intensity contrast. The result will be a chromatic map corresponding to a set of maximum five frames, buid from separate values each corresponding to one region:

$$CM_i^f = \sum_{j=1}^{j=N} k_{op}^{i,j} k_d^{i,j} k_{wc}^{i,j} k_s^{i,j} k_I^{i,j} \left(1 + \Delta S_i^f + \Delta P_i^j\right)$$

A region size map is included in the color contrast map due to the presence of the perimeter factor ΔP_i^j, evaluated for de region R_i^f and its neighbor R_j^f.

Figure 1. Saliency map for "parrots". Original image taken from [12].

B. *Modified SSIM metric*

The distorted sequence analysis is realised in three steps: at block level, frame level and at the entire video sequence level. Our modifications regard the first level, where blocks of 8x8 pixels are extracted from the salient regions in the previous map. This is completely different from the algorithm SSIM for static images described in [6], where a window was sliding pixel by pixel over the entire image. Our solution leads to less computational effort without accuracy loss, since we analyse only the frame parts that actually are viewed by the human observers. Classic SSIM algorithm for video sequences, without the attention model, is presented in [7], including the frame level and the entire video sequence processing level.

IV. Metrics For Comparison

A. *VQM*

VQM has been presented in detail in [9]. It has been developped for many types of video coding and transmission systems, beeing a reduced-reference metric. This method defines and computes several parameters for the original video sequence, using only a subset of spatio-temporal regions from this sequence. VQM needs a supplementary data of 14% from the transmission band filled by the uncompressed video sequence.

In VQM, the original and distorted video sequences need to be spatially realigned. Then, the algorithm tries to give an estimate for the regions in the distorted video corresponding to the regions from the reference video. The calibration step includes a process of computing the gain and amplitude offsets between the two video sequences and another process for temporal alignment. The objective VQM score is computed from the model quality parameters, which in turn depend on several perceptual features.

 ISBN: 978-1-61208-315-5

B. PDM

PDM stands for Perceptual Distortion Metric and was initially proposed by Winkler in [11]. It is a metric based on a model of the human visual system and needs the reference as well as the distorted videos. The main aspects integrated in PDM are the perception of colours and the opponent colour theory, multi-channel decomposition corresponding to the neural spatio-temporal mechanisms, contrast sensitivity, texture masking and the excitation/inhibition behaviour of the neurons in the primary visual cortex. The algorithm that we used is modified from the original one and has several adaptations. A complete presentation of the modified PDM metric that we used can be found in [12].

V. Results and Conclusions

In order to assess the metrics performances, we used a video sequences database developped by the LIVE laboratory and described in [10] and [11]. Each original video in this database has been subjected to four types of distortions and then subjectively assessed by a group of human observers. The subjectives scores called Difference Mean Opinion Score (DMOS) are then plotted against the objective scores estimated with each algorithm from the ones presented before. The results have been analyzed in Figures 1, 2, and 3 in the form of scatter-plots corresponding to a distortion type and a specific quality evaluation metric. The more condensed are the points in the scatter-plot forming a band or ideally a curve, the more accurate are the metric objectives estimates.

From Figures 1, 2, and 3, it is obvious that no metric performs as well for all distortion types. Although, it can be observed that scatter-plots for PDM and the proposed modified SSIM with attention modelling resemble more to a thin and long cloud. At this point, it is not easy to distinguish which metric performs better. Anyway, PDM is a very complex algorithm that needs a significant amount of computation resources. On the other hand, SSIM with attention modelling is more simple.

TABLE I. Correlation Coefficients

	VQM	**SSIM-attention model**	**PDM**
Pearson coefficient	0.73	0.70	0.85
Spearman coefficient	0.72	0.79	0.87

The prediction accuracy can be evaluated through the Pearson correlation coefficient, while the prediction monotony is analysed with the Spearman correlation coefficient. Their values are presented in Table 1 for the wireless distorted videos. It can be observed that our proposed modification for the SSIM has a good outcome, resulting in a greater monotony coefficient than VQM and a similar accuracy coefficient value. In terms of correlation values, PDM remains the best metric of all three serving as a model for comparison. The inconvenience with PDM is that it is not adaptable nor designed for actual use because of its complexity, while our proposed metric has the advantage of beeing practical.

Acknowledgment

This research activity was supported by UEFISCDI Romania under the grant PN-II-RU-TE-2011-3-0261 / No.10 / 2011.

References

[1] A. M. Treisman and G. Gelade, “A feature-integration theory of attention”, Cogn. Psychol., vol. 12, 1980, pp. 97-136.

[2] L. Itti, “Automatic foveation for video compression using a neurobiological model of visual attention”, IEEE Trans. Image Process., 2004, vol.13, (10), pp. 1304 – 1318.

[3] O. L. Meur, P. L. Callet, D. Barba, and D. Thoreau, “A coherent computational approach to model bottom-up visual attention”, IEEE Trans.Pattern Anal. Mach. Intell., 2006, vol. 28, pp. 802-817.

[4] C. Oprea, I. Pirnog, C. Paleologu, and M. Udrea, “Perceptual Video Quality Assessment Based on Salient Region Detection”, Proceedings of AICT 2009, May 2009, pp. 232-236.

[5] T1.801.03, American National Standard for Telecommunications – “Parameters for Objective Performance Assessment”. s.l. : American National Standards Institute, 2003.

[6] Z. Wang, L. Lu, and Al. C. Bovik, “Video quality assessment based on structural distortion measurement”, Signal Processing: Image Communication, 2004, Vol. 19 No.2, pp. 121-132.

[7] Al. Bovik, “Handbook of image and video processing”, s.l. : Elsevier Academic Press, 2005. ISBN:0121197921.

[8] S. Winkler, “Digital video quality: vision models and metrics”, s.l. : John Wiley & Sons, 2005.

[9] C. Oprea, C. Paleologu, I. Pirnog, and M. Udrea, ”Saliency detection making use of human visual perception modelling”, International Journal on Advances in Life Sciences, vol. 2, nr. 3&4, 2010, ISSN: 1942-2660, pp. 200-208.

[10] K. Seshadrinathan, et al. “A subjective study to evaluate video quality assessment algorithms”, s.l. : SPIE Proceedings Human Vision and Electronic Imaging, 2010.

[11] http://live.ece.utexas.edu/research/quality/live_video.html, accesses September 2012.

[12] G. Kootstra, A. Nederveen, and B. Leeds, “Paying Attention to Symmetry” UK. : Proceedings of the British Machine Vision Conference (BMVC2008), 2008. pp. 1115-1125.

 ISBN: 978-1-61208-315-5

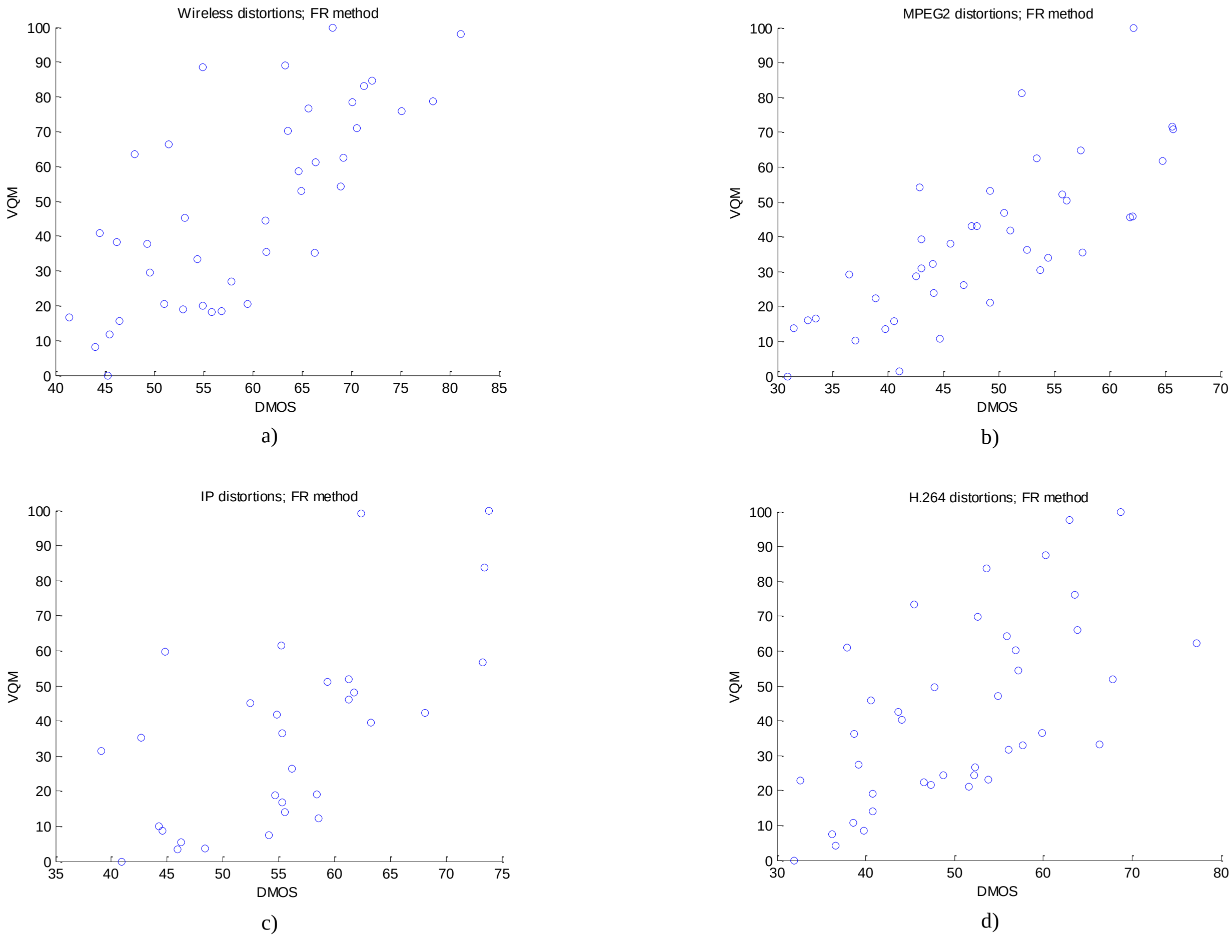

Figure 1. Scatter-plots obtained for the LIVE video database and VQM metric: a) videos with wireless distortions; b) MPEG2 distortions; c) IP distortions; d) H.264 distortions.

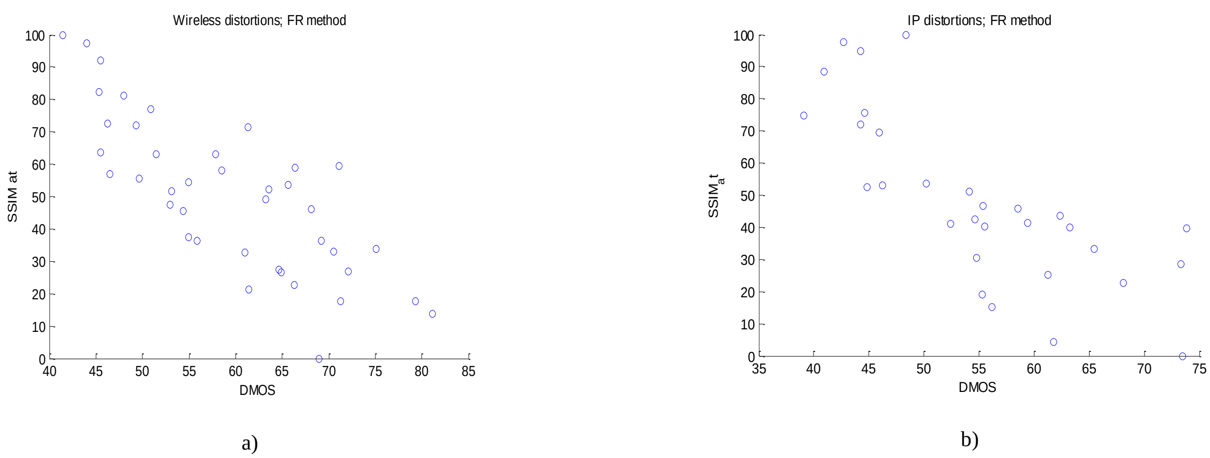

Figure 2. Scatter-plots obtained for the LIVE video database and the proposed metric, SSIM with attention modelling: a) videos with wireless distortions; b) IP distortions.

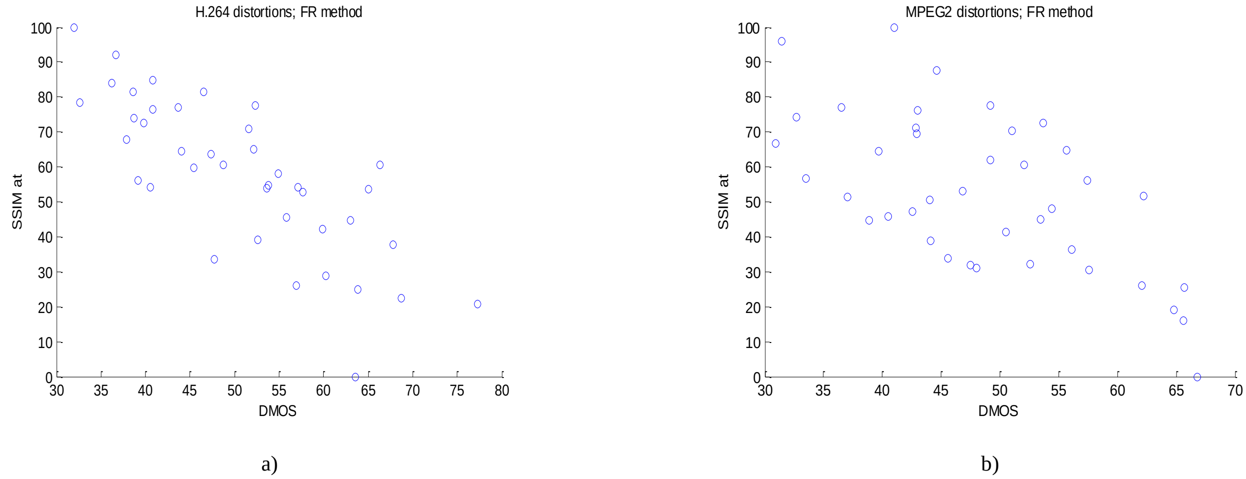

Figure 3. Scatter-plots obtained for the LIVE video database and the proposed metric, SSIM with attention modelling: a) videos with H.264 distortions; b) MPEG2 distortions.

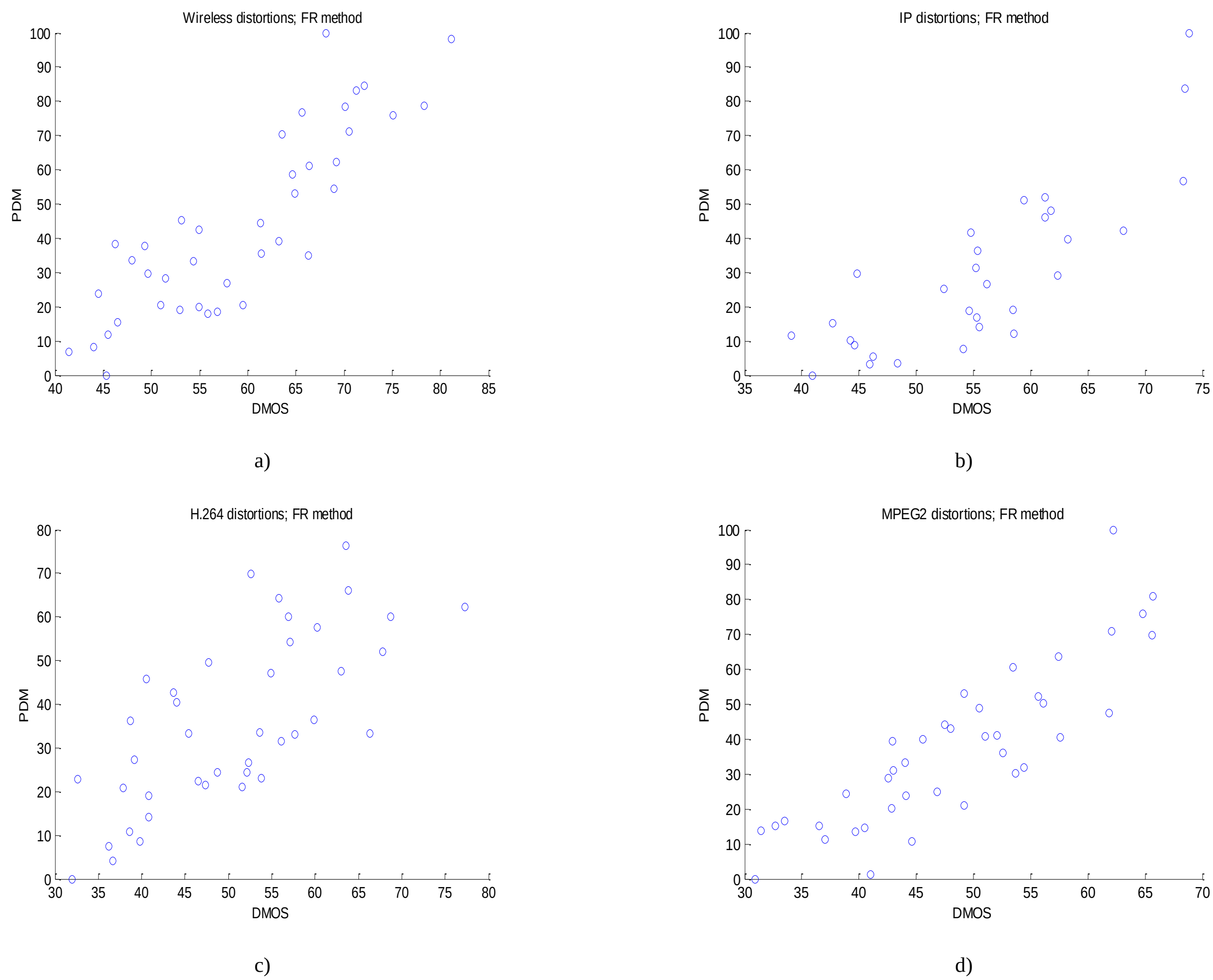

Figure 4. Scatter-plots obtained for the LIVE video database and the adapted PDM metric: a) videos with wireless distortions; b) IP distortions; c) H.264 distortions; d) MPEG2 distortions.

Topology and Network Resources Discovery Protocol for Content-Aware Networks

Vlad Poenaru, Eugen Borcoci, Marius Zotea
Telecommunication Department
University POLITEHNICA of Bucharest
Bucharest, Romania
vlad.poenaru@elcom.pub.ro, eugen.borcoci@elcom.pub.ro, marius.zotea@elcom.pub.ro

***Abstract* —Cooperation of several network providers to offer a common infrastructure to support overlay virtualized networks is of interest in the context of content/information orientation of the current and Future Internet. This paper is part of a work on management framework of a complex system, aiming to construct multi-domain overlay Virtual Content Aware Networks (VCAN), QoS enabled. A protocol (TNRDP). is proposed, evaluated and basically implemented, running between domain managers, to discover inter-domain topology (represented with different degrees of abstraction) and available network resources of different IP domains. The scalability and efficiency is preliminary analyzed.**

Keywords — Multi-domain, Topology and Resources Discovery, Content-Aware Networking, Management, Future Internet.

I. INTRODUCTION

The current, and especially Future Internet, has a strong content/information orientation, including multimedia flows distribution, [1-3]. Customizing the media transport can be done by creating virtualized *Content Aware Networks* (VCAN) on top of the IP level and coupling the network layer more strongly with the upper layers by having *Network Aware Applications* (NAA). The Content Awareness (CA) novel concept means that new intelligent routers will process and forward the data, based on *content type* recognition or, even more, treating the data objects based on their *name* and not based on *location address*, [4], [5]. The VCANs can be constructed based on virtualization techniques [6], agreed to be used to overcome the ossification of the current Internet [1], [2]. The VCANs should be finally mapped onto networking infrastructures, while respecting the requirements of Service Providers exploiting these VCANs.

A complex system architecture, CAN/NAA oriented, is proposed in ALICANTE European FP7 ICT research project, "Media Ecosystem Deployment Through Ubiquitous Content-Aware Network Environments", [7]. It works on top of multi-domain IP networks, to offer services for different business actors playing roles of consumers and/or providers. The architecture defines several cooperating environments: *User (UE), Service (SE)* and *Network (NE).* The UE contains the End-Users (EU) terminals; SE contains High Level Service Providers (SP) and Content Providers (CP). The NE contains a novel entity CAN Provider (CANP) to manage and offer VCANs and traditional Network Providers (NP/ISP) - managing the network at IP level. On demand of SP, the CANP (represented by CAN Managers, each being associated to a network domain) creates unicast or multicast VCANs (QoS enabled) over multi-domain, multi-provider IP networks. VCANs are realized as parallel planes as in [8], but additionally being content aware. The network resources are provided by the NPs. They are managed quasi-statically (provisioning phase) and also dynamically (during delivery phase) by using media flows adaptation. The management is based on dynamic Service Level Agreements/Specifications (SLA/SLSs) negotiated and concluded between providers (e.g., SP, CANP). In the Data Plane, content/service description information (metadata) can also be inserted in the media flow packets by the Content Servers, then recognized and treated appropriately by the intelligent routers of the VCAN (called Media Aware Network Elements – MANE).

This paper proposes a *topology and network resource discovery protocol (TNRDP)* running between the CAN Managers, to collect inter-domain and abstracted intra-domain information, to support the VCAN mapping algorithms onto network resources. Section II presents samples of related work. Section III summarizes the overall ALICANTE architecture and VCAN management. Section IV presents assumptions, requirements and main characteristics of the TNRDP protocol. Implementation is summarized in Section V. Specific aspects, evaluation and optimizations are discussed in Section VI. Section VII contains some conclusions and future work outline.

II. RELATED WORK

The TNDRP belongs to the management framework, aiming to construct QoS enabled VCANs, , over several independent but interconnected network domains..Given that VCAN solutions are based on the overlay concept, inter-domain QoS peering and routing are of interest, based on the overlay network ideas [9], [10]. An overlay network is defined, which first, abstracts each domain with a node, represented by the domain resource manager, or more detailed with several nodes represented by the egress routers from that domain. There exist protocols to transport QoS and other information between nodes and, based on this information, QoS routing algorithms can choose some QoS capable path.

In [9], a Virtual Topology (VT) is defined by a set of virtual links that map the current link state of the domain without showing internal details of the physical network

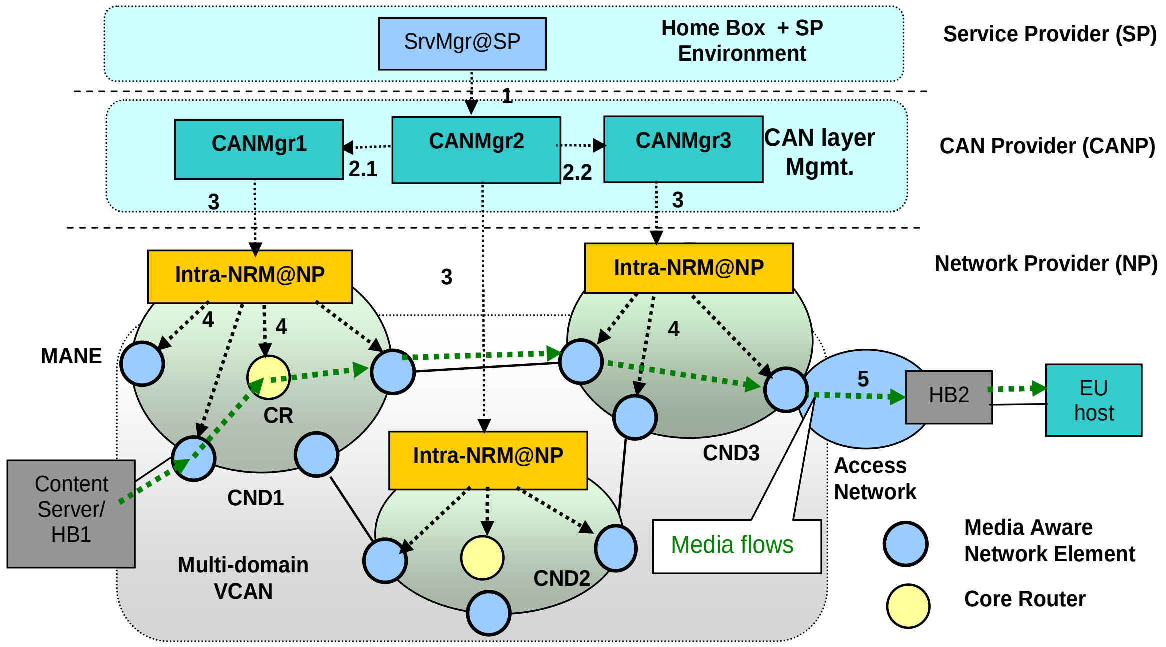

Figure 1. Example of network infrastructure (three domains) and management actions

Notations: SP - Service Provider; HB - Home Box; SrvMgr@SP – Service Manager at SP; CANMgr- Content Aware Network Manager at CAN Layer; Intra-NRM@NP – Intra-domain Network Resource Manager at Network Provider; CND – Core Network Domain; MANE – Media Aware Network Element; CR – Core Router; EU – End User Terminal

topology. Then *Push* and *Pull* models for building the VT at each node are considered and analyzed.

In the *Push* model each AS advertises its VT to their neighbor ASes. This model is suited for small topologies. In the *Pull* model, the VT is requested when needed, and only from the ASes situated along the path between given source and destinations; the path itself is determined using BGP. Also, the solution in [10] is based on BGP protocol. However we need a protocol to run between managers, so, although some ideas of [9] and [10] are valuable, our context is different. So, in Section IV we will develop our solution adapted to our system architecture to find overlay topology.

The system architecture is briefly presented in Section III to prepare extraction of TNDRP requirements. It is similar to Software Defined Networking, [11], [12]: evolutionary architecture; Control Plane / Data Plane separation; network control intelligence can be (logically) centralized in SW - based SDN controllers, which maintain a global view of the network; execution of the Control Plane infrastructure SW can be done on general purpose HW; the intelligent layers are decoupled from specific networking HW. Actually the Alicante CAN Manager and Intra-domain Network Resource Manager of the network Provider constitute together the "controller" defined in SDN.

III. System Architecture and VCAN Management

A. *System Architecture*

Figure 1 presents a simplified partial view on the ALICANTE architecture, with emphasis on the CAN layer and management interaction. The network contains several Core Network Domains (CND), belonging to NPs (they can be also seen as Autonomous Systems - AS) and access networks (AN). The ANs are out of scope of VCANs. One *CAN Manager* (CANMgr) exists for each IP domain to assure the consistency of VCAN planning, provisioning, advertisement, offering, negotiation, installation and exploitation. Each domain has an *Intra-domain Network Resource Manager* (Intra-NRM), as the ultimate authority configuring the network nodes. The CAN layer cooperates with some local entities called Home Boxes (HB) and SE by offering them CAN services. Details on this architecture are found in [7], [13], [14].

B. *VCAN Management*

The actions 1, 2, 3, 4 (Figure 1) are a simplified set performed in M&C Plane in order to negotiate, agree and install a VCAN:

1- Request (SLA/SLS) for a VCAN from SP - to an *initiator iCAN Manager* (e.g., CANMgr2);

2.1, 2.2 – "Horizontal" negotiations between the iCAN Manager and others, associated to different CNDs – for multi-domain VCAN mapping and construction;

 ISBN: 978-1-61208-315-5

3- "Vertical" requests from each CANMgr to its Intra-NRM, to install VCAN;

4- Commands given by each Intra-NRM (involved in the multi-domain VCAN) to configure its MANE and Core Routers (install ingress policies, and egress rules).

The number 5 action represents a data flow, transported on a unicast path over several network domains inside the VCAN, from Content server (or HB if it generates content).

The SP VCAN requests are expressed in the SLS. The SP knows the edge points of this VCAN, i.e., the MANEs IDs where different sets of HB will be connected. The requested VCAN belonging to a given QoS class, should be mapped onto real network topologies, while considering available transport resources of one or several network domains. The solution adopted (for unicast VCANs) was to run on the initiator CAN Manager a combined algorithm capable of computing QoS constrained paths, making logical resource reservation, and finally mapping the VCAN onto such resources [14]. The algorithm provides a global optimum and is appropriate for VCANs mapped onto MPLS paths. The latter are pre-provisioned by each Intra-NRM and then offered as available to the associated CAN Manager.

The *VCAN mapping problem is:* given a resource availability graph and a *Traffic Demand Matrix (TDM),* how to map it onto a real graph while respecting the minimum bandwidth constraints and also optimize the resource usage.

The mapping algorithm needs input data at the iCAN Manager level, i.e. an "international" graph, plus link capacities, where the intra-domain paths are abstracted as traffic trunks.

The TNDRP should collect the inter-domain graph information(it is assumed that each CAN Manager has knowledge on its domain and inter-domain links with neighbors). In such a way, the VCAN initiator CAN Manager can learn the inter-domain *Overlay Network Topologies (inter-ONT)* and inter-domain available link capacities, and can map VCANs onto network resources. At domain level, a summary only of the intra-domain topology is uploaded by each Intra-NRM to the CAN Manager, represented in an abstract way by sets of virtual links (called *Traffic Trunks*), belonging to a given QoS class. This information is called *Resource Availability Matrix (RAM*). Each domain is free to upload whatever RAM it wants, depending on its own policy. This is important from a business point of view, given that it preserves each CND independence.

A partial description of the VCAN mapping algorithm is given in Figure 2, in order to derive the requirements of the TNRDP protocol. Assume that SP issues a request for *VCAN (I1, O1, O2, O3)* to CANMgr2 (this is the VCAN initiator). Here the generic *Ix, Oy* represent inputs and respective output points in the VCAN. Actually these will be mapped onto MANEs interfaces. Note that SP does not know which (if they exist) some transit domains are needed for the requested VCAN. It is not its job but the initiator CAN Manager must determine the best transit domains.

In order to run the mapping algorithm, the initiator iCANMgr should discover (viaTNRDP) the complete inter-domain and intra-domain abstracted summarized graphs

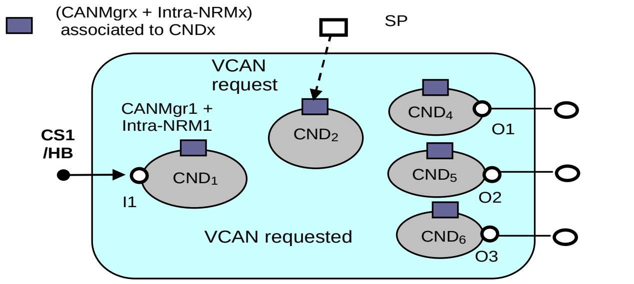

Figure 2. Example of a requested VCAN (I1, O1, O2, O3)

(within a given region of the world).

Figure 3 shows as an example the context associated to the VCAN request addressed to CANMgr2 playing the role of iCANMgr (see Figure 2). After running the TNRDP iCANMgr2 will know the inter-domain graph and the RAM abstractions of each Core Network Domain (CND1 … CND6). Then iCANMgr runs the VCAN mapping algorithm and determines all CNDs involved in the requested VCAN. Then it splits the SLS parameters, thus preparing new requests for each of these CNDs and negotiates *with each CAN Manager* involved, to agree and reserve resources for the VCAN. The main TNRDP design decision is on how a CAN Manager can obtain the topology and resources information? One can use several approaches.

On demand variant: the iCANMgr knows initially only a pure inter-domain graph (where each domain is abstracted with a node). Then it can determine (first step) via a routing algorithm, which domains could be candidates to participate to the VCAN. Then iCANMgr asks from them their RAMs and obtains the complete abstracted graph as in Figure 3. On this graph iCANMGr applies the VCAN mapping algorithm (second step). While being hierarchical, this solution does not provide a global optimum, given that the intra-domain RAMs are not known in the first step.

Proactive variant: the TNRDP works periodically or in event-triggered mode. Every iCANMgr has in any moment the image of the graph as presented in Figure 3, then it applies the VCAN mapping algorithm and finds the best paths by using an appropriate metric and constraints expressed in the TDM.

The VCAN mapping algorithm is described in [14]. Its details are out of scope of this paper. Shortly, the cost of a link (i,j) in the ONT can be *C(i,j) = Breq/Bij= Breq/Bavail*, where Bij is the available bandwidth on this link and *Breq* is the bandwidth requested for that link. The metric is additive, so one can apply a modified Dijkstra algorithm to compute the *Shortest Path Trees (SPT),* i.e one tree for each ingress node where the traffic flows will enter. A basic simpler metric of 1/Bij as an additive link metric can be used.

This paper considers the proactive variant of TNRDP. The advantages are similar to other proactive protocols: it provides a global optimum of the paths; at any moment each CAN Manager is ready to receive VCAN requests and to serve them based on updated information produced by TNRDP. This is similar to the routes computation in

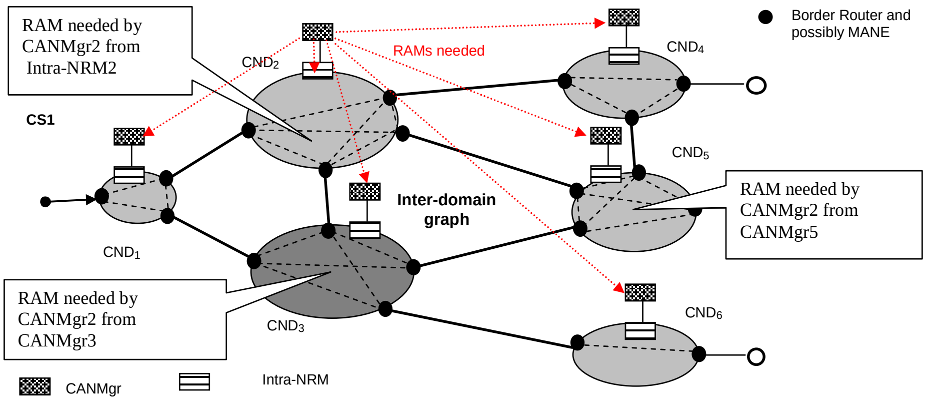

Figure 3. The graph information necessary to be known by CANMgr2 in order to be able to map the VCAN requested in Figure 1

proactive style and then their usage for forwarding the data packets. Actually, TNRDP decouples the topology and resources discovery process from the actual VCAN resource reservation and allocation processes. However, different from the routing protocols context, the VCAN reservation means a *stateful* approach. So, this solution creates some problems, e.g., a decision on VCAN-k mapping is taken at iCANMgr_k, based on the known graph at that instant, while some other SPs VCAN requests can arrive at some other CAN Managers, for resource overlapping with those used by the iCANMgr_k.

Another problem of the proactive TNRDP can be related to the signaling overhead, due to broadcast style, similar to OSPF working mode. These aspects will be discussed in more depth in Section VI.

IV. Topology and Network Resources Discovery Protocol

A. *Assumptions*

TNRDP is an application layer protocol.The folowing assumptions are considered valid for the context of TNRDP:

- Each Intra-NRM knows its physical connections and internal paths and uploads its RAM information to its CAN Manager by using a mechanism external to TNRDP. The RAM can be expressed in the most simple form as a set of tuples : *{(Router_in, Router_out, Capacity) (.. .)..}.*
- The CANMgrs exchange TNRDP messages with other CANMgrs.
- TNRDP does not solve itself security issues; reliable transport is supposed over TCP.
- The Managers Identities are statically known, by each manager in a given region of a larger network. Also the associated domains IDs (e.g., AS numbers) are known, or at least the ASes of neighbours.
- A simplifying assumption for the basic TNRDP version is that all domains are VCAN capable.
- VCAN Mapping algorithm itself is out of TNRDP scope.

B. *TNRDP Requirements*

The essential TNRDP requirements resulted from the previous sections are:

- To accommodate any number of CAN Managers;
- Scalability in terms of signalling traffic overhead;
- To serve any CANMgr to find out the topological and resources graph in a given network region;
- To support parallelism, i.e., it should leverage topological and resurces changes that can appear in several places in the same time;
- To disseminate in a stable way the topology and capacity information when changes appear in some domains.
- Run in “soft realtime” conditions, given that VCANs are established in aggregated way and not per-path. Also they are established and terminated by SPs not so frequently (tens-hundreds, at most thousands per day).

The main TNRDP design decisions are described below:

- Any CANMgr communicates directly (via TCP) only to its neighbors; two managers are defined as neighbors if their associated network domains have at least an inter-domain link. Before exchanging NSA (*“Network State Advertisments”)* a CANMgr must establish logical connection to its neighbors;
- Each CANMgr receiving a NSA message, aggregates this information to its own ONT and then broadcasts a new NSA to its neighbors;
- TNRDP is a simple stateful protocol. The Finite State Machine states of a given CANMgr related to one of its neighbors are: *Listen, Waiting for connection or disconnection confirmation, Connected* (NSAs can be exchanged);
- Any CANMgr may initiate Connect/Disconnect;
- If a CANMgr receives a Disconnect message form a neighbor, it will exclude from its known graph all information related to its links to that neighbor. If the disconnected domain is a leaf one, then the CANMgr observing this disconnection will also erase all information related to that disconnected domain;

 ISBN: 978-1-61208-315-5

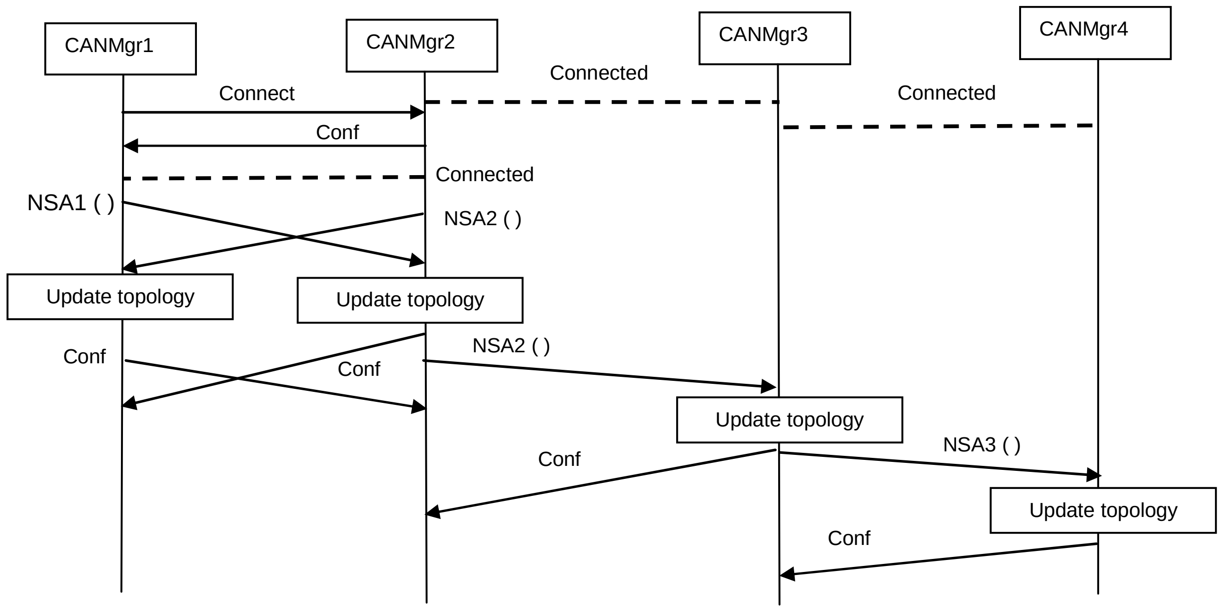

Figure 4. Message Sequence Chart example; Neighbors (CANMgr1-CANMgr2, CANMgr2-CANMgr3, CANMgr3-CANMgr4)

- NSA messages can be exchanged in triggered event style or periodically, in asynchronous mode (this simple assumption could be modified in the section related to optimizations)
- "Hello" messages will maintain the connection alive.

C. TNRDP main characteristics and basic operation

The TNRDP messages types are:

Connect – to initiate connection between two neighbors;

Disconnect – to ask the end of a connection to a neighbor; the NSA exchange will be stopped;

Network State Advertisments – broadcasted to neighbors to update the graphs. It contains parts of the global graphs, known by the sender;

Confirmation- sent in unicast mode as confirmation to "active messages";

Error- to signal synthax or semantic errors;

Hello - to maintain/confirm the connection.

V. IMPLEMENTATION

A *Java basic implementation* has been realized, as *proof of concept* and for preliminary evaluation/validation of TNRDP. The program has five classes:

Runner – is the starting point of the program. Here a configuration file representing the managers set data, topology and resources information, etc., is loaded as input data;

SimFactory – performs configuration file analysis and generates the topology;

SimHost – generates the manager instances and processes the states and mesasages of the managers;

Graph – contains the graph and also the operations to be performed on the graph;

The message format is the following: *Type, Seq_no, Src_Mgr_Id, Dest_Mgr_Id, Data_length, Data*, where the fields semantics is straightforward. The sequence numbers are inserted by each sender and incremented at each new message. They allow to control the pairs message/ confirmation and can also be used in mimimizing the amount of control traffic.

Figure 4 presents an example of success simple scenario in which the neighbors are: CANMgr1 - CANMgr2, CANMgr2 - CANMgr3, CANMgr3 - CANMgr4. It is supposed that connections already exist: 2-3 and 3-4. Then CANMgr1 wants to join the "community" and initiates a connection to CANMgr2. After connection is established the two partners exchange topology and resources information, via NSA1 and NSA2 messages. After updating its Database, CANMgr2 informs CANMgr3 about changes and CANMgr3 performs a similar action for CANMgr4. Here it is supposed an asynchronous mode, event-triggered style for communication.

Message – executes transmission of messages between the managers.

Each manager has three message queues: *input_Q, processing_Q*, and *output_Q*. The program runs in each (iterative) step all managers participating to topology. Each manager processes all messages received from other managers in the previous step, makes updates and sends messages to the output queue to be sent to other managers in the next iteration.

The topological graph and resource information has been represented in matrix form, considering all uni-directional logical pipes (intra-domain or inter-domain segments) of the graph. The simple entry in such a matrix could be *(Input_i, Output_j, Available_Bandwidth).*

VI. SPECIFIC VCAN-RELATED ASPECTS, EVALUATION RESULTS AND OPTIMIZATIONS

The TNRDP protocol runs in a specific environment given by the ALICANTE architecture. The full implementation should be compliant with CAN architectural requirements. One major TNRDP aspect is related to the *stateful* characteristic of a VCAN (i.e. working based on logical reservation and then resource allocation following QoS constrained paths determination). From this point of view the requirements are harder than in the classic routing protocols (OSPF, BGP); there, a router is free to make fast path changes (for routing and consequently – forwarding) if new conditions arise. In our case, a first precaution is that VCAN mapping algorithm is atomically performed at a given iCANMgr; i.e., the TNRDP is not allowed to change the topology and resource information until the VCAN mapping calculation is ended at this iCANMgr. However, this computation is done relatively fast (see [1]) in comparison with communication delays between CAN Managers. Another issue is that several SPs can concurrently request different VCANs, to different CAN Managers. Therefore, some competition for the same resources could appear, given the fact that two iCANMgrs could possibly compute and try to reserve the same path segments in un-coordinated way. This requires an additional confirmation/negotiation (after the initial VCAN mapping has been computed), done in hub style between iCANMgr and the others ones CAN Managers involved, but this step (i.e., unicast communication between iCANMgr and the other CAN Managers) existed anyway in the original approach of VCAN negotiation..

Scalability of TNRDP in terms of number of domains to be accommodated and their dimension is of interest. It is estimated that a VCAN-capable region could involve tens or at most hundreds (less probable) of CAN Managers. On the other side the system architecture supposed that each Intra-domain Network Resource Manager is aware of its paths and resources. Therefore such issue is not a bottleneck of the TNRDP protocol.

The proactive style implies broadcasting NSAs to neighbors and further in an area. The TNRDP messages overhead is of interest. The total number of messages following a given event depends on the number of domains and on the topology. On the JAVA implementation the following numbers have been measured, for a linear network topology. The complexity order is $O(D^2)$, where D is the number of domains, as it can be seen in Table I.

A general approximate estimation (for an arbitrary topology) of the message numbers can be done, considering that we have D domains (managers) and each manager has a connection order of n (it has to broadcast to n other domains a given message), while the diameter of the region is d (measured in number of domains). The overhead traffic produced by one event would be OVH = O(2*D*n*d) messages. If we include confirmation then for D = 100, n = 10, d=5, we get OVH = 10000. A reduction can be achieved if after receiving a mesasage from a neighbor, a manager will wait (based on a timer) some time (e.g., ~100ms) for other possible messages to be received , then aggegating them and re-transmiting its new NSA.

TABLE I. COMMUNICATION COMPLEXITY

No. of domains	5	10	20	50	100	250
No. of messages	24	99	399	2499	9999	62499

VII. CONCLUSIONS AND FUTURE WORK

This work proposed a novel *topology and network resource discovery protocol* (TNRDP) usable between several domain managers, aiming to produce information to serve *Virtual Content Aware Networks* mapping onto network resources in an optimal way. Solutions analysis is done in the architectural context of a complex system aiming to media distribution. Proof of concept preliminary implementation has been done and a summary of performance analysis and optimization measures. Currently a full version of this protocol is in development in the framework of the ALICANTE project.

VIII. AKNOWLEDGEMENTS

This work was supported partially by the EC in the context of the ALICANTE project (FP7-ICT-248652) and partially by the project POSDRU/89/76909.

REFERENCES

[1] J. Schönwälder, M. Fouquet, G. Dreo Rodosek, and I. C. Hochstatter, “Future Internet = Content + Services + Management”, IEEE Communications Magazine, vol. 47, no. 7, Jul. 2009, pp. 27-33.

[2] C. Baladrón, “User-Centric Future Internet and Telecommunication Services”, in: G. Tselentis, et. al. (eds.), Towards the Future Internet, IOS Press, 2009, pp. 217-226.

[3] J.Turner and D. Taylor, “Diversifying the Internet,” Proc. GLOBECOM ‘05, vol. 2, St. Louis, USA, Nov./Dec. 2005, pp. 760-765.

[4] J. Choi, J. Han, E. Cho, T. Kwon, and Y. Choi, A Survey on Content-Oriented Networking for Efficient Content Delivery, IEEE Communications Magazine, March 2011.

[5] V. Jacobson et al., “Networking Named Content,” CoNEXT ’09, New York, NY, 2009, pp. 1–12.

[6] 4WARD, “A clean-slate approach for Future Internet”, http://www.4ward-project.eu/.

[7] G. Xilouris, et al., FP7 ICT project, Deliverable D2.1: "ALICANTE Overall System and Components Definition and Specification", September 2011, http://www.ict-ALICANTE.eu/

[8] M. Boucadair, et al., “A Framework for End-to-End Service Differentiation: Network Planes and Parallel Internets”, IEEE Communications Magazine, Sept. 2007, pp. 134-143.

[9] Fabio L. Verdi and Maurıcio F. Magalhaes “Using Virtualization to Provide Interdomain QoS-enabled Routing”, Journal of Networks, April 2007, pp. 23-32.

[10] Z. Li, P. Mohapatra, and C. Chuah, Virtual Multi-Homing: On the Feasibility of Combining Overlay Routing with BGP Routing, University of California at DavisTechnical Report: CSE-2005-2, 2005.

[11] T. Koponen, et al., Architecting for Innovation. SIGCOMM CCR, 41(3), 2011.

[12] M. Casado, T. Koponen, S. Shenker, and A. Tootoonchian, Fabric: A Retrospective on Evolving SDN. In Proc. Of HotSDN, August 2012.

[13] E. Borcoci, and R. Iorga, “A Management Architecture for a Multi-domain Content-Aware Network” TEMU 2010, July 2010, Crete, http://www.temu.gr/2010/.

[14] R. Miruta and E. Borcoci (UPB) "Optimization of Overlay QoS Constrained Routing and Mapping Algorithm for Virtual Content Aware networks" ICNS 2013 - Lisbon, Portugal. http://www.thinkmind.org/index.php?view=article&articleid=icns_2013_4_30_10174.

The Typology and the Current Perspectives of Network Organizations

Karolína Kolerová, Tereza Otčenášková
Faculty of Informatics and Management
University of Hradec Králové
Hradec Králové, Czech Republic
e-mail: karolina.kolerova@uhk.cz, tereza.otcenaskova@uhk.cz

***Abstract*—The competitiveness within as well as among organizations represents a topical and challenging issue. Therefore, network organizations should pay considerable attention to both researchers and practitioners in order to gain various advantages. Particular models of network organizations are introduced and their implications for potential improvements are discussed.**

Keywords-communication mechanism; coordination mechanism; organizational model; network organization; typology

I. INTRODUCTION

A lot of current organizations might be considered as network organizations, at least to a certain extent. A network organization can be defined as any collection of actors ($N \geq 2$) that pursue repeated, enduring exchange relations with one another and, at the same time, lack a legitimate organizational authority to arbitrate and resolve disputes that may arise during the exchange [1]. The basic features include the presence of independent teams and departments sharing common values, interconnected projects supporting each other, links among projects through Information and Communication Technology and the presence of the role of key coordinator for communication and coordination managing the units [2].

The necessity and the pursuit of the employment of the network principles might be beneficial for the purposes of the communication and coordination enhancement. Nevertheless, the organizations do not realize and utilize the potential of such structures. The main benefit of this work in progress is the initial introduction of the network organizations and their implications. The models of the network organizations are reviewed and the practical advice is outlined. The paper also aims to serve as an inspiration for the development of the organizations aiming to enhance their competitive advantage

This paper firstly describes the methodology. Secondly, the basic concepts of network organizations are introduced in Section III. Partners within models of networks and the significant aspects of coordination and communication are identified. Furthermore, advantages and disadvantages are outlined and theoretical assumptions are discussed. Various models, including their description, are introduced in Section IV. Finally, the most important issues are concluded, the paper limitations are outlined and significant areas for further research are mentioned in Section V.

II. METHODOLOGY

The methodology is mainly based on the literature review. Various resources are examined, analyzed and compared. These provide an initial overview of the potential models which can be subsequently employed within various organizations. The variety of resources includes journal paper, web resources and books. Moreover, the effort to use relevant experts and their opinions was pursued. For instance Jashapara, Desouza or Prusak are encompassed.

Moreover, the practical perspective is included through the observation of and experience with specific organizations with a network structure. Mostly, clusters were analyzed. Currently, only the practices of companies from the Czech Republic were observed. Nevertheless, this was sufficient from the perspective of this paper.

III. MODELS OF NETWORKS – IDENTIFICATION AND CHARACTERIZATION

A network can be characterized as a set of linkages among a defined set of people in which the character of the linkages is specified [3].

A. Partners within Models

Two crucial components - organization and infrastructure - are the most important for networks to be effective. Moreover, reliable partners in a network should be identified and involved to ensure the continuous improvement and innovation. Dedina and Odchazel [4] define three main roles of partners within various models. These encompass the following:

- Trading partners (distributors, producers, service providers).
- Nominal trading partners (these partners ensure mainly physical distribution represented by transportation providers, bank institutions, the Internet providers).

 ISBN: 978-1-61208-315-5

- Organizers with key activities including, among other things, the future development prediction, the network support and assistance, recruitment of new partners or development management.

Moreover, Desouza and Awazu [5] emphasize the role of so called gatekeepers - people who, especially within larger teams and networks, screen and select relevant information before making it available to the group. This eliminates the distraction of the group with useless, irrelevant and low-quality knowledge.

Nevertheless, such important role is usually neglected, because all the potential candidates with relevant experience and optimal involvement sometimes lack time, are not willing to provide their know-how or face various obstacles that prevent them from participating in these activities.

Another model of partnership stressing the continuous improvement and innovation network was introduced Clark et al. [6]. Continuous improvement and innovation requires people willing to share knowledge and ideas and utilize specialized knowledge, skills and support. Fig. 1 indicates categories of people that may be targeted to form a sustainable partnership and network.

B. Aspects of Coordination and Communication

One advantage of network organizations is the improvement in cooperation and communication improvements leading to better flow of information and knowledge [7]. Support of coordination and communication in organizations [8] is represented mainly by the following attributes:

- Events: enhancements of the sense of purpose.
- Leadership: role of community coordinators.
- Connectivity: enhancement of dialogues, trust and relationships.
- Membership: optimal amount of members.
- Learning projects: learning agenda.
- Artifacts: documents, tools, stories, symbols, websites.

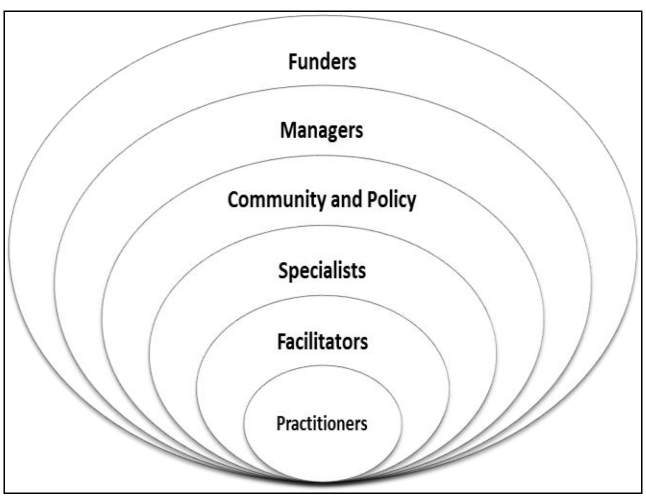

Figure 1. A diagram of potential partners in a continuous improvement and innovation network (Adapted according to [6]).

Apparently, every organization varies in the extent and inclusion of the mentioned components. These differ on the basis of the demands and needs of the internal and external context of the given subject. Nevertheless, Jashapara [8] claims that it is really crucial to find the ideal extent of cooperation and competition within the coordination mechanisms to support the optimal "zone" which results in knowledge creation, better performance and better results.

C. Advantages and Disadvantages

The advantages and disadvantage of both the individual organizations within the networks, and performance of such organizations are strongly interconnected. These sometimes influence each other significantly and therefore these will be discussed together. As discussed in Dedina and Odchazel [4], the benefits include income increase, costs decrease, higher flexibility and efficiency, higher innovation potential together with more prompt, efficient and desirable ways of knowledge and skills spread and transfer. The flexible and organic structure might also cause some problems. Struggles with the determinations of clear responsibilities or unstable organizational structure linked with this approach can serve as an example. Furthermore, the transactional costs should not be omitted in the budget planning. These include for example necessarily needed costs for the communication among the member organizations.

Nevertheless, the extent of informal connections and trust among members remain important. Some organizations promote familiar environment and personal relations. On the other hand, within other companies only strict and formal relations among employees occur. Both extremes are perceived as not very worthy especially from the long term perspective. To find the right balance of the supportive initiatives is relatively hard task.

The advantages of the network organizations support their importance as well as the usability of such organizational structure. In comparison with hierarchical and market forms of organization, the network structure possess particular advantages (more details in Table 1). The relational nature of the membership and the way of communication evokes the emphasis on reputation among member organizations as well as with the external institutions. Moreover, these ensure usually higher commitment and more open climate enabling better cooperation and results. The stress on mutual benefits among member organization is promoted strongly and therefore the complementary strengths are present. These can provide the competitive advantage and enhance the efficiency of the network as a whole. Nevertheless, these are not as flexible as hierarchal structures. The reason is in reaching consensus through the communication with all the members [9].

IV. MODELS OF NETWORKS

There are many diverse models available in the literature [9]. Some of them are completely different. However, similar aspects are sometimes found when comparing two or more models.

TABLE I. COMPARISON OF VARIOUS FORMS OF ORGANIZATIONS (ADAPTED ACCORDING TO [9])

Key Features	**Forms**		
	Market	***Hierarchy***	***Network***
Normative Basis	Contract	Employment relationship	Complementary strengths
Means of Communication	Prices	Routineso	Relational
Methods of Conflict Resolution	Haggling	Administrative fiat - supervision	Norm of reciprocity - reputational concerns
Degree of Flexibility	High	Low	Medium
Amount of Commitment among Parties	Low	Medium to high	Medium to high
Tone of Climate	Precision and/or suspicion	Formal, bureaucratic	Open-ended, mutual benefits
Actor Preferences or Choices	Independent	Dependent	Interdependent

The network and the social network analyses are needed for identifying roles and experts for particular areas and specifying the organizational flow ([5] and [10]). This also helps to determine the employed network model. A more appropriate model can possibly be revealed and considered for further use.

A. Fundamental Typology

Fundamental typology [4] can be used in many cases, even for the models of networks. This way, models can be differentiated based on production, distribution and industry.

There are three most typical models:

- Vertical networks represented by independent specialized companies.
- Cross-sectorial networks.
- Opportunity networks.

B. Typology according to the Dominance

This typology [4] distinguishes two models. The first one is a network with one dominant partner who communicates with all other partners in the network. Nevertheless, individual partners do not have to communicate with each other. The other one is a network in which all partners are equal. None of individual partners has power to change regulations and activities of the whole network. They have a limited authority and responsibility. The decision-making power is delegated and changes constantly.

C. Layered Model of Networks Based on Five Perspectives

The layered model as an example of a conceptual model of networking is based on the following five interdependent perspectives:

- Use of technology (mobile phones, voice recorders, camcorders, cameras, email, instant messaging, NAS servers, cloud computing).
- The nature of the work process.
- Organizational functions.
- Associational perspective.
- Societal perspectives.

The interdependency among all of the components is depicted in Fig. 2.

Obviously, the network organizations seem to penetrate to various areas of business. Nevertheless, not all organizations realize that they can be perceived as part of the network. The principles of networks should be emphasised and promoted by the idea of sharing and supported by the use of technology. There is a variety of technologies and platforms possibly beneficial for the communication and coordination of network organizations. Some examples include Microsoft SharePoint, Dropbox, IBM Connections or Microsoft SkyDrive. Nonetheless, these are mostly very robust or useless especially for the purposes of small- and medium-sized or non-profit organizations. Hence, the basic principles and approaches are introduced in this paper to cover the mentioned issues at a general level.

D. Structure Model of Networks

The structure model of networks as in Qureshi [11] differentiates two approaches – a relational approach focused on relations among actors (both individual and social actors) and a positional approach focused on their attributes (attributes such as relationships, people characteristics and behaviors). The main differences of these two types of models are represented in Fig. 3.

V. RESULTS, LIMITATIONS, AND FURTHER RESEARCH

Considering the appropriate utilization of the introduced typology and the general concept of the network organizations, the provided assumptions should be verified in practice. Moreover, the discussion with professionals from the network-based organization as well as from these not aware about this concept might be useful to reveal the point of view of the current and prospective end users. This would also provide the scope for the comparison and adjustment of the discussed models and issues for various purposes and types of organizations (based on their size, sector, financial means, technological readiness and the like).

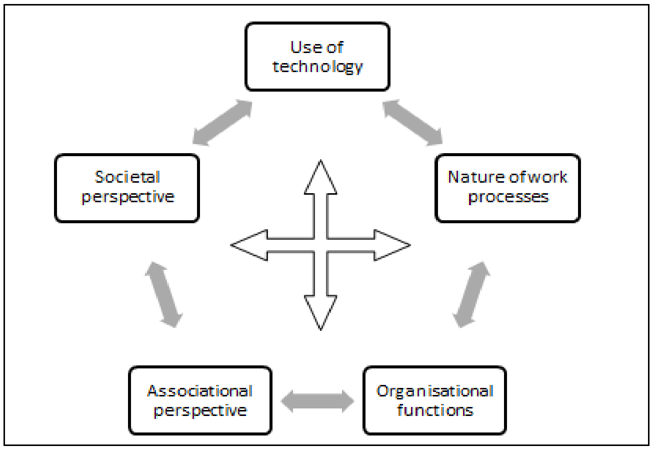

Figure 2. Layered model based on five perspectives [11].

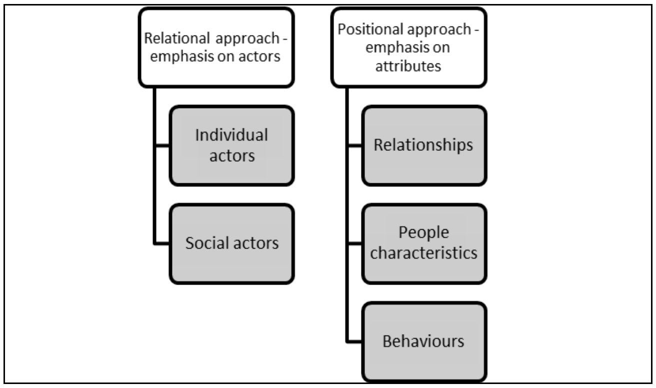

Figure 3. Structure model of networks (adapted according to [11])

Nevertheless, there are specific problems connected with networks. These can be exemplified by the lack of trust, weak engagement and willingness of all involved parties to participate, insufficient or unreliable technical equipment negatively influencing the successfulness and the benefits of the networks. This paper does not aim to cover these issues, because these should be addressed more in detail during the organizational processes at the operational level.

Furthermore, the general model may be created on the basis of the discussed ones. This should be more complex and flexible for the use under different conditions. As mentioned above, the proper testing should follow in pursuit to increase the relevancy and extent of practical implications.

VI. CONCLUSION

Network organizations represent a way to utilize advantageous practices within and among organizations. The employment of the suggested communication and coordination mechanisms and network principles might be beneficial from the internal as well as external environment of the organization. Therefore, particular suggestions which should be verified in practice are provided.

ACKNOWLEDGMENT

This paper is supported by the project No. CZ.1.07/2.2.00/28.0327 Innovation and support of doctoral study program (INDOP) financed from EU and Czech Republic funds. It is also written with the support of the specific research project 2/2013 “Cooperation mechanisms of network organizations” funded by the University of Hradec Králové, Czech Republic.

REFERENCES

[1] J. M. Podolny and K. L. Page, „Network forms of organization,“ in Annual Review of Sociology, vol. 24, no. 1, pp. 57-76, 1998.

[2] D. Parrish, “Ideas and tools – the network organisation,” 2013. http://www.d.parrish.dial.pipex.com/network.html. [retrieved: 6, 2013]

[3] M. Hahn, ”Characteristics of the network organization,” in Article Gratuits. 2007. http://en.articlesgratuits.com/characteristics-of-the-network-organization-id1486.php. [retrieved: 6, 2013]

[4] J. Dědina and J. Odcházel, “Management a moderní organizování firmy,” 1st ed. Grada Puublishing, Prague, 2007.

[5] K. C. Desouza and Y. Awazu, “Engaged knowledge management - engagement with new realities,” Palgrave Macmillan, New York, 2005.

[6] R. Clark, J. Timms, A. MacCartney, K. Egerton-Warburton, N. O’Dempsey, and B. Radokovich, “Achieving and enabling continuous improvement and innovation,” 10th Australian Agronomy Conference, pp. 501-508, 2001.

[7] K. Kolerová and T. Otčenášková, “The purpose and potential of clusters: theoretical background and a case study,” in 3rd World Conference on Information Technology, Barcelona, Spain, 14-16 Nov, 2012.

[8] A. Jashapara, “Kowledge management - an integrated approach,” 2nd ed. Prentice Hall, Pearson Education Limited, London, 2011.

[9] W. W. Powell, “Neither market nor hierarchy: network forms of organization,” in Research in Organizational Behavior, vol. 12. JAI Press Inc., pp. 295-336, 1990.

[10] L. Prusak and E. Matson, “Knowledge management and organisational learning,” Oxford University Press, Oxford, 2006.

[11] S. S. Qureshi, “Organisations and networks: theoretical considerations and a case study of networking across organisations,”. Doctoral Thesis, Univesity of London, 1995. http://is2.lse.ac.uk/research/theses/qureshi.pdf. [retrieved: 7, 2013]

Performance Evaluation of Lateration, KNN and Artificial Neural Networks Techniques Applied to Real Indoor Localization in WSN

Leomário Machado, Mauro Larrat and Dionne Monteiro
Research Group on Computer Networks and Multimedia Communication
Federal University of Para – UFPA
Belém/PA - Brazil
leomariomachado@gmail.com, maurolarrat@gmail.com, dionne@ufpa.br

***Abstract*-In Wireless Sensor Networks, several protocols and algorithms seek to prolong the network lifetime; among them, the localization algorithms are used as an accessory to provide the smallest distances for sending messages. This paper compares the Lateration, KNN and ANN as localization techniques to estimate planar coordinates using the RSSI in an indoor environment using a real WSN based on IRIS motes. The results show that a well worked out ANN is superior to Lateration and KNN.**

Keywords-Wireless Sensors Network; RSSI; Artificial Neural Network; localization algorithms; Lateration; KNN.

I. INTRODUCTION

Wireless Sensors Networks (WSNs) have many applications in various segments of society, including areas such as military surveillance, industrial and agricultural monitoring and residential automation [1]. They operate with low-power consumption devices, low hardware processing, and higher autonomy [2].

There are various protocols, algorithms and techniques in which the objective is to extend the network lifetime as security systems [1][2], data fusion [3], tracking [4], etc. Several metrics are approached to improve network routing accuracy considering reducing energy consumption. The sensor node localization in WSN is noteworthy as one of the more effective metrics to reduce energy consumption, since the longer the distance the greater the energy consumed by a node for routing information.

GPS (Global Positioning System) usage is a very popular method when it comes to sensors localization. However, it has some disadvantages in WSN, such as the use of up to four satellite signals to achieve a good localization, the need of high power consumption due to the receiving of satellite signals communication and the high cost of using a GPS in each sensor node [3]. Leaving this method aside, other alternative techniques are able to compose a WSN localization system [4].

Some well-known localization techniques are Lateration, Nearest Neighbor (NN), K-Nearest Neighbor (KNN), Min-Max [4], ANN [5], Kalman Filter [6], Time of Arrival (TOA) [3], Time Frequency of Arrival (TFOA) [7] and Least-Square Support Vector Regression (LSSVR) [8].

The Lateration, Min-Max, KNN and ANN is a set of techniques that use the Received Signal Strength Indicator (RSSI) for estimating node localization in WSN. It is a measure of the signal power on the radio link while a message is being received. Thus, the power of RSSI is inversely proportional to the distance (1).

Featuring the RSSI, it is possible to:

- Calculate in matrix the location of each point (Lateration),
- Estimate the coordinates as weights in the arithmetic average of its nearest neighbors (KNN),
- Insert inputs in an Artificial Neural Network (ANN) or,
- Obtain the geographical location of a network node using other techniques and filters.

This paper presents an experimental evaluation that compares the following techniques: Lateration [9], KNN [10] and ANN designed for localization. All the experiments were performed based on a set of RSSI collected and applied to localization techniques in the study in a real environment.

The paper is divided into seven sections including this introduction. In Section 2, a survey of related work is performed, showing localization techniques without the use of GPS, using the RSSI as a metric for estimating planar coordinates. In Section 3, the methodology presenting the test environment, hardware and software is discussed, and a brief introduction about the Lateration, KNN and ANN techniques is given. Section 4 presents the benchmark results of comparing the Lateration, KNN and ANN techniques. Later, in Section 5, we present our final remarks, followed by our suggestions for future work in Section 6.

II. RELATED WORKS

This section introduces some researches about performance evaluations of WSN, highlighting comparisons between the localization techniques.

Priwgharm and Chemtanomwong [4] performed a comparative study of the Lateration, Min-Max, NN and KNN techniques in an experimental indoor environment measuring 3x3 meters using XBee modules. The article shows Lateration as the technique with a smaller margin of error than the KNN technique.

 ISBN: 978-1-61208-315-5

Rice and Harle [11] compared the localization algorithms Non-Linear Regression (NLR), Iterative Non-Linear Regression (INLR), Least Squares (LS), Random Sample Consensus (RANSAC) and Trilaterate on Minima (ToM) is shown. Data were collected from real environments in an area of 550 m^2. The radios use Ultra-Wide Band (UWB) (IEEE 802.15.4a) in their physical layer for transmitting and receiving signals. The INLR (Iterative Non-Linear Regression) technique gives the smallest number of estimate of location errors.

Shareef et al. [12] made a comparative work in a real environment is performed among several ANN based on methods such as Radial Basis Function (RBF), Multi-Layer Perceptron (MLP), Recurrent Neural Networks (RNN), Position-Velocity (PV), Position Velocity Acceleration (PVA) and Reduced Radial Basis Function (RRBF). These methods evaluate the location errors in centimeters using these techniques in a 3x3 meters indoor environment. However, they do not compare the performance between some of the ANN methods with the common localization methods applied in a real environment; their results do not show a benchmark among the solutions.

Tian and Xu [13] performed in a real environment a comparison between a Multi-Layered Perceptron Neural Network (MLPNN) model and two Kalman Filter models, namely PV and PVA techniques. The environment measures 3x3 meters, marked in grid spacing of 0.30 meters with four beacon nodes located in the vertices of a square. The mobile nodes were placed on each intersection of the grid to collect the data. The experimental results indicate that the MLPNN neural network has the best performance, but there is a potential retraining or redesign cost associated with the use of MLP, which is not associated with the Kalman's filter [13]. This work only compares Kalman's filter with MLPNN, not with tests performed using mathematical techniques.

Rahman et al. [14] made a simulation is performed in an environment measuring 200x200 meters, where they implement an ANN with the Levenberg-Marquardt algorithm [14] using an error minimization function. The simulation shows that the location accuracy improves with the increase of grid sensors density and Access Points (APs). The work is restricted to simulation and does not present the benchmark among the techniques analyzed, disadvantaging a clearer comparative preview.

Langendoen and Reijers [15] simulated Lateration and Min-Max algorithms together with the sum of the distances, DV-Hop and Euclidean algorithms. They perform only simulation and it is described as an ideal environment, without discussing the effect of path loss or variable depending on the Lateration.

Zheng and Dehghani [16] made a new ANN method known as Location Neural Network Ensembles (LNNE) is compared with DV-Hop and a Localization Signal Neural Network (LSNN). The experimental results demonstrate that LNNE consistently outperforms the other three algorithms in localization accuracy. This work employs a simple scenario controlled by simulation, but it does not perform a validation of the theoretical aspects in real environments.

Maheshwari and Kemp [17] compared three methods of localization, namely Optimal Multi-Lateration (OML), Sub-optimal Blind Trilateration (SBT) and Geometric Dilution of Precision (GDOP). They provide the benchmark in terms of achievable accuracy. However, they only compare the performance among Lateration variants and not techniques of a different nature, and their approach only uses simulation without conducting any real experiment.

According to an analysis of related work, many of the studies involve the use of Lateration or some of its aspects [4][11][15][17], while other studies [12][13][14][16] compare the various topologies of ANNs. An experimental approach is presented in [4], in which Lateration shows the best results compared with three other techniques (KNN, NN and Min-Max). Based on the results presented in the above references, this paper conducts a performance evaluation using the two best methods presented in [4] (Lateration and KNN) to compare with ANN.

Since most of the current works are restricted to the use of simulation, and experimental works are much scarcer, the main contribution of this paper consists in establishing a practical comparison between some common classification techniques employed to use locating sensors in real environments.

III. METHODOLOGY

This section presents the methodology to acquire the data and characteristics from the environment of the experiment, besides the approach to determine the architecture of the artificial neural network used in this research.

A. Environment

The tests were conducted in an indoor environment grid of 12x12 meters, as seen in Fig 1. The beacon nodes (red) are positioned in the corners of the scenario, points (0, 0), (0, 12), (12, 0) and (12, 12), and the nodes to be localized (blue) are positioned in the internal area of the grid.

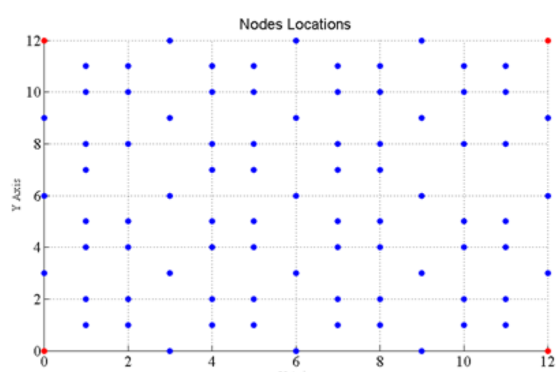

Fig. 1 Points in which the RSSI signals are collected through the target node

This area is configured in an indoor environment free of obstacles or barriers that may degrade the radio signal propagation.

B. Hardware

For real tests, this research uses the IRIS motes of MEMSIC, which implement the IEEE 802.15.4 radio standard in the range of 2.4 to 2.48 GHz and a maximum data rate of 250 kbps. This hardware can transmit data at a distance of up to 500 meters in direct view. The IRIS hardware is based on XM2110CB radio and ATmega1281 CPU, both from ATMEL. To support the WSN node gateway, a MIB520 is used, which consists of a USB adapter that enables two serial communications, one for control and one for data.

C. Middleware

We used the TinyOS platform as a software component running on the sensor nodes. This is an operating system with open source BSD license devices designed for low power consumption such as WSN, ubiquitous computing, personal networks, home automation and industrial [18].

A Java-based RSSI application was developed to extract the data in real time. The data are recorded in a file with the following fields: actual location, node identifier and the RSSI value obtained by target node.

D. Lateration Algorithm for Localization

A target node receives a message from each beacon node, in which it can get the value of the power in dBm from the signal, as shown in (1), or as a function of distance (in meters) according to (2).

$$RSSI[dbm] = -10 * n * \log(d)_{10} + A \quad (1)$$

$$d = 10^{\frac{Rssi-A}{-10*n}} \quad (2)$$

The values of A (1 meter RSSI distance) and n (path loss) are estimated before testing localization, thereby obtaining the value of the distance in meters from the anchor nodes to the target node. For each message received by an anchor node, the distance d_i is determined from the (3).

$$d_i^{\,2} = (x - x_i)^2 + (y - y_i)^2 \quad (3)$$

where x and y are coordinate values related to the target, x_i and y_i are coordinate values related to the i-th beacon node, and d is the distance between the target and the i-th beacon node.

For each i visible anchor nodes, there will be one distance and one equation as a function of x and y coordinates. They can be summarized in matrix form as shown in (4). B is the estimated localization vector from coordinates x and y. Equations (4) – (7) are found in [4].

$$\mathrm{B} = \mathrm{A}\,\mathrm{X}^{-1} \quad (4)$$

where,

$$\mathrm{B} = \begin{bmatrix} x \\ y \end{bmatrix} \quad (5)$$

$$A = \begin{bmatrix} x_1^2 - x_n^2 + y_1^2 - y_n^2 - d_1^2 + d_n^2 \\ \vdots \\ x_{n-1}^2 - x_n^2 + y_{n-1}^2 - y_n^2 - d_{n-1}^2 + d_n^2 \end{bmatrix} \quad (6)$$

and,

$$\mathrm{X} = \begin{bmatrix} 2(\mathrm{x_1} - \mathrm{x_n}) & 2(\mathrm{y_1} - \mathrm{y_n}) \\ \vdots & \vdots \\ 2(\mathrm{x_{n-1}} - \mathrm{x_n}) & 2(\mathrm{y_{n-1}} - \mathrm{y_n}) \end{bmatrix} \quad (7)$$

E. KNN Algorithm for Localization

KNN is one of the most popular methods of classification due to its simplicity and reasonable effectiveness. It requires the assembling of a specific model and has shown good performance for the classification of various types of data [19].

As shown in (1), the RSSI is inversely proportional to the distance. To calculate the distance, we use the largest RSSI values that correspond to the closest nodes. Subsequently, we calculate the weighted average to estimate the values of x and y coordinates from the target node, as shown in (8).

$$\begin{cases} x = \dfrac{\sum_{i=1}^{i \le k}(RSSI_i * x_i)}{\sum_{i=1}^{i \le k} RSSI_i} \\ y = \dfrac{\sum_{i=1}^{i \le k}(RSSI_i * y_i)}{\sum_{i=1}^{i \le k} RSSI_i} \end{cases} \quad (8)$$

where x and y are the coordinate values of the target node, $RSSI_i$ is the power value of the received signal from the i-th node, x_i and y_i are fixed coordinate values of the i-th node, k is the limit number of neighboring nodes. Then, the weighted average is calculated to estimate the values of the x and y coordinates of the target node.

F. Artificial Neural Network

Our application of ANN used for the localization of nodes in WSN is accomplished through the acquisition of RSSI from the messages sent from each node. Subsequently, these data are used to train and validate the network.

For each location point, the mean and standard deviation are calculated to identify the 10% worst samples. With this, only accurate samples are used to train the ANN.

To train the ANN, the collected points (x, y) are normalized maintaining those data output values between 0 and 1.

To define the architecture of the ANN, several tests are done to achieve a satisfactory configuration of the ANN, i.e., number of hidden layers, number of neurons in each layer, activation function for each layer or neuron and training algorithm.

The architecture used in the ANN is a feed-forward network with five layers: the input layer with 4 neurons (RSSI values of each beacon nodes), three hidden layers with 10 neurons in each and the output layer with two neurons (x and y). The learn rate used is 0.7. The parameters of the ANN are summarized in Table 1.

TABLE I. ANN CONFIGURATIONS

Training	SARPROP
Activation Function for output layer	Symmetric Gaussian
Activation Function for hidden layers	Symmetric Cosine
Training error function	Linear
Training stop function	MSE
Input Layer	4 neurons + 1 bias
Hidden Layers	3 layers with 10 neurons + 1 bias for each
Output Layers	2 neurons
Total neurons	40

The network training is preceded by the choice of training method, activation function, number of hidden layers and the number of neurons in each layer. The training stops after 500 000 epochs.

IV. RESULTS

The absolute values of positioning errors in the coordinates x (x_{error}), y (y_{error}) and the Localization error (l_e) are used as a metric for comparing the techniques. The position error is the distance between the actual point and the estimated value as shown in equations (9), (10) and (11), respectively.

$$x_{error} = |x_{real} - x_{estimated}| \quad (9)$$

$$y_{error} = |y_{real} - y_{estimated}| \quad (10)$$

$$l_e = \sqrt{(x_{error})^2 + (y_{error})^2} \quad (11)$$

To compare the results between Lateration, KNN and ANN techniques, we use the equations (12) and (13) to generate the graphs of the benchmark shown in Figs. 3 and 4, respectively.

$$bm_{ANN\times LAT} = 100 * \frac{l_{e_{LAT}} - l_{e_{ANN}}}{l_{e_{LAT}}} \quad (12)$$

$$bm_{ANN\times KNN} = 100 * \frac{l_{e_{KNN}} - l_{e_{ANN}}}{l_{e_{KNN}}} \quad (13)$$

where,

$bm_{ANN\times LAT}$: Benchmark of ANN versus Lateration
$bm_{ANN\times KNN}$: Benchmark of ANN versus KNN
$l_{e_{ANN}}$: Localization error obtained from ANN in meters
$l_{e_{LAT}}$: Localization error obtained from Lateration in meters
$l_{e_{KNN}}$: Localization error obtained from KNN in meters

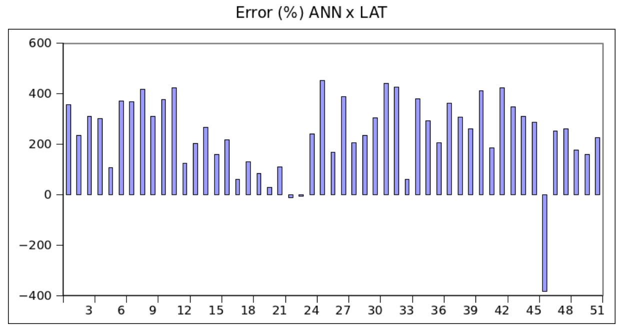

Fig. 2 Percentage comparison between the Artificial Neural Network and Lateration

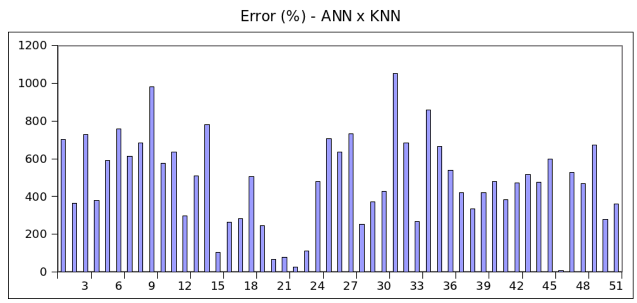

Fig. 3 Percentage comparison between the Artificial Neural Network and KNN

In Fig. 2, we can observe the comparison between ANN and Lateration, in percentage, where positive values indicate that ANN has better results than Lateration. Similarly, Fig. 3 shows a comparison between ANN and KNN. To illustrate these graphs, only 50 points were selected from a total of 255 used to estimate the localization (blue points in Fig. 1). Negative values occur whenever the ANN technique does not yield better results in the comparison.

The ANN achieved better results in the localization with average 232.43% better than Lateration and 470.45% better than KNN. Table 2 shows the percentage of cases in which each technique has the best performance for the aforementioned metrics.

The x and y values were estimated to apply filters in future works that allow reducing the error estimated by a coordinated reducing of the distance error.

The data in Table 2 show that ANN obtains 76.0784% and 79.6078% of the best samples in the x and y coordinates, respectively, when compared to Lateration. In the comparison of KNN and ANN, Table 2 shows that 85.8824% and 88.6275% of samples have better results for x and y coordinates, respectively. Therefore, the ANN shows better results when compared with Lateration and KNN.

TABLE II. BEST PERFORMANCE: ANNxLAT AND ANNxKNN

Comparison	**ANN x LAT**		**ANN x KNN**	
Metric	*%ANN*	*%LAT*	*%ANN*	*%KNN*
x	76,0784	23,9216	85,8824	14,1176
y	79,6078	20,3922	88,6275	11,3725
Localization error	94,902	5,098	98,0392	1,9608

When we consider the distance of the estimated localization and real localization, ANN has better results than Lateration and KNN, i.e., the ANN estimates the localization near the real localization in 94.902% compared to Lateration and 98.0392% compared to KNN.

Table 3 shows the average (Avg) and variance (Var) of the positioning error of each technique. The mean and variance of the positioning error in the ANN are lower than those calculated for Lateration and KNN.

TABLE III. MEAN AND VARIANCE OF THE POSITION ERROR IN METERS FOR EACH TECHNIQUE

	LAT		**KNN**		**ANN**	
	Avg	Var	Avg	Var	Avg	Var
error *x*	1,76	1,27	2,68	3,55	0,80	0,51
error *y*	2,02	1,34	3,77	4,68	0,88	0,88
error distance	2,94	1,18	5,01	4,49	1,30	1,12

TABLE IV. QUANTITATIVE ANALYSIS OF EVERY 1 METER IN DISTANCE ERROR

Localization error	**%LAT**	**%KNN**	**%ANN**
$0 \leq l_e < 1$	5,09	1,96	48,62
$1 \leq l_e < 2$	19,21	6,27	29,80
$2 \leq l_e < 3$	22,74	8,23	15,29
$3 \leq l_e < 4$	35,29	14,50	3,52
$4 \leq l_e < 5$	17,64	21,17	2,74
$l_e \geq 5$	0	47,84	0

In Table 4, samples are organized by zones of error. For errors which are smaller than 1 meter, the ANN shows the best results with a large advantage because 48.62% of the estimated positions' locations are within the lower zone of error, i.e., error < 1 meter.

Figs. 4 and 5 show some samples of estimated locations. Each real point is illustrated by a square connected to its three estimates by a dashed line.

Fig. 4 shows samples located in the periphery of the scenario. With this figure, we can observe the behavior of these techniques near the borders. Fig. 5 shows the internal points situated in the innermost region of the experiment.

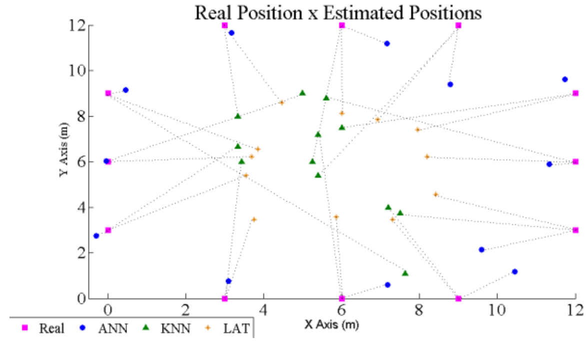

Fig. 4 Real and estimated positions by technique for peripheral points

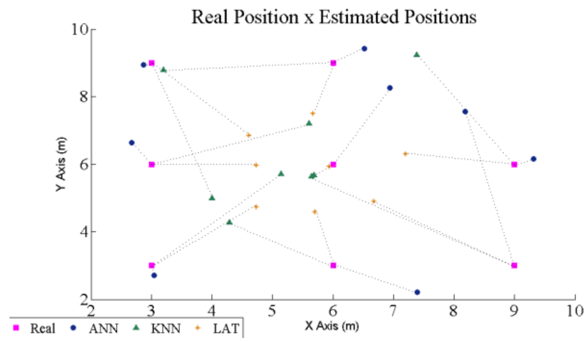

Fig. 5 Real and estimated positions by technique for internal points

To compare the localization techniques presented here, we estimate the size of the accumulated error indicating how each technique deviated from the actual value for each record as illustrated in Fig. 6.

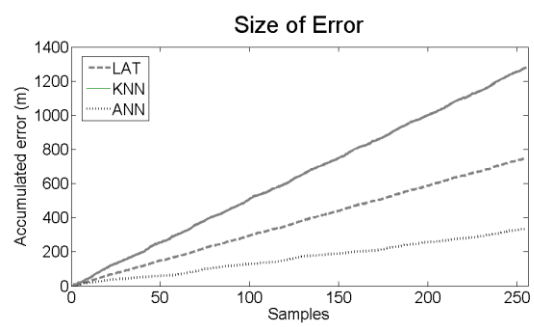

Fig. 6 Size of accumulated error (m)

Fig. 6 shows the near-linear behavior resulting from the accumulation of error over the samples. The final accumulated error of Lateration and KNN is 225.05% and 383.61% greater than the final accumulated error observed in ANN, respectively.

V. FINAL CONSIDERATIONS

The application of a neural network to resolve the localization problem shows satisfactory results in the estimation of a node within the limits of the beacon nodes. This method had lower mean and error variance when compared to the techniques of Lateration and KNN, therefore having increased stability. However, the number of inputs is fixed and directly related to the number of beacon nodes, resulting in the need to retrain for different amounts of network beacon nodes.

The Lateration is a technique for easy adjustment, depending on the number of beacon nodes maintaining stability similar to ANN. This technique requires an accuracy of the path loss of the environment and the value of the constant *A*, otherwise, the error can be increased and may exceed the expected limits.

In this experiment, the KNN technique was performed by the weighted sum of each coordinate by dividing the sum of their weights. In KNN, the estimated location tends to be found within the boundaries of the scenario, due to the weighted sum of RSSIs from their neighbors.

In [4], the performance evaluation indicates the Lateration technique has better precision, followed by KNN. In this study, both techniques were compared to ANN for localization. We observed that ANN had lower error and, consequently, better precision in locating the target node, as show in the analysis of the mean and variance of the error for both x and y coordinates and localization error, creating good reliability in the use of ANN compared to classical techniques for locating wireless sensor networks.

VI. FUTURE WORK

As future work, we intend to perform the target tracking in real time by applying this current artificial neural network to perform this task, test for different amounts of beacon nodes and test these techniques with other hardware, and apply filters to minimize error for coordinate and distance.

ACKNOWLEDGMENTS

The team is grateful to Amazon Research Foundation, CNPq (National Council for Scientific and Technological Development) and FINEP (Financing Agency for Studies and Projects), and UFPA (Federal University of Pará).

REFERENCES

[1] L. Chunxia, F. Chen, Y. Zhan and L. Wang, “Security verification of localization estimate in wireless sensor networks,” 6th International Conference on Wireless Communications Networking and Mobile Computing (WiCOM 10), IEEE Press, Sept. 2010, pp. 1-4, doi:10.1109/WICOM.2010.5601183.

[2] A. Gaur, S. Toshniwal, A. Prakash and D. Agrawal, “Enhanced localization based key pre-distribution scheme for secure communication in wireless sensor network (WSN),” 7th International Conference on Mobile Adhoc and Sensor Systems (MASS 10), IEEE Press, Nov. 2010, pp. 8–12, doi:10.1109/MASS.2010.5663897.

[3] L. Lichuan and E. Manli, “Improve the positioning accuracy for wireless sensor nodes based on TFDA and TFOA using data fusion,” International Conference on Networking Sensing and Control (ICNSC 10), IEEE Press, Apr. 2010, pp. 32–37, doi:10.1109/ICNSC.2010.5461554.

[4] R. Priwgharm and P. Chemtanomwong, “A comparative study on indoor localization based on RSSI measurement in wireless sensor network,” Eighth International Joint Conference on Computer Science and Software Engineering (JCSSE 11), IEEE Press, May 2011, pp. 1–6, doi:10.1109/JCSSE.2011.5930074.

[5] S. Rajaee, S. AlModarresi, M. Sadeghi and M. Aghabozorgi, “Energy efficient localization in wireless ad-hoc sensor networks using probabilistic neural network and independent component analysis,” International Symposium on Telecommunications (IST 08), IEEE Press, Aug. 2008, pp. 365–370, doi:10.1109/ISTEL.2008.4651329.

[6] Y. Jieyang and Z. Liang, “Enhanced localization algorithm with received signal strength using fading Kalman filter in wireless sensor networks," International Conference on Computational Problem-Solving (ICCP 11), IEEE Press, Oct. 2011, pp. 458–461, doi:10.1109/ICCPS.2011.6089930.

[7] H. Hsi-Chou, W. Jyh-Horng, C. Chia-Hsin, E. Yi-Shiang and W. Tian-Yue, “TOA estimation with DLC receivers for IEEE 802.15.4a UWB systems,” Fifth International Conference on Innovative Mobile and Internet Services in Ubiquitous Computing (IMIS 11), IEEE Computer Society, pp. 424–428, doi:10.1109/IMIS.2011.119.

[8] H. Xuewen, X. Yong and W. Yanmeng, “A LSSVR three-dimensional WSN nodes localization algorithm based on RSSI,” International Conference on Electrical and Control Engineering (ICECE 11), IEEE Press, Sept. 2011, pp. 1889–1895, doi:10.1109/ICECENG.2011.6058085.

[9] J. Neto, J. Neto, Y. Yang and I. Glover, “Plausibility of practical low-cost localization using WSN path-loss law inversion,” Proceedings of IET International Conference on Wireless Sensor Network (IET-WSN 10), IEEE Press, Nov. 2010, pp. 15-17, doi:10.1049/CP.2010.1064.

[10] H. Chun-Chieh, S. Yi-Jing, L. Seng-Yong, C. Chia-Hui, H. Fei-Hsiu, C. Hao-Hu and H. Polly, “Towards long-term mobility tracking in NTU hospital’s elder care center,” IEEE International Conference on Pervasive Computing and Communications Workshops (PERCOM 11), IEEE Press, March 2011, pp. 649–654, doi:10.1109/PERCOMW.2011.5766969.

[11] A. Rice and R. Harle, “Evaluating lateration-based positioning algorithms for fine-grained tracking,” Proceedings of the 2005 joint workshop on Foundations of mobile computing (DIALM-POMC 05), ACM, Sept. 2005, pp. 54–61, doi:10.1145/1080810.1080820.

[12] A. Shareef, Y. Zhu and M. Musavi, “Localization using neural networks in wireless sensor networks,” Proceedings of the 1st international conference on MOBILe Wireless MiddleWARE, Operating Systems, and Applications (MOBILWARE 08), ACM, Feb. 2008, pp. 1–7.

[13] J. Tian and Z. Xu, “RSSI localization algorithm based on RBF neural network,” 3rd Int. Conf. on Software Engineering and Service Science (ICSESS 12), IEEE Press, Jun. 2012, pp. 321–324, doi:10.1109/ICSESS.2012.6269470.

[14] M. Rahman, P. Youngil and K. Ki-Doo, “Localization of wireless sensor network using artificial neural network,” 9th International Symposium on Communications and Information Technology (ISCIT 09), IEEE Press, Sept. 2009, pp. 639–642, doi:10.1109/ISCIT.2009.5341165.

[15] K. Langendoen and N. Reijers, “Distributed localization in wireless sensor networks: A quantitative comparison,” Computer Networks: The International Journal of Computer and Telecommunications Networking - Special Issue: Wireless Sensor Networks, ACM, vol. 43, Nov. 2003, pp. 499–518, doi:10.1016/S1389-1286(03)00356-6.

[16] J. Zheng and A. Dehghani, “Range-free localization in wireless sensor networks with neural network ensembles,” Journal of Sensor and Actuator Networks, vol. 1, no. 3, Sept. 2012, pp. 254–271, doi:10.3390/JSAN1030254

[17] H. K. Maheshwari and A. H. Kemp, “Comparative performance analysis of localization using optimal and sub-optimal lateration in

WSNs," Third International Conference on Next Generation Mobile Applications, Services and Technologies (NGMAST 09), IEEE Press, Sept. 2009, pp. 369–374, doi:10.1109/NGMAST.2009.97.

[18] TinyOs. TinyOs Home Page [Online]. Available: http://www.tinyos.net/.

[19] R. Min, D. A. Stanley, Z. Yuan, A. Bonner and Z. Zhang, "A deep non-linear feature mapping for large-margin KNN classification," Proceedings of the 2009 Ninth IEEE International Conference on Data (ICDM 09), ACM, Dec. 2009, pp. 357–366, doi:10.1109/ICDM.2009.27.

[20] Tool for the fast artificial neural network library [Online]. Available: https://code.google.com/p/fanntool/.

[21] Fast artificial neural network library [Online]. Available: http://leenissen.dk/fann/wp/.