EMERGING 2010

The Second International Conference on Emerging Network Intelligence

October 25-30, 2010 - Florence, Italy

Editors

Tulin Atmaca

Michael D. Logothetis

EMERGING 2010

Foreword

The Second International Conference on Emerging Network Intelligence (EMERGING 2010) held from October 25 to October 30, 2010 in Florence, Italy, constituted a stage to present and evaluate the advances in emerging solutions for next-generation architectures, devices, and communications protocols. Particular focus was aimed at optimization, quality, discovery, protection, and user profile requirements supported by special approaches such as network coding, configurable protocols, context-aware optimization, ambient systems, anomaly discovery, and adaptive mechanisms.

Next-generation large distributed networks and systems require substantial reconsideration of existing 'de facto' approaches and mechanisms to sustain an increasing demand on speed, scale, bandwidth, topology and flow changes, user complex behavior, security threats, and service and user ubiquity. As a result, growing research and industrial forces are focusing on new approaches for advanced communications considering new devices and protocols, advanced discovery mechanisms, and programmability techniques to express, measure, and control the service quality, security, environmental and user requirements.

We take here the opportunity to warmly thank all the members of the EMERGING 2010 technical program committee as well as the numerous reviewers. The creation of such a broad and high quality conference program would not have been possible without their involvement. We also kindly thank all the authors that dedicated much of their time and efforts to contribute to the EMERGING 2010. We truly believe that thanks to all these efforts, the final conference program consists of top quality contributions.

This event could also not have been a reality without the support of many individuals, organizations and sponsors. We also gratefully thank the members of the EMERGING 2010 organizing committee for their help in handling the logistics and for their work that is making this professional meeting a success. We gratefully appreciate to the technical program committee co-chairs that contributed to identify the appropriate groups to submit contributions.

We hope Florence provided a pleasant environment during the conference and everyone saved some time for exploring this historic city.

EMERGING 2010 Chairs

Tulin Atmaca, IT/Telecom&Management SudParis, France
Raj Jain, Washington University in St. Louis, USA
Michael D. Logothetis, University of Patras, Greece
Naoki Wakamiya, Osaka University, Japan
Robert Forster, Edgemount Solutions - Plano, USA

Corrado Moiso, Telecom Italia, Italy
Krishna Murthy, Infosys, USA
David Carrera, Barcelona Supercomputing Center (BSC) / Universitat Politecnica de Catalunya (UPC), Spain
Peter Deussen, Fraunhofer Research Institute for Open Communication Systems - Berlin, Germany
Daniel Scheibli, SAP Research, Germany

EMERGING 2010

Committee

EMERGING Advisory Chairs
Tulin Atmaca, IT/Telecom&Management SudParis, France
Raj Jain, Washington University in St. Louis, USA
Michael D. Logothetis, University of Patras, Greece
Naoki Wakamiya, Osaka University, Japan

EMERGING 2010 Industry Liaison Chairs
Robert Forster, Edgemount Solutions - Plano, USA
Corrado Moiso, Telecom Italia, Italy
Krishna Murthy, Infosys, USA

EMERGING 2010 Research/Industry Chairs
David Carrera, Barcelona Supercomputing Center (BSC) / Universitat Politecnica de Catalunya (UPC, Spain
Peter Deussen, Fraunhofer Research Institute for Open Communication Systems - Berlin, Germany
Daniel Scheibli, SAP Research, Germany

EMERGING 2010 Technical Program Committee
Ozgur B. Akan, Middle East Technical University - Ankara, Turkey
Khalid Al-Begain, University of Glamorgan - Pontypridd, UK
Artur Andrzejak, Zuse Institute Berlin (ZIB), Germany
Richard Anthony, The University of Greenwich, UK
Tadashi Araragi, Nippon Telegraph and Telephone Corporation - Kyoto, Japan
Tulin Atmaca, IT/Telecom&Management SudParis, France
Borbala Katalin Benko, Budapest University of Technology and Economics, Hungary
Andreas Berl, University of Passau, Germany
Robert Bestak, Czech Technical University in Prague, Czech Republic
Christian Blum, Universitat Politècnica de Catalunya, Spain
David Carrera, Barcelona Supercomputing Center (BSC) / Universitat Politecnica de Catalunya - UPC, Spain
Chih-Yung Chang, Tamkang University, Taiwan
Dong-Ho Cho, KAIST, Korea
Reuven Cohen, Technion-Israel Institute of Technology/Haifa, Israel
Alberto Dainotti, University of Napoli "Federico II", Italy
Carl James Debono, University of Malt, Malta
Peter Deussen, Fraunhofer Research Institute for Open Communication Systems - Berlin, Germany
Rolf Drechsler, University of Bremen, Germany
Tarek El-Bawab, Jackson State University, USA
Mohamed Eltoweissy, Virginia Tech, USA
Henrique Ferreiro, University of Coruña, Spain
Robert Forster, Edgemount Solutions - Plano, USA
Anna Förster, Networking Laboratory, SUPSI, Switzerland
Niloy Ganguly, Indian Institute of Technology, India

Nuno M. Garcia, Universidade Lusófona de Humanidades e Tecnologias - Lisbon, Portugal
Costas Georghiades, Texas A&M University, USA
Yan Gong, Beijing University of Posts & Telecommunications (BUPT) - Beijing, P. R. China
Christophe Guéret, Vrije Universiteit Amsterdam, The Netherlands
Go Hasegawa, Osaka University, Japan
Eva Hladka, Masaryk University, Czech Republic
Li-Ling Hung, Aletheia University, Taiwan
Raj Jain, Washington University in St. Louis, USA
Michael D. Logothetis, University of Patras, Greece
Radu Lupu, Politehnica University of Bucharest, Romania
Ahmed M. Mahdy, Texas A&M University - Corpus Christi, USA
Josemaria Malgosa Sanahuja, Polytechnic University of Cartagena, Spain
Zoubir Mammeri, IRIT - Toulouse, France
Anna Medve, University of Pannonia, Hungary
Andrej Mihailovic, King's College London, UK
Corrado Moiso, Telecom Italia, Italy
Maurice Mulvenna, University of Ulster, UK
Juan Pedro Muñoz-Gea, Polytechnic University of Cartagena, Spain
Krishna Murthy, Infosys, USA
Tadashi Nakano, University of California, Irvine, USA
Oznur Ozkasap, Koc University - Istanbul, Turkey
Euthimios (Thimios) Panagos, Telcordia Applied Research - Piscataway, USA
Gianluca Reali, Università degli Studi di Perugia, Italy
Joel Rodrigues, Institute of Telecommunications / University of Beira Interior, Portugal
Daniel Scheibli, SAP Research, Germany
Thomas C. Schmidt, Hamburg University of Applied Sciences, Germany
Patrick Sénac, ISA/ Université de Toulouse, France
Dimitrios Serpanos, ISI/R. C. Athena and University of Patras, Greece
Yutaka Takahashi, Kyoto University, Japan
Jim Torresen, University of Oslo, Norway
Davide Tosi, University of Insubria - Como, Italy
Phuoc Tran-Gia, University of Wuerzburg, Germany
Athanasios Vasilakos, Technical University of Athens (NTUA), Greece
Manuel Villen-Altamirano, Universidad Politécnica de Madrid, Spain
Naoki Wakamiya, Osaka University, Japan
Maarten Wijnants, Hasselt University, Belgium
Moustafa Youssef, Nile University, Egypt

Copyright Information

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

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

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

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

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

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

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

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

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

Table of Contents

An Application of Pattern Matching for the Adjustment of Quality of Service Metrics

Doug Legge and Atta Badii
IMSS
University of Reading
Reading, England
{d.j.s.legge, atta.badii}@reading.ac.uk

Abstract—**Quality of Service is an important component of Internet Protocol traffic as it allows a prioritisation of designated applications during periods of high utilisation or where there is restricted resource, such that the end-user experience can be optimised. Typically, Quality of Service is defined through manual policies with expert human input required. In this paper, we present initial support for the hypothesis that the definition and on-going change management of manual Quality of Service policies can be replaced through the use of pattern matching techniques, which classify traffic *in real-time*. This paper serves as an introduction to concepts, which may be new to many Internet Protocol-based network engineers, and as a motivation for those in the field of Artificial Intelligence, machine-learning, as to where advanced learning functions could be applied.**

Keywords-Quality of Service (QoS); self organizing map (SOM); k Nearest Neighbour (kNN); agent

I. CONCEPT

Quality of Service (QoS) policies provide for the prioritisation of packets on IP networks, partly motivated in the early 2001 by the convergence of voice and data traffic.

Despite the increasing availability of high-speed consumer (e.g., xDSL ≈50Mbps) and corporate (e.g., Ethernet ≈10Mbps) data links, Internet Protocol (IP) engineers now face the conundrum, once seen in the Personal Computer (PC) world, where a faster network resource is rapidly consumed at an increasing rate by bandwidth hogging applications, such as Videoconferencing or Tele-presence. As a result little planned capacity overhead remains and QoS mechanisms continue to be required.

With no end in sight, the dependence of QoS implementation upon expert judgement leaves organisations exposed to high salary costs and a potential loss of critical knowledge resulting from staff churn. In addition, these existing optimisation techniques lead to increasingly complex network operations regards the service mechanisms required to deliver appropriate application performance. Including those supporting activities such as 'requirements analysis' and 'policy change management control'. Verma [1] states "as networks make the transition from *all traffic is equal* to the new model in which some *traffic is more equal than others*", a way must be found in which to specify differentiate and service traffic types on the network whilst maintaining a simplified abstraction.

Whilst optimisation of traffic-flow is a valid and well researched field, simplification of its engineering and on-going management has, in the first author's experience as the IT Operations Manager for a UK FTSE 250 company, been long overdue at the coalface of network support.

For some years now machine learning, and in particular 'neural-networks', have been used within many industries [2][3][4]. However, it is in the field of data mining that neural networks have been most productive, being used to "extract new information, from existing data, thus providing innovative insights and tactical commercial benefit" [5].

The authors' previous work [6][7] highlighted those issues corporate organisations face regarding this 'requirements analysis phase' necessary to collate that information required to build network traffic services (e.g., QoS policy statements). This work demonstrated how these policies in practice remain sub-optimally implemented through lack of a facility to allow their dynamic adaptation responsive to changing circumstances and thus changing priorities of different business data traffic types.

In this report the authors' present paper results from initial experimentation regards how varying traffics differ in *sensitivity*, and how that 'footprint' within IP packets could be used to characterise data for machine-learning.

This shows how autonomous agents could be devised such that they were capable of traffic categorisation and dynamic differential, reallocation of computer network bandwidth, to various business data streams according to their relative dynamic priorities. Such agents could then reduce the reliance and complexity of current (human) expert QoS policy definition through the deployment of machine-learning techniques for the re-classification of the IP traffic. This paper also introduces the platform on which further experimentation will be completed.

II. APPLICATION SENSITIVITY, DELAY AND PROCESSING

Certain Internet traffic applications are time-critical. The stutter arising from delayed packets often renders

 ISBN: 978-1-61208-103-8

videoconferencing unusable. For any System for Intelligent Network Control (SINC) to be adopted it must satisfy the end-user delivery requirements [6]. A summary of sensitivities is given in Table 1.

TABLE I. APPLICATION SENSITIVITIES

Traffic Type	Sensitivities			
	Bandwidth	***Loss***	***Delay***	***Jitter***
Voice (set-up)	Very low	Medium	High	High
eCommerce	Low	High	High	Low
Transactions	Low	High	High	Low
Email	Low	High	Low	Low
Telnet	Low	High	Medium	Low
Casual browsing	Low	Medium	Medium	Low
Serious browsing	Medium	High	High	Low
File transfer	High	Medium	Low	Low
ICA	Medium	Medium	High	Medium
Video conferencing	High	Medium	High	High
Multicast	High	High	High	High

Any rule-based system must therefore respect such end-user observable concerns such as: *delay*, defined as *a lapse of time*, which includes *jitter*, often defined as a *packet delay variation* (PDV) used as a measure of the 'variability over time' of the packet latency across a network. In traditional QoS deployments a set of common applications, group of users, or business performance requirements, can be profiled and a template developed for that application and each resulting flow. Thus each flow which fits that profile can be treated the same, reducing the cost of replicating flow information for similar flows. The authors' previous paper [7] drew on an observation-action pair mapping, or "*policy of an agent*" [8], shown in equation 1 below:

$$F\ \theta_t = \alpha_t \tag{1}$$

in which a stateless function F maps its current observation (of network traffic and available resource) to a new action, representing a classification of the data, and $_t$ is the budget (e.g., time) in which the observation is made and the mapping completed. As an example of relevance to this paper, consider the *rule* for an attribute found within IP traffic, such as packet length, shown in equation 2 below:

$$packet\_length: \begin{matrix} \leq y \rightarrow a \\ > y \rightarrow b \end{matrix} \tag{2}$$

where y, in this instance is packet length in Bytes and a and b exhaust all possible classifications of that packet. An illustration being packets ≤ 80 Bytes are classified a, where a equals a classification of 'Expedite Forward', and all other packets (e. g., > 80 Bytes) are classified b, where b equals a classification of, in this instance, 'AF21'.

Such tests are, however, dependent on a known *typing* of net packets: given variability of the packet structure, simple rules are likely to misclassify traffic, with a resulting incorrect prioritisation. This motivates our research to classify packets based on a "black-box" (unsupervised) approach, by which *a priori* unknown packet structures can be presented to the learning-function for classification based on *learned characteristics* (e.g., voice packets which have a high QoS priority, have these known sensitivities). Where this characterisation is completed using some or all of those attributes available within an IP packet, but where the attribute choice is not fixed, thus allowing for an assessment of belief in a hypothesis to be updated with new data at each *observation epoch*. This set of attributes provides a 'frame of discernment' Θ [9].

III. LEARNING FOR AN INTELLIGENT NETWORK CONTROL

There are many well-known mechanisms for the useful characterisation of data [10][11][12] including:

- Supervised learning
- Unsupervised learning
- Reinforcement learning

Evidence from the authors' background research indicated that the use of pattern matching is effective at finding previously unseen patterns within the dataset, and given IP networks have vast sums of data traversing networks, with each packet or frame having an inherent *footprint* resultant from its header(s), this would appear to offer a suitable mechanism for intelligent network control.

There is no lack of data in the typical network [13] making statistical analysis techniques of traffic a relatively easy task. The challenge to the data classification is to accurately classify traffic in real-time. Initial experimentation with the first author's corporate network has focused on the classification of traffic flows using a *k*-Nearest Neighbour (*k*NN) clustering feature. Tarassenko [14] defines the objective of any clustering as:

"Given P patterns in n-dimensional space, find a partition of the patterns into K groups, or clusters, such that the patterns in a cluster are more similar to each other than to patterns in different clusters."

Clustering requires the characterisation of input data within a multidimensional space; in the context of this research we characterise data as the five-tuple:

- source IP address;
- destination IP address;
- source port number;
- destination port number;
- protocol;

- plus additional attributes including length and frame_time_delta.

The authors' have adapted the framework of [15] to identify a classification process that isolates differences within a population of network traffic, each having different a *model* (or description). The process by which this is achieved is defined below, and whilst the final realisation of this research is expected to be deployed using Application Specific Interface Card (ADIC) or Programmable Logic modules, integrated within internetwork devices in much the same way as the WAN Interface Cards (WIC) seen in routers, Fig. 3 shows the current system in development:

1. **Sensing:** input to the system, the packets arriving on the ingress interface
2. **Pre-process:** signals are pre-processed such that they can be transposed for subsequent operations without loosing relevant processing information. This may use a *segmentation* function to isolate the *features* of the data from each other or from background noise. One such segmentation would be to separate UDP from TCP traffic as each has a differing underlying network requirement.
3. **Feature Extraction:** whose purpose is to reduce the data by measuring certain features or properties, which in turn are passed to a
4. **Classifier:** which evaluates the evidence and makes a decision as to the queue in which the traffic will be transmitted
5. **Post-processing:** Are those processes engaged after the classifier required to return the newly classified traffic to a network egress interface

Of course, the use of *k*NN is not new, however, with networks and the data they handle within a QoS setting highly time critical, we cannot allow the real-time classification of data detrimentally to slow it. In particular, there are industry standards for node processing: the optimal decision boundary which accepts a level of error within the classification to ensure any process keeps within any *overall budget* defined (e.g., RTD) should be less than 150ms according to ITU-T G.114 [16]. The novelty in our research includes, therefore, the engineering of a classification algorithm which will prevent the modelling of extremely complex dimensional dependencies. It was this requirement of visualisation that led the authors to the thought of Pattern Matching for QoS, and which the authors' can now model using that framework of [15] where:

1. **Sensing:**
   - **a.** Traffic is generated from a production network or simulated on a laboratory environment, using client>server transactions, or application simulation software
   - **b.** That traffic generated is captured using network protocol analysis software (e.g., Wireshark; Etherpeek) and saved as a .cap file for analysis
2. **Pre-process:**
   - **a.** The traffic is exported as an .XML compliant .pdlm file such that it can be imported into spread-sheet applications for statistical analysis
   - **b.** The resultant .xml file is loaded into a text reader (e.g., textpad) and searched for non ASCII characters, which would prevent import to those spread-sheet applications
   - **c.** At this stage the .xml data requires to be transposed such that each *attribute* (e.g., *frame_length*) is a column header, and each instance of that attribute (per packet or frame) is a row of known variable type (e.g., nominal, string etc.). This is completed by a SQL transpose routine, the output of which is a reordered .xml file.
   - **d.** That .xml file is opened within Microsoft Excel to ensure consistency of data. This includes missing or errored cells, inconsistent format or corrupt file.
   - **e.** This data file is presented to SPSS PASW, Weka, and Matlab software which enables a number of statistical analysis visualisations to be completed, such as a scatterplot of Frame_length over Time, shown at Fig. 1, "Visualisation of Captured Data in WEKA" below.
   - **f.** A suitable file (e.g., Weka .arf; Matlab .mat) can now be built such that varying learning processes can be explored and evaluated.
3. **Feature Extraction:**
   - **a.** Feature exaction then is the reduction of data to its principle features. In the case of Internetwork traffic this has previously been defined as the five tuple with additional attributes, including but not exclusive to:
     - ip_src
     - ip_dst
     - ip_srcport (tcp or udp)
     - ip_dstport (tcp or udp)
     - prot
     - frame_pkt_length
     - frame_time_delta
     - frame_time_relative

 ISBN: 978-1-61208-103-8

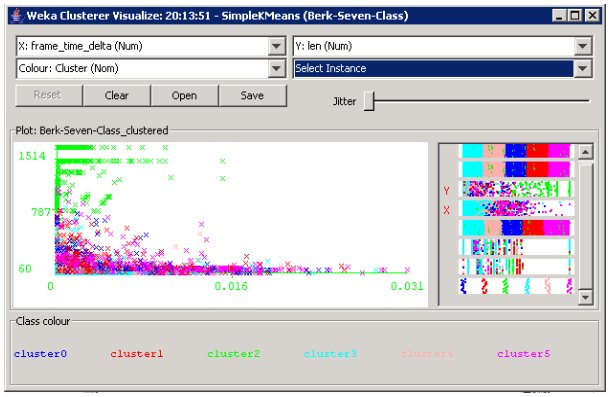


Figure 1. Visualisation of Captured Data in WEKA.

Equation 3 illustrates how three 'characterisation features', say packet_length (x_a), time_frame_delta (x_b) and ip_dstport (x_c), could be used to characterise a packet as distinct from any other packet not within the same cluster. This would present a feature vector in a two-dimensional feature space [14], where:

$$\begin{pmatrix} x_a \\ x_b \\ x_c \end{pmatrix} \quad (3)$$

This use of feature extraction can be illustrated with an adaptation of [14] shown in Fig. 2, where these three attributes are presented to the clustering function (F) which has then defined the packets in two distinct clusters with an individual mean of X, where:

X_n = *the packets captured within time t*

(where time t is from the start of capture to the end of the capture processed) with features

$X = \{x_a, x_b, x_c\}$

C_n = *the clusters of the set* $K = \{C_1, C_2, \ldots\ldots, C_n\}$

$\otimes$= *the cluster centre characterised in this instance using the mean vector* m_k

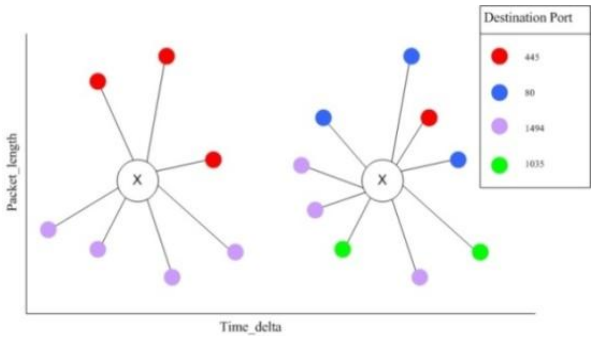


Figure 2. Adaption of Features of a Cluster [14]

This difference between clusters typically measured by the *distance between them* (such as Euclidean distance or Pearson correlation) to define a *closeness* or interrelationship of the traffic. The known issue of Euclidean distance, which assigns more *weight* (i.e., preference) to features with large range could be further optimised through the use of *Manhattan* (or block) distance, such that higher powers increase the influence of large (neighbour) differences at the expense of small ones. An example being, if presented with two features: ip_dstport and frame_pkt_length, with Euclidean distance there would likely be a preference to the former. This is due to the legal range of IP ports being in the range is 0 to 65,535, compared to a standard Ethernet frame *Maximum Transmission Unit* (MTU) of only 1500 (bytes).

That traffic, which is clustered based on its perceived characteristic, as demonstrated in Fig. 2, can then be marked using coding such as Differentiated Services Code Point (DSCP). That cluster of traffic which is defined as being the most *sensitive* (such as voice traffic) or of having the *highest business importance* would then be marked with an appropriate classification, such as Expedite Forward (EF). Each of the six clusters (0-5) would then be forwarded to one of six pre-configured virtual hardware queues within the network devices egress interface.

IV. SUMMARY AND CURRENT RESEARCH

This paper built on an early "feasibility study" to investigate how the agent models, as conceived in the authors' previous papers [6][7], would be implemented to influence internetwork traffic management. Whilst the technique is promising, the current manual transposition of data for the machine-learning application means that only off-line processing is possible. However a realistic implementation requires a platform with the ability to perform online, real-time automated transposition of QoS targets for various data steams intended for presentation to the machine-learning mechanism.

Such a system, initially online if not real-time, has been implemented on a Linux based server running the Snort Intrusion Detection System (IDS) [17]. Ingress network traffic is *sensed* (captured) and updated within a MySQL database table. From the database, rather than using the traditional Snort rule-base, the instances (packets) within the various tables are JOINED and a Python language script presents the data to an Orange k-means clustering algorithm [18]. This algorithm performs the function F described, completing the *pre-processing, feature extraction, and classification* including *visualisation*. Further research investigates those activities related to *post-processing*, such as the *marking* of that traffic based on Differentiated Services Code Point (DSCP), and *dispatch* routines which, at present, is the re-population of the database with the newly reordered traffic ready for network transmission on the appropriate egress interface.

 ISBN: 978-1-61208-103-8

In addition, further investigation into other Python-based pattern classification applications, such as pyMVPA [19], PyBrain [20], and OpenElectrophy [21] will be progressed to see whether they can perform the work more efficiently.

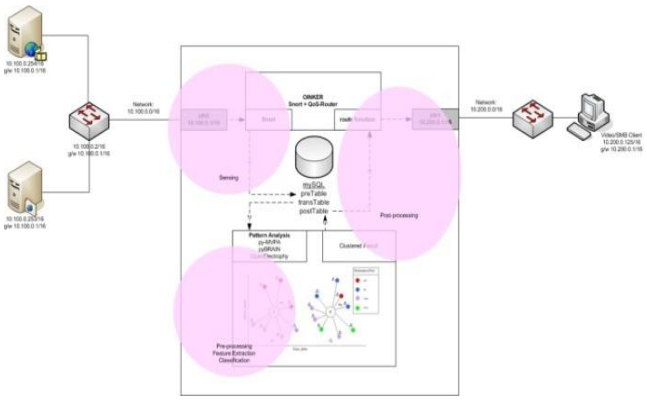


Figure 3. Oinker: Snort-based Systen for QoS Route Forwarding using Pattern Classification

ACKNOWLEDGMENT

The authors' would like to express their sincere thanks to Dr. Jon Hall, Open University, for his assistance in the completion of this paper.

REFERENCES

[1] D. C., Verma, "Policy-based networking architecture and algorithms", Technology Series, New Riders, Indianapolis, Indiana, pp. 6, November 2000.

[2] S. De Lurgio, "Predicting micro-loan defaults using probabilistic neural networks", Credit & Financial Management Review, (Internet), 2002, http://www.allbusiness.com/finance/1049916-1.html, (accessed 19th July 2010).

[3] M. Buchanan, "Why complex systems do better without us", New Scientist, Reed Business Information Ltd., Issue 2668, pp. 28-31, August 2008.

[4] C. J., Tebelskis, "Speech recognition using neural networks", an unpublished thesis, (Internet), 1995, http://portal.acm.org/citation.cfm?id=239333&dl=GUIDE&coll=GUIDE&CFID=97551353&CFTOKEN=37801874, (accessed 26th July 2010).

[5] I. H, Whitten and E. Frank, "Data mining: practical machine learning tools & techniques, Second Edition, Elsevier, San Francisco, 2005

[6] D. Legge and A. Badii, "Conceptualisation of an application of adaptive synthetic socioeconomic agents for intelligent network control", 2nd PERADA Workshop on Pervasive Adaptation, Edinburgh, AISB, pp. 14-21, April 2009.

[7] D. Legge and A. Badii, "A primer for an application of adaptive Synthetic Socioeconomic Agents for Intelligent Network Control", in press, 2nd School of Systems Engineering Conference, University of Reading, December 2009.

[8] N. Vlassis, "A concise introduction to multi-agent systems and distributed artificial intelligence, Morgan & Claypool, California, pp. 55-57, 2007.

[9] R. Callan, "Artificial Intelligence", Palgrave Macmillan, Basingstoke, Hampshire, pp. 163, 2003.

[10] J. P, Bigus, "Data mining with neural networks: Solving business problems-from application development to decision support", Computing McGraw-Hill, New York, pp. 6-29, 1996.

[11] J. F, Sowa, "Conceptual structures: InformationpProcessing in mind and machine", The Systems Programming Series, Addison-Wesley Publishing, Reading, Massachusetts, pp. 281-292, 1986.

[12] G. Marshall, "Advanced students' guide to expert systems", Heinemann Newnes, Oxford, pp. 128-142, 1990.

[13] S. Sui, and C. Zhixiong, "Adaptive network flow clustering", IEEE International Conference on Networking, Sensing and Control, pp. 596-601, April 2007.

[14] L. Tarassenko, "A guide to neural computing applications", Arnold, London, pp. 20-23, 1998.

[15] R. O. Duda, P. E. Hart, and D. G., Stork, "Pattern classification", Chichester, Wiley, New York, pp. 3-23, 2000.

[16] ITU-T G.114 "One-way transmission time, series G: Transmission systems and media, digital systems and networks international telephone connections and circuits–general recommendations on the transmission quality for an entire international telephone connection", Telecommunication Standardization Sector of the International Telecommunication Union, pp. 2-12, May 2003.

[17] Snort, "a lightweight open source network intrusion prevention and detection system (IDS/IPS)", (Internet), 2010, http://www.snort.org, (accessed 26th July 2010).

[18] Orange, Laboratory of Artificial Intelligence, Faculty of Computer and Information Science, University of Ljubljana, Slovenia, http://www.ailab.si/orange/doc/modules/orngClustering.htm, (accessed 20th July 2010).

[19] M. Hanke, Y. O. Halchenko, P.B. Sederberg, S.J. Hanson, J.V. Haxby, and S.Pollmann, "PyMVPA: A Python toolbox for multivariate pattern analysis of fMRI data", Neuroinformatics, volume 7, pp. 37-53, 2010.

[20] T. Schaul, J. Bayer, D. Wierstra, and Y. Sun, "PyBrain", Journal of Machine Learning Research 11, pp. 743-746 Submitted 11/09; Published 2/10, (Internet), 2010, http://www.idsia.ch/~tom/publications/pybrain.pdf, ((accessed 26th July 2010).

[21] S. Garcia and N. Fourcaud-Trocmé, "OpenElectrophy: an electrophysiological data- and analysis-sharing framework", Frontiers in Neuroinformatics, Volume 3:14, (Internet), 2010, http://frontiersin.org/neuroinformatics/10.3389/neuro.11.014.2009/full,(accessed 26th July 2010).

 ISBN: 978-1-61208-103-8

Gaussian Fitting of Multi-scale Traffic Properties for Discriminating IP Applications

Eduardo Rocha, Paulo Salvador and António Nogueira
University of Aveiro/Instituto de Telecomunicações
Aveiro, Portugal
e-mails: {eduardorocha, salvador, nogueira}@ua.pt

***Abstract*—In the last years, there has been an increasing need to accurately assign traffic to its originating application or protocol. Several new protocols and services have appeared, such as VoIP or file sharing, creating additional identification challenges due to their peculiar behaviors, such as the use of random ports or ports associated to other protocols. The number and variety of security vulnerabilities and attacks that are carried out over the Internet has also drastically increased in recent years. Besides, privacy and confidentiality are also growing concerns for Internet users: traffic encryption is becoming widely used and, therefore, access to the user payload is more and more difficult. Therefore, new identification methodologies that can be accurate when applied to different types of traffic and be able to operate in cyphered traffic scenarios are needed. In this paper, we present an identification methodology that relies on a multiscale analysis of the traffic flows, differentiating them based on the probability that their characteristic multiscale behavior estimators belong to specific probability distributions whose parameters are inferred from traffic flows of real applications. The classical concept of traffic flow was replaced by the definition of *data stream*, which consists of all traffic (in the upload or download directions) of a local IP address that is univocally identified by a numeric identifier. The results achieved so far show that the proposed methodology is able to accurately classify licit traffic and also identify some of the most common Internet security attacks. Besides, this approach can also circumvent some of the most important drawbacks of existing identification methodologies, namely their inability to work under strict confidentiality restriction scenarios.**

Keywords: Application identification, multiscale analysis, wavelets, licit and illicit applications.

I. INTRODUCTION

Classifying Internet traffic is a critical task for many areas, such as traffic engineering, Quality-of-Service (QoS), access control and security/intrusion detection. In recent years, the emergence of diversified and demanding applications made some of the mostly used classification methodologies (like port-based classification or payload inspection) inadequate. Besides, the number and diversity of attacks to hosts and services in the Internet increased in a dramatic way. Among these new threats, *botnets* are some of the most severe and dangerous [24], being responsible for some of the most stealth attacks, such as Distributed Denial-of-Service, Spam and phishing e-mails [4], [7], [6]. A *botnet* is a network of compromised computers under the control of a master, the *bot* master, which issues commands to the compromised hosts. Usually, these communications are encrypted, which poses a significant obstacle for Intrusion Detection Systems (IDSes). Moreover, the distributed nature of these attacks and the evolving (from centralized to distributed) structure of the botnets [16] also makes them extremely difficult to prevent.

This paper presents a new technique for identifying licit and illicit traffic flows based on the classification of different multi-scale behavior estimators. The classification methodology relies on the probability that these estimators belong to a Gaussian distribution whose parameters are inferred from traffic flows of the real applications. This approach presents several advantages over existing ones, namely its compliance with privacy issues since only packet headers at the IP and/or IP security protocols levels are analyzed. This work is an extension of a previous work [28] that also analyzed the multi-scale behavior of sampled flows generated by different applications using a kind of ”blind” clustering to classify the multi-scale coefficients’ estimators. Here, we assume that these estimators follow a Gaussian distribution and use a probabilistic methodology to classify them, thus being able to discriminate their underlying generating applications. Besides the three widely used Internet applications that were also considered in [28] (web-browsing, video streaming and BitTorrent), we also include two of the most common attacks that are used by *botnets*: (i) *port scanning* and (ii) *snapshots* of the users’ desktops. The classification results that have been already obtained show that the proposed approach is very promising, while being immune to some of the main disadvantages of current detection methodologies.

In order to be able to classify the different interactions that an application creates, which may consist of several sessions with different end-hosts/servers (and we strongly believe that the analysis of these interactions as a whole provides a deeper insight into how the applications behave and can assist in traffic discrimination), the restrictive classical definition of flow was replaced by the definition of *data stream*, which consists of all traffic (in the upload or download directions) of a local IP address that is univocally identified by a numeric identifier.

The remaining part of this paper is organized as follows: Section II presents some related work in the fields of traffic classification and attacks identification, Section III presents some background on wavelets and multiscale analysis, Section IV presents the details of the identification methodology; Section V presents some identification results that were already obtained in order to evaluate the efficiency of the proposed methodology and, finally, Section VI presents some brief

 ISBN: 978-1-61208-103-8

conclusions about the conducted work.

II. Related Work

The issue of traffic classification has been studied for many years and many techniques have been proposed to address this problem. In an early stage, traffic was classified according to the ports used for communication. However, this analysis became inaccurate when new protocols, such as BitTorrent or VoIP protocols, started to use random ports or ports associated to other applications. In fact, in a study conducted by [21], port-based techniques were unable to classify most of the network traffic that was generated by Peer-to-Peer (P2P) protocols. Payload analysis was one of the techniques proposed to overcome this limitation. It consists on the inspection of the packet's payload searching for characteristic signatures that can identify the generating protocol. A study carried out in [14] used this technique to identify P2P traffic and the results achieved were very accurate. In another work [30], digital signatures were also used to classify P2P traffic. The results achieved were very accurate and the authors proved that the proposed methodology can be effective in high-speed networks. However, in recent years, traffic encryption is becoming widely used to guarantee the confidentiality of the exchanged data in the Internet and, therefore, in these scenarios the packet payload is no longer accessible. Besides, when traffic is not encrypted the access to the packet's payload may not be allowed due to privacy restrictions.

Statistical analysis of traffic flows appeared as a solution that could overcome these restrictions, since only the headers of the packets are analyzed. The main concept of this approach is that traffic generated by the same protocol will present the same profile. Karagiannis *et al.* tried to identify P2P traffic based on a three-level analysis: social, functional and application levels. The accuracy of the obtained results was very high [15]. In another work [13], the authors built behavioral profiles that describe dominant patterns of the studied applications and the results showed that this approach was quite promising. In [21], the authors only analyzed the TCP SYN, FIN and RST flags in order to obtain connection-level information about P2P traffic. This technique has several inherent drawbacks: traffic presenting unknown behavior cannot be classified; when traffic is transported through a secure tunnel, the port numbers and the TCP flags may not be available and, consequently, classification is not possible.

In the last years, the number of security vulnerabilities and attacks increased at a dramatic rate [29]. *Botnets* have emerged and became one of the most dangerous threats to on-line security, being used for a wide variety of illegal activities such as DDoS, Spam, flooding attacks and exploit scanning, just to name some of them [22]. Besides, they are undetected by *anti-virus* software and IDSes [4]. Most IDSes, such as Snort [2], perform intrusion detection based on the recognition of signatures and known patterns from security attacks. This can constitute an accurate detection methodology, but these defense mechanisms cannot detect *zero-day* threats and attacks with unknown profiles [17]. Of course, IDSes can protect their networks by classifying any traffic pattern that deviates from an already known normal profile as an attack. Although this strategy could make them able to detect *zero-day* attacks, the detection accuracy would decrease since some of these "abnormal" profiles may be originated by legitimate user actions.

The structure of the botnets is also evolving, becoming more complex and distributed. For instance, the C&C infrastructure evolved from a centralized one, in which IRC protocols were used for communication, to a distributed one where P2P protocols and networks are used. Moreover, these communications can also be embedded in the HTTP protocol. Therefore, the detection of these networks is becoming more difficult and new methodologies are needed for their accurate detection.

Several studies have been conducted in order to collect, analyze and understand how *botnets* work: [5] studies the communications between the Command and Control (C&C) server and the infected machines; [25] analyzed the network behavior of spammers; [8] conducted several basic studies of *botnet* dynamics; [9] proposed to use DNS sink holing technique for *botnet* study and pointed out the global diurnal behavior of *botnets*; finally, [6] studied the relationship between *botnets* and scanning/spamming activities.

Based on this knowledge, different approaches have been proposed to solve the *botnet* detection problem: in [26], the authors used DNS-based black hole list counter-intelligence to find *botnet* members that generate spam; in [27], the authors proposed a system to detect malware (including *botnets*) by aggregating traffic that shares the same external destination, have a similar payload and involves internal hosts with similar OS platforms; [20] proposed a machine learning based approach for *botnet* detection using some general network-level traffic features of chat-like protocols, such as IRC; finally, [12] describes BotHunter, which is a passive *botnet* detection system that uses dialog correlation to associate IDS events to a user-defined *bot* infection dialog model.

III. Wavelets and Multiscale Analysis

A wavelet $\psi(t)$ can be defined as a pass-band function oscillating at a central frequency f_0. By performing a scaling change, which may consist of an expansion or a compression, and a temporal shift, we obtain $\psi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t - k)$, that is the oscillating central frequency moves to $2^{-j}f_0$ and the origin of the temporal reference to $2^j k$. Note that *j* represents the temporal scale, *k* represents the k^{th} coefficient corresponding to scale *j*, with j_0 being the larger time scale. Wavelet decomposition also uses a low-pass function, $\phi_{j_0,k}(t)$, known as scaling function, that can be scaled and temporarily shifted in a similar way to function $\psi_{j,k}(t)$. Therefore, the definition of the Discrete Wavelet Transform (DWT) of a stochastic process $X(t)$ is [11]:

$$X(t) = \sum_k c_X(j_0,k)\phi_{j_0,k}(t) + \sum_{j=j_0}^{\infty}\sum_k d_X(j,k)\psi_{j,k}(t) \quad (1)$$

where $c_X(j_0,k)$ are the scaling coefficients and $d_X(j,k)$ are the wavelet coefficients. The estimators for the first order moment of the wavelet coefficients can be defined as:

 ISBN: 978-1-61208-103-8

$$\mu_j = \frac{1}{n_j} \sum_{k=1}^{n_j} |d_X(j,k)| \quad (2)$$

where n_j is the number of coefficients to be analyzed at scale j. The scaling behavior of any stochastic process can then be studied by an analysis of the Logscale diagrams, which consist of logarithm plots of these estimators with the scales [3].

As mentioned in Section I, phenomena such as Short-Range Dependence (SRD) and Long-Range Dependence (LRD) have been studied in several works. In [18] can be fond the first evidence that network traffic has self-similar characteristics. In [23], several TCP statistics, such as session and connection arrivals, were analyzed and self similarity was found in many traces. In [32], the authors provided several measurements which showed that network traffic exhibits self-similar behavior. Physical features of communication networks were also presented to explain such behavior. In [31], time-series extracted from network traffic were proven to exhibit LRD. Feldmann *et al.* investigated several aspects of user and network behaviors contribute to the scaling regimes in WAN traffic [10].

IV. Identification Methodology

Our work aims at classifying the several interactions that an application creates, which may consist of several sessions with different end-hosts/servers. We believe that the analysis of such interactions as a whole can provide a deeper insight into how the applications behave and can assist in traffic discrimination. To be able to perform such study, the classical definition of a flow, the *5-tuple*, becomes too restrictive since it does not capture all the mentioned interactions. Therefore, we used the definition of *data stream*, which consists of all traffic (in the upload or download directions) of a local IP address and univocally identified by a numeric identifier. This *data stream* numeric identifier is: (i) for unencrypted traffic, a specific TCP/UDP (local or remote) port number and (ii) for encrypted traffic, the Security Parameters Index (SPI) in ESP headers in case of IPsec tunnels or any other specific identifier of IP-level encrypted tunnel technology. Therefore, *data streams* are uniquely identified by a *2-tuple* (IP address, unique identifier). Other important definitions in our work are the *known data streams* which consist of *streams*, as previously defined, analyzed *a priori* to determine its origin application(s). On the other hand, let us define the *unknown data streams* as a traffic *stream* created by an unknown application. Several stochastic processes (and respective statistics) can be extracted from these *data streams*, which, in this work, will be processed by a DWT, as described in Section III, in order to obtain the estimators defined in (2). Since the applications that generated the analyzed traffic might have different network conditions, these estimators were normalized to zero mean:

$$\hat{\mu}_j = \mu_j - \sum_{j=1}^{J} \frac{\mu_j}{J} \quad (3)$$

in which J represents the number of scales considered for analysis. In the following lines we will present some more definitions. For instance, let A represent the number of known applications, M represent the number of *unknown data streams* that we want to classify and N correspond to the number of *known data streams*. Let $p_{i,a}, a = 1, ..., A$ designate the probability that the *unknown stream* i belongs to the Gaussian distribution inferred from the *known streams* of the application a. Let $E_{a,j} = \{e^i_{a,j}, i = 1, \ldots, N\}$ and $U_j = \{u^i_j, i = 1, \ldots, M\}$ represent the normalized estimators, as defined in (3), for the first order moment of the wavelet coefficients of a stochastic process, respectively, extracted from a *known data stream* i of the application a at the scale j and extracted from a *unknown data stream* i at the scale j. The proposed methodology assumes that $E^i_{a,j}$ and U_j, for all j and a, follows a Gaussian distribution. Therefore, let

$$P_{i,a,j} = \int_{u^i_j - \Delta}^{u^i_j + \Delta} \frac{1}{\sqrt{2\pi\sigma^2_{a,j}}} e^{\left(\frac{-(u-\overline{e}_{a,j})^2}{2\sigma^2_{a,j}}\right)} du \quad (4)$$

represent the probability that the estimator of the *unknown stream* i, of the scale j, is within a neighborhood of width 2Δ, centered on itself originated by a distribution whose parameters, $\overline{e}_{a,j}$ and $\sigma^2_{a,j}$, are empirically inferred from the *known data streams* of an application a:

$$\overline{e}_{a,j} = \frac{1}{N} \Sigma_{i=1}^{N} e^i_{a,j} \quad (5)$$

$$\sigma^2_{a,j} = \frac{1}{N-1} \Sigma_{i=1}^{N} (e^i_{a,j} - \overline{e}_{a,j})^2 \quad (6)$$

The probability $P_{i,a,j}$ is then computed for all *unknown streams* and for all distributions inferred from the *known streams* studied applications, for each scale of analysis.

Subsequently, it is possible to compute $P_{i,a}$ as:

$$P_{i,a} = \prod_{j=1}^{J} P_{i,a,j}, a = 1, ..., A; i = 1, ..., M \quad (7)$$

Finally, an *unknown data stream* i, $i = 1, \ldots, M$, is associated with application $\alpha, \alpha = 1, \ldots, A$, such that

$$\exists \alpha, P_{i,\alpha} = \max_a [P_{i,a}]. \quad (8)$$

V. Results

In this Section we present the obtained results from several traffic *data streams* extracted from: (i) licit TCP and UDP traffic traces passively collected at the University of Aveiro network on September 15, 2008 and (ii) illicit traces experimentally generated in laboratory simulating some of the most relevant *botnet* uses. The licit applications *data streams* extracted (and classified *a priori*) from the traffic collected were file-sharing (BitTorrent), video streaming and HTTP (browsing). Figures 1 to 3 present the variation of the number of bytes in the upload and download directions for the mentioned applications. The illicit traffic was experimentally generated in our lab in an attempt to simulate some of the most relevant reconnaissance attacks. The NMAP [1] flows were generated using a discrete scan profile in order to replicate a typical *botnet* port scan that tries to evade IDS detection and

 ISBN: 978-1-61208-103-8

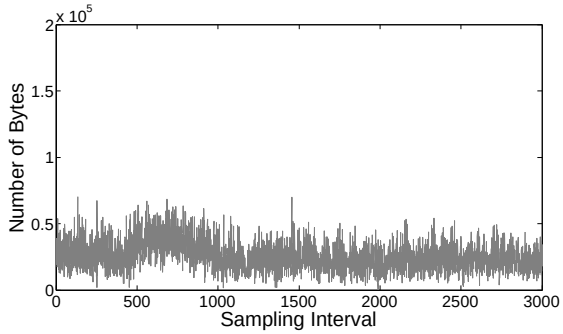


Figure 1. Number of bytes for a Torrent flow.

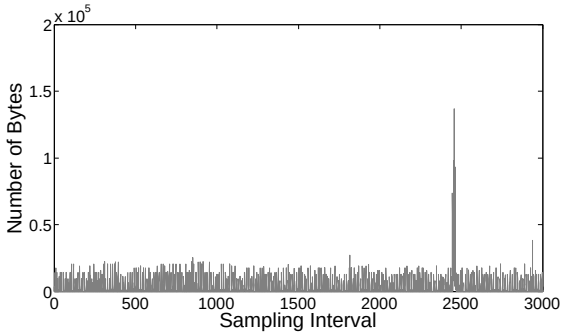


Figure 2. Number of bytes for a Streaming flow.

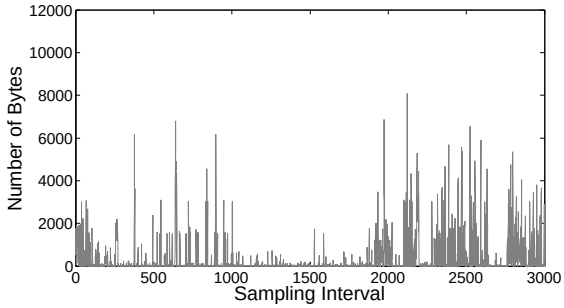


Figure 3. Number of bytes for an HTTP flow.

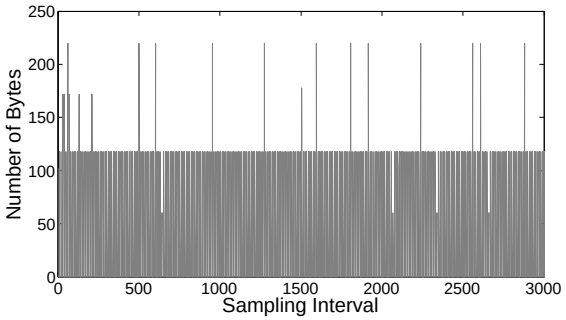


Figure 4. Number of bytes for a NMap flow.

scan hosts and networks, bypassing their firewalls and proxies. Therefore, we performed a sequential port scan with one second of interval between (SYN) probes and a waiting time of 15 seconds before start scanning a new machine. The Snapshot flows were generated by emulating the capture of a fixed size small image (335x180 pixels, 120KBytes) of the user's desktop around the cursor every time the user performed a click. We assumed that the user was browsing the Internet and performed a click with an exponentially distributed interval with average equal to 120 seconds [33]. The flows of these applications are presented in Figs. 4 and 5, respectively.

In this case, the values analyzed were the overall number of transmitted bytes, independently of direction. The extracted *data streams* were 5 and 15 minutes long and were divided in *known* and *unknown streams*, however, the real classification of all *streams* was kept for validating the classification results of the *unknown streams*. Now let us present the values of the several variables defined in Section IV. The number of application considered was 5 ($A = 5$). The number of *known streams*, N, used for inferring the parameters of the Gaussian distributions was 30 and the number of *unknown streams*, M, was 80, for each application. The value of the interval Δ used was 0.1.

The *known* and *unknown streams* were analyzed via a DWT in order to obtain the estimators for the first order moment of the wavelet coefficients. The first mentioned values were then used to validate the assumption that the estimators for the first order moment of the wavelet coefficients, for each application and scale, follow a Gaussian distribution. The test used was the Lilliefors goodness-of-fit test which verifies the null hypothesis that the sample in a vector comes from a distribution in the Gaussian family, against the alternative that it does not [19]. All the tests did not reject the null hypothesis, that is, all the estimators can be approximated by a Gaussian distribution.

The classification results were computed by comparing the classification achieved with the proposed methodology with the real application. In the first part of our results, we considered the 5 minutes long *data streams*. We only used the first 5 scales since at higher scales the estimators of all applications tend to converge. Figures 6 and 7 show box plots with 25%, 50%, 75% and 95% quantiles, for the estimators of the first order moment of the wavelet coefficients of the 5 minutes and 15 minutes *data streams*, respectively. We can observe that the distributions of the estimators of the HTTP and Snapshot *streams* almost overlap in all scales. This suggests that some HTTP and Snapshot *streams* might be misclassified. However, for the 15 minutes *data streams* (Figure 7) the Snapshot traffic estimators are now more concentrated around the mean, which suggests that the accuracy will be higher. For the remaining estimators' distributions we can observe that, at least in one scale, they are very separated and therefore they will not be misclassified.

The numerical results obtained, for the 5 minutes traffic traces, are presented in Table I and it is possible to observe that these are relatively accurate for all applications. With the exception of HTTP 5 minutes *data streams*, the obtained percentage of correctly identified *data streams* is between 73% and 100%. For HTTP traffic, the correct classification percentage is lower, as some of these *data streams* were misclassified as Snapshot, which is in accordance with the previous analysis. These result can be explained by the fact

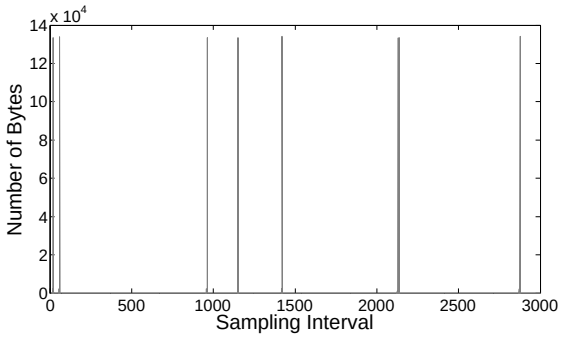


Figure 5. Number of bytes for a Snapshot flow.

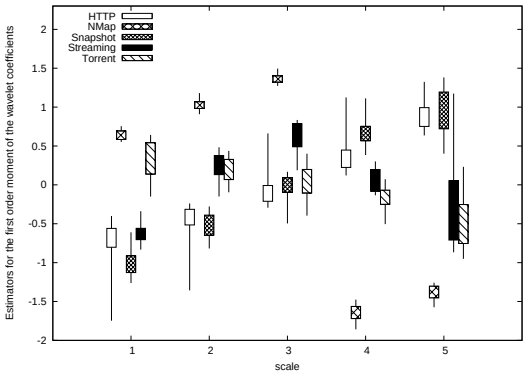


Figure 6. Distributions for 5 minutes traces.

that HTTP *data streams* multiscale estimators have an higher variance, resulting from the various and heterogeneous user behaviors, making this distribution partially overlap the snapshot estimators distribution (which has a much lower variance) in all scales. Moreover, several protocols, such as file sharing and video streaming, run on top of HTTP communications which justifies the large variance the estimators of these *streams* present and some classification mistakes. The classification results for the 15 minutes *data streams* are presented in Table II and we can observe that the accuracy of the results for all applications is higher. This can be explained by the fact that traces are longer, contain more information and more differentiating characteristics. This allows a deeper decomposition of each signal and therefore, a better analysis of their unique behaviors and leads to better classification results.

Table I
RESULTS FOR 5 MINUTES TRACES USING 5 SCALES

Data Streams	Classified as *NMap*	*Snapshot*	*HTTP*	*Streaming*	*Torrent*
NMap	**100%**	0%	0%	0%	0%
Snapshot	0%	**72.7%**	22.7%	3.1%	1.5%
HTTP	0%	29.4%	**64.7%**	2.9%	2.9%
Streaming	0%	0%	3.6%	**96.4%**	0%
Torrent	0%	3.1%	1.6%	1.5%	**93.8%**

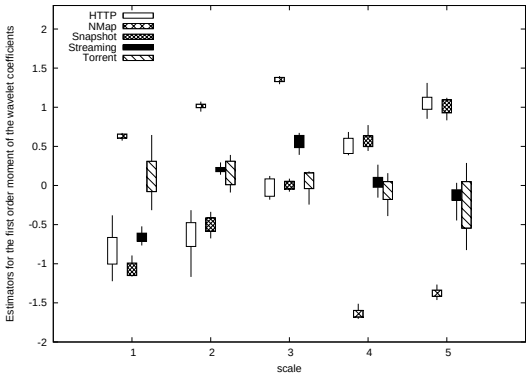


Figure 7. Distributions for 15 minutes traces.

Table II
RESULTS FOR 15 MINUTES TRACES USING 5 SCALES

Data Streams	Classified as *NMap*	*Snapshot*	*HTTP*	*Streaming*	*Torrent*
NMap	**100%**	0%	0%	0%	0%
Snapshot	0%	**95.2%**	4.8%	0%	0%
HTTP	0%	15.4%	**76.9%**	0%	7.7%
Streaming	0%	0%	0%	**100%**	0%
Torrent	0%	0%	0%	0%	**100%**

VI. CONCLUSIONS

The last years have witnessed the appearance of several new protocols and services, a huge increase on the number and variety of security vulnerabilities and attacks that are carried out over the Internet and the growth of the privacy and confidentiality concerns of Internet users. Thus, new identification methodologies that can be accurate when applied to different types of traffic and be able to operate in cyphered traffic scenarios are needed. This paper proposed an identification methodology that relies on a statistical multiscale analysis of the traffic flows, differentiating them based on the probability that their characteristic multiscale behavior estimators belong to Gaussian probability distributions whose parameters are inferred from traffic flows of real applications. The results obtained show that the proposed methodology is able to accurately classify licit traffic and also identify some of the most common Internet security attacks. Besides, the approach can also avoid some of the most important drawbacks presented by existing identification methodologies, namely their inability to work under strict confidentiality restriction scenarios. Finally, the definition of *data stream* also proved to be adequate for discriminating between several IP applications, constituting an important step towards a complete understanding of their behaviors.

ACKNOWLEDGEMENTS

This research was supported in part by Fundao para a Cincia e a Tecnologia, grant SFRH/BD/33256/2007.

REFERENCES

[1] Nmap: Free security scanner for network exploration and security audits. http://nmap.org/, March 2009.

[2] Snort :: Home page. http://www.snort.org/, May 2010.
[3] P. Abry, P. Flandrin, M. Taqqu, and D. Veitch. Wavelets for the analysis, estimation, and synthesis of scaling data. In K. Park and W. Willinger, editors, *Self-Similar Network Traffic and Performance Evaluation*, pages 39–88. Wiley, 2000.
[4] P. Barford and V. Yegneswaran. An inside look at botnets. *Springer Verlag*, 2006.
[5] K. Chiang and L. Lloyd. A case study of the rustock rootkit and spam bot. In *HotBots'07: Proceedings of the first conference on First Workshop on Hot Topics in Understanding Botnets*, pages 10–10, Berkeley, CA, USA, 2007. USENIX Association.
[6] M. Collins, T. Shimeall, S. Faber, J. Janies, R. Weaver, M. Shon, and J. Kadane. Using uncleanliness to predict future botnet addresses. In *ACM/USENIX Internet Measurement Conference IMC'07*, 2007.
[7] E. Cooke, F. Jahanian, and D. Mcpherson. The zombie roundup: Understanding, detecting, and disrupting botnets. pages 39–44, June 2005.
[8] E. Cooke, F. Jahanian, and D. McPherson. The zombie roundup: Understanding, detecting, and disrupting botnets. In *USENIX SRUTI'05*, pages 39–44, 2005.
[9] D. Dagon, C. Zou, and W. Lee. Modeling botnet propagation using timezones. In *13th Annual Network and Distributed System Security Symposium NDSS'06*, January 2006.
[10] A. Feldmann, A. Gilbert, P. Huang, and W. Willinger. Dynamics of IP traffic: A study of the role of variability and the impact of control. In *SIGCOMM*, pages 301–313, 1999.
[11] A. Feldmann, A. C. Gilbert, and W. Willinger. Data networks as cascades: investigating the multifractal nature of internet wan traffic. In *SIGCOMM '98: ACM SIGCOMM '98 conference on Applications, technologies, architectures, and protocols for computer communication*, pages 42–55, New York, NY, USA, 1998.
[12] G. Gu, P. Porras, V. Yegneswaran, M. Fong, and W. Lee. Bothunter: Detecting malware infection through ids-driven dialog correlation. In *16th USENIX Security Symposium*, 2007.
[13] Y. Hu, D.-M. Chiu, and J. Lui. Application identification based on network behavioral profiles. *Quality of Service, 2008. IWQoS 2008. 16th International Workshop on*, pages 219–228, June 2008.
[14] T. Karagiannis, A. Broido, N. Brownlee, K. Claffy, and M. Faloutsos. Is p2p dying or just hiding? [p2p traffic measurement]. *Global Telecommunications Conference, 2004. GLOBECOM '04. IEEE*, 3:1532–1538 Vol.3, Nov.-3 Dec. 2004.
[15] T. Karagiannis, K. Papagiannaki, and M. Faloutsos. Blinc: multilevel traffic classification in the dark. In *SIGCOMM '05: 2005 conference on Applications, technologies, architectures, and protocols for computer communications*, pages 229–240, New York, NY, USA, 2005. ACM.
[16] A. Karasaridis, B. Rexroad, and D. Hoeflin. Wide-scale botnet detection and characterization. In *HotBots'07: First Workshop on Hot Topics in Understanding Botnets*, Berkeley, CA, USA, 2007. USENIX Association.
[17] O. Kolesnikov, D. Dagon, and W. Lee. Advanced polymorphic worms: Evading ids by blending in with normal traffic. Technical report, 2004.
[18] W. E. Leland, M. S. Taqqu, W. Willinger, and D. V. Wilson. On the self-similar nature of ethernet traffic (extended version). *IEEE/ACM Transactions on Networking*, 2(1):1–15, 1994.
[19] H. Lilliefors. On the kolmogorov-smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, 1967.
[20] C. Livadas, R. Walsh, D. Lapsley, and W. Strayer. Using machine learning techniques to identify botnet traffic. In *2nd IEEE LCN Workshop on Network Security*, 2006.
[21] A. Madhukar and C. Williamson. A longitudinal study of p2p traffic classification. *Modeling, Analysis, and Simulation of Computer and Telecommunication Systems, 2006. MASCOTS 2006. 14th IEEE International Symposium on*, pages 179–188, Sept. 2006.
[22] A. H. Nicholas Ianelli. Botnets as a Vehicle for Online Crime. *The International Journal of Forensic Computer secience*, 2(1), 2007.
[23] V. Paxson and S. Floyd. Wide-area traffic: The failure of Poisson modeling. *IEEE/ACM Transactions on Networking*, 3(3):226–244, June 1995.
[24] M. Rajab, J. Zarfoss, F. Monrose, and A. Terzis. A multi-faceted approach to understanding the botnet phenomenon. In *ACM SIGCOMM/USENIX Internet Measurement Conference IMC'06*, October 2006.
[25] A. Ramachandran and N. Feamster. Understanding the network-level behavior of spammers. *SIGCOMM Comput. Commun. Rev.*, 36(4):291–302, 2006.
[26] A. Ramachandran, N. Feamster, and D. Dagon. Revealing botnet membership using dnsbl counterintelligence. In *USENIX SRUTI*, 2006.
[27] M. Reiter and T. F. Yen. Traffic aggregation for malware detection. In *Fifth GI International Conference on Detection of Intrusions and Malware, and Vulnerability Assessment*, 2008.
[28] E. Rocha, P. Salvador, and A. Nogueira. Detection of illicit traffic based on multiscale analysis. In *Software, Telecommunications Computer Networks, 2009. SoftCOM 2009. 17th International Conference on*, pages 286 –291, September 2009.
[29] S. E. Security. Symantec Global Internet Security Threat Report: Trends for 2008. Technical report, Symantec, April 2009.
[30] S. Sen, O. Spatscheck, and D. Wang. Accurate, scalable in-network identification of p2p traffic using application signatures. In *WWW '04: Proceedings of the 13th international conference on World Wide Web*, pages 512–521, New York, NY, USA, 2004. ACM.
[31] M. Taqqu, V. Teverovsky, and W. Willinger. Is network traffic self-similar or multifractal? 5:63–73, 1997.
[32] W. Willinger, V. Paxson, and M. Taqqu. *Self-similarity and Heavy Tails: Structural Modeling of Network Traffic*. A Practical Guide to Heavy Tails: Statistical Techniques and Applications. Birkhauser, 1998.
[33] I. Zukerman, D. W. Albrecht, and A. E. Nicholson. Predicting users' requests on the www. In *Seventh International Conference on User Modeling*, pages 275–284, 1999.

 ISBN: 978-1-61208-103-8

Blocking Equalization in the Erlang Multirate Loss Model for Elastic Traffic

Ioannis D. Moscholios
Dept. of Telecommunications Science and Technology,
University of Peloponnese,
22 100, Tripoli, Greece
idm@uop.gr

Vassilios G. Vassilakis, Michael D. Logothetis
WCL, Dept. of Electrical & Computer Engineering,
University of Patras,
26 504, Patras, Greece
vasilak@wcl.ee.upatras.gr
m-logo@wcl.ee.upatras.gr

Anthony C. Boucouvalas
Dept. of Telecommunications Science and Technology,
University of Peloponnese,
22 100, Tripoli, Greece
acb@uop.gr

***Abstract*—We consider a single-link loss system of fixed bandwidth capacity, which accommodates *K* service-classes of Poisson traffic with different bandwidth-per-call requirements. Depending on the occupied link bandwidth, in-service calls can tolerate bandwidth compression while increasing their service time (elastic calls). In this system, we study the effect of the bandwidth reservation (BR) policy on various performance measures and mainly on Call Blocking Probabilities (CBP). The BR policy can achieve CBP equalization among service-classes, or, alternatively, guarantee a certain quality of service for each service-class. We provide a recurrent formula for the calculation of the link occupancy distribution. Based on it we determine CBP, link utilization and average number of calls in the system. The accuracy of the proposed formula is verified by simulation and is found to be quite accurate.**

Keywords-loss system; blocking probability; reservation; elastic traffic; Markov chain.

I. INTRODUCTION

The classical Erlang Multi-rate Loss Model (EMLM) is used to analyze the call blocking behavior of a single-link loss system that accommodates *K* service-classes with different and fixed bandwidth-per-call requirements. Calls of each service-class arrive to the system according to a Poisson process and compete for the available link bandwidth under the complete sharing (CS) policy (calls of all service-classes compete for all available bandwidth resources). Calls are blocked and lost only if their required bandwidth is higher than the available link bandwidth. Otherwise, they are accepted in the system for a generally distributed service time [1]. Note that while in service, calls cannot alter their assigned bandwidth.

In the EMLM, exploiting the fact that the steady state distribution of the number of calls in the link has a product form solution (PFS) [2], an accurate recursive formula (known as Kaufman-Roberts formula, KR formula) has been separately proposed by Kaufman [1] and Roberts [3] which determines the link occupancy distribution and simplifies the determination of call blocking probabilities (CBP). This simplification resulted in a large amount of extensions of the EMLM and applications of the KR formula both in wired (e.g., [4]-[7]) and wireless networks (e.g., [8]-[11]). Among other EMLM extensions, Roberts proposed in [12] an approximate recursive formula for calculating CBP in the EMLM under the Bandwidth Reservation policy (EMLM/BR). The BR policy is used in order to achieve CBP equalization among service-classes, or guarantee a certain quality of service (QoS) for each service-class. Note that contrary to the CS policy where the stationary probabilities have a PFS, the BR policy cannot be analyzed by the use of a PFS. This is because one-way transitions appear in the state space, which destroy reversibility [13].

In this paper, we apply the BR policy and study its effects in another extension of the EMLM proposed in [14] by Stamatelos and Koukoulidis who incorporate elastic traffic in the EMLM. We name, herein, the model of [14] Extended EMLM (E-EMLM) and our proposed model E-EMLM/BR. In the E-EMLM, calls of each service-class arrive to the system according to a Poisson process with different and elastic bandwidth-per-call requirements. As long as the occupied link bandwidth does not exceed the capacity of the link, all in-service calls use their peak-bandwidth requirement. When a new call arrives and its required peak-bandwidth is higher than the available link bandwidth, the system accepts the new call (contrary to the EMLM where this call is blocked) by compressing not only the bandwidth of all in-service calls (of all service-classes) but also the initial peak-bandwidth of the new call. On the other hand when an in-service call, whose bandwidth is compressed, departs from the system then the remaining in-service calls (of all service-classes) expand their bandwidth. It is worth mentioning that the allocated bandwidth to elastic in-service calls alters (becomes compressed or expanded) in proportion to their peak-bandwidth requirement and that their service time is adjusted accordingly (becomes expanded or compressed) so that the product (service time) by (bandwidth per call) remains constant. A new call is blocked and lost when the compressed bandwidth should be less than a minimum proportion ($r_{\min}$) of its required peak-bandwidth. Note that $r_{\min}$ is common for all service-classes.

The compression/expansion of bandwidth destroys reversibility in the E-EMLM and therefore no PFS exists. However, in [14] an approximate recursive formula is proposed which determines the link occupancy distribution. Before we proceed to the E-EMLM/BR note that extensions of the E-EMLM which study the co-existence of elastic and adaptive traffic (in-service calls can alter their bandwidth but not their service time) can be found in [15], [16]. Potential applications of the E-EMLM (and the E-EMLM/BR) are mainly in emerging wireless networks supporting elastic traffic (e.g., [17], [18]).

 ISBN: 978-1-61208-103-8

Since the proposed E-EMLM/BR does not have a PFS we provide an approximate recursive formula for the calculation of the link occupancy distribution that simplifies the determination of various performance measures including: a) CBP, b) link utilization, c) average number of calls of each service-class in the system and d) delay of calls due to their bandwidth fluctuation.

The remainder of this paper is as follows: In Section II, we review the EMLM, the EMLM/BR and the E-EMLM. In Section III, we propose the E-EMLM/BR. In Section IV, we present numerical results where the new model is compared to the existing models and evaluated through simulation results. We conclude in Section V.

II. REVIEW OF THE EMLM, EMLM/BR AND E-EMLM

A. Review of the EMLM

Consider a link of capacity C bandwidth units (b.u.) that accommodates calls of K service-classes. A call of service-class k (k =1,…,K) arrives in the system according to a Poisson process with rate λ_k, requests b_k b.u. and if these b.u. are available it remains in the system for an exponentially distributed service time with mean μ_k^{-1}. While in service, the call cannot alter its assigned bandwidth. If the b_k b.u. are not available the call is blocked and lost. Let j be the occupied link bandwidth (j=0,…,C) then the link occupancy distribution, $G(j)$, is given by the accurate and recursive KR formula [1] ,[3]:

$$G(j) = \left\langle \begin{array}{c} 1 \quad for\ j = 0 \\ \frac{1}{j}\sum_{k=1}^{K} a_k b_k G(j - b_k)\ for\ j = 1,\ldots, C \\ 0 \quad otherwise \end{array} \right\rangle \tag{1}$$

where: $\alpha_k = \lambda_k \mu_k^{-1}$ is the offered traffic load of service-class k calls.

The proof of (1) is based on the fact that the steady state distribution of the number of calls in the link has a PFS. If n_k is the number of calls of service-class k in the steady state and $\boldsymbol{n}=(n_1, n_2,\ldots,n_k,\ldots, n_K)$ then the steady state distribution, $P(\boldsymbol{n})$, is given by [2]:

$$P(\boldsymbol{n}) = G^{-1}\left(\prod_{k=1}^{K}\frac{a_k^{n_k}}{n_k!}\right) \tag{2}$$

where: G is the normalization constant given by $G \equiv G(\boldsymbol{\Omega}) = \sum_{\boldsymbol{n}\in\boldsymbol{\Omega}}\left(\prod_{k=1}^{K}\frac{a_k^{n_k}}{n_k!}\right)$ and $\boldsymbol{\Omega}=\{\boldsymbol{n}: 0 \le \boldsymbol{nb} \le C\}$ is the state space with $\boldsymbol{b}=(b_1, b_2,\ldots,b_k,\ldots, b_K)$ and $j = \boldsymbol{nb} = \sum_{k=1}^{K} n_k b_k$.

Having determined the values of $G(j)$'s we can calculate various performance measures, including:

1) The CBP of service-class k, denoted as B_k, is calculated by the formula:

$$B_k = \sum_{j=C-b_k+1}^{C} G^{-1}G(j) \tag{3}$$

where $G = \sum_{j=0}^{C} G(j)$ is the normalization constant.

2) the link utilization, denoted as U :

$$U = \sum_{j=1}^{C} jG(j) \tag{4}$$

3) The average number of service k calls in the system, denoted as $\bar{n}_k$:

$$\bar{n}_k = \sum_{j=1}^{C} y_k(j)G(j) \tag{5}$$

where $y_k(j)$ is the average number of service-class k calls given that the system state is j, and can be determined by (proof is similar to [15] and thus is omitted):

$$y_k(j) = \frac{1}{jG(j)}\left[a_k b_k G(j-b_k)(1+y_k(j-b_k))\right] + \frac{1}{jG(j)}\sum_{\substack{i=1 \\ i\neq k}}^{K} a_i b_i G(j-b_i) y_k(j-b_i) \tag{6}$$

where j = 1,…,C while $y_k(x)$ = 0 for $x \le 0$ and k = 1, …, K.

B. Review of the EMLM/BR

If we apply the BR policy to the EMLM according to Roberts [12], then the formula for the approximate calculation of $G(j)$ takes the form:

$$G(j) = \left\langle \begin{array}{c} 1 \quad for\ j = 0 \\ \frac{1}{j}\sum_{k=1}^{K} a_k D_k(j-b_k) G(j-b_k)\ for\ j = 1,\ldots, C \\ 0 \quad otherwise \end{array} \right\rangle \tag{7}$$

where: $D_k(j-b_k) = \left\langle \begin{array}{l} b_k \ for\ j \le C - t(k) \\ 0 \ for\ j > C - t(k) \end{array} \right\rangle$ (8)

and $t(k)$ is the reserved bandwidth (BR parameter) for service-class k calls.

Note that (7) is recursive, although the EMLM/BR is a non PFS model. This feature is based on the assumption (approximation) that calls of service-class k do not exist (are negligible) in states $j > C$-$t(k)$ and is incorporated in (7) by the variable $D_k(j$-$b_k)$ of (8). The BR policy is used to attain

CBP equalization among different service-classes that share a link by a proper selection of the BR parameters. If, for example, CBP equalization is required between calls of two service-classes with b_1=1 and b_2=10 b.u., respectively, then $t(1)$ = 9 b.u and $t(2)$ = 0 b.u. so that $b_1 + t(1) = b_2 + t(2)$. Note that $t(1)$ = 9 b.u means that 9 b.u. are reserved to benefit calls of the 2nd service-class. The CBP of service-class k, B_k, in the EMLM/BR is given by:

$$B_k = \sum_{j=C-b_k+1}^{C} G^{-1} G(j) \quad (9)$$

If $t(k) = 0$ for all k (k=1,…,K) then the EMLM results.

Having obtained the values of $G(j)$'s according to (7) we calculate the link utilization and the average number of service-class k calls in the system according to (4) and (5), respectively.

C. Review of the E-EMLM

Consider again a link of capacity C b.u. that accommodates calls of K service-classes. A call of service-class k (k =1,…,K) arrives in the system according to a Poisson process with arrival rate λ_k and requests b_k b.u. (peak-bandwidth requirement). If $j + b_k \le C$, the call is accepted in the system with its peak-bandwidth requirement and remains in the system for an exponentially distributed service time with mean μ_k^{-1}. If $T \ge j + b_k > C$ the call is accepted in the system by compressing not only its peak-bandwidth requirement but also the assigned bandwidth of all in-service calls. The compressed bandwidth of the new service-class k call is:

$$b_k' = r b_k = \frac{C}{j'} b_k \quad (10)$$

where $r \equiv r(\boldsymbol{n}) = C/j'$, $j' = j + b_k = \boldsymbol{nb} + b_k$ and T is the limit (in b.u.) up to which bandwidth compression is permitted.

Similarly, the bandwidth of all in-service calls will be compressed and become equal to $b_i' = \frac{C}{j'} b_i$ for i = 1,…,K. After the compression of both the new call and the in-service calls the state of the system is $j = C$. The minimum bandwidth that a call of service-class k (either new or in-service) can tolerate is given by the expression:

$$b_{k,\min}' = r_{\min} b_k = \frac{C}{T} b_k \quad (11)$$

where $r_{\min} = C/T$ is the minimum proportion of the required peak-bandwidth and is common for all service-classes.
This means that if upon arrival of a service-class k call, with peak-bandwidth requirement b_k b.u., we have $j' = j + b_k > T$ (or equivalently, $j' > T$ or $C/j' < r_{\min}$) then the call is blocked and lost without further affecting the system.

After the bandwidth compression, calls increase their service time so that the product (service time) by (bandwidth per call) remains constant. Thus, due to bandwidth compression calls of service-class k may remain in the system more than μ_k^{-1} time units. Increasing the value of T, decreases $r_{\min}$ and increases the delay of calls of service-class k (compared to the initial service time μ_k^{-1}). Therefore the value of T can be chosen so that this delay remains within acceptable levels.

To illustrate the previous compression mechanism consider the following simple example. Let C = 3 b.u., T = 5. b.u., K = 2 service-classes, $\alpha_1 = \alpha_2 = 1$ erl, b_1 = 1 b.u., b_2 = 2 b.u and $\mu_1^{-1} = \mu_2^{-1} = 1$ time unit. The permissible states $\boldsymbol{n} = (n_1, n_2)$ of the system are 12 and are presented in Table I together with the occupied link bandwidth, $j = n_1 b_1 + n_2 b_2$, before and after compression has been applied. Note that compression is applied if $T \ge j > C$ (bold values of the 3rd column of Table I). After compression has been applied, we have that $j = C$ (bold values of the 4th column of Table I). For example, assume that a new 2nd service-class call arrives while the system is in state $(n_1, n_2) = (1, 1)$ and $j = C$ = 3 b.u. The new call is accepted in the system, since $j' = j + b_2 = T$ = 5 b.u., after bandwidth compression has been applied to all calls (new and in-service calls). The new state of the system is now $(n_1, n_2) = (1, 2)$. In this state, and based on (11), calls of the 1st and 2nd service-class compress their bandwidth to:

$$b_{1,\min}' = r_{\min} b_1 = \frac{3}{5} b_1 = 0.6, \quad b_{2,\min}' = r_{\min} b_2 = \frac{3}{5} b_2 = 1.2$$

so that $j = n_1 b_{1,\min}' + n_2 b_{2,\min}' = 0.6 + 2.4 = 3 = C$.

Similarly, the values of μ_1^{-1}, μ_2^{-1} become $\frac{\mu_1^{-1}}{r_{\min}}, \frac{\mu_2^{-1}}{r_{\min}}$ so that $b_1 \mu_1^{-1}$ and $b_2 \mu_2^{-1}$ remain constant.

Consider now that the system is in state $(n_1, n_2) = (1, 2)$ and a 2nd service-class call departs from the system. Then, its assigned bandwidth $b_{2,\min}' = 1.2$ is shared to the remaining calls in proportion to their peak-bandwidth requirement. Thus, in the new state $(n_1, n_2) = (1, 1)$ the 1st service-class call expands its bandwidth to $b_1' = b_1 = 1$ b.u. and the 2nd service-class call to $b_2' = b_2 = 2$ b.u. Thus, $j = n_1 b_1 + n_2 b_2 = C$ = 3 b.u. Furthermore, the service times of both calls are decreased to their initial values $\mu_1^{-1} = \mu_2^{-1} = 1$ time unit.

The compression/expansion of bandwidth destroys reversibility in the E-EMLM and therefore no PFS exists. However, in [14] an approximate recursive formula is proposed which determines $G(j)$'s:

$$G(j) = \left\langle \begin{array}{c} 1 \quad for \ \ j = 0 \\ \frac{1}{\min(j, C)} \sum_{k=1}^{K} a_k b_k G(j - b_k) \ \ for \ \ j = 1, \ldots, T \\ 0 \quad otherwise \end{array} \right\rangle \quad (12)$$

Equation (12) is based on a reversible Markov chain which approximates the bandwidth compression/expansion mechanism of the E-EMLM, described above. The local balance equations of this Markov chain are of the form [14]:

$$\lambda_k P(\boldsymbol{n}_k^-) = n_k \mu_k \phi_k(\boldsymbol{n}) P(\boldsymbol{n}) \quad (13)$$

where $P(\boldsymbol{n}) = (n_1, n_2,\ldots,n_k,\ldots, n_K)$, $P(\boldsymbol{n}_k^-) = (n_1, n_2,\ldots,n_{k-1},n_k-1,n_{k+1}\ldots, n_K)$ and $\varphi_k(\boldsymbol{n})$ is a state dependent factor which describes: i) the compression factor of bandwidth and ii) the increase factor of service time of service-class k calls in state $\boldsymbol{n}$, so that (service time) by (bandwidth per call) remains constant. In other words, $\varphi_k(\boldsymbol{n})$ has the same role with $r(\boldsymbol{n})$ in (10) or $r_{\min}$ in (11) but it may be different for each service-class. It is apparent now why the model of (12) approximates the E-EMLM. The values of $\varphi_k(\boldsymbol{n})$ are given by:

$$\phi_k(\boldsymbol{n}) = \begin{cases} 1 & , \text{ when } \boldsymbol{nb} \le C \text{ and } \boldsymbol{n} \text{ in } \boldsymbol{\Omega} \\ \dfrac{x(\boldsymbol{n}_k^-)}{x(\boldsymbol{n})} & , \text{ when } C < \boldsymbol{nb} \le T \text{ and } \boldsymbol{n} \text{ in } \boldsymbol{\Omega} \\ 0 & , \text{ otherwise} \end{cases} \quad (14)$$

where $\boldsymbol{\Omega}=\{\boldsymbol{n}: 0 \le \boldsymbol{nb} \le T\}$ and $\boldsymbol{nb} = \sum_{k=1}^{K} n_k b_k$.

In (14), $x(\boldsymbol{n})$ is a state multiplier, associated with state $\boldsymbol{n}$, whose values, are chosen so that (13) holds, [14]:

$$x(\boldsymbol{n}) = \begin{cases} 1 & , \text{ when } \boldsymbol{nb} \le C, \ \boldsymbol{n} \text{ in } \boldsymbol{\Omega} \\ \dfrac{1}{C}\sum_{k=1}^{K} n_k b_k x(\boldsymbol{n}_k^-), & \text{when } C < \boldsymbol{nb} \le T, \ \boldsymbol{n} \text{ in } \boldsymbol{\Omega} \\ 0 & , \text{ otherwise} \end{cases} \quad (15)$$

Table II shows, for our simple example, the values of $r(\boldsymbol{n})$ (common for both service-classes), $\varphi_1(\boldsymbol{n})$ and $\varphi_2(\boldsymbol{n})$.

TABLE I. STATE SPACE AND OCCUPIED LINK BANDWIDTH

n_1	n_2	j (before compression) $0 \le j \le T$	j (after compression) $0 \le j \le C$
0	0	0	0
0	1	2	2
0	2	**4**	**3**
1	0	1	1
1	1	3	3
1	2	**5**	**3**
2	0	2	2
2	1	**4**	**3**
3	0	3	3
3	1	**5**	**3**
4	0	**4**	**3**
5	0	**5**	**3**

TABLE II. VALUES OF STATE DEPENDENT FACTORS

n_1	n_2	$r(n)$	$\varphi_1(n)$	$\varphi_2(n)$
0	0	1.00	1.00	1.00
0	1	1.00	1.00	1.00
0	2	0.75	0.00	0.75
1	0	1.00	1.00	1.00
1	1	1.00	1.00	1.00
1	2	0.60	0.75	0.5625
2	0	1.00	1.00	1.00
2	1	0.75	0.75	0.75
3	0	1.00	1.00	1.00
3	1	0.60	0.67	0.50
4	0	0.75	0.75	0.00
5	0	0.60	0.60	0.00

Having determined the values of $G(j)$'s we can calculate various performance measures, including:

1) The CBP of service-class k, B_k :

$$B_k = \sum_{j=T-b_k+1}^{T} G^{-1} G(j) \quad (16)$$

where $G = \sum_{j=0}^{T} G(j)$ is the normalization constant.

2) the link utilization, denoted as U :

$$U = \sum_{j=1}^{C} jG(j) + \sum_{j=C+1}^{T} CG(j) \quad (17)$$

3) The average number of service-class k calls in the system, $\bar{n}_k$:

$$\bar{n}_k = \sum_{j=1}^{T} y_k(j) G(j) \quad (18)$$

where $y_k(j)$ is the average number of service-class k calls given that the system state is j, and is given by [15]:

$$y_k(j) = \frac{1}{\min(j,C)G(j)}\left[a_k b_k G(j-b_k)(1+y_k(j-b_k))\right] + \frac{1}{\min(j,C)G(j)}\sum_{\substack{i=1 \\ i\ne k}}^{K} a_i b_i G(j-b_i) y_k(j-b_i) \quad (19)$$

where $j = 1,\ldots,T$ while $y_k(x) = 0$ for $x \le 0$ and $k = 1, \ldots, K$.

4) The average delay of service-class k calls, denoted by D_k, given by Little's formula, [19]:

$$D_k = \frac{\bar{n}_k}{\lambda_k (1 - B_k)} \quad (20)$$

As T increases, B_k decreases and D_k increases. Therefore, the choice of T can be a trade-off between B_k and D_k. Before we proceed to the application of the BR policy in the E-EMLM we give the accurate and approximate CBP results for our simple example in the E-EMLM, and the corresponding CBP results for the EMLM, when C = 3:

E-EMLM
Accurate CBP: B_1 = 17.48%, B_2 = 35.74%
Approx. CBP (based on (12), (16)): B_1=17.00%, B_2=36.04%
EMLM
Accurate CBP (based on (1), (3)): B_1= 25.00%, B_2 = 57.14%

In the E-EMLM, the accurate CBP results are based on the numerical calculation of the irreversible Markov chain. The comparison shows that even in a small example, the approximation of [14] is quite well. Furthermore, compared to the EMLM we see a substantial CBP decrease due to the existence of a compression/expansion mechanism.

III. THE E-EMLM UNDER THE BR POLICY

If we apply the BR policy to the E-EMLM (E-EMLM/BR) according to [12], then (12) takes the form:

$$G(j)=\left\langle \begin{array}{c} 1 \quad for\ j=0 \\ \dfrac{1}{\min(j,C)}\sum_{k=1}^{K} a_k D(j-b_k)G(j-b_k)\ for\ j=1,\ldots,T \\ 0 \quad otherwise \end{array} \right\rangle \quad (21)$$

where: $$D_k(j-b_k)=\left\langle \begin{array}{l} b_k \ \ for\ j \le T-t(k) \\ 0 \ \ for\ j > T-t(k) \end{array} \right\rangle \quad (22)$$

and $t(k)$ is the reserved bandwidth (BR parameter) for service-class k calls.

The CBP of service-class k, B_k, in the E-EMLM/BR is given by:

$$B_k = \sum_{j=T-b_k-t(k)+1}^{T} G^{-1}G(j) \quad (23)$$

If $t(k)$ = 0 for all k (k=1,…,K) then the E-EMLM results. Having obtained the values of $G(j)$'s according to (22) we can calculate the link utilization, the average number of service-class k calls in the system and their average delay according to (18), (19) and (20), respectively. Note that in (19), $y_k(j)$ = 0 if $j>T-t(k)$ as (22) implies. The accurate and approximate equalized CBP results for our simple example in the E-EMLM/BR, and the corresponding equalized CBP results for the EMLM/BR, when C = 3 are:

E-EMLM/BR ($t(1)$ = 1, $t(2)$=0)
Accurate CBP: $B_1 = B_2$ =32.34%
Approx. CBP (based on (21)-(23)): $B_1 = B_2$ = 31.71%
EMLM/BR ($t(1)$ = 1, $t(2)$=0)
Accurate CBP: $B_1 = B_2$ = 54.5%
Approx. CBP (based on (7)-(9)): $B_1 = B_2$ = 52.0%

In the E-EMLM/BR and the EMLM/BR, the accurate CBP results are based on the numerical calculation of the corresponding irreversible Markov chains. In the E-EMLM/BR, the comparison between the accurate and approximate CBP results shows that even in this small example, the proposed formulas ((21)-(23)) are valid.

IV. EVALUATION

In this section, we compare the analytical CBP and link utilization results obtained by the EMLM, EMLM/BR, E-EMLM and E-EMLM/BR via a numerical example. Due to space limitations we present simulation CBP results (mean values of 7 runs) only for the E-EMLM and the E-EMLM/BR. Simulation is based on Simscript II.5 [20].

We consider a single link of capacity C = 60 b.u. that accommodates calls of two service-classes, with the following traffic characteristics:
1st service-class: α_1 = 24 erl, b_1 = 1 b.u.
2nd service-class: α_2 = 6 erl, b_2 = 4 b.u.

The value of T = 70, and $r_{min} = C/T$=6/7 is the minimum proportion of the required peak-bandwidth. In the case of the BR policy, we choose $t(1)$=3 and $t(2)$=0 in order to achieve CBP equalization between the two service-classes since: $b_1 + t(1) = b_2 + t(2)$. In the x-axis of all figures, α_1 increases in steps of 1 erl while α_2 is constant. So Point 1 is (α_1, α_2) = (24.0, 6.0) while Point 8 is (α_1, α_2) = (31.0, 6.0). In Fig. 1 and 2, we present the analytical and the simulation CBP results of the 1st and the 2nd service-class calls, respectively, in the case of the E-EMLM. For comparison, we give the corresponding analytical CBP results of the EMLM. In Fig. 3, we present the analytical and simulation CBP results (equalized CBP) in the case of the E-EMLM/BR policy. For comparison, we give the corresponding analytical results for the EMLM/BR. All figures show that: i) analytical and simulation CBP results are very close and ii) the compression/expansion mechanism of the E-EMLM and the E-EMLM/BR, reduces the CBP compared to those obtained by the EMLM and the EMLM/BR, respectively. Finally in Fig.4, we present the link utilization (analytical results) for all models. The compression/expansion mechanism increases the link utilization since it decreases CBP.

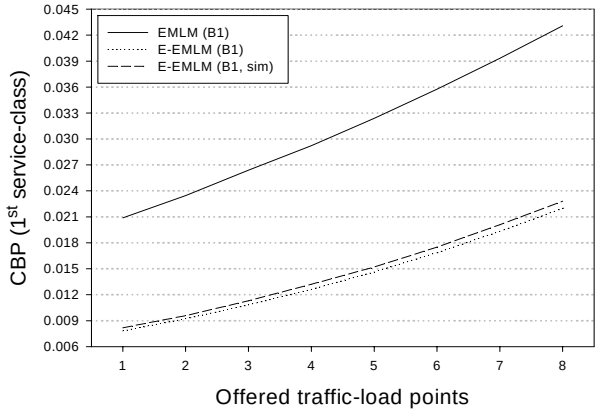


Figure 1. CBP of the 1st service-class (EMLM, E-EMLM).

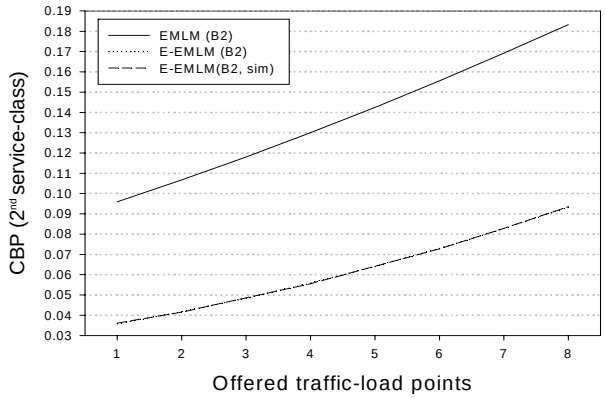


Figure 2. CBP of the 2nd service-class (EMLM, E-EMLM).

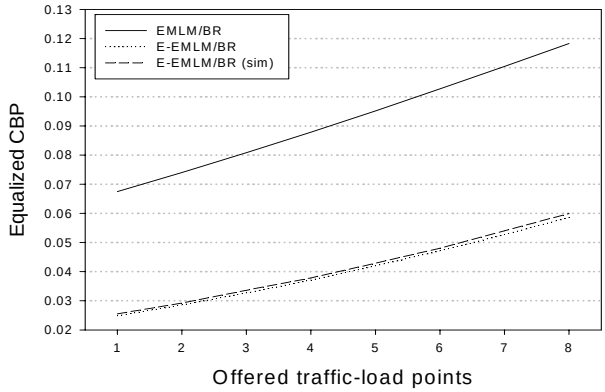


Figure 3. Equalized CBP (EMLM/BR, E-EMLM/BR).

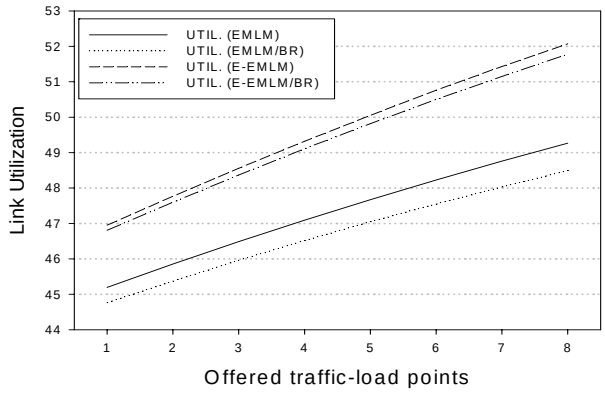


Figure 4. Link utilization for all models.

V. CONCLUSION

We propose an analytical model for the recursive calculation of various performance measures in the E-EMLM/BR. The BR policy guarantees a certain QoS among elastic calls of different service-classes. Simulation results verify the analytical results and prove the accuracy and the consistency of the proposed model. Potential applications of the proposed model are in emerging wireless networks that support elastic traffic. A future extension, is the application of our model in such networks based on the reduced load approximation method which has been extensively used for the CBP calculation in multirate loss networks.

REFERENCES

[1] J. Kaufman, “Blocking in a shared resource environment”, IEEE Trans. Commun. vol. 29, Oct. 1981, pp. 1474-1481.

[2] K. Ross, “Multiservice Loss Models for Broadband Telecommunication Networks”, Springer, 1995.

[3] J. Roberts, “A service system with heterogeneous user requirements”, in: G. Pujolle (Ed.), Performance of Data Communications systems and their applications, North Holland, Amsterdam, 1981, pp.423-431.

[4] A. Greenberg and R. Srikant, “Computational Techniques for Accurate Performance Evaluation of Multirate, Multihop Communication Networks”, IEEE/ACM Trans. on Networking, Vol. 5, April 1997, pp.266-277.

[5] M. Logothetis and G. Kokkinakis, “Path Bandwidth Management for Large Scale Telecom Networks”, IEICE Trans. Commun., vol.E83-B, Sept. 2000, pp.2087-2099.

[6] H. Shengye, Y. Wu, F. Suili, and S. Hui, “Coordination-based optimisation of path bandwidth allocation for large-scale telecommunication networks”, Computer Communications, vol. 27, Jan. 2004, pp.70-80.

[7] I. Moscholios, M. Logothetis, and G. Kokkinakis, “Call-burst blocking of ON-OFF traffic sources with retrials under the complete sharing policy”, Performance Evaluation, vol. 59, March 2005, pp.279-312.

[8] P. Fazekas, S. Imre, and M. Telek, “Modeling and Analysis of Broadband Cellular Networks with Multimedia Connections”, Telecommunication systems, vol. 19, 2002, pp. 263-288.

[9] D. Staehle and A. Mäder, “An Analytic Approximation of the Uplink Capacity in a UMTS Network with Heterogeneous Traffic”, Proc. 18th ITC, Berlin, Sept. 2003, pp. 81-90.

[10] A. Mäder and D. Staehle, “Analytic Modeling of the WCDMA Downlink Capacity in Multi-Service Environments”, ITC Specialist Seminar on Performance Evaluation of Wireless and Mobile Systems, Sept. 2004, pp.217-226.

[11] M. Glabowski, M. Stasiak, A. Wisniewski, and P. Zwierzykowski, “Blocking Probability Calculation for Cellular Systems with WCDMA Radio Interface Servicing PCT1 and PCT2 Multirate Traffic”, IEICE Trans. Commun., vol.E92-B, April 2009, pp.1156-1165.

[12] J. Roberts, “Teletraffic models for the Telecom 1 Integrated Services Network”, Proceedings of ITC-10, Mondreal, Canada, 1983.

[13] F. Kelly, “Reversibility and Stochastic Networks”, Wiley Series in Probability and Mathematical Statistics. Wiley: New York, 1979.

[14] G. Stamatelos and V. Koukoulidis, “Reservation – Based Bandwidth Allocation in a Radio ATM Network”, IEEE/ACM Trans. Networking, vol. 5, June 1997, pp.420-428.

[15] S. Racz, B. P. Gero, and G. Fodor, “Flow level performance analysis of a multi-service system supporting elastic and adaptive services”, Performance Evaluation, vol. 49, Sept. 2002, pp.451-469.

[16] V. Vassilakis, I. Moscholios, and M. Logothetis, “Call-level Performance Modelling of Elastic and Adaptive Service-classes with Finite Population”, IEICE Trans. Commun., vol. E91-B, Jan. 2008, pp.151-163.

[17] G. Fodor and M. Telek, “Bounding the Blocking Probabilities in Multirate CDMA Networks Supporting Elastic Services”, IEEE/ACM Trans. on Networking, vol. 15, Aug. 2007, pp.944-956.

[18] V. Vassilakis, G. Kallos, I. Moscholios, and M. Logothetis, “Call-Level Analysis of W-CDMA Networks Supporting Elastic Services of Finite Population”, IEEE ICC, May 2008, pp.285-290.

[19] V. Koukoulidis, “A Characterization of Reversible Markov Processes with Applications to Shared-Resource Environments”, Phd Thesis, Concordia University, 1993.

[20] Simscript II.5, http://www.simscript.com

 ISBN: 978-1-61208-103-8

Estimation of Traffic Amounts on all Links by Using the Information From a Subset of Nodes

Yuya Tarutani*, Yuichi Ohsita†, Shin'ichi Arakawa*, and Masayuki Murata*
*Graduate School of Information Science and Technology, Osaka University, Osaka, Japan
{y-tarutn, arakawa, murata}@ist.osaka-u.ac.jp
†Graduate School of Economics, Osaka University, Osaka, Japan
y-ohsita@econ.osaka-u.ac.jp

***Abstract*—Traffic information is required to perform traffic engineering. However, as the network that require traffic engineering becomes large, the overhead to collect the traffic amount information required for traffic engineering becomes large. In this paper, we propose a method to reduce the overhead to collect the traffic amount information. In our method, we select a subset of nodes and collect the traffic amount information only from the selected nodes. Then, we estimate the traffic amount on each link by using the information collected from the selected nodes. According to simulation results, our method can estimate the traffic amount on each link required for traffic engineering accurately by monitoring 30% of all nodes.**

***Keywords*-Estimation; Selection of Monitoring Nodes; Traffic Matrix; Traffic Engineering.**

I. INTRODUCTION

In recent years, various applications are deployed and their traffic is carried over the Internet. The traffic in the Internet still doubles each year, the network providers are required to accommodate their traffic in a cost effective way.

Traffic engineering is a method to optimize the performance of networks by dynamically changing the topology and/or route of traffic [1-5]. The topology and route of traffic are calculated and controlled with a server called Path Computation Element (PCE). To perform the traffic engineering at the PCE, we may need to know the information of traffic in the network. Depending on the granularity of traffic engineering, we may require different degree of traffic information. For example, when we apply some optimization technique to determine the topology and route of traffic, we need the traffic matrix that expresses the traffic amount for each edge-to-edge traffic in the network.

However, collecting all the edge-to-edge traffic requires much overhead: one reason is the monitoring overhead at each router. Header inspection is necessary at the router to identify the edge-to-edge traffic that monitored packets belong to. However, the header inspection caused the overhead at each router; another reason is collecting overhead at the PCE server. To collect all the edge-to-edge traffic, the PCE server has to query the nodes monitoring edge-to-edge traffic and obtain the information of the traffic amounts. Especially, as the number of routers increases, the number of nodes the PCE server has to query and the size of information to be collected become large, which causes the significant overhead at the PCE server.

To overcome these overheads, traffic engineering using only the information of traffic amount on each link has been investigated. Juva [5] calculates the range of each edge-to-edge traffic by using the information of traffic amount on each link, and optimizes the traffic routes that minimize the worst-case link utilization. Roughan et al. [2] and Ohsita et al. [3, 4] use the traffic matrices estimated from the information of traffic amount on each link. The traffic amount on a link can be easily counted at the node connected to the link, and the PCE can obtain the information of traffic amount on the link by querying the node.

However, the granularity of traffic information depends on the application of them. Recently, network virtualization [6] to support deployments of various network services, such as P2P services and cloud computing services, has been investigated. Virtual networks are prepared and reconfigured for each service. In this case, traffic engineering is required for each virtual networks and the PCE has to collect the information of traffic amount on each virtual link of each virtual network. It will be thought that diversification of the service in the network advances more and the number of virtual networks that require traffic engineering remarkably increases in future. That is, the number of links whose information must be collected by the PCE increases in future as the number of service increases. This may cause the heavy collecting overhead at the PCE.

In this paper, we propose a method to reduce the overhead to collect traffic information necessary for traffic engineering by estimating traffic amounts on all links from the traffic information collected from a subset of nodes. In our method, we first select the nodes we collect the traffic information from, and collect the information of traffic amount on each link from the selected nodes. Then, we estimate the traffic amounts of all links by using only the information collected from the selected nodes. Throughout this paper, we call the selected node *monitoring node*.

The rest of this paper is organized as follows. Section II explains the existing methods to estimate traffic matrices. In Section III, we propose a method to select monitoring nodes and estimate traffic amount on each link by using

 ISBN: 978-1-61208-103-8

the information collected from the monitoring nodes. In Section IV, we evaluate our method by simulation and clarify that our method can estimate traffic amount on each link accurately by selecting the monitoring node properly. Finally, Section V provides a conclusion.

II. Overview of Traffic matrix estimation

Traffic matrix is the matrix of $T_{s,d}$ that represents the traffic amount from node s to node d. Let N be the number of nodes in the network. Then, the traffic matrix is represented as,

$$T = \begin{bmatrix} T_{1,1} \\ T_{1.2} \\ \vdots \\ T_{N,N} \end{bmatrix}. \quad (1)$$

As this equation indicates, obtaining the traffic matrix requires the traffic information between all nodes and requires more overhead as the number of nodes increases. Therefore methods to estimate a traffic matrix from the traffic amount on each link have been investigated. The traffic amount on each link is determined from the routing information A, which is known to the network administrator, and traffic matrix T, which is unknown to the network administrator. That is, the following equation is hold;

$$AT = X, \quad (2)$$

where X is a matrix of X_i that represents the traffic amount that pass through the link i. That is,

$$X = \begin{bmatrix} X_1 \\ \vdots \\ X_L \end{bmatrix}. \quad (3)$$

In the above equation, L is the number of links in the networks. A is a matrix that has an element $A_{s,d,l}$ that represents the route of flow between node s and d and when the flow passes through the link l, $A_{s,d,l}$ takes one, otherwise takes zero. Note that when we consider the splittable flow, $A_{s,d,l}$ is the rate of end-to-end traffic $T_{s,d}$ flows the link l. A is called *routing matrix*.

$$A = \begin{bmatrix} A_{1,1,1} & A_{1,2,1} & \cdots & A_{N,N,1} \\ A_{1,1,2} & A_{1,2,2} & \cdots & A_{N,N,2} \\ \vdots & \vdots & \ddots & \vdots \\ A_{1,1,L} & A_{1,2,L} & \cdots & A_{N,N,L} \end{bmatrix} \quad (4)$$

Traffic matrix estimation is an approach to estimate T that satisfies the Eq. 2, based on the monitored traffic amount X and the routing matirx A. However, we cannot obtain the unique traffic matrix that satisfies Eq. 2 since the number of equations in Eq. 2 is usually less than the number of elements in T. That is, there are several candidates for the traffic matrix to satisfy the Eq. 2.

Many approaches have been considered to obtain the true traffic matrix from the candidates. One of approaches is to use the model of traffic matrix [7-12]. Zhang et al. [11] proposed the estimation method called *tomogravity method* that estimates traffic matrix so as to follow the gravity model where the traffic amount between two nodes is proportional to the product of the traffic of the two nodes. The tomogravity method works as follows. At first, the method estimates the edge-to-edge traffic $T^{grav}_{s,d}$ based on monitored traffic in the ingress and egress links to follow the gravity model by the following equations;

$$T^{grav}_{s,d} = X_{l^{in}_s} \frac{X_{l^{out}_d}}{\sum_k X_{l^{out}_k}}, \quad (5)$$

where l^{in}_s is the ingress link at node s and l^{out}_d is the egress link at node d. We denote T^{grav} as the matrix in which each entry is $T^{grav}_{s,d}$. Then, the tomogravity method estimates a traffic matrix $\hat{T}$ by following equations;

$$\min \|\hat{T} - T^{grav}\|, \quad (6)$$
$$s.t.\ A\hat{T} = X.$$

That is, the tomogravity method calculates $\hat{T}$ that satisfies Eq. 2 and minimize the difference between $\hat{T}$ and T^{grav}. Although the traffic information required for the tomogravity method is much smaller than the case of collecting traffic matrix information directly, L numbers of traffic information are still required to be collected to estimate traffic matrices.

One approach to reduce the collecting overhead is to collect traffic amount information only from a subset of nodes and estimate the uncollected traffic amount information from the collected information. Zhang et al. [12] proposed a method to estimate the uncollected traffic amount information from the traffic amount information that was monitored and collected before at the same point or currently at the different points. In this method, we calculate the correlation between each traffic amount information monitored at different times or different points by using the traffic amount information collected before. Then, we estimate the uncollected traffic amount information by using the correlation. However, this method cannot estimate the uncollected traffic amount information accurately when the traffic change that is different from a past tendency occurs.

In this paper, we investigate the method to estimate the uncollected traffic amount information of each link from the information collected from a subset of monitoring nodes without using the past information. In addition, because the accuracy of the estimation depends on selection of the monitoring nodes, we also propose a method to select the monitoring nodes.

III. Estimation of the traffic amounts on all links from the information of a subset of nodes

In this section, we propose a method to estimate traffic amounts on all links by using the traffic amounts monitored at a subset of nodes. In addition, since the accuracy of the estimated traffic amounts depends on the selection of the

 ISBN: 978-1-61208-103-8

monitoring nodes, we also propose a method to select the monitoring nodes. Our method works as the following steps.

Step. 1 Select monitoring nodes and collect the information of traffic amount from the selected monitoring nodes.

Step. 2 Estimate the traffic amount on each link X' roughly by using the number of edge-to-edge traffic passing the link.

Step. 3 Estimate the traffic matrix $\hat{T}$ from the traffic amount on each link X' and the routing matrix A.

Step. 4 Estimate the traffic amount on each link $\hat{X}$ from the estimated traffic matrix $\hat{T}$ and the routing matrix A.

After performing the above steps, we designate $\hat{X}$ as the final estimation results for the traffic amount on each link. The details of the above steps are described below.

A. Selecting monitoring nodes

In this subsection, we propose a method to select monitoring nodes so as to estimate the traffic amounts on all links accurately. The edge-to-edge traffic whose amount is not monitored at any monitoring nodes is difficult to estimate and may also cause large estimation errors on the traffic amount on each link. Thus, in our method, we select monitoring nodes so as to cover as many edge-to-edge traffic as possible. In addition, when no nodes can increase the number of edge-to-edge traffic covered by the selected monitoring nodes, we select the nodes where the number of edge-to-edge traffic passing the node is the largest so as to increase the accuracy of as many edge-to-edge traffic as possible.

In our method, initially we regard all nodes as the candidates for the monitoring nodes. Then, we eliminate the selected nodes from the candidate until the number of remaining candidates becomes the target number of monitoring nodes H.

To select the nodes eliminated from the candidates for the monitoring nodes, we use the number of edge-to-edge traffic monitored by node i (Q_i), the number of edge-to-edge traffic that cannot be monitored at any other candidates than node i (P_i), and the number of candidates passed by the edge-to-edge traffic from node n to node m ($R_{n,m}$). Our method selects the monitoring nodes by the following steps.

Step. 1.1 Select all nodes as candidates for the monitoring nodes .

Step. 1.2 Initialize P_i to 0, Q_i to the number of edge-to-edge traffic passing the node i and $R_{n,m}$ to the number of nodes passed by the edge-to-edge traffic from node n to node m.

Step. 1.3 If there exists the node whose P_i is 0, eliminate the node whose Q_i is the smallest among the candidates whose P_i is 0 from the candidates,and then go to Step 1.5. Otherwise, go to step 1.4.

Step. 1.4 If $P_i > 0$ for all nodes, eliminate the node whose P_i is the smallest from the candidates.

Step. 1.5 If the number of candidates is larger than the threshold H, update $R_{n,m}$ and P_i for all candidates and go back to Step 1.3. Otherwise, go to step 1.6.

Step. 1.6 Designate the remaining candidates as the monitoring nodes.

In the Step. 1.5 of the above steps, $R_{n,m}$ is updated by decrementing its value if the edge-to-edge traffic from node n to node m passes the node eliminated from the candidates. Then, P_i is updated by counting the elements of $R_{n,m}$ where the edge-to-edge traffic from node n to node m passes through the node i and $R_{n,m} = 1$.

B. Estimation of traffic amounts by using the number of edge-to-edge traffic

In our method, we use only traffic amount information monitored at the selected monitoring nodes. However, the lack of traffic amount information causes the difficulty in estimating traffic matrices. Thus, we estimate the uncollected traffic amount information before estimating the traffic matrix.

To estimate the traffic amounts, we use the relation between the number of edge-to-edge traffic passing a link and the traffic amounts on the link. We investigate this relation by simulation. In this simulation, we use AT&T's router-level topology (523 nodes and 1304 links) measured in Ref. [13]. We add one ingress link and one egress link for all nodes in the AT&T topology, and generate traffic between each pair of ingress and egress links.

According to Ref. [11], actual traffic matrices follow the gravity model. In addition, according to Ref. [14], each element of actual traffic matrices obeys a lognormal distribution. Thus, in this simulation, we generate traffic matrix T indicating traffic amounts between each ingress and egress links so as to follow both the gravity model and a lognormal distribution. The traffic matrix T used in this simulation is generated as

$$T = T^{grav} + \Delta, \tag{7}$$

where T^{grav} is a traffic matrix generated so as to follow both the gravity model and a lognormal distribution, and Δ is a matrix indicating the white Gaussian noise with the mean of 0 and the variance of 1. We generate $T^{grav}_{i,j}$ as

$$T^{grav}_{i,j} = G_i * G_j, \tag{8}$$

where G_i is the weight for node i. We generate G_i based on the lognormal distribution with a mean of $e^{4.8}$ and the variance $e^{9.7}$ so as to match the results described in Ref. [14]. In this simulation, the unit of the traffic amount of the edge-to-edge traffic generated the above steps is Mbps.

Fig. 1 shows the relation between the number of edge-to-edge traffic passing a link and the traffic amount on the link obtained by our simulation. According to Fig. 1, we can model the relation as

$$W_i = \alpha Z_i + \beta, \tag{9}$$

 ISBN: 978-1-61208-103-8

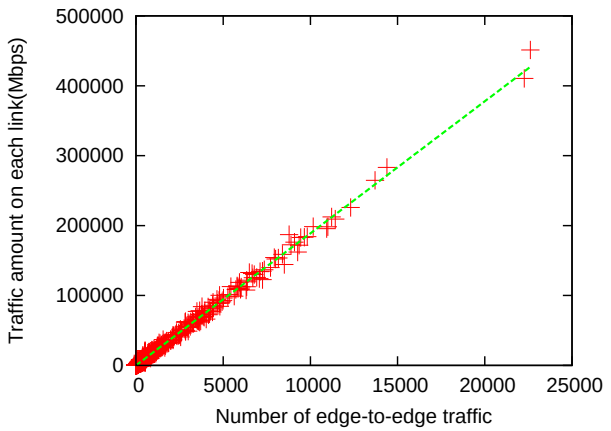


Figure 1. Relations of the number of edge-to-edge traffic and the traffic amount on each link

where W_i is the traffic amount of the link i, Z_i is the number of edge-to-edge traffic passing the link i, and α and β are the constant parameters. Z_i for any node i can be calculated from the routing matrix.

By using this relation, we estimate the traffic amount on each link as following steps. First, we calculate the constant parameters, α and β by using traffic amount on each link collected from the selected monitoring nodes. To calculate α and β, we use the least-square method. That is,

$$\alpha = \frac{|S| \sum_{i \in S} Z_i W_i - \sum_{i \in S} Z_i \sum_{i \in S} W_i}{|S| \sum_{i \in S} Z_i^2 - \left(\sum_{i \in S} Z_i\right)^2}, \tag{10}$$

$$\beta = \frac{\sum_{i \in S} Z_i \sum_{i \in S} W_i - \sum_{i \in S} Z_i W_i \sum_{i \in S} Z_i}{|S| \sum_{i \in S} Z_i^2 - \left(\sum_{i \in S} Z_i\right)^2}, \tag{11}$$

where S is the set of links connected to the monitoring nodes. Then, we estimate the traffic amount U_j on the link j that is not collected from the monitoring nodes as

$$U_j = \alpha Z_j + \beta. \tag{12}$$

Finally, we define the matrix X' which is a matrix indicating the roughly estimated traffic amount on each link as

$$X' = \begin{bmatrix} X'_1 \\ \vdots \\ X'_L \end{bmatrix}, \tag{13}$$

where

$$X'_l = \begin{cases} X_l & \text{if } l \text{ is the link connected to the monitored nodes,} \\ U_l & \text{otherwise.} \end{cases} \tag{14}$$

C. Estimating traffic matrices

We estimate the traffic matrix from the roughly estimated traffic amount on each link. If we apply the tomogravity method to estimate traffic matrix from the estimated traffic amount on each link, the estimation errors may become large, because the estimation errors included in the traffic amounts on ingress and egress links cause the inaccurate estimation of T^{grav} and large estimation errors of the tomogravity method even when traffic amounts on other links are estimated accurately. Therefore, we need a traffic matrix estimation method where estimation errors included in the traffic amounts on particular links do not affect the estimation results significantly.

Though there may be more sophisticated estimation method, in our evaluation described in Section IV, we use the simple approach to estimate the traffic matrix by minimizing the following equation;

$$\min \|X' - A\hat{T}\|. \tag{15}$$

The results shown in Section IV clarifies that we can estimate the traffic amount on each link accurately even when we use this simple approach to estimate the traffic matrix.

D. Estimating traffic amount from estimated traffic matrices

Once we obtain the estimated traffic matrix $\hat{T}$, we calculate the matrix $\hat{X}$ that represents the traffic amount on each link as

$$\hat{X} = A\hat{T}. \tag{16}$$

Then, we designate $\hat{X}$ as the final estimation results for the traffic amount on each link. By estimating $\hat{X}$ as Eq. 16, even when significant traffic changes on a small number of edge-to-edge traffic occurs and the changes are not captured by X', $\hat{X}$ may follow the traffic changes since $\hat{X}$ is estimated so as to fit the current traffic amount information collected from the monitored nodes.

IV. Numerical Evaluations

In this section, we evaluate our method by simulation. In this evaluation, we use the same topology and traffic matrix as Section III-B

In this evaluation, we investigate the accuracy of the estimation of the traffic amount on each link, because the traffic amount on each link is important information for traffic engineering and the estimation errors of the traffic amount on each link may cause the misidentification of the congested links.

To evaluate the accuracy of the estimation of the traffic amount on each link, we use the Root Mean Squared Error (RMSE) and the Root Mean Squared Relative Error (RMSRE). The RMSRE (X_{RMSRE}) and the RMSE (X_{RMSE}) are defined as

$$X_{RMSRE} = \sqrt{\frac{1}{L} \sum_{k=1}^{L} \left(\frac{\hat{X}_k - X_k}{X_k}\right)^2}, \tag{17}$$

$$X_{RMSE} = \sqrt{\frac{1}{L} \sum_{k=1}^{L} \left(\hat{X}_k - X_k\right)^2}, \tag{18}$$

 ISBN: 978-1-61208-103-8

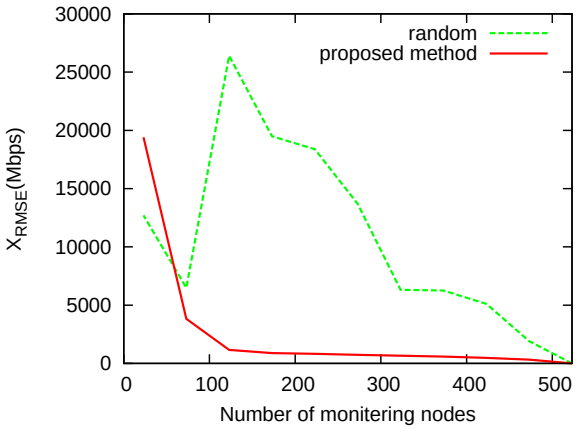


Figure 2. RMSE of traffic amount on each link

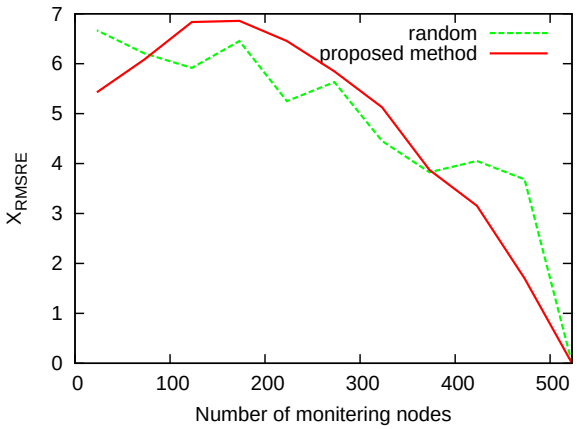


Figure 3. RMSRE of traffic amount on each link

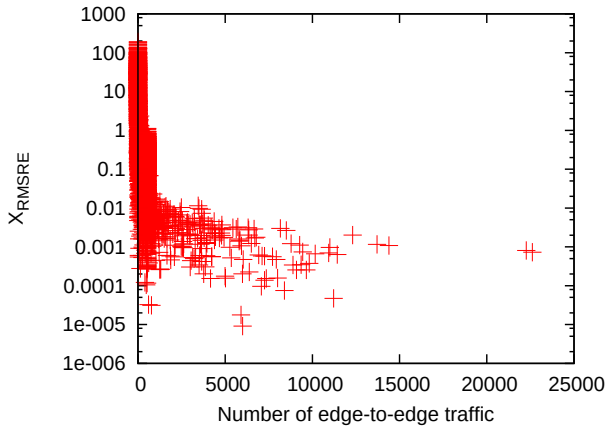


Figure 4. Relations of the number of edge-to-edge traffic and RMSRE when our method selectes 173 nodes

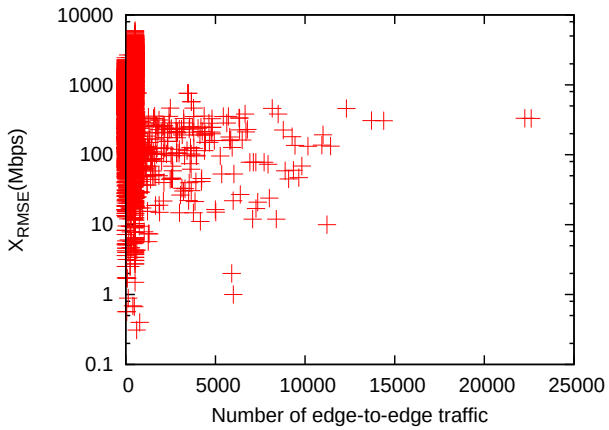


Figure 5. Relations of the number of edge-to-edge traffic and RMSE when our method selectes 173 nodes

where L is the number of links in the network, $\hat{X}_k$ is the estimated traffic amount of link k, and X_k is the actual traffic amount of the link k.

Figures 2 and 3 show X_{RMSE} and X_{RMSRE} respectively when we change the number of monitoring nodes. In these figures, the vertical axis is X_{RMSE} or X_{RMSRE}, and the horizontal axis is the number of monitoring nodes. In these figures, "proposed method" indicates the case that we select the monitoring nodes by our method and "random" indicates the case that we select the monitoring nodes randomly.

According to Fig. 2, we can estimate the traffic amount of each link accurately by selecting more than 173 monitoring nodes, while the RMSE of the traffic amount on each link become significantly large if the number of monitoring nodes is less than 173 since the number of traffic amount information is too small to estimate the parameters of Eq. 12.

Fig. 2 also shows that we can estimate traffic amounts much more accurately in the case of selecting monitoring nodes by our method, compared with the case of selecting monitoring nodes randomly. This is because our method selects the monitoring nodes so as to cover as many edge-to-edge traffic as possible. Therefore, most of edge-to-edge traffic pass at least one of monitoring nodes selected in our method and can be estimated from the information of the traffic amounts collected from the monitoring nodes. On the other hand, in the case of selecting monitoring nodes randomly, several edge-to-edge traffics pass no monitoring nodes. Since we cannot obtain the traffic amount information corresponding to such edge-to-edge traffics from the monitoring nodes, such edge-to-edge traffics cannot be estimated accurately. As a result, the estimation errors of traffic amounts on links passed by such edge-to-edge traffic whose estimation error is large become also large.

However, according to Fig. 3, unlike the RMSE, the RMSRE of our method is still large and close to the RMSRE of the case of selecting monitoring nodes randomly. To investigate this in more detail, we show the estimation error of traffic amount on each link and discuss whose estimation

 ISBN: 978-1-61208-103-8

error is large.

Figs. 4 and 5 show the relations between the number of edge-to-edge traffic passing a link and RMSRE or RMSE of the traffic amount on the link when we select 173 monitoring nodes by our method. According to Fig. 4, the RMSREs only for the traffic amounts on the links where the number of edge-to-edge traffic is small become large. The actual traffic amount on the link where the number of edge-to-edge traffic is small may be small. The small actual traffic amount makes the value of relative error quite large even when the estimation error is not large.

In addition, according to Fig. 5, the estimation errors for the traffic amount on the links where the number of edge-to-edge traffic is small also become large. This is because one-hop traffics whose source and destination nodes are both ends of a link cannot be estimated from the traffic amounts monitored at any other links and their estimation errors become large. The ratio of one-hop traffic among the total traffic on the link increases as the number of edge-to-edge traffic passing the link becomes small. Thus, the estimation errors of one-hop traffic cause the large estimation errors of traffic amounts on the links where the number of edge-to-edge traffic is small.

However, routes of one-hop traffic are rarely changed by traffic engineering. In addition, the links where the number of edge-to-edge traffic is small are on the edge of the network. Thus, the estimation errors on the traffic amount on such links may have only little impact on the traffic engineering. According to Figs. 4 and 5, most of the traffic amount on each link passed by many edge-to-edge traffic can be estimated accurately. That is, our method can estimate the traffic amount on each link required for traffic engineering accurately by using only the information from a subset of nodes.

V. Conclusion

In this paper, we proposed a method to select the monitoring nodes and estimate the traffic amounts on all links from the traffic information collected from the selected monitoring nodes. Through the simulation, we clarified that our method can estimate the traffic amount on each link required for traffic engineering accurately by monitoring 30% of all nodes.

One of our future research topics is to evaluate the performance of traffic engineering using the traffic amount on each link estimated by our method.

Acknowledgments

This work was partly supported by Grant-in-Aid for Scientific Research (B) 22300023 and Grant-in-Aid for Young Scientists (B) 21700074 of the Ministry of Education, Culture, Sports, Science and Technology in Japan.

References

[1] B. Fortz and M. Thorup, “Internet traffic engineering by optimizing OSPF weights,” in *Proceedings of IEEE INFOCOM*, vol. 2, pp. 519–528, Mar. 2000.

[2] M. Roughan, M. Thorup, and Y. Zhang, “Traffic engineering with estimated traffic matrices,” in *Proceedings of the 3rd ACM SIGCOMM Conference on Internet Measurement*, pp. 248–258, Nov. 2003.

[3] Y. Ohsita, T. Miyamura, S. Arakawa, S. Ata, E. Oki, K. Shiomoto, and M. Murata, “Gradually reconfiguring virtual network topologies based on estimated traffic matrices,” *IEEE/ACM Transactions on Networking*, vol. 18, pp. 177–189, Feb. 2010.

[4] Y. Ohsita, T. Miyamura, S. Awakawa, E. Oki, K. Shiomoto, and M. Murata, “Estimation of current traffic matrices from long-term traffic variations,” *IEICE Transactions on Communications*, vol. E92-B, pp. 171–183, Jan. 2009.

[5] I. Juva, “Robust load balancing,” in *Proceedings of GLOBECOM*, pp. 2708–2713, Nov. 2007.

[6] T. Miyamura, Y. Ohsita, E. Oki, S. Arakawa, Y. Koizumi, A. Masuda, K. Shiomoto, and M. Murata, “Network virtualization server for adaptive network control,” in *Proceedings of 20th ITC Specialist Seminar on Network Virtualization - Concept and Performance Aspects*, May 2009.

[7] Y. Vardi, “Network tomography: Estimating source-destination traffic intensities from link data.,” *Journal of the American Statistical Association*, vol. 91, pp. 365–377, Mar. 1996.

[8] J. Cao, D. Davis, S. Wiel, and B. Yu, “Time-varying network tomography: Router link data,” *Journal of the American Statistical Association*, vol. 95, Feb. 2000.

[9] I. Juva, S. Vaton, and J. Virtamo, “Quick traffic matrix estimation based on link count covariances,” in *Proceedings of IEEE ICC*, vol. 2, pp. 603–608, June 2006.

[10] A. Soule, A. Nucci, R. Cruz, E. Leonardi, and N. Taft, “Estimating dynamic traffic matrices by using viable routing changes,” *IEEE/ACM Transactions on Networking*, vol. 15, pp. 485–498, June 2007.

[11] Y. Zhang, M. Roughan, N. Duffield, and A. Greenberg, “Fast accurate computation of large-scale IP traffic matrices from link loads,” *ACM SIGMETRICS Performance Evaluation Review*, vol. 31, pp. 206–217, June 2003.

[12] Y. Zhang, M. Roughan, W. Willinger, and L. Qiu, “Spatio-temporal compressive sensing and Internet traffic matrices,” *ACM SIGCOMM Computer Communication Review*, vol. 39, pp. 267–278, Aug. 2009.

[13] N. Spring, R. Mahajan, D. Wetherall, and T. Anderson, “Measuring ISP topologies with Rocketfuel,” *IEEE/ACM Transactions on networking*, vol. 12, pp. 2–16, Feb. 2004.

[14] A. Nucci, A. Sridharan, and N. Taft, “The problem of synthetically generating IP traffic matrices: Initial recommendations,” *ACM SIGCOMM Computer Communication Review*, vol. 35, pp. 19–32, July 2005.

 ISBN: 978-1-61208-103-8

Emulation Environment for Ground Truth Establishment

Carlos Miranda, Paulo Salvador, António Nogueira, Eduardo Rocha
University of Aveiro / Instituto de Telecomunicações, Campus de Santiago, 3810-193 Aveiro, Portugal
E-mail:{cmiranda, salvador, nogueira, erocha}@av.it.pt
Rui Valadas
Instituto Superior Técnico - UTL / Instituto de Telecomunicações, Av. Rovisco Pais, 1049-001 Lisboa, Portugal
E-mail:rui.valadas@ist.utl.pt

***Abstract*—Network security is a hot topic to network users and managers, whether they are institutional, enterprise or domestic. New threats or mutations of existing ones appear at a very fast rate and the solutions that are nowadays used to fight them frequently require a real-time analysis of the network traffic or a previous training based on real data. Most of the times, this training must be supervised by humans that, depending on their experience, can create security breaches in the system without knowing it. Since existing anomaly detection methodologies have to be trained and tested in order to validate their efficiency, there is an increasing need for trustworthy network traffic data that can be used without compromising users confidentiality, obeys to some pre-established criteria and is completely known in terms of its underlying protocols. In fact, the effectiveness of network anomaly detectors cannot be fully evaluated without having a complete control of the entire evaluation experiment, which requires that it should be possible to change the location, magnitude and type of individual anomalies and background traffic. In this work, we propose an emulation environment that can be used to obtain trustworthy network data both in the presence of licit and illicit applications. We also present some topological and traffic scenarios that were already defined to start gathering network data and make it immediately available to the scientific community. The emulation environment was built in an evolutionary way, enabling the easy introduction of new network scenarios and services and/or the refinement of the existing ones.**

***Keywords*-Ground truth, emulation, licit and illicit applications.**

I. Introduction

Most of the research activities on network traffic modeling, application identification and anomaly detection methodologies require the association of application and protocol ground truth information with traffic traces [1]. There are several approaches that can be used to link ground truth data to Internet traffic traces. The first approach is to create a trace manually by instantiating a realistic pool of applications on many machines, but in this case the captured traffic typically lacks characteristics that only human intervention can induce. The second methodology is to record traffic on a live network and perform deep packet inspection to each packet in order to identify its underlying generating protocol/application [2], [3], [4], [5]. However, deep packet inspection is ineffective when traffic is encrypted and ambiguous when different protocols exhibit similar signatures. Finally, the last approach works by probing a set of monitored host kernels in order to obtain information on active Internet sessions, thus gathering ground truth at the application level [6]. However, this methodology is intrusive and requires the agreement of the different monitored users.

In order to circumvent some of these limitations, we propose in this paper an emulation approach that can be used to generate trusty network data. Basically, we have created a flexible emulation environment that can be used to obtain trustworthy data for traffic scenarios that include both licit and illicit applications. The details of the emulation platform will be presented in this paper, as well as the main development options that were adopted. Besides, the paper will also present the topological and traffic scenarios that were already defined to start the process of gathering trusty network data in order to make it immediately available to the scientific community.

By defining the topologies of the scenarios, the protocols to be used and the anomalies that will be ”injected”, a vast amount of trustworthy data can be obtained. We can see the ”before and after” the injection of a specific anomaly in order to see the behavior over the network, while having a high accuracy in the obtained data without compromising privacy. Simulators and emulators are a very effective and cheap way to create environments that otherwise could be very expensive to deploy with real equipment. They are also very useful to test new networking protocols or to change existing ones in a controlled and reproducible manner, saving a lot of time and money. Currently, there is a huge variety of simulators and emulators, ranging from payware to freeware solutions, destined to educational or to research purposes. Network emulators work in a similar way as simulators. However, emulators are able to run in real time while network simulators are not. In network emulators, the network behaves like a real network and it is possible to monitor, measure and test its performance. Real equipment can be connected to the emulation (computers, routers and other network entities), behaving exactly as if they were in a real network. Thus, no approximations are necessary to describe real processes, like routing and queuing mechanisms for example.

The emulation methodology that is proposed in this paper is completely generic, enabling the integration of automatic traffic generation tools with the normal human usage of different network services. In this way, we are able to capture

 ISBN: 978-1-61208-103-8

the different components and nuances of the network traffic characteristics.

The rest of the paper is organized as follows. Section II will briefly present the most important approaches to ground truth establishment that have been published so far; Section III will describe the emulation framework that was created and the network and service scenarios that were already defined, including their main configuration steps; Section IV describes the data that can be collected from the emulated scenarios and, finally, Section V presents the main conclusions of this work and points out some topics for future research.

II. Related work

In the context of network anomaly detection, ground truth requires a complete list of all anomalies existing in the data traces. Some previous works [4], [7] have injected anomalies in real traces taken from operational networks but this technique cannot provide truthful data because it relies on existing traces and there are no guarantees that they are free of anomalies and, if they are not, on the number and type of anomalies they include.

Several other techniques have been used to obtain trusty data. Payload inspection and port-based mechanisms [8], [9], [10], [11], [12] have been recurrently used to establish a form of protocol ground truth but, due to their uncertainty, they can only provide an estimate of the protocol that is being used. Besides, port-based analysis became quite an obsolete technique for protocol identification.

Manual generation of network traffic can provide ground truth at the application level [13] but the presence of background applications can lead to the generation of additional traffic that can not be accurately tagged. However, this approach is able to deal with encrypted traffic, which is one of its main advantages.

The approach presented by Szabo et al. [14] offers an advanced approach in application detection and tagging by embedding ground truth information directly into IP packets. However, this approach creates some methodological problems (like lowering the precision of the recorded traces or being unable to mark packets with sizes closer to the MTU), their implementation is Windows XP-specific and the created tool provides application but no protocol information for a given flow.

Trestian et al. [15] describe an heuristic that combines information freely available on the web (through Google) to retrieve the class of applications that generated a specific flow. This approach is interesting but it relies on external resources.

In [16], the authors present a platform, named Ground Truth Verification System (GTVS), that uses a combination of heuristics at different levels (host, flow, packet) to improve the quality of ground truth associated with packet traces, but does not include application labels with guaranteed accuracy.

Finally, in reference [6] the authors present a new mechanism to provide ground truth at the application level. The architecture of the proposed tool is based on a client tool that, by monitoring the kernel of a host, associates each packet flow with the name of its controlling application and transmits the collected information to a back-end. The post-processing toolset analyzes the traffic captured at the network border by an independent probe and associates each flow with its application label, being able to establish ground truth for that flow. This tool works on many operating systems and it is freely available under an Open Source (BSD) license [17]. However, this approach relies on monitoring the kernels of the different network hosts, which can be an intrusive task and requires the agreement of the corresponding network users.

III. Emulation environment

The main objective of this work is to create a real-time emulator environment that can be used to obtain trustworthy traffic traces and network data, without having to concern on all issues related to the usage of public traces. In this type of environment, the entire emulation process can be controlled, thus achieving a higher confidence degree on the collected data. The Graphical Network Simulator, version 3 (GNS3), was the chosen network emulator.

GNS3 [18] is a multi-platform, open-source Graphical Network Emulator primarily developed by Jeremy Grossman. GNS3 allows the emulation of complex network topologies by emulating many Cisco IOS router platforms, IPS, PIX and ASA firewalls, and JunOS with the help of Dynamips and Dynagen. Dynamips is the core program behind the emulation process and the Dynagen tool runs on top of it to create a user-friendly, text-based environment. GNS3 provides the graphical front-end for Dynagen, so that users can create the topologies in a graphical and user-friendly environment. GNS3 also allows the emulation of ATM and Frame Relay switches, enables packet capture using Wireshark and, most important, allows the connection of the emulated network to the real world. So, using this flexible tool, it is possible to create network scenarios with the desired degree of complexity.

A. Topological scenarios

Two network scenarios were already defined and emulated. The first scenario, illustrated in Figure 1, represents the interaction between two main routers of two departments of a small enterprise network and a backbone router that is connected to a server. The departments can also interact between them and if one of the links between the three routers goes down, there is a backup route passing through the link of the other department, so that access to the server is guaranteed. The clouds represent the LANs located inside each department. For this scenario, the LANs have a total of five emulated end-hosts performing random requests to the server and an also emulated end-host that can be connected in any part of the network. This movable end-host is used only for triggering security attacks to the network. By using this end-host, which is fully dedicated to network attacks, it is possible to easily identify the packets generated by the attacks and analyze them separately from the rest of the network traffic that is generated by the emulation scenario.

In this scenario, the end-hosts can send DHCP requests to obtain their IP address and can also send random HTTP and FTP requests to the server. The emulated routers belong to the

 ISBN: 978-1-61208-103-8

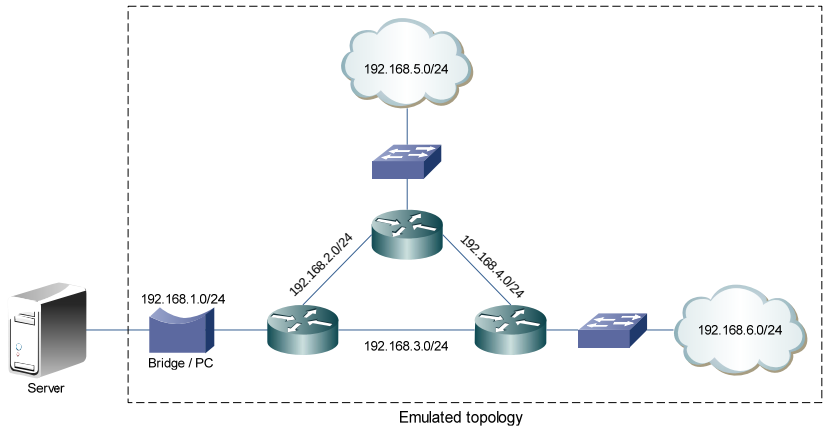


Fig. 1. Network Topology for Scenario A.

TABLE I
SERVICES CONFIGURED ON SCENARIO A

Routers	Server
DHCP Server	Apache
OSPF	BIND9
SNMP Agent	ProFTPD
	NMS

3745 series. The services that were configured at the routers and the server are listed in Table I.

This simple scenario was mainly used to test the Personal Computers (PCs) that host the emulated devices in order to look for problems in the emulation and to check if these hosts were able to handle all the processing that is needed for the emulation to run. By capturing only the first 100 bytes of each packet, this scenario generates, approximately, 5GB of network traffic per day.

The topology of scenario B, represented in Figure 2, represents the network infrastructure of a medium/large enterprise network. The network backbone part of this topology is composed by four real Catalyst 3750 Metro series layer 3 switches (IOS version 12.1(14r)) and three emulated 3745 Cisco routers. One of the layer 3 switches is directly connected to the server in order to allow communication between each point of the network and the server. In this topology, each part of the network is composed by three departments that are connected to the server via different points of the backbone network. Each router, representing the connection to a department network, holds a LAN composed by 6 end-hosts where 5 of them generate normal network traffic (DHCP, HTTP, FTP, SMTP and POP3 requests) and the sixth one is used to trigger network security attacks. The services that were configured at the routers and the server are the same of the previous scenario, except that now the server additionally includes the POSTFIX service. This scenario can generate 40GB of network traffic per day, capturing only the first 100 bytes of each packet.

B. *Traffic scenarios*

The above network topologies were already tested using different network traffic scenarios, each one having a 24 hours duration. These scenarios can be divided into three main groups:

- Clean scenario: this scenario generates network traffic without anomalies (that is, using only licit applications);
- Port-Scan Attack scenario: the network traffic generated in this scenario is a mixture of normal network traffic with port-scan security attacks made from a single or multiple end-hosts to the server;
- Snapshot Attack scenario: the network traffic generated in this scenario is a mixture of normal network traffic with snapshot security attacks made from a single end-host to the server.

The Port Scan attack is made using Nmap (Network Mapper) [19], an open source tool used for network exploration and security auditing. Nmap uses raw IP packets to obtain several characteristics of the scanned host, like the services that are offered by the host, the OS that is running on it or the type of firewall that is used. Nmap also retrieves a list with the used port numbers, the service that is running on those ports and their states. Port scans are normally used by hackers to discover open ports on the target end-host in order to gain access to it or to trigger Distributed Denial of Service (DDoS) attacks.

The Port-Scan attack traffic scenario is further sub-divided into three types of port scans, depending on the timing profile that is used:

- Sneaky: used for Intrusion Detection System (IDS) evasion. A complete port scan with this timing template can take up to five hours to complete because scans are not made in parallel and the waiting time between each probe that is sent to the server is 15 seconds.
- Normal: this is the default timing template used by Nmap. It allows the port scan parallelization process.
- Aggressive: in this case, the scan delay of the probes does not exceed 10 milliseconds (for TCP ports).

For each type of Port Scan attack that is used in this work there is also a division of the different types of attacks based

 ISBN: 978-1-61208-103-8

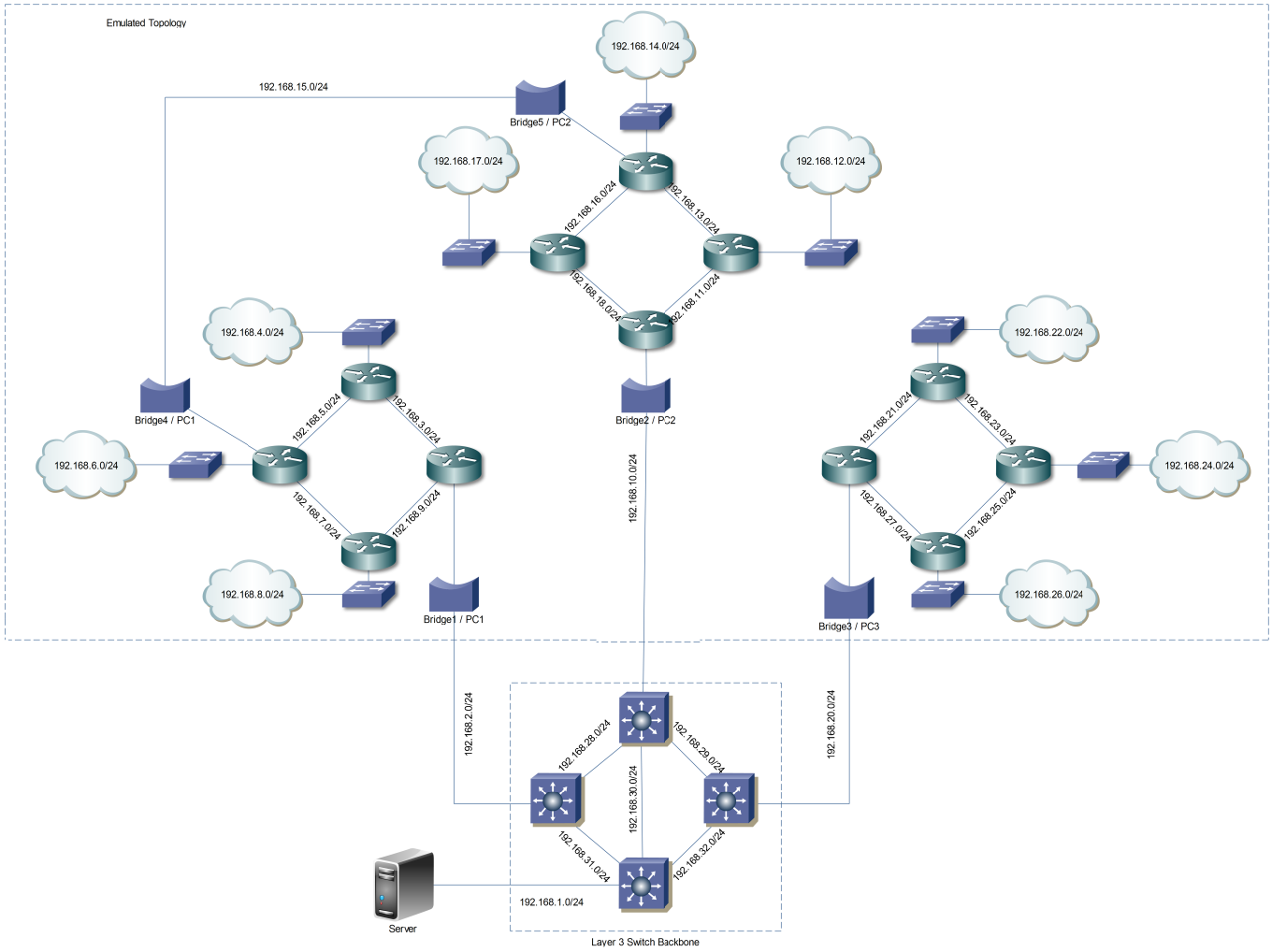


Fig. 2. Network Topology for Scenario B.

TABLE II
DIFFERENT TYPES OF PORT SCAN ATTACKS BASED ON THE NUMBER OF ATTACKERS AND TARGETS.

Port Scan attack	Description
1-to-1	Single attacker against a single target.
1-to-N	Single attacker against multiple targets.
N-to-1	Multiple attackers against a single target.
N-to-N	Multiple attackers against multiple targets.

on the number of attackers and targets. This division and a brief description of each sub-type are given in Table II.

The Snapshot Attack scenario, as the name suggests, consists of performing snapshots of the target monitor content. These images are then uploaded to a destination of the attacker preference and can be used to discover confidential information about the target. Some ”worms” that can be found on the Internet act like this: while consuming the target resources, they can ”catch up” snapshots of the target. In this scenario, the snapshot attack is divided into two categories:

- Exponential: in this attack, each time a users ”clicks” in some part of the monitor, a snapshot of a small area around the pointer is taken. The name of this category derives from the fact that intervals between user clicks are exponentially distributed, in order to model human behavior in the best possible way.
- Periodical: in this kind of attack, a snapshot of the entire screen is taken in periodical time intervals.

Because these attacks are uploads of relatively small images, in this work this attack is emulated by using a virtual end-host inside the emulation that uploads image files to the server using an exponentially distributed time interval, in the first case, and a periodic time interval in the second case. The tool that was used to emulate this behavior (and also for the FTP download and upload traffic that will be explained in the following sub-section) is called cURL, a command line tool for transferring files with a URL syntax. It currently supports FTP, HTTP, FTPS, HTTPS, TELNET and many others protocols that use the URL syntax.

 ISBN: 978-1-61208-103-8

C. Server configuration

Three services are implemented at the server: HTTP, FTP and Email (POP3 to retrieve mail messages and SMTP to send them). In order to have these services, the Network Interface Card (NIC) of the server is configured with one real IP address and several Virtual IP Addresses (VIPA), in such a way that the DNS server will be associated to the real IP address and the first VIPA will be associated with the domain configured for the APACHE, ProFTPD and POSTFIX servers. The other VIPAs are used for testing part of the attacks against multiple IP addresses that are inside the server IP range in the emulated scenario. All server IP addresses are statically configured.

D. Network device configuration

The routers' interfaces have to be configured as DHCP servers for their LANs, so that the end-hosts can obtain their IP addresses via DHCP. All routers involved in the emulation (real and emulated) were configured with the OSPF protocol, so that they can construct their routing tables and the topology map of the network. Finally, in order to gather the MIB information about all the interfaces of the different network devices that are present in the emulation, SNMP Agents were enabled at each device. For example, the number of incoming, outgoing and dropped packets per router interface is some of the information that can be gathered using SNMP agents.

E. PC configuration

In order to emulate a multiple end-host environment, several TAP interfaces were configured in each one of the host PCs. A TAP interface is a virtual Ethernet interface created in the system kernel that can behave as a real Ethernet interface and can be connected to the emulated topology in order to generate all the traffic requests for the emulated scenarios.

In the first tests that were made, we noticed that the only responses received from the server were those corresponding to the requests that were made by the first end-host. This is due to the fact that the Linux routing table establishes priorities for the configured routes. In order to solve this problem, the routes for each end-host need a set of rules for incoming and outgoing traffic, thus forming a routing table for each one of them. When packets arrive, the system will look for matching the set of rules in order to properly deliver the packets to the end-hosts. Besides, the *arp_filter* file corresponding to each interface (the interfaces of the end-hosts and the bridge), found in the "/proc/sys/net/ipv4/conf/interface_name/" directory, needs to be enabled so that all the interfaces' ARP requests can be correctly answered.

F. Operational procedure

After creating all the necessary virtual end-hosts and bridges, configuring the default route at the server and connecting all emulated and real network devices, the following procedure must be followed to start the emulation:

- Start the emulation inside GNS3 and configure all the network devices (emulated and real).

```
192.168.4.2 - - [21/Sep/2009:14:33:45 +0100] "GET /CNN.htm HTTP/1.0" 200 103483 "-" "Wget/1.11.4"
192.168.4.2 - - [21/Sep/2009:14:33:48 +0100] "GET /CNN.htm HTTP/1.0" 200 103483 "-" "Wget/1.11.4"
192.168.4.2 - - [21/Sep/2009:14:33:48 +0100] "GET /CNN.htm HTTP/1.0" 200 103483 "-" "Wget/1.11.4"
192.168.4.2 - - [21/Sep/2009:14:33:56 +0100] "GET /CNN.htm HTTP/1.0" 200 103483 "-" "Wget/1.11.4"
192.168.4.5 - - [21/Sep/2009:14:34:04 +0100] "GET /cyberhome.htm HTTP/1.0" 200 25247 "-" "Wget/1.11.4"
```

Fig. 3. APACHE log file.

- Run the DHCP script file so that all end-hosts can obtain their IP addresses.
- Configure the different routing tables and rules for each virtual end-host.
- Start the *tshark* application at the server in order to capture all the packets and execute the SNMP script file for gathering the MIB data.
- At the host PC, start the *tshark* application to capture all packets in all virtual interfaces.
- Execute the HTTP, FTP, SMTP and POP3 traffic generators.
- For the scenarios that include security attacks, execute the script file corresponding to the attack that is going to be launched.

It was assumed that a typical user makes a new HTTP or FTP request in time intervals of 120 seconds, in average. This value will be used by the emulation tool as the mean of the exponential distribution that will rule the timing profiles of the HTTP and FTP requests. It was also assumed that a typical user sends an E-mail every 10 minutes [20], [21], in average. This value is used to create the exponential distribution that rules the generation process of the SMTP requests. The average time interval used in this work to check for new E-mail messages is 10 minutes. This is the default time value that is used by almost all E-mails clients, like Thunderbird or Microsoft Outlook.

IV. COLLECTED DATA

A lot of data can be gathered from the virtual interfaces (in order to characterize the behavior of single users) or from the server (enabling the characterization of traffic aggregates). As previously explained, virtual interfaces are used in this work with two purposes: the interconnection between real equipment and the emulated topologies and the emulation of end-hosts connected to the network. In real networks, a user can randomly ask for any type of request to the server: for example, a user can see a web page while waiting for the e-mail client to verify his e-mail accounts. The end-hosts used for this work can emulate this behavior with the help of the traffic generator program that was previously explained. So, in the first scenario the end-hosts can obtain their addresses via DHCP, visualize websites and can upload and download files to/from a FTP server; in the second scenario, the end-hosts can also send and receive e mails.

Log files can play a vital role in a server and provide crucial information regarding the services that are running on it. The log files that can be gathered in this framework correspond to the logs of the HTTP (Apache), FTP (ProFTPD) and e-mail (Postfix) servers. Figure 3 shows an example of the APACHE server log file.

 ISBN: 978-1-61208-103-8

```
IF-MIB::ifInOctets.1 = Counter32: 50118
IF-MIB::ifInOctets.2 = Counter32: 54919
IF-MIB::ifInOctets.3 = Counter32: 117948
IF-MIB::ifInOctets.4 = Counter32: 0
IF-MIB::ifInOctets.5 = Counter32: 0
IF-MIB::ifInUcastPkts.1 = Counter32: 407
IF-MIB::ifInUcastPkts.2 = Counter32: 451
IF-MIB::ifInUcastPkts.3 = Counter32: 1328
IF-MIB::ifInUcastPkts.4 = Counter32: 0
IF-MIB::ifInUcastPkts.5 = Counter32: 0
IF-MIB::ifInNUcastPkts.1 = Counter32: 104
IF-MIB::ifInNUcastPkts.2 = Counter32: 111
IF-MIB::ifInNUcastPkts.3 = Counter32: 8
IF-MIB::ifInNUcastPkts.4 = Counter32: 0
IF-MIB::ifInNUcastPkts.5 = Counter32: 0
```

Fig. 4. Example of data collected using SNMP.

To collect SNMP data and store it at the server, the "Net-SNMP" utility (available in the Linux repositories) was installed. This utility allows the communication between SNMP clients and agents. The requests made to the agents are triggered from a terminal, so a bash file was constructed in order to perform this task automatically while emulation was running. Figure 4 shows an example of the data that can be collected using SNMP.

The different traffic profiles that were already collected from the emulation environment and processed *a posterior* are very similar to service profiles obtained from real traffic measurements. Due to lack of space, the comparison plots are not presented in this paper but all collected data will become available on a website that will be used to share all the emulation environment developments and results with the scientific community (the URL of the website is *groundtruth.av.it.pt*).

Regarding the computational requirements of the emulation environment, we can say that for the medium complexity scenarios that were considered in this paper the (quadcore) CPU usage of each one of the PCs hosting part of the topology went up to 40%; in terms of RAM, GNS3 has some memory optimitazion features (the "ghost IOS" and the "sparse memory" featureS) that allowed memory requirements to be kept to a minimum.

V. Conclusions and Future Work

Since most of the research activities on network traffic modeling, application identification and anomaly detection methodologies require the association of application and protocol ground truth information with traffic traces, this paper proposed an emulation environment that can be used to obtain trustworthy data, both in the presence of licit and illicit applications. Besides presenting the details of the created emulation environment, some topological and traffic scenarios that were already used to start gathering network data and make it immediately available to the scientific community were also presented. In order to complement this work, there are some interesting issues that will be addressed in a near future: define and include more complex topologies, emulate more services per interface (P2P file sharing, for example), implement other types of security attacks (DDoS, worm propagation and botnets), study the effects of Spam e-mail, correlate all the gathered data and upgrade the traffic generation tools in order to generate traffic according to different probability distributions and daily patterns.

VI. Acknowledgments

This work was part of the Euro-NF project, funded by the European Union.

References

[1] H. Ringberg, M. Roughan, and J. Rexford, "The need for simulation in evaluating anomaly detectors," *ACM SIGCOMM Computer Communication Review*, vol. 38, no. 1, pp. 55–59, 2008.

[2] Marco Canini, Wei Li, Martin Zadnik, and Andrew W. Moore, "Experience with high-speed automated application-identification for network-management," in *Proceedings of the 5th ACM/IEEE Symposium on Architectures for Networking and Communications Systems (ANCS'09)*, Oct 2009.

[3] Martin Zadnik, Marco Canini, Andrew W. Moore, David J. Miller, and Wei Li, "Tracking elephant flows in internet backbone traffic with an FPGA-based cache," in *Proceedings of the 19th International Conference on Field Programmable Logic and Applications (FPL'09)*, Aug 2009.

[4] A. Lakhina, M. Crovella, and M. Diot, "Diagnosing network-wide traffic anomalies," in *ACM SIGCOMM*, 2004, pp. 219–230.

[5] H. Ringberg, A. Soule, J. Rexford, and C. Diot, "Sensitivity of PCA for traffic anomaly detection," in *ACM SIGMETRICS*, 2007.

[6] F. Gringoli, L. Salgarelli, M. Dusi, N. Cascarano, F. Risso, and K. C. Claffy, "GT: picking up the truth from the ground for internet traffic," *ACM SIGCOMM Computer Communication Review*, vol. 39, no. 5, pp. 13–18, 2009.

[7] A. Soule, K. Salamatian, and N. Taft, "Combining filtering and statistical methods for anomaly detection," in *ACM Internet Measurement Conference*, 2005.

[8] Jeffrey Erman, Martin Arlitt, and Anirban Mahanti, "Traffic classification using clustering algorithms," in *MineNet '06: Proceedings of the 2006 SIGCOMM workshop on Mining network data*, New York, NY, USA, 2006, pp. 281–286, ACM.

[9] H. Kim, K. Claffy, M. Fomenkova, D. Barman, and M. Faloutsos, "Internet traffic classification demystified: The myths, caveats and best practices," in *ACM CoNEXT*, 2008.

[10] T. Karagiannis, A. Broido, N. Brownlee, K.C. Claffy, and M. Faloutsos, "Is P2P dying or just hiding?," *IEEE Global Telecommunications Conference (GLOBECOM'04)*, vol. 3, pp. 1532–1538 Vol.3, Nov.-3 Dec. 2004.

[11] P. Haffner, S. Sen, O. Spatscheck, and D. Wang, "Acas: Automated construction of application signatures," in *MineNet '05: Proceedings of the 2005 ACM SIGCOMM workshop on Mining network data*, Philadelphia, Pennsylvania, USA, August 2005, pp. 197–202, ACM Press.

[12] Subhabrata Sen, Oliver Spatscheck, and Dongmei Wang, "Accurate, scalable in-network identification of P2P traffic using application signatures," in *Proceedings of the 13th international conference on World Wide Web (WWW'04)*, New York, NY, USA, 2004, pp. 512–521, ACM.

[13] M. Dusi, M. Crotti, F. Gringoli, and L. Salgarelli, "Tunnel hunter: Detecting application-layer tunnels with statistical fingerprinting," *Elsevier Computer Networks*, vol. 53, no. 1, pp. 81–97, 2009.

[14] G. Szabo, D. Orincsay, S. Malomsoky, and I. Szabo., "On the validation of traffic classification algorithms," in *Proceedings of the Passive and Active Measurement Conference (PAM'08)*, 2008.

[15] I. Trestian, S. Ranjan, A. Kuzmanovi, and A. Nucci, "Unconstrained endpoint profiling (googling the internet)," *ACM SIGCOMM Computer Communication Review*, vol. 38, no. 4, pp. 279–290, 2008.

[16] Marco Canini, Wei Li, Andrew W. Moore, and Raffaele Bolla, "GTVS: Boosting the collection of application traffic ground truth," in *Proceedings of the First International Workshop on Traffic Monitoring and Analysis (TMA'09)*, May 2009.

[17] "The ground truth software tools, available at http://www.ing.unibs.it/ntw/tools/gt," 2009.

[18] "Graphical network emulator - GNS3, available at http://www.gns3.net/," 2009.

[19] "NMAP - network mapper, available at http://nmap.org," 2009.

[20] K. Mandia, C. Prosise, and M. Pepe, *Incident Response & Computer Forensics*, Mcgraw Hill, 2003.

[21] Luiz Henrique Gomes, Cristiano Cazita, Jussara M. Almeida, Virgílio Almeida, and Wagner Meira, Jr., "Characterizing a spam traffic," in *IMC '04: Proceedings of the 4th ACM SIGCOMM conference on Internet measurement*, New York, NY, USA, 2004, pp. 356–369, ACM.

 ISBN: 978-1-61208-103-8

Simple Storage Replication Protocol (SSRP) for Intercloud

David Bernstein
Huawei Technologies, Ltd.
North America R&D Division
Santa Clara, California, USA
email: dbernstein@huawei.com

Deepak Vij
Huawei Technologies, Ltd.
North America R&D Division
Santa Clara, California, USA
email: dvij@huawei.com

Abstract – **Working groups have proposed building a layered set of protocols to solve Cloud Computing interoperability challenges. This is not the problem of application portability which calls for standardized programmer interfaces. This is the problem of generalized, transparent, back-end cloud-to-cloud federation. Current Intercloud designs do not envision each cloud provider establishing connectivity with another cloud provider in a Point-to-Point manner, as this will result into the n^2 complexity problem. Instead, Intercloud Directories, Exchanges, and Gateways will help facilitate as mediators for enabling connectivity and collaboration among disparate cloud providers in a manner analogous to the Internet itself. These Intercloud elements would implement a profile of protocols becoming known as Intercloud protocols. Researchers and working groups have proposed a layered set of such protocols using Extensible Messaging and Presence Protocol for transport, and Semantic Web with Resource Description Framework for resource matching. This paper builds on that work and discusses the next layer up, where a specific use case of federation is explored. We find that the back-end implementation of cloud-distributed, unstructured storage can be extended to generalized cloud-to-cloud federation using a Simple Storage Replication Protocol (SSRP), which is built on top of the base Intercloud protocol set. This paper details SSRP.**

Keywords-component: Intercloud; Cloud Computing; Cloud Storage; Cloud Federation; XMPP; RDF; SSRP.

I. INTRODUCTION

Cloud Computing has a well accepted terminology [1], and Use Cases and Scenarios for Cloud IaaS and PaaS interoperability [2][3] have been detailed in the literature along with the challenges around actually implementing standards-based Intercloud federation and hybrid clouds. Work detailing high level architectures for Intercloud interoperability have been published [4][5].

Specific implementation approaches for Intercloud protocols [6][7] have been proposed, including specifically Extensible Messaging and Presence Protocol (XMPP) [8][9] for transport, and using Semantic Web [10] techniques such as Resource Description Framework (RDF) [11] to specify resources.

Detailed approaches outlining the use of these technologies to implement the base Intercloud protocols; have been published first on the feasibility of XMPP as a control plane operations for Intercloud [12], and next how Cloud Computing resources can be described, cataloged, and mediated using Semantic Web Ontologies, implemented using RDF techniques [13]. The base topology and the base transport and resource framework provide a foundation for a specific use case of Intercloud federation, which is the subject of this work.

Here, we outline the problem of federating cloud storage services. We look at the simplest of cloud storage, which is distributed, unstructured storage, best exemplified by the Amazon S3 [14] commercial offering. Cloud Storage challenges for structured data such as relation databases are equally important but of a slightly different scope and dimensions, and not addressed in this paper. We will cover the cloud storage challenges and proposed solutions for structured data storage at a later time.

Our work has detailed a specific use case of distributed unstructured cloud storage federation. We find that the back-end implementation of cloud-distributed, unstructured storage can be extended to generalized cloud-to-cloud federation using a Simple Storage Replication Protocol (SSRP), which is built on top of the previously referenced base Intercloud protocol set. This paper details SSRP.

The rest of the paper is organized as follows: Section II outlines the proposed overall "Intercloud Topology". Section III briefly describes an overview of cloud storage, in general. Section IV specifically describes an overview of unstructured cloud storage. Section V briefly describes typical storage interoperability challenges across various cloud providers. Section VI outlines an overview of Intercloud topology. Section VII briefly delves into brief overview of major constituents of proposed Intercloud topology. Section VIII briefly describes Ontology based cloud resources catalog. Section IX briefly describes cloud resources Ontology itself. Section X describes an overview of XMPP based Intercloud negotiation and services framework proposed as part of the Intercloud standards and protocols. Section XI outlines the sequencing of Intercloud protocols for federated unstructured storage use case. Next Section XII delves deep into proposed Intercloud Enabled Federated unstructured storage system architecture. Section XIII describes an actual use case for Intercloud enabled federated unstructured cloud storage. And finally, Section XIV presents our conclusions.

 ISBN: 978-1-61208-103-8

II. REVIEW OF INTERCLOUD TOPOLOGY

Cloud instances must be able to dialog with each other. One cloud must be able to find one or more other clouds, which for a particular interoperability scenario is ready, willing, and able to accept an interoperability transaction with and furthermore, exchanging whatever subscription or usage related information which might have been needed as a pre-cursor to the transaction.

Thus, an Intercloud Protocol for presence and messaging needs to exist which can support the 1-to-1, 1-to-many, and many-to-many Cloud to Cloud use cases. The vision and topology for the Intercloud we will refer to is as follows. At the highest level, the analogy is with the Internet itself: in a world of TCP/IP and the WWW, data is ubiquitous and interoperable in a network of networks known as the "Internet".

In a world of Cloud Computing, content, storage and computing is ubiquitous and interoperable in a network of Clouds known as the "Intercloud"; this is illustrated in Figure 1.

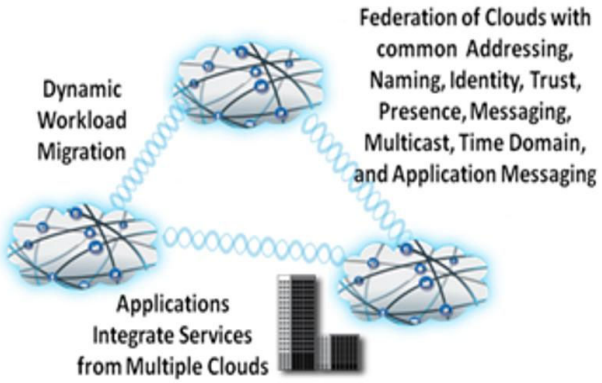


Figure 1. The Intercloud Vision

The reference topology for realizing this vision is modeled after the public Internet infrastructure. Again, we use the generally accepted terminology.

There are Public Clouds, which are analogous to ISP's and Service Providers offering routed IP in the Internet world.

There are Private Clouds which are simply a Cloud which an organization builds to serve itself.

There are Intercloud Exchanges (analogous to Internet Exchanges and Peering Points) where clouds can interoperate.

Finally, there is an Intercloud Root, containing services such as Naming Authority, Trust Authority, Directory Services, and other "root" capabilities. It is envisioned that the Intercloud root is of course physically not a single entity, a global replicating and hierarchical system similar to DNS [15] would be utilized.

All elements in the Intercloud topology contain some gateway capability analogous to an Internet Router, implementing Intercloud protocols in order to participate in Intercloud interoperability. We call these Intercloud Gateways. The entire topology is detailed in Figure 2.

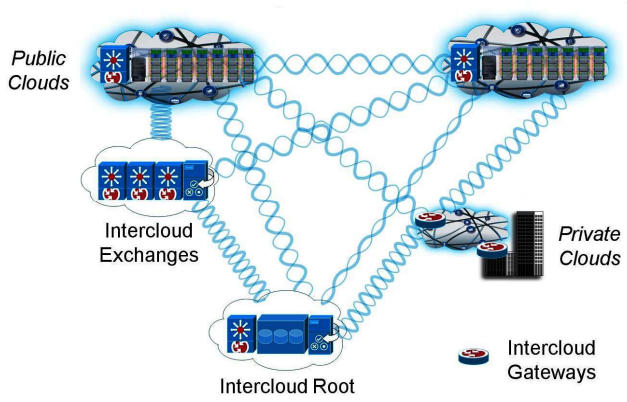


Figure 2. Reference Intercloud Topology

The Intercloud Gateways would provide mechanism for supporting the entire profile of Intercloud protocols and standards. The Intercloud Root and Intercloud Exchanges would facilitate and mediate the initial Intercloud negotiating process among Clouds. Once the initial negotiating process is completed, each of these Cloud instance would collaborate directly with each other via a protocol and transport appropriate for the interoperability action at hand; for example, a reliable protocol might be needed for transaction integrity, or a high speed streaming protocol might be needed optimized for data movement over a particular link.

III. OVERVIEW OF CLOUD STORAGE

At a very high level, cloud storage solves the following issues faced by IT organizations:

- Dynamic capacity – storage capacity is fixed once purchased/leased. Cloud storage provides an almost infinite amount of storage for data. One pays for this storage, in GB or TB per month increments, with added storage services (multi-site replication, high availability, etc.) at extra charge. Such capacity can be reduced or expanded at a moments notice.
- Offsite DR – disaster recovery for many small shops is often non-existent or rudimentary at best. Using cloud storage, data can be copied to the cloud and accessed anywhere via the internet. Such data copies can easily support rudimentary DR for a primary data center outage.
- Access anywhere – storage is typically local to the IT shop and can normally only be accessed at that location. Cloud storage can be accessed from any internet access point. Applications that are designed to operate all over the world can easily take advantage of such storage.
- Data replication – storage data is replicated for high availability. Cloud storage providers replicate

 ISBN: 978-1-61208-103-8

storage data to multiple sites so that if one site goes down other sites can still provide service transparently to the consumer.

Typically, users are provided with a very simplistic set of APIs (RESTful Web Services call) in order to be able to PUT and GET data into a big cloud of storage with the following high-level characteristics:

- End users Customers need not be aware about exact physical location of the data
- End users can advise the system of their physical location requirements for data (limiting it to a particular country or region) and the Cloud may be able to support this requirement
- The storage needs to be transparently reliable (and replicated) as far as the user is concerned
- Users access the storage over the public Internet through a Storage API

Functionally, the basic intent of the Storage API is to provide a static URI storage repository for files which will be referenced as a whole. Often, this is called "BLOB" storage, meaning "Binary Large Object", as it is not providing any structure like a file system or a database. The most successful commercial implementation of such a BLOB storage service is Amazon S3.

In contrast to Amazon S3, some cloud storage solutions provide various other access points, in addition to HTTP based web services APIs (SOAP or REST based). Examples of these access points may include: NFS like, FTP like, WebDAV like, CIFS like, iSCSI like, BitTorrent like, etc. Essentially, these access points are a veneer or an abstraction adapter layer on top of underlying BLOB storage access mechanism (web services in the case of Amazon S3). Additionally, there are cloud storage solutions which provide underlying structures beyond BLOB (such as Block Storage, Queue Storage etc.). These storage structures, in turn, may provide all of the above access points on top of the underlying access mechanism for the respective storage structure supported by the cloud storage provider.

This paper does not study the interoperability challenges of these examples. We look specifically at the problem of interoperability across cloud providers of unstructured (BLOB) storage.

IV. UNDERSTANDING UNSTRUCTURED CLOUD STORAGE MECHANISMS

For cloud storage providers such as Amazon S3, users do not get direct access to the cloud storage. Instead, they need to invoke HTTP based web services call in order to perform data storage functions such as store, remove or retrieve data etc. However, there are various third party tools such as "JungleDisk" which provides WebDAV based file mount so that the Amazon S3 cloud storage looks like a local device to the end users (*storage-as-storage)*.

As mentioned earlier that enabling cloud storage typically does not require any manual intervention and is essentially self-services. While provisioning storage, users may be asked to choose preferences for geographic regions in which primary copy of their content may be stored. The preference may be one or more regions. If more than one region is selected, clients may optionally choose a geographic location while uploading (PUT) the content.

If no geographic preference is provided during PUT, the system will choose from one of the preferences. If no preference is selected, the system can pick the region on its own discretion. It is typically recommended that clients choose a few regions for optimal performance and Business Continuity Planning. Essentially, the storage provisioning process is "a-la-carte menu" approach to provisioning.

V. STORAGE INTEROPERABILITY CHALLENGES ACROSS CLOUD PROVIDERS

Let us consider an interoperability use case scenario. A user is performing a function that utilizes Cloud based storage capabilities. In Cloud Computing, storage is not like disk access, there are several parameters around the storage which are inherent to the system, and one decides if they meet certain needs or not. For example, storage is typically replicated to several places in the cloud, In AWS [16] and in Azure [17] it is replicated three places. The storage API is such that, a *write* will return as successful when one replicate of the storage has been affected, and then a "lazy" internal algorithm is used to replicate the storage object to two additional places. If one or two of the storage object replicates are lost the cloud platform will replicate it to another place or two such that it is now in three places.

A user has some control over where the storage is, physically, for example, one can restrict the storage to replicate entirely in North America or in Europe. There is no ability to vary from these parameters; that is what the storage system provides.

We do envision other providers implementations might say, five replicates, or a deterministic replication algorithm, or a replicated (DR) *write* which doesn't return until and unless *n* replicates are persisted. One can create a large number of variations around "quality of storage" for Cloud.

In the interoperability scenario, suppose AWS is running short of storage, or wants to provide a geographic storage location for an AWS customer, where AWS does not have a datacenter, it would be sub-contracting the storage to another cloud service provider. In either of these scenarios, AWS would need to find another cloud, which was ready, willing, and able to accept a storage subcontracting transaction with them. AWS would have to be able to have a reliable conversation with that cloud, again exchanging whatever subscription or usage related information which might have been needed as a pre-cursor to the transaction, and finally have a reliable transport on which to move the storage itself.

 ISBN: 978-1-61208-103-8

Note, the S3 storage API is not guaranteed to succeed, if there is a failed "*write"* operation from AWS to a subscriber request, the subscriber code is supposed to deal with that (perhaps, via an application code level retry). However Cloud to Cloud, a target cloud "*write"* failing is not something the subscriber code can take care of. That needs to be reliable.

Currently, due to the proprietary nature of each cloud, every cloud environment is a silo in itself. There are no formal protocols and standards established in order to address the above mentioned issues. The intent of our work is to address storage interoperability issues such as naming, discovery, conversation setup items challenges etc. End goal is to provide a common set of standards, protocols and new components (such as cloud exchange providers role as intermediaries) to address all these interoperability issues in a seamless manner for enabling "Federated Cloud Storage" environment.

VI. OVERVIEW OF INTERCLOUD SYSTEM ARCHITECTURE

As shown in the Figure 3, we envision storage in the Intercloud environment to be federated among disparate and heterogeneous cloud environments. "Intercloud Exchanges" would be a key component for enabling the seamless federated storage environment. These exchanges would facilitate and mediate initial negotiating process among clouds.

Once the initial negotiating process is completed, each of these cloud environments, in turn, would collaborate with each other via "Intercloud Gateways". Intercloud Gateways would provide mechanism for authentication, support for various storage replication and storage access protocols and standards, presence/collaboration, and common "Cloud Ontology" set among heterogonous cloud environments.

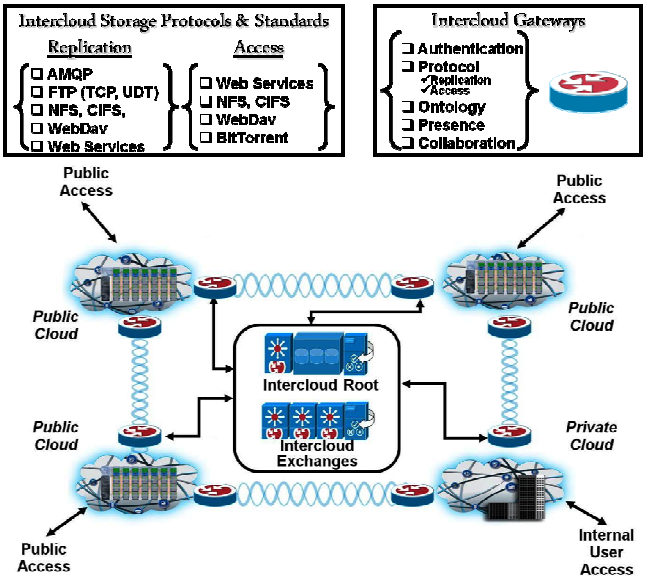


Figure 3. Intercloud enabled Federated Storage Architecture

As shown in the schematic above that cloud-to-cloud storage replication process might leverage application level protocols such as FTP or messaging protocol such as AMQP [18]. These cloud-to-cloud application protocols could be delivered over underlying transport protocol such as UDT (UDP based Data Transfer) [19] instead of traditional TCP protocol. This is mainly due to the fact that TCP protocol becomes very inefficient in a high bandwidth environment. On the other hand, public access to the federated cloud storage may still be delivered via Web Services APIs (RESTful or SOAP based) etc.

VII. INTERCLOUD ROOT, EXCHANGES AND GATEWAYS

As mentioned earlier the various providers will emerge in the enablement of the Intercloud. We first envision a community governed set of Intercloud Root providers who will act as brokers and host the Cloud Computing Resource Catalogs for the Intercloud computing resources. They would be governed in a similar way in which DNS, Top Level Domains [20] or Certificate Authorities [21] are, by an organization such as ISOC [22] or ICANN [23]. They would also be responsible for mediating the trust based federated security among disparate clouds by acting as Security Trust Service providers using standards such as SASL [24] and SAML [25].

The Intercloud Root instances will work with Intercloud Exchanges to solve the n^2 problem by facilitating as mediators for enabling connectivity among disparate cloud environments. This is a much preferred alternative to each cloud vendor establishing connectivity and collaboration among themselves (point-to-point), which would not scale physically or in a business sense.

As we mentioned earlier that all elements in the Intercloud topology contain some gateway capability analogous to an Internet Router, implementing Intercloud protocols in order to participate in Intercloud interoperability. We call these Intercloud Gateways.

Intercloud Gateways would provide mechanism for supporting the entire profile of Intercloud protocols and standards. The Intercloud Root and Intercloud Exchanges would facilitate and mediate the initial Intercloud negotiating process among Clouds.

Once the initial negotiating process is completed, each of these Cloud instance would collaborate directly with each other via a protocol and transport appropriate for the interoperability action at hand.

VIII. ONTOLOGY BASED RESOURCES CATALOG

In order for the Intercloud capable Cloud instances to federate or otherwise interoperate resources, a Cloud Computing Resources Catalog system is necessary infrastructure. This catalog is the holistic and abstracted view of the computing resources across disparate cloud environments. Individual clouds will, in turn, will utilize this catalog in order to identify matching cloud resources by applying certain Preferences and Constraints to the

 ISBN: 978-1-61208-103-8

resources in the computing resources catalog. The technologies for this use the Semantic Web which provides for a way to add "meaning and relatedness" to objects on the Web. To accomplish this, one defines a system for normalizing meaning across terminology, or Properties. This normalization is called Ontology.

Comprehensive semantic descriptions of services are essential to exploit them in their full potential. That is discovering them dynamically, and enabling automated service negotiation, composition and monitoring. The semantic mechanisms currently available in service registries such as UDDI [26] are based on taxonomies called "tModel" [27]. tModel fails to provide the means to achieve this, as they do not support semantic discovery of services [28][29].

We propose a new service directory along the lines of UDDI but based on RDF/OWL [30] ontology instead of current tModel based taxonomy. This catalog is illustrated in Figure 4. As can be seen, the catalog captures the computing resources across all clouds in terms of "Capabilities", "Structural Relationships" and Policies (Preferences and Constraints).

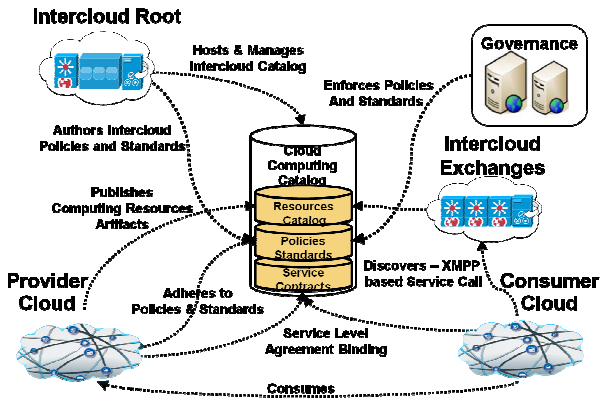


Figure 4. Cloud Computing Catalog

IX. RDF BASED RESOURCES ONTOLOGY

In order to ensure that the requirements of an intercloud enabled cloud provider are correctly matched to the infrastructure capabilities in an automated fashion, there is a need for declarative semantic model that can capture both the requirements and constraints of computing resources.

The chief objectives of the planned configuration are to provide cost effective use of computing resources and to meet the business objectives of the enterprise. In order to automate an environment whereby software agents versus traditional human users discover and consume services, intelligent ontology based service registries are needed for dynamically discovering and provisioning computing resources across various computing cloud environments (Amazon, Azure etc. etc.). We are proposing ontology based semantic model that captures the features and capabilities available from a cloud provider's infrastructure. These capabilities are logically grouped together and exposed as standardized units of provisioning and configuration to be consumed by another cloud provider/s. These capabilities are then associated with policies and constraints for ensuring compliance and access to the computing resources.

This model not only consists of physical attributes but quantitative and qualitative attributes such as "Service Level Agreements (SLAs)", "Disaster Recovery" policies, "Pricing" policies, "Security and Compliance" policies, and so on.

Our earlier work [13] explains how resources can be described, cataloged, and mediated using Semantic Web Ontologies with RDF. Although the terms "taxonomy" and "ontology" are sometimes used interchangeably, there is a critical difference. Taxonomy indicates only class/subclass relationship whereas Ontology describes a domain completely. The essential mechanisms that ontology languages provide include their formal specification (which allows them to be queried) and their ability to define properties of classes. Through these properties, very accurate descriptions of services can be defined and services can be related to other services or resources. Figure 5 shows a high level schematic of proposed ontology based semantic model.

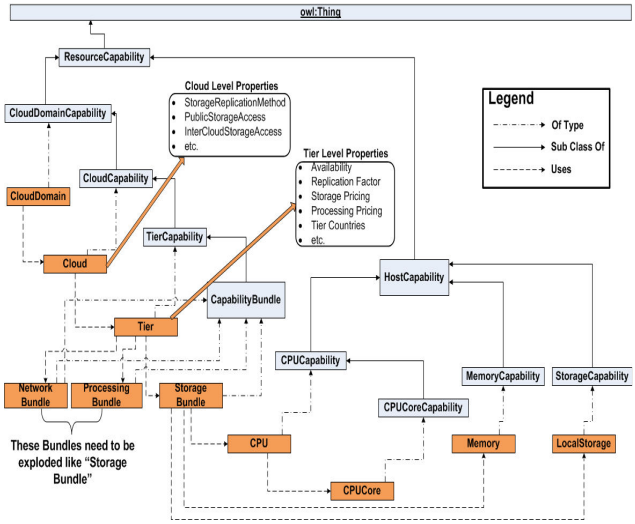


Figure 5. Cloud Computing Resources Ontology

X. XMPP BASED INTERCLOUD PROTOCOLS NEGOTIATION AND SERVICES FRAMEWORK

Part of interoperability is that cloud instances must be able to conduct dialog with each other. As part of the "Federated Storage" use case, one cloud must be able to find another cloud, which for a particular interoperability scenarios, is ready, willing, and able to accept an interoperability transaction with and furthermore, exchanging whatever subscription or usage related information which might have been needed as a pre-

 ISBN: 978-1-61208-103-8

cursor to the transaction. Thus, an Intercloud Protocol for presence and messaging needs to exist.

Extensible Messaging and Presence Protocol (XMPP) is exactly such a protocol. XMPP is a set of open XML technologies for presence and real-time communication developed by the Jabber open-source community in 1999, formalized by the IETF in 2002-2004, continuously extended through the standards process of the XMPP Standards Foundation. XMPP supports presence and structured conversation of XML data.

Our earlier work [12] explains in great detail as far as feasibility of XMPP as control plane operations protocol for Intercloud.

XI. SEQUENCING THE PROTOCOLS FOR INTERCLOUD ENABLED FEDERATED UNSTRUCTURED CLOUD STORAGE

The following is a high level sequence for "Intercloud Enabled Federated Unstructured Cloud Storage" amongst disparate cloud storage providers. "Inter-Cloud Exchanges" facilitate the negotiation process among disparate heterogeneous clouds in order to enable a seamless federated storage environment.

Detailed code samples, and how Cloud Computing resources can be described, cataloged, and mediated using RDF techniques for the various steps in this sequence diagram are outlined in our earlier work.

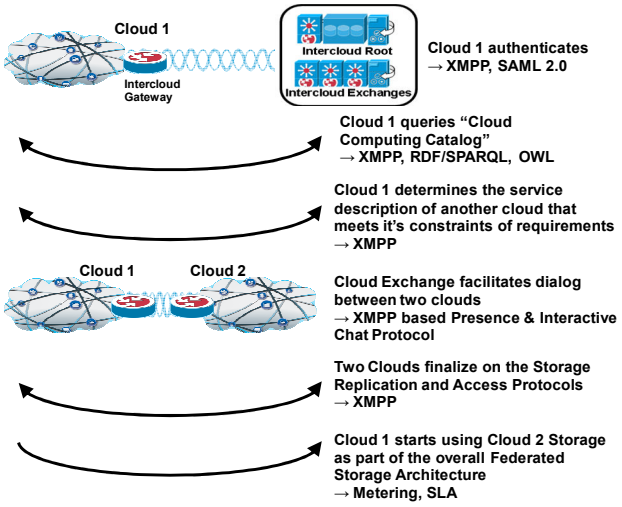


Figure 6. SSRP Unstructured Cloud Storage Sequence Diagram

XII. SSRP ENABLING CLOUD STORAGE ARCHITECTURE

To describe it in very simple terms, unstructured BLOB storage is a large-scale URI storage system. The goal of unstructured BLOB storage architecture is to provide a scalable mass storage solution. The system is designed to be scalable both in terms of total data stored, as well as the number of requests per second for that data.

In a cloud storage environment, due to the sheer number of data storage objects that must be stored and managed, it is very cumbersome and inefficient for traditional file systems supported by network storage models such as SAN and NAS to store and organize these data storage objects. These objects are stored near the users that will access them. Cloud vendors leverage their own proprietary storage solutions in order to efficiently mange the scale and QoS for data storage.

At its core, it is a middleware layer which virtualizes mass storage, allowing the underlying physical storage to be SAN, NAS, or DAS etc. The architecture also manages the replication of data between storage clusters in geographically distributed datacenters. The application can specify fine-grained replication policies, and the architecture layer replicates data according to the policies.

Data stored is typically organized over a two-level namespace. At the top level are buckets–similar to folders or containers–which have a unique global name and serve several purposes: they allow users to organize their data, they identify the user to be charged for transfers, and they serve as the unit of aggregation for audit reports. Each bucket, in turn, can store an unlimited number of data objects. Each object has a name, an opaque blob of data, and metadata consisting of a small set of predefined entries of user-specified name/value pairs. From optimization standpoint, it strives to locate data close to users to reduce latency.

Users can create, modify and read objects in buckets through the REST interface, subject to access control restrictions. End users typically create collections of files, and each file is identified with a URL. This URL can be embedded directly in a web page, enabling the user's browser to retrieve files from the cloud storage system directly, even if the web page itself is generated by a separate HTTP or application server. URLs are also virtualized, so that moving or recovering data on the back end file system does not break the URL.

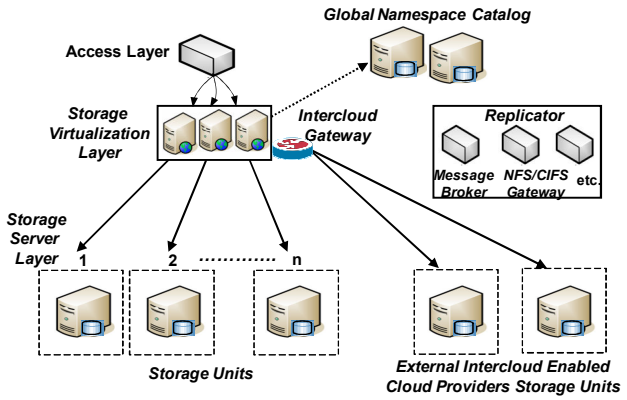


Figure 7. SSRP Enabled Cloud Storage Architecture

Figure 7 illustrates an SSRP enabled federated cloud storage architecture works. Following is a brief description of various components of this architecture.

- Access Layer: This layer is responsible providing various access methods to the underlying cloud based virtualized storage.

Protocols supported are HTTP based web services APIs (RESTful and SOAP based), *storage-as-storage* methods such as NFS, FTP, WebDAV, CIFS and iSCSI etc.

- Global Namespace Catalog: This catalog maintains a Global Namespace across multiple heterogeneous and distributed storage systems, data centers, and administrative domains across the storage Federation (across clouds), independent of their physical storage infrastructure. This catalog is used by the "Storage Virtualization Layer" to determine where a given "Storage Object" is physically located. It maps the Logical Namespace of the storage object to the Physical Storage resources within the overall storage federation. Logical Resources are used to group one or more replicas of Physical Resources, making it transparent to the user where the storage is ultimately stored. Logical resources, are used for load balancing among several storage replicas.
- Storage Virtualization Layer: This layer is essentially the heart of a typical cloud storage system and is responsible for virtualization of the underlying raw storage. This layer determines what to do with the request. For example, if there is a GET request it will first determine which storage instance this needs to be routed to in order for processing. In the event of storage instance unavailability, this layer will determine the next best storage replica the GET call needs to be forwarded to. It uses Global Namespace Catalog to determine the overall routing process. In the case of federated storage where replicas might be dispersed across multiple cloud systems, "InterCloudStorageAccess" is a cloud resources ontology element initially negotiated among storage cloud systems. It determines how the storage is accessed across cloud system boundaries. The "InterCloudStorageAccess" methods might include: NFS/sNFS, CIFS, iSCSI, WebDAV, Web Services (RESTful or SOAP based), BitTorrent, etc.
- Replicator: This layer is responsible for keeping multiple copies of data across cloud storage clusters, storage Replication is the sole purpose of this module. The write operations (PUT/DELETE) are propagated to other physical locations when they are committed to the primary storage location. It allows for replication across datacenters for a cloud. This component is responsible for N-way replication. In the case of federated storage where replicas might be replicated across multiple cloud systems, "StorageReplicationMethod" is a cloud resources ontology element initially negotiated among storage cloud systems. It determines how the data storage is replicated across system boundaries. The "StorageReplicationMethod" methods might include: AMQP based Message Broker delivered over UDT or TCP, FTP/sFTP delivered over UDT or TCP, NFS/sNFS, CIFS, iSCSI, WebDAV, Web Services (RESTful or SOAP based), BitTorrent, etc. These storage replication protocols are negotiated as part of the initial conversational dialog (using XMPP protocol) between cloud exchanges and participant cloud vendors.
- Storage Server Layer: This layer actually handles the reads/writes/deletes for the data storage objects to their physical location.

XIII. THE CONSUMER VISIBLE USE CASE – MOBILE STORAGE ROAMING

Once a federated storage cloud system has been enabled, we can envision the experience for an end user. A typical application of cloud storage will be to provide a transparent persistence for a mobile user.

As the user makes and receives calls, adds to their address book, snaps photos, and takes video, all of this content will immediately be streamed up to the cloud, persisting the data in a reliable way for the user. The mobile service provider will arrange for a cloud entry point which is "closest" to the mobile user so that they get the best uploading performance possible.

The cloud will use its internal replication algorithms and mechanisms to copy the data to several geographically storage nodes, most likely distributed around the geography for which the mobile service provider has coverage. If the user flies across the country, they will find their content replicated close to them and be able to experience superior download performance for their content as well. The user will also immediately be able to access or share their content from alternative channels, such as a web browser, or an Internet connected television, in real-time, as the cloud makes this content available.

Next, the user flies outside of the coverage area of the home service provider such that he is roaming for voice and messaging and data traffic. The mobile network already is designed to allow the roaming provider to service the user by providing the handling of this voice, messaging, and data traffic through an agreement with the home service provider such that the user does not realize he is being serviced by a roaming provider (perhaps other than the fees he will end up paying). However, his content is on a different continent that he is. It would be optimal for the user to provide an equivalent service for all of his data, such that for uploading and for downloading, some working set of his data was persisted in a nearby replicate.

This requires that the clouds which are implementing the storage for the mobile users, from the home and the roaming provider, operated in a federated manner, as we have described with SSRP. If this was the case, the mobile user would experience reliable and "nearby" storage, with all of the replication and performance

 ISBN: 978-1-61208-103-8

capabilities he enjoys with his home service provider. This scenario is illustrated in Figure 8.

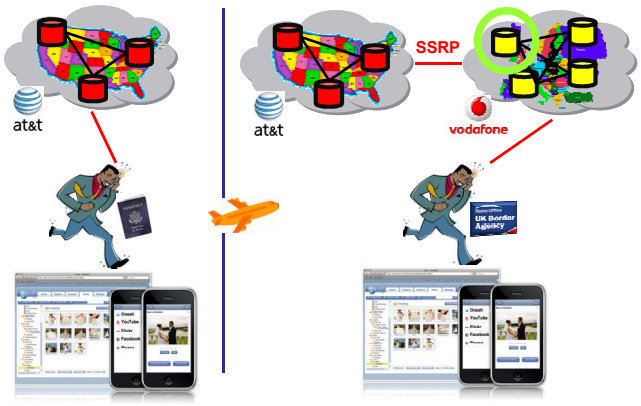


Figure 8. SSRP Enabled Cloud Storage Roaming use case

Here we show the user traveling from the USA, and ending in the UK, and the replicate which the roaming provider serves up to him, because they know where he is by virtue of his connectivity, is served up nearby to him, as illustrated.

XIV. CONCLUSIONS

The conclusion is that we have gone into great detail to test the proposal that Intercloud Topology and Protocols are suitable for a federated cloud storage use case scenario. We call the collection of these protocols and the resource definitions, SSRP. We have also described a meaningful real-world use case for this.

The next stages of our work are to develop details for Intercloud Topology and the governance process in support of proposed Intercloud Topology.

REFERENCES

[1] Youseff, L. and Butrico, M. and Da Silva, D., *Toward a unified ontology of cloud computing*, GCE'08 Grid Computing Environments Workshop, 2008.

[2] Lijun Mei, W.K. Chan, T.H. Tse, *A Tale of Clouds: Paradigm Comparisons and Some Thoughts on Research Issues*, APSCC pp.464-469, 2008

[3] *Cloud Computing Use Cases Google Group (Public)*, at http://groups.google.com/group/cloud-computing-use-cases, http://www.scribd.com/doc/18172802/Cloud-Computing-Use-Cases-Whitepaper ,

[4] Buyya, R. and Pandey, S. and Vecchiola, C., *Cloudbus toolkit for market-oriented cloud computing,* 1st International CloudCom, 2009

[5] Yildiz M, Abawajy J, Ercan T., Bernoth A., *A Layered Security Approach for Cloud Computing Infrastructure*, ISPAN, pp.763-767, 2009

[6] Bernstein, D., Ludvigson, E., Sankar, K., Diamond, S., and Morrow, M., *Blueprint for the Intercloud - Protocols and Formats for Cloud Computing Interoperability*, ICIW '09. Fourth International Conference on Internet and Web Applications and Services, pp. 328-336, 2009

[7] Bernstein, D., *Keynote 2: The Intercloud: Cloud Interoperability at Internet Scale*, NPC, pp.xiii, 2009 Sixth IFIP International Conference on Network and Parallel Computing, 2009

[8] *Extensible Messaging and Presence Protocol (XMPP): Core,* and related other RFCs at http://xmpp.org/rfcs/rfc3920.html

[9] *XMPP Standards Foundation* at http://xmpp.org/

[10] *W3C Semantic Web Activity*, at http://www.w3.org/2001/sw/

[11] *Resource Description Framework (RDF*), at http://www.w3.org/RDF/

[12] Bernstein, D., Vij, D., *Using XMPP as a transport in Intercloud Protocols,* Proceedings of CloudComp 2010, the 2nd International Conference on Cloud Computing, Barcelona, Spain, 2010.

[13] Bernstein, D., Vij, D., *Using Semantic Web Ontology for Intercloud Directories and Exchanges*, Proceedings of ICOMP'10, the 11th International Conference on Internet Computing, Las Vegas, USA, 2010.

[14] *James Murty, programming Amazon Web Services; S3, EC2, SQS, FPS, and SimpleDB, O'Reilly Press, 2008*

[15] *Domain Names – Concepts and Facilities,* and related other RFCs, at http://www.ietf.org/rfc/rfc1034.txt

[16] *Amazon Web Services*, at http://aws.amazon.com

[17] *Microsoft Azure*, at http://www.microsoft.com/azure/default.mspx

[18] *Advanced Message Queuing Protocol*, at http://jira.amqp.org

[19] *UDT – UDP-based Data Transfer*, at http://udt.sourceforge.net

[20] *Domain Name System Structure and Delegation*, at http://www.ietf.org/rfc/rfc1591.txt

[21] *Internet X.509 Public Key Infrastructure, Certificate Policy and Certification Practices Framework,* at http://tools.ietf.org/html/rfc3647

[22] *The Internet Society,* at http://www.isoc.org/

[23] *The Internet Corporation for Assigned Names and Numbers*, at http://www.icann.org/

[24] *Simple Authentication and Security Layer (SASL*), at http://tools.ietf.org/html/rfc4422

[25] *Security Assertion Markup Language (SAML*), at http://saml.xml.org/saml-specifications

[26] *OASIS UDDI Specification TC,* at http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=uddi-spec

[27] *UDDI Registry tModels,* at http://www.uddi.org/taxonomies/UDDI_Registry_tModels.htm

[28] Paolucci, M., Kawamura T., Payne T., and Sycara K., *Importing the Semantic Web in UDDI*, Web Services, E-Business and Semantic Web Workshop, 2002.

[29] Moreau, L. and Miles, S. and Papay, J. and Decker, K. and Payne, T., *Publishing semantic descriptions of services*, First GGF Semantic Grid Workshop, held at the Ninth Global Grid Forum, Chicago IL, USA, 2003

[30] Web Ontology Language*,* at http://www.w3.org/TR/owl-features/

[31] *SPARQL Query Language for RDF*, at http://www.w3.org/TR/rdf-sparql-query/

 ISBN: 978-1-61208-103-8

A Novel Dynamic Bandwidth Allocation Algorithm Based on Half Cycling for EPON

Özgür Can TURNA [1], Muhammed Ali AYDIN [1], Tülin ATMACA [2] , A. Halim ZAİM [1]
{ ozcantur, aydinali, ahzaim@istanbul.edu.tr} [1]
Department of Computer Engineering, Istanbul University, 34320 Istanbul, Turkey
{tulin.atmaca@it-sudparis.eu} [2]
Department RST, Institut Telecom/ Telecom SudParis, 91011 Evry, France

***Abstract*— The access network solutions based on the fiber infrastructure are examined and developed in the last decade. Ethernet Passive Optical Network (EPON), which has only passive optical units in its' infrastructure, comes front in cost and deliverability of service with high bandwidth and long haul access. In upstream direction, EPON needs a multiple access control mechanism to control the bandwidth allocation among Optical Network Units (ONUs) where Multi-Point Control Protocol (MPCP) is responsible for. In this article we propose a novel dynamic bandwidth allocation algorithm which can increase the link utilization with a fair distribution among ONUs. Our algorithm uses half cycle stops thereby we don't have to wait for calculation while waiting report messages from the entire ONUs. Finally, we simulate an EPON network with mono-service and multi-service traffic in two cases to compare our algorithm with Interleaved Polling with Adaptive Cycle Timing (IPACT) and offline Dynamic Bandwidth Allocation (oDBA) algorithms. Our algorithm gives better performance in byte loss ratio and mean access delay values compared to IPACT and oDBA.**

Keywords*— *Ethernet Passive Optical Network (EPON), Dynamic Bandwidth Allocation (DBA), Half Cycle Dynamic Bandwidth Allocation (hcDBA), Performance Evaluation.

I. INTRODUCTION

By the incredible development in internet and computer technology, users' demands of more bandwidth increase rapidly. While the backbone and local area networks have very fast infrastructures (such as 10Gbit Ethernet LAN), access networks drop down the total network capacity for users while they try to access remote sources [1][2]. In early years of Internet, carried traffic was comprised of plain text pages and images which can be carried by a limited bandwidth capacity. However, nowadays mostly carried traffic over internet is comprised of peer-to-peer file and video sharing, online real-time gaming, video streaming, on demand video and education, IP telephony and IPTV. These applications need more bandwidth in access network area and some also need quality of service in packet delay variation (PDV), packet loss and end-to-end delay cases. To overcome such demands of future internet applications, service providers always research new access technologies. Most of the service providers and network infrastructure designers start to study Fiber-to-the-Home (FTTH) architectures [3]. The most popular FTTH architecture is passive optical network (PON) architecture which has the best cost-effectiveness among fiber access architectures. In PONs for downstream, the data packets are broadcasted from the central office part of the PON, namely Optical Line Terminal (OLT) and the subscriber part, namely Optical Network Unit (ONU) collects the packets sent to itself. In downstream direction, the messages must be encrypted for undesirable access of other subscribers. For upstream direction, since ONUs are connected to the OLT over a single fiber line, a multiple access technology must be used to overcome the congestion conditions. In PON, two different multiple access types are in use; Time Division Multiple Access (TDMA) and Wavelength Division Multiple Access (WDMA).

In TDMA, there are two main standardization branches exist in network area. The first one is Gigabit PON (GPON) which is standardized by ITU-T. The second one is EPON which is standardized by IEEE 802.3ah Task Force [4]. They have published EPON standard in 2000 at 1Gbps up/down transmission capacity. By 2009, 10Gbps up/down transmission capacity has been standardized for EPON architecture [5].

The ease of implementation and cost effectiveness of EPON makes it more popular than GPON in academic studies and industrial world [1]. There are lots of studies to improve the performance, access capacity and service quality in EPON.

The scheme of bandwidth allocation in EPON can be either static or dynamic. In static allocation, a fixed-size transmission window is allocated by OLT, to each ONU, regardless of the traffic requirements at each ONU. On other hand, in the dynamic allocation, a variable-size transmission window is dynamically allocated to the different ONUs, taking into account their traffic which is expressed explicitly by each ONU. The communication between OLT and ONUs is achieved by the multipoint-control protocol (MPCP) which is defined by the IEEE 802.3ah Task Force [4].

In this paper, we present a novel Dynamic Bandwidth Allocation (DBA) algorithm for EPON. We show the basic DBA algorithms for EPON and compare our algorithm with them in terms of mean access delay, byte loss ratio and packet delay variation.

This paper is organized as follows. In section II, existing dynamic bandwidth allocation algorithms have been summarized. In section III, our proposed algorithm is introduced. In section IV, our simulation environment is described and the simulation results are presented. Finally, section V concludes the paper.

 ISBN: 978-1-61208-103-8

II. DBA ALGORITHMS

There is lots of dynamic bandwidth algorithms developed for EPON. The early solution for dynamic bandwidth allocation "Interleaved Polling with Adaptive Cycle Time (IPACT)" is proposed by G. Kramer et al. In [6] there are five different conditions which can be used over IPACT and according to the authers, the best variation of IPACT algorithm is to use a maximum window limit approach. In IPACT, OLT sends gate messages to the ONUs one by one in an interleaved fashion without waiting the next report message to arrive from other ONUs. If we increase the maximum window size, this can cause longer waiting time for packets in ONUs local buffers to be sent in next cycle. On the contrary, if we set the window size shorter, this will cause more GATE and REPORT transmission which bring extra overhead to the system.

For fair bandwidth distribution over highly loaded ONUs, another DBA algorithm presented which is based on Interleaved Polling with Stop (also known as offline DBA and here after called as oDBA in this paper). In this scheme, OLT waits for report messages from the entire ONUs in each cycle before it starts to send gate messages to ONUs for next cycle. By doing this, OLT can know the entire bandwidth request from the entire ONUs before it starts to grant bandwidth for ONUs. Thus, OLT can distribute the excess bandwidth fairly among highly loaded ONUs. However, oDBA inserts an idle time ***(T_{idle})*** in upstream channel which consists of Computation Time for the algorithm and Round Trip Time (RTT) between OLT and ONUs (assumed that all ONUs have the same RTT).

oDBA algorithms collect all the bandwidth demands in a cycle. An Excess Bandwidth Distribution (EBD) mechanism allots the excess bandwidth collected from lightly loaded ONUs, among highly loaded ONUs. For EBD, firstly minimum guaranteed bandwidth in a cycle for each ONU B_i^{MIN} is computed for N ONUs as in formula 1.

$$B_i^{MIN} = \frac{(T_{cycle} - N \times T_g)}{8 \times N} \times R \quad (1)$$

where T_{cycle} is cycle time, T_g is the guard time and R is the upstream channel capacity. Then, the bandwidth B_i^g needs to attribute to ONU_i is computed as in formula 2.

$$B_i^g = \begin{cases} R_i & if \quad R_i < B_i^{MIN} \\ B_i^{MIN} & if \quad R_i \geq B_i^{MIN} \end{cases} \quad (2)$$

where R_i is the bandwidth requested by the ONU_i.
After, the excess bandwidth which is not yet attributed in the current cycle is fairly distributed among all highly loaded ONUs.

Some previous works for oDBA have been carried out to fill the idle time period. In [7], the authors developed an algorithm that schedule lightly loaded ONUs in idle time period without waiting for entire ONUs to send their REPORT messages for next cycle timing. This approach is good for increasing the throughput in low loads. If the entire ONUs are highly loaded the idle period is still wasted. In [8] authors improved the idle time usage by adding a case to choose one highly loaded ONU to use idle time if no lightly loaded ONU exist in current cycle. Also in [9], another algorithm has been proposed for using idle time period which is capable of highly loaded cases. Contrary to the previous two approaches this one does not use an early allocation; instead it uses the scheme that OLT calculates supplementary granted bandwidth by using the cycle-based arrival rate of client packets.

These proposed algorithms are designed to solve the idle time problem in offline DBA. They change ONUs servicing order which can cause PDV. Proposed algorithms in [9] and [8] send extra control messages which cause extra overhead.

Another approach to decrease delay of packets in ONUs local buffers is to use queue size prediction algorithms. If an ONU is able to predict its buffer size for next cycle then it can demand the necessary bandwidth without waiting for a cycle period [10][11]. However the bursty nature of local traffic sources, the queue size prediction can waste the bandwidth by faulty predictions.

In [6], an exhaustive summary of DBA algorithms for EPON has been presented. For grant sizing there are two main approaches have been studied in literature; online DBA (Interleaved polling with adaptive cycle time approaches) and offline DBA (Interleaved Polling with Stop). Online DBA algorithms give better bandwidth utilization results because they have the capability to allocate all the bandwidth without idle time periods. On the contrary, offline DBA approaches have fair allocation among highly loaded ONUs. However, it introduces an idle time problem which can decrease the bandwidth utilization.

Our motivation to do this work is to develop a middle approach between online and offline DBA algorithms that is able to behave fairly among highly loaded ONUs and provide maximum bandwidth utilization. A grant sizing approach has been developed that switches between online and offline mode dynamically.

III. HALF CYCLING DBA ALGORITHM (HCDBA)

In this section, we present a novel DBA algorithm which is based on Interleaved Polling with Stop and use a different cycle timing control for transmission in upstream channel. We named the proposed algorithm as "Half Cycling Dynamic Bandwidth Allocation Algorithm" and hcDBA abbreviation is used in this article to identify our algorithm. hcDBA algorithm works in two modes according to the load of the upstream channel. In low loads the algorithm switch into online DBA mode which is similar to IPACT algorithm and in high loads it switches into offline DBA mode. The working mode changes respectively according to the incoming upstream bandwidth demands to the OLT.

In offline DBA (Interleaved Polling with Stop) algorithm, if OLT is able to know the bandwidth demands from entire ONUs before idle period start time as earlier as the length of idle period "T_{idle} ($T_{computation}$ + RTT)", GATE messages can be sent without any idle period in upstream channel. In hcDBA, to send GATE messages, OLT

 ISBN: 978-1-61208-103-8

calculates bandwidth amount to be given for half of ONUs instead of entire list, if more than half of ONUs have reported their demands after the last gated ONU. Otherwise, OLT directly sends a GATE message to the following ONU in polling list. In this case, OLT is in online DBA mode which will continue until number of reported ONUs is more than half of the total. When reported ONU count reaches this amount, OLT again starts working in offline DBA mode and jump in half cycle algorithm to distribute bandwidth among ONUs. The algorithm switches into online DBA mode when the OLT cannot collect enough REPORT messages. If time slots for ONUs are so short, the upstream channel is lowly loaded. Thereby, in online mode we do not have to care about fair distribution because we know that the system have a cycle time below the desired cycle time limit. The maximum window size can be held much bigger than IPACT (limited) approach. Even the cycle time becomes longer; the algorithm changes its form to offline DBA which suppose to give a fair bandwidth distribution among ONUs in a cycle.

hcDBA uses the MPCP in EPON standardization without any upgrade necessity in control messages and ONU side implementation. hcDBA changes the OLT side algorithm for polling. In two cases, operation of the algorithm will be explained. First, the work flow diagram of GATE Timer Expire function is given in Figure 1. Second, the EBD algorithm is going to be introduced.

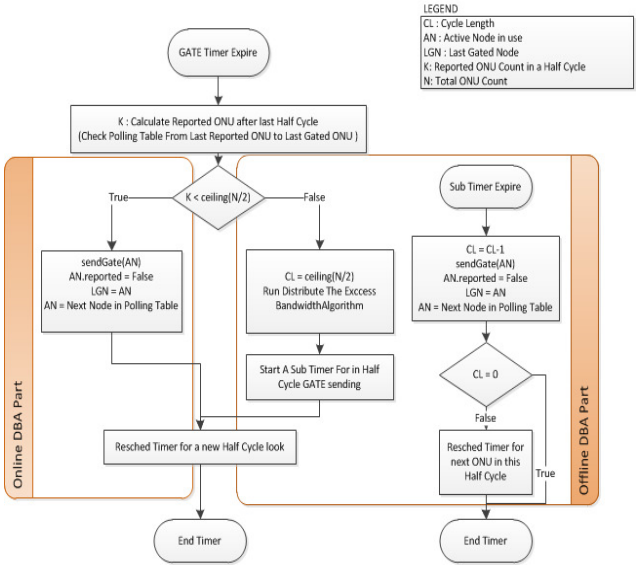


Figure 1. GATE Timer work flow in OLT

GATE timer is responsible for GATE messaging process in OLT. Each time the GATE timer expires, it calculates and prepares the parameters for novel GATE messages.

Addition to offline DBA algorithm, hcDBA algorithm needs demanded and given bandwidth information in previous cycle. Besides, for hcDBA algorithm, the instant monitoring of the last gated ONU and the last reported ONU has to be done. In hcDBA, the servicing order of ONUs during a cycle doesn't change. Thus, the last gated ONU and last reported ONU can be monitored over the polling table simply. The algorithm switches between online and offline mode according to the number of reported ONUs from last gated ONU to gate timer expiration. This parameter is related to the cycle time, requested window size and RTT of ONUs. If the window sizes of half of the ONUs are very small, they can be served in time less than RTT. Thus, when the gate timer expires, if the granted total window size is not more than RTTs for packets, the OLT cannot collect enough report to compute a new half cycle.

The main case to think about in our algorithm is the EBD in offline mode. hcDBA uses a similar method like the one used in offline DBA algorithm for EBD. However, in hcDBA the OLT has to make the EBD process with the ***K*** report messages instead of entire list (***N***). For (***N-K***) nodes, the OLT has not received the bandwidth request yet. If OLT distributes the excess bandwidth according to the ***K*** report information, the algorithm may misjudge the bandwidth demand of ONUs in a full cycle and EBD can be unfair. The half cycle that OLT is going to give grants can take more or less "excess bandwidth-highly loaded node" ratio than consecutive cycles. In a situation like this, unfairness takes places between two ONU groups. For this reason, in hcDBA algorithm, while the excess bandwidth distribution is being calculated for a new half cycle for ***N/2*** ONU, the algorithm does not make the decision just over ***K*** report messages (note that always, $K \geq \lceil N/2 \rceil$). It also includes the excess bandwidth and highly loaded ONUs information for (***N-K***) ONUs from the previous bandwidth requests and grants. If the bandwidth demands are less than excess bandwidth in previous half cycle, more bandwidth can be used in current half cycle. Otherwise, the excess bandwidth is distributed in fair for next half cycle according to the situation of current half cycle.

ONUs are examined in two groups in EBD algorithm such as reported and unreported. The algorithm needs entire ONUs requests to distribute the bandwidth fairly. For ONUs of which reports have not arrived to OLT, the needed information will be generated based on their previous cycle. The details about the process of excess bandwidth algorithm are given below. Minimum bandwidth is calculated same as oDBA as in formula 1.

Excess bandwidth calculation is a bit more different than oDBA approach. The usable excess bandwidth in a half cycle cannot be measured just with the ONUs requests in current half cycle. To distribute the excess bandwidth fairly to entire ONUs in PON, OLT also take into consideration the bandwidth requests of ONUs served in previous half cycle. Since the algorithm cannot use all the excess bandwidth (B^{EXCESS}), it will calculate the usable excess bandwidth B^{USABLE} by using B^{EXCESS} values. B^{EXCESS} is calculated as the sum of excess bandwidth amount of reported (***K***) and previous excess bandwidth amount of unreported (***O: N-K***) nodes. The excess bandwidth values are calculated separately for ***K*** ONUs and ***O*** ONUs. Unused bandwidth of ***K*** ONUs from last report information is given in formula 3. Unused bandwidth of ***O*** ONUs (***O: N-K***), last ***O*** ONUs from previous report information kept in polling table is given in formula 4.

 ISBN: 978-1-61208-103-8

$$B_K^{EXCESS} = \sum_{i=1}^{K} [B_i^{MIN} - B_i^{REQ}] \quad (B_i^{MIN} > B_i^{REQ}) \tag{3}$$

$$B_O^{EXCESS} = \sum_{j=K}^{N} [B_j^{MIN} - B_j^{REQ}] \quad (B_j^{MIN} > B_j^{REQ}) \tag{4}$$

B_K^{USABLE} will show the total maximum excess bandwidth available to use for ***K*** ONUs. In processing half cycle, bandwidth arrangement will be done just for $\lceil N/2 \rceil$ ONUs. In a half cycle, despite EBD is done through ***K*** ONUs, just the distribution of $\lceil N/2 \rceil$ ONUs will be determined. B_K^{USABLE} calculation is done as follows.

$$B_K^{USABLE} = \begin{cases} B_K^{EXCESS} + B_L^{UNUSED} & if \quad \frac{R_K^H}{R_K^H + R_O^H} \times B^{EXCESS} \geq B_K^{EXCESS} \\ \frac{R_K^H}{R_K^H + R_O^H} \times B^{EXCESS} + B_L^{UNUSED} & if \quad \frac{R_K^H}{R_K^H + R_O^H} \times B^{EXCESS} < B_K^{EXCESS} \end{cases} \tag{5}$$

Here, B_L^{UNUSED} is the unused excess bandwidth according to extra bandwidth demand and excess bandwidth of ***N/2*** nodes served just before. (This calculation should be done by checking the polling table each time needed, because of the dynamically switching between online DBA and offline DBA modes. There may be some ONUs served according to online DBA between the previous and current half cycle. If the bandwidth demand of overloaded amount exceeds the excess bandwidth, they will be assumed as zero.) R_K^H indicates the total bandwidth demand of highly loaded ones of ***K*** ONUs and R_O^H indicates the total bandwidth demand of highly loaded ones in previous cycle of ***O*** ONUs.

If the excess bandwidth amount of ***K*** nodes is lower than the bandwidth amount portion of ***K*** nodes in total excess bandwidth (this means that ONUs are overloaded in processing half cycle), algorithm marks the whole excess bandwidth for ***K*** ONUs as usable. If it results in other way, it means that ONUs in previous half cycle are overloaded. In this case, the algorithm will distribute the assigned excess bandwidth to current half cycle considering total needs of entire ONUs in PON, in order to let ONUs in previous cycle to have more excess bandwidth in following cycle. Besides, if unused excess bandwidth exists for ***N/2*** ONUs from previous half cycle, this unused value is also added to excess bandwidth. With combination of these calculations, hcDBA tries to guarantee fair distribution between respective half cycles.

After calculation of B_K^{USABLE}, for each half cycle (just for ***N/2*** ONUs, always ***K*** ≥ ***N/2***), the bandwidth assigned for each ONU will be calculated as below:

$$B_i^{g} = \begin{cases} R_i & if \quad R_i \leq B_i^{MIN} \\ R_i & if \quad R_i > B_i^{MIN} \wedge B_K^{USABLE} \geq (R_K^H - B_i^{MIN} \times K) \\ B_i^{MIN} + \frac{R_i}{R_K^H} \times B_K^{USABLE} & if \quad R_i > B_i^{MIN} \wedge B_K^{USABLE} < (R_K^H - B_i^{MIN} \times K) \end{cases} \tag{6}$$

IV. PERFORMANCE EVALUATION

In this section, we present simulation results to verify our analysis and demonstrate the performance of the proposed hcDBA algorithm. We compare the results obtained from the hcDBA algorithm with IPACT and offline DBA algorithms. We use the same basis for each algorithm on the simulation.

We consider an EPON access network consisting of 16 ONUs connected to an OLT through a passive coupler. All ONUs are assigned a downstream and an upstream propagation delay (from ONU to OLT). We fix the distance between the coupler and OLT and distances between ONUs and the coupler about 10 km (about 0.05 ms). We compare the algorithms in two cases; 1Gbps upstream channel for EPON and 10Gbps upstream channel for 10G-EPON. The algorithms are compared with four different priority classes described as below.

TABLE I. TRAFFIC HYPOTHESIS [9]

	CoS1 Premium	***CoS2 Silver***	***CoS3 Bronze***	***CoS4 BE***
Traffic Ratio	10 %	10 %	30 %	50 %
Packet size (in Bytes)	70	70	50,500,1500	50,500,1500
Source and Burstiness	CBR	CBR	PPBR/ μ=1.4	PPBP/μ=1.4
Burst Length (# of Packets)	CBR	CBR	10	20

TABLE II. SIMULATION PARAMETERS

Parameter	***Value(Case1)***	***Value(Case2)***
No. of ONUs	16	16
Upstream Bandwidth, R	1 Gbit/s	10 Gbit/s
Maximum cycle time for hcDBA and oDBA	2ms	2ms
Maximum transfer window size for IPACT and hcDBA	15 KB (IPACT) 30 KB (hcDBA)	150 KB (IPACT) 300 KB (hcDBA)
Guard Time	5μs	5μs

For Premium and Silver traffic, we use CBR (Constant Bit Rates) sources. To generate self-similar traffic of Ethernet LAN (Bronze and Best Effort BE classes), we use an aggregation of multiple sources of Poisson Pareto Burst Process (PPBP), so called Pareto-distributed ON-OFF [12]. In hcDBA, since our first goal is to improve the bandwidth utilization with fairness among ONUs, we also give some results without service classes to show the overall utilization performance of hcDBA.

Simulations were done using discrete event network simulation tool (ns2.34). Table II shows the simulation parameters for each algorithm.

We shall start the performance comparison of hcDBA algorithm with others, considering byte loss ratio. In Figure 2a the byte loss ratios for mono-service traffic are shown. Only oDBA algorithm has byte loss in 0.9 offered load in mono-service traffic condition. Since oDBA provides bad bandwidth utilization compared to others. In Figure 2b the byte drop results of three algorithms for multi-service traffic is given. Drops occur only in the lowest service class in each algorithm after 0.7 offered load. hcDBA gives the best performance while considering byte loss ratio.

 ISBN: 978-1-61208-103-8

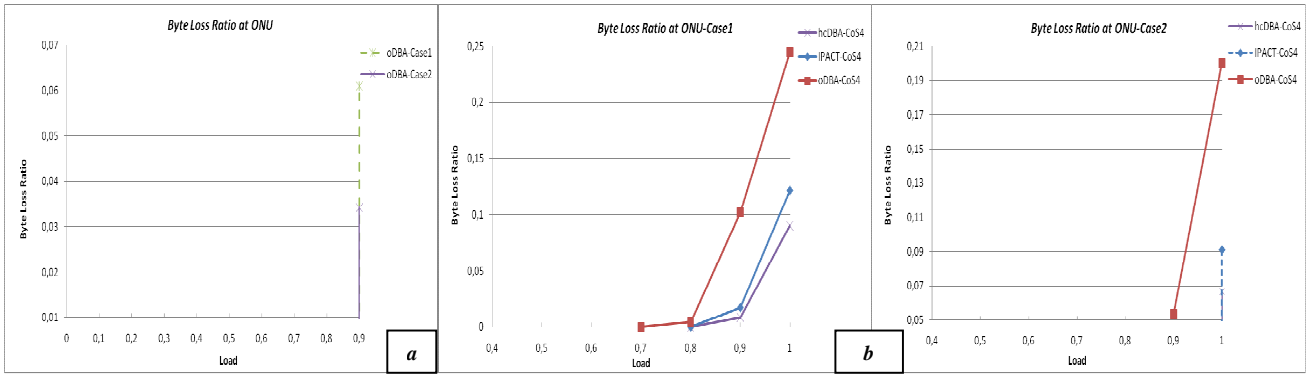


Figure 2. a) Byte Loss Ratio in Case1 and Case2 with mono-service class b) Byte Loss Ratio in Case1 and Case2 with multi-service classes

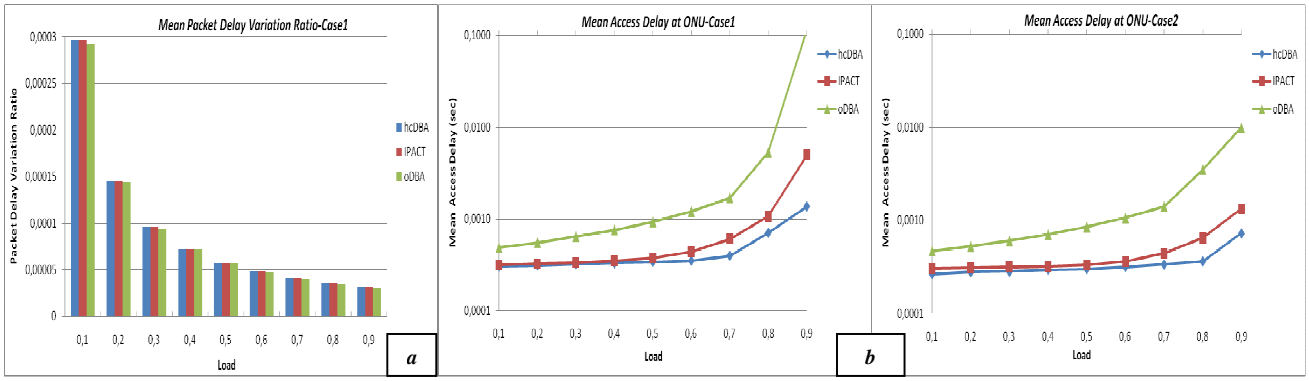
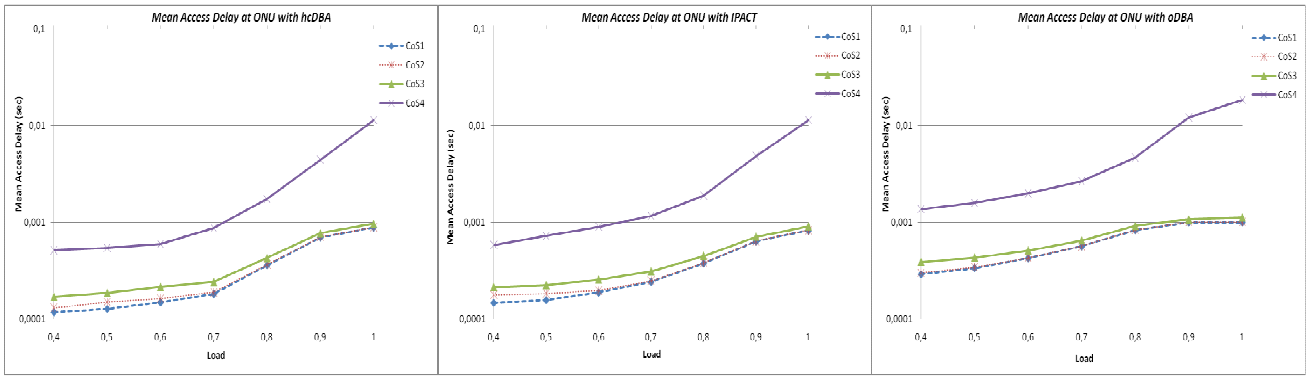
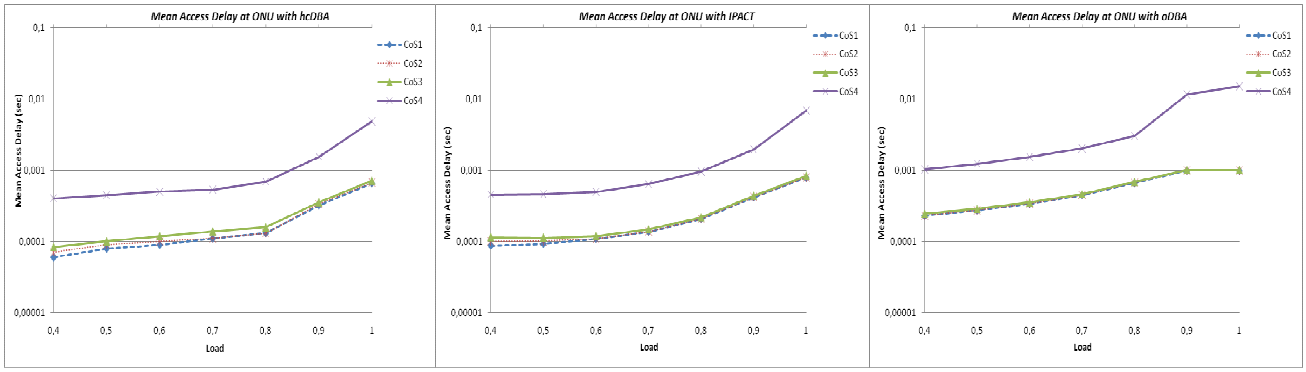


Figure 3. a) Packet Delay Variations in Case1 b) Mean Access Delay with Mono-Service Traffic



Case1



Case2

Figure 4. Mean Access Delay with Multi-Service Traffic

Figure 3a shows the PDV of each algorithm in Case1. From the figure it can be seen that all the algorithms have similar PDV values in the same loads. The used algorithm has no significant affect over PDV in PON network. Thus, Case2 results are not given.

Figure 3b shows mean access delay values for mono-service traffic of both cases. hcDBA and IPACT algorithms are better than oDBA. In all algorithms, the mean access delay is below 1ms in low loads. When the load increases, the mean access delays increase as expected. hcDBA gives better performance in both cases in terms of access delay. When the bandwidth rate is 10Gbps, hcDBA algorithm success increases as seen in Figure 4, and hcDBA algorithm gives mean access delay below 1ms at 0.9 load.

Figure 4 shows mean access delays in Case1 and Case2 with multi-service traffic. For each priority classes except CoS4 (Best Effort traffic) hcDBA and IPACT stays under 1ms in every offered load values. hcDBA algorithm is better in all cases of each service class. The lowest priority class is worst in all conditions. In simulations, we use Head of Line (HoL) priority scheduling at ONUs. Thus, lower priority traffic has to wait in buffers each time if there is not enough bandwidth has been given by the OLT. hcDBA gives better mean access delay results than IPACT because that it can give more bandwidth to the highly loaded ONU if there is excess bandwidth exists thanks to the low demands of other ONUs.

We also check fairness of the proposed algorithm among ONUs. When hcDBA works in offline mode, in each cycle time, ONUs are separated into two groups. We must be sure that the algorithm distributes the excess bandwidth fairly between two ONU groups. For this reason, in each cycle, we check the difference of EBD values of hcDBA algorithm with the EBD values if it distributes excess bandwidth as standard offline DBA. This difference ratio is 0 in low loads. Since, in low loads excess bandwidth is enough for all highly loaded nodes. There is only a small difference occurs among 0.7 to 0.9 offered loads (%1 at 0.9 load, %0.01 at 0.8 load and %0.0006 at 0.7 load).

V. CONCLUSION

In this paper, we have proposed a novel dynamic bandwidth allocation algorithm that stays between offline DBA and online DBA algorithms. Our first aim is to eliminate idle time problem in offline DBA algorithm while keeping fair EBD scheme of offline DBA algorithm. In hcDBA, we distribute the excess bandwidth in two half cycle and we also switch to online DBA mode according to the incoming traffic load in each cycle time. Besides, by simulation results, we evaluate the performance of the proposed algorithm with mono-service and multi-service traffic under two cases as 1Gbps and 10Gbps upstream channel bandwidth rates. The performance improvement is measured in terms of mean access delay and byte loss ratio.

We have compared our DBA algorithm with IPACT and oDBA algorithms. hcDBA shows better performance both in mean access delay and byte loss ratio values. The simulation studies provide that hcDBA is almost as fair as offline DBA and has better bandwidth utilization than IPACT algorithm.

Our algorithm's advantages compared to other algorithms proposed to eliminate idle time problem in offline DBA algorithm can be listed as;

- Uses the standard MPCP control messages defined in EPON standard and does not need any change in ONUs.
- Can be combined with different QoS approaches.
- Does not change the service order in polling table therefore, it does not cause additional PDV.
- While it eliminates idle time period in offline DBA, hcDBA does not need extra GATE/REPORT messages, as a result hcDBA introduce less overhead in upstream and downstream channels.

As a future work the hcDBA algorithm can be improved with addition of intra-ONU and inter-ONU quality of service approaches to obtain better results in multi-service environments.

ACKNOWLEDGMENT

This work is supported partially by the European EURO-NF project. This work is also a part of ongoing PhD thesis titled "Next Generation Optical Access Networks" at Istanbul University, Institute of Physical Sciences.

REFERANCES

[1] C. Lam, "Passive Optical Networks: Principles and Practice", Burlington, MA: Academic, 2007.

[2] F. Effenberger and T.S. El-Bawab, "Passive Optical Networks (PONs): Past, present, and future", Elsevier, Optical Switching and Networking 6, pp 143-150, 2009.

[3] F. Effenberger, D.Clearly, O. Haran, G.Kramer, R.D. Li, M. Oron, and T. Pfeiffer, "An introduction to PON technologies", IEEE Commun. Mag., vol.45, no.3, pp.s17–s25, Mar. 2007.

[4] IEEE 802.3ah task force webpage: http://www.ieee802.org/3/efm [accessed on February 2010]

[5] G. Krammer, "10G-EPON: Drivers, Challenges, and Solutions", Optical Communication, 2009. ECOC '09. 35th European Conference on , vol., no., pp.1-3, 20-24 Sept. 2009, Vienna Austria.

[6] G. Kramer, B. Mukherjee, and G. Pesavento, "IPACT: a dynamic protocol for an Ethernet PON (EPON)", IEEE Commun. Mag., 2002, 40, (2), pp. 74–80.

[7] C.M. Assi, Y.Ye, S. Dixit, and M.A. Ali, "Dynamic bandwidth allocation for quality-of-service over Ethernet PONs", IEEE Journal on Selected Areas in Communications, Vol. 21, No. 9, Nov. 2003, pp. 1467-1476.

[8] J. Zheng, "Efficient bandwidth allocation algorithm for Ethernet passive optical networks" Communications, IEE Proceedings- , vol.153, no.3, pp.464-468, 2 June 2006

[9] T.D. Nguyen, T. Eido, and T. Atmaca, "An Enhanced QoS-enabled dynamic bandwidth allocation mechanism for Ethernet PON", Emerging Network Intelligence, pp.135-140, 11-16 Oct. 2009

[10] H.J. Byun, J.M. Nho, and J.T. Lim, "Dynamic bandwidth allocation algorithm in Ethernet passive optical networks," Electronics Letters, vol. 39, no. 13, June 2003, pp. 1001–2.

[11] Y. Luo and N. Ansari, "Limited sharing with traffic prediction for dynamic bandwidth allocation and QoS provisioning over Ethernet passive optical networks," OSA J. Opt. Net., vol. 4, no. 9, Sept. 2005, pp. 561–72.

[12] W. Willinger, M.S. Taqqu, R. Sherman, and D.V Wilson, "Self-Similarity through High-Variability: statistical analysis of Ethernet LAN traffic at the source level", Networking, IEEE/ACM Transactions on , vol.5, no.1, pp.71-86, Feb 1997

 ISBN: 978-1-61208-103-8

Hot-Spot Blob Merging for Real-Time Image Segmentation for Privacy Protection

Florian Matusek
KiwiSecurity Software GmbH
Vienna, Austria
e-mail: matusek@kiwi-security.com

Abstract—**One of the major, difficult tasks in automated video surveillance is the segmentation of relevant objects in the scene. This is important for various tracking tasks. Especially in the emerging field of privacy protection in video surveillance systems it is imperative that objects are accurately separated and shadows removed. Current implementations often yield inconsistent results on average from frame to frame when trying to differentiate partly occluding objects. This paper presents an efficient block-based segmentation algorithm, which is capable of separating partly occluding objects and detecting shadows. It has been proven to perform in real-time with a maximum duration of 47.48 ms per frame (for 8x8 blocks on a 720x576 image) with a true positive rate of 89.2%. The flexible structure of the algorithm enables adaptations and improvements with little effort. Most of the parameters correspond to relative differences between quantities extracted from the image and should therefore not depend on scene and lighting conditions. Thus, our proposal is presenting a performance-oriented segmentation algorithm, which is applicable to all critical real-time scenarios.**

Keywords-image segmentation; privacy protection; region growing; blob analysis; occlusion; shadow detection; intelligent video surveillance.

I. INTRODUCTION

Image segmentation algorithms used to partition a digital image into multiple regions are essential in numerous applications including security relevant domains, medical imaging, face recognition, fingerprint recognition and machine vision. Also, a new field where such algorithms are employed is video surveillance with special features for automatic privacy protection. Privacy enhanced video surveillance works under the assumption that only a small portion of recorded people pose a threat to security. Therefore it is imperative to protect the identity and person relevant information of innocent persons as best as possible. Current state-of-the-art systems work on the whole image and use background models to identify changes in the scene. Those areas are then masked by applying a transformation function.

Several general-purpose segmentation algorithms have been developed. These algorithms can be divided into categories depending on the technique used for segmentation.

One category is based on density or color histograms where peaks and valleys in the histogram distribution are used to locate clusters [1]. Another method called region growing starts by so called seeds and iteratively grows regions by comparing all unallocated neighboring pixels to the seed value. Edge detection is used as well in this field as objects tend show strong differences in intensity at the region boundaries. Model based segmentation on the other hand works on the assumption that domain-relevant objects show - within some minor variations - a unique form of geometry. Apart from these techniques other algorithms have been proposed including level-set [2], graph partitioning [3], and watershed [4].

Despite introducing general-purpose algorithms, no optimal solution for image segmentation has been found. The algorithms heavily depend on domain knowledge and problem-specific optimizations. A combination of methods seems promising to minimize the inherent disadvantages of each algorithm. On the one hand, region growth suffers from the tendency to cover multiple overlapping objects whereas model based algorithms tend to become unreliable when unwanted artifacts, for example shadows and reflections, show up in the scene. Also the mean shift clustering algorithm [5], used to find the local maxima, is susceptible for covering multiple overlapping objects especially in crowded scenes [6]. Furthermore the mean shift results have been shown to be inconsistent in subsequent frames; this deteriorates post processing steps like object tracking. State-of-the-art automated video surveillance systems have increasing complexity to meet the high standards expected by today's customers, with respect to true- and false-positive rates, while being robust against environment changes. With the rapid increase of processing power of modern computers image analysis algorithms, which have been too complex can now be implemented in real time.

However when implementing the whole algorithmic pipeline needed for state-of-the-art automated video surveillance systems the resources for a single algorithm are still limited. Currently, available closed circuit television (CCTV) systems are based on traditional per pixel analysis, which limits either the image resolution or the count of operations per frame that can be performed. Recent publications on the topic on background modeling proposed pre-defined image regions (hence called blocks) to allow more complex analysis; the same methodology can be applied for image segmentation. Due to the reduction of resolution, the algorithm presented in this

 ISBN: 978-1-61208-103-8

work can be used in conjunction with a pixel-based segmentation algorithm (e.g., mean shift).

This work is structured as follows: Section II explains the motivation for developing the proposed algorithm. Section III introduces the proposed hotspot blob image segmentation algorithm, while Section IV shows how shadow cancellation can be performed with it. In Section V, test results are presented and Section VI gives an outlook to future work.

II. SELECTIVE PRIVACY PROTECTION

Privacy protection is a feature used in video surveillance systems to mask foreground areas of the image with the goal of protecting the privacy of people seen in the video. The first iterations of algorithms where masking whole image regions, such as desks of employees or entrances. This proves to be very ineffective as soon as persons start moving out of masked areas. The next step in privacy protection in video surveillance used algorithms that mask all movement in an image [7]. This has several disadvantages: First, all movement is masked, including background movement, shadows and highlights. Second, it is not possible to distinguish between persons. Thus, either all or no person in the image can be masked. If the original, unmasked, video material of a covert operation has to be used, e.g., in court, covert agents would be seen and their identity released. Many other such examples exist, including in office buildings where only unauthorized personnel should be unmasked, at security critical infrastructures where only VIPs should be masked and situations where only persons who triggered an alarm should be unmasked.

Selective privacy protection aims to remedy this shortcoming by using tracking and matching algorithms to identity persons in the video. Using this information authorized users can choose to unmask only offending persons without compromising the privacy of possible bystanders. This unmasking process can be implemented securely by using personal chip cards and asymmetric encryption. Selective privacy protection increases the complexity of "whole image privacy protection" schemes by adding domain problems relevant to object association, e.g., object occlusions. A possible solution is to use a half automatic tracking process in which the authorized user is asked for assistance if the association confidence is deemed too low.

Apart from association problems one of the main issues of automatic privacy protection are lighting changes and shadows. Both conditions occur frequently in outdoor scenarios resulting in large areas of the video being privacy protected. Accordingly, these considerations were part of the design process of the proposed algorithm.

III. HOTSPOT BLOB IMAGE SEGMENTATION

This work presents an image segmentation algorithm, which uses a block based method to reduce image resolution (while keeping all relevant information) and in turn down-scales the problem complexity and processing performance.

In this work, the term "block" is used in a very general way and stands for a certain image area. It can range from a single pixel to a square or even rectangular image part containing multiple pixels. The block size should be chosen to be significantly smaller than then the expected object size to have sufficient resolution for analysis and tracking. A block size of 8x8 pixels was found to be the optimal trade-off between loss of resolution and computing performance determined by empirical tests. Furthermore the block size is kept constant within the whole image and over time. This method can be integrated into background models (used to distinguish between fore- and background image areas) that also commonly use blocks to improve - in the same way - the performance of the foreground detection in complex scenes including lighting changes and/or moving objects in the background.

Each block is represented by the following data:

I: The index of the block. It holds the unique position within the image similar to an index in a one-dimensional array.

Sb: The state of the block. It influences the algorithmic behaviour and can change throughout the algorithm. Possible states can be seen in Figure 1.

Wb: The weight of the block corresponding to the integrated intensity within the block's area.

Ab: The covered image area in units of blocks. It starts with one and is incremented for blocks with certain states as the blob size increases.

Rb: The reference to another block. It can link one block to another block.

For the unprocessed image, a block starts out with Sb = unassigned, Wb = 0, Ab = 1, Rb = 'no reference'; this is called the pre-processing stage. Throughout the stages of the algorithm Sb can change to one of the following states: irrelevant, relevant, assigned, center, joined center and junction, where all states but relevant are possible final states (see Figure 1). When the background model designates a block as background, this block is no longer relevant for the algorithm and thus labeled as irrelevant, on the other hand if the background model flags the block as foreground it is tagged as relevant. Only relevant blocks are considered for further calculations and can either become center blocks if a certain amount of neighboring blocks has the correct state, which is an indication that the location of the block may be part of a new blob within the image, or assigned if the block is in close proximity to another block that belongs to a center. Furthermore blocks that connect areas of different assignments will be labeled as junction. Finally, the different parts connected by junctions can be bridged or separated due to certain rules derived from their characteristics forming a bigger blob or splitting blobs into smaller segments.

A. The algorithmic stages

The algorithm is performed in stages numbered from one to six (see Figure 1 and Figure 2). A pre-processing stage is also introduced, which resets any information contained in the blocks used in a previous frame. This allows the minimization of allocations, which improves the performance.

 ISBN: 978-1-61208-103-8

Stage 1: The algorithm starts by calculating the integral sum of intensities of all blocks (SoI) deemed relevant by

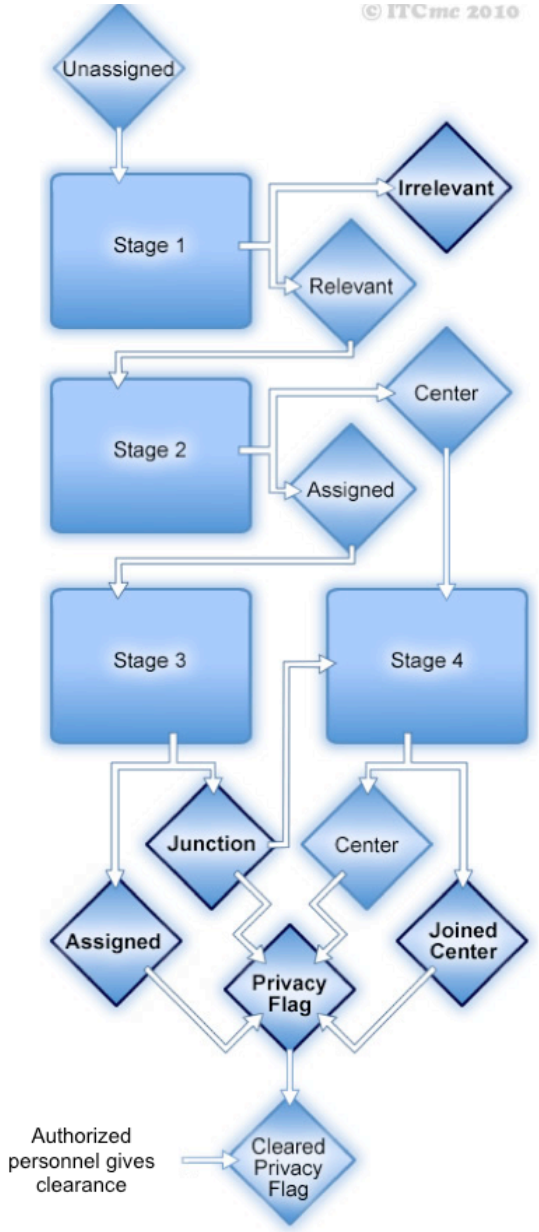


Figure 1: Possible algorithmic stages of the method.

the background model, places the SoI into the Wb variable for each block and builds a list for these blocks. The list is sorted by Wb, where the highest Wb is the first element the second the second highest and so on. If Wb is below a certain threshold tI a block is completely discarded and sets Sb = irrelevant, therefore the list only contains blocks with Sb = relevant.

Stage 2: For each block in the list, starting by the first, Sb is checked. If Sb ≠ relevant it means that the block has already been assigned to a center and doesn't need to be processed in this step. Otherwise the block is processed and all block states in the neighborhood are checked.

The neighborhood is a possible design parameter of the algorithm and can include only the adjacent blocks (as implemented in this work) or also blocks farther away. Depending on the implementation the algorithm does an iterative check of how many neighbors are found with Sb = relevant. In the first iteration it checks if it finds a block within the image where all neighbors are in relevant state. If this holds true the block is labeled as center (Sb = center) and all neighboring block states are changed to Sb = associated. Furthermore, Wb of every associated neighbor is added to the weight of the center block Wc.

If one or more blocks are found to already be associated the algorithm proceeds by finding all corresponding centers and associates the current block to the center with the highest Wc. This corresponds to setting Sb = associated, storing the center's address in Rb and add Wb of the current block to Wc.

Should the first iteration yield no centers at all (and therefore no associations as well) the number of neighbors needed to form a center decreases and the iterative search for centers and associations continues until all blocks are either center or associated.

Stage 3: After labeling and associating the blocks, possible borderlines (junctions) between the regions of different centers have to be found. The list containing the relevant blocks is traversed once more and all blocks that are in associated state and have one or more blocks with different center references in their neighborhood are marked as junction. To manage the weight of the junctions a junction object is introduced; it holds references Rj,1 , Rj,2 to two center blocks, a weight Wj and an area Aj. The junction objects are identified by the two references and stored in a list. If a block is part of a junction, the list of junctions is iterated to find the corresponding object. If no corresponding references are found in the list, a new junction is created with Wj = 0 and Aj = 0 and appended to the list. In either case, the weight and area of the current block is added to the values of the junction.

Stage 4: After finishing the search for blocks being part of a junction the list of junctions is sorted according to Wj. If the junction is found to be of relevance (e.g., by comparing to a threshold tJ or by analyzing the balance of weights of the two centers with respect to the junction weight) the centers shall be joined. In this case the state of the center with less weight (the weak center) is changed to joined center and the reference is updated to point to the second center (the strong center), which effectively merges the two centers in an efficient way. Now the final center of a blob can be found simply by traversing the center reference chain from any block until the reference doesn't change anymore.

Stage 5: A new object called "blob" is introduced, which essentially holds the relevant data of one segmented region within the image. A blob object consists of the following data:

Lblob: A list of blocks belonging to it (and sharing the same center).

Cblob: The final center of the blob.

 ISBN: 978-1-61208-103-8

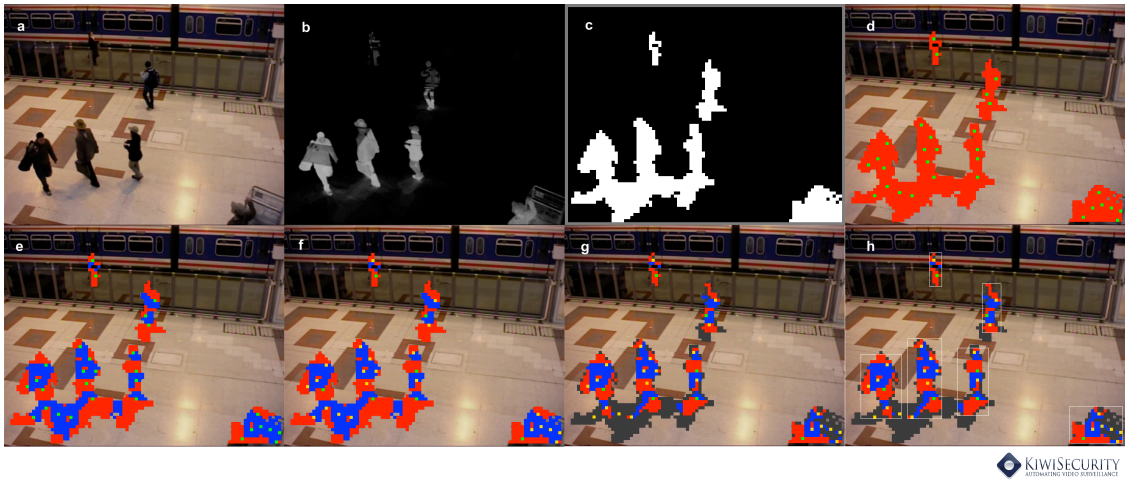


Figure 2: Block states for the different algorithmic stages: a, original image; b, differential image calculated by the background model; c, classification of blocks into relevant (white) and irrelevant (black) blocks; d, labeling centers (green) and associated (red) blocks; e, labeling junctions (blue); f, conversion of centers to joined centers (yellow); g, cancellation of shadow blocks (gray); h, the final bounding boxes of the objects.

BBblob: The coordinates of the final bounding box (left, right, lower, upper border).

Wblob: The total weight of the blob.

Ablob: The total area of the blob.

A last time the list of blocks is traversed to create one blob per center and store all associated blocks in the reference list.

Stage 6: Due to static occlusions within the scene or object parts with very similar color to the background image, an object can be split into two or more blobs. To avoid this unwanted behavior we implemented a simple model-fitting algorithm based on the shape of a human approximated by a rectangle. The dimensions of the model are manually calibrated at three distinct positions in the image and interpolated in between for every other position (barycentric interpolation). As the head (or top) regions of the objects are the most stable areas (generally fixed with respect to the object's center and mostly free from shadows) we sort the list of blobs according to their y-coordinate starting with the uppermost blobs (low y-coordinate). A rectangular shaped acceptance area is positioned with congruent upper border to the bounding box BBblob. Furthermore the acceptance area is placed horizontally with an offset to the center of BBblob. The offset depends on the perspective of the scene, which yields shear/rotation of the objects and the size of the acceptance area. The offset value is calibrated by hand at the left and right border of the scene and linearly interpolated between these positions.

If any other blob has ample overlap with the acceptance area, this blob is joined to the "accepting" blob and then deleted from the list of blobs. Ample overlap is given if kO percent of the blob's bounding box is within the acceptance area (kO = 50% was chosen in the current implementation).

Post-processing stage: Here a final filtering of the remaining blobs is performed. Currently two strategies are applied: Firstly the size of blobs being much smaller than the size of a human estimated by the model and secondly an approximated width:height ratio being larger than 1 can yield to the deletion of a blob. The more the computed ratio and size differs from the constraints the lower is the confidence rating; candidates with a confidence rating below a threshold tC are removed. After this stage the remaining blobs can be visualized with a rectangular outline.

The blocks remaining after the post-processing stage represent the relevant foreground areas. In the best case foreground areas are regions of movement including persons. Therefore all blocks are flagged as privacy enabled blocks and a transformation function t(x) is applied on the corresponding areas in the source image.

In the case of emergency authorized personnel may want to reverse the transformation function on a certain blob which represents a person. Depending on the implementation t(x) may be one or two way. In the case of a one-way function the original source image must be saved to restore the assigned portion of the image. In case of a two way function the inverse of t(x) can be applied on the blob to restore the image to the original state. In every following frame afterwards the user either associates the selected blob automatically by an association algorithm or manually. The unmasking process is then reapplied to the new frame.

IV. SHADOW CANCELLATION

One of the well-known problems common in the area of region growth techniques is the tendency to cover multiple independent objects. This characteristic is further enhanced in the case of inaccurate background images containing strong shadows. This behavior is also present in the proposed algorithm and became evident in the scenes tested; shadows were present in the computed difference image, which resulted in the merging of multiple persons to a single object. To cancel the perturbing shadows standard shadow cancellation algorithms including Horprasert et al. [8] were considered and tested but proved to provide minimal success. In general either too many areas where eliminated, which resulted in the deletion of complete, valid objects because of the different lighting

 ISBN: 978-1-61208-103-8

conditions present in different parts of the view, or too few shadows were removed depending on the parameters of the algorithm.

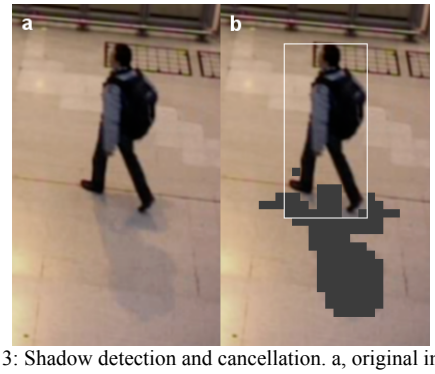


Figure 3: Shadow detection and cancellation. a, original image. b, detected shadow (gray blocks) and blob bounding box (white).

Following these considerations a new shadow detection based on the already existing block data was used. The block density:

$$d_b = W_b / A_b \ , \qquad (1)$$

where the area is measured in units of blocks, is compared to the density of the center

$$d_c = W_c / A_c \qquad (2)$$

of every block within a blob.

If

$$d_b > d_c \cdot k_d \qquad (3)$$

where kd is a constant factor (0.95 in this work), the block's coordinates are used to update the bounding box to accommodate this block.

The same applies, if the maximum intensity value within the block is higher than dc divided by the number of pixels of a block. This ensures that blocks, which hold small but bright details are not labeled as shadows (e.g., at object borders or within small objects). Figure 3 shows the effect of this mechanism. It significantly reduces the perturbing amount of shadows while keeping objects with a generally low density in the difference image. After processing all blocks in this way, the bounding box is defined for this blob. This procedure offers the advantage of using the already computed values also needed for the main algorithm, which results in an easy to implement and efficient way to detect shadows. Compared to the algorithm defined by Horprasert et al. using an YUV image, which needs about 12 ms on the test system the performance impact of this implementation is on average much less with approximately 1.5 ms and maximum 6 ms. It should be stated, that the mechanism can lead to unwanted results when the intensity of the object is generally lower in the difference image than the intensity of the shadow.

V. TEST SEQUENCE AND EVALUATION

To validate the blob-merging approach the PETS 2006 [9] sequence was chosen, as it is known to show a lot of typical situations in video surveillance including problems like shadows, reflections and occlusions. The annual PETS workshop is organized in conjunction with IEEE Computer Society Conference on Computer Vision and Pattern Recognition. It should be mentioned that there is currently no implementation of tracking, object association and occlusion handling. Thus the results cannot be directly compared to the officially available ground truth data. To provide a useful measure of the performance of the algorithm, each single frame was checked by hand for false positives. The checks were performed beginning with frame 349 (initialization of background model ended at this point) until frame 2224. After tuning the parameters 202 false positives were found in 1875 frames of the sequence. This corresponds to a true positive rate of 89.2% (see Table 1 for details). Most parameters of the algorithm are not depending on absolute quantities and thus should be relatively independent on the chosen test sequence for achieving best results.

The high number of 'object not found' errors is due to occlusions with static objects in the scene, which are in front of relevant objects and cover a large part of them.

The 'shadow interpreted as object' errors come from the constraint that low intensity objects are not removed from the scene, as this would lead to more 'object not found' errors. Therefore, all shadows that get separated from their originator and are big enough in size are interpreted as objects.

TABLE I. DETAILED DISTRIBUTION OF ERRORS IN THE PETS TEST SEQUENCE.

Error	Count
Object not found (too small)	89
Shadow interpreted as object	80
Split object	31
Object too large	2

The 'split object' errors arise from unwanted separations of junctions within an object (often due to low intensity areas in the difference image). On the contrary the 'object too large' errors originate from unwanted bridging to artifact objects.

Generally, it has to be mentioned that the obvious next stage in the algorithmic pipeline – the object associator, which essentially takes the history of objects into account – would eliminate a lot of the errors that have been found in the evaluation. For example, often shadow objects or other artifacts were only present for one single frame.

The algorithm needed a maximum computation time of 47.48 ms for about 1100 relevant blocks (8x8 pixels per block) present in the image with a resolution of 720x576. The computation time of the algorithm on whole sequence (3021 frames) was 5.063 s, which corresponds to 1.655 ms on average per frame. The tests were performed on a 2.13 GHz Intel Core 2 Duo machine with 1GB RAM.

Unfortunately, details about optimizations that bring faster performance and make it possible to perform the algorithm in real-time cannot be published in this work since they tough sensitive confidential information.

 ISBN: 978-1-61208-103-8

VI. PROPOSED ALGORITHM VARIATIONS AND OUTLOOK

Although the current implementation of the proposed algorithm already achieves good results, there are a lot of possibilities to further improve the capabilities, performance and computing time (generally, the algorithm has not yet been optimized). Besides code and performance optimization the following improvements are planned:

To reduce the 'object not found' error count, a self-learning static occlusion detection algorithm is planned to be implemented, which should inform if an object is touching an image area that is in front of it. Looking at the intensity histogram of an object might diminish the 'shadow interpreted as object' error count. Shadows tend to have very little structure and should have a very narrow distribution in the histogram. It is planned to substantially improve the model-fitting algorithm with a more complex shape, where different regions are weighted with different strength. In turn the shadow detection rules can be optimized; the block elimination threshold could be varied according to the chosen appearance model. Thus, a much stricter threshold value could be chosen until unwanted elimination sets in. Furthermore, it is planned to use an appearance model for defining the shape of the 'neighborhood', which is responsible for the positioning of centers in the image. In this way, centers should only be set within blobs with the right size and shape.

Possible variations due to the block-based nature of the algorithm: The size of the blocks can be used to meet the required frame rate. In scenes with high degree of object size variations due to perspective the block size could be changed for certain areas of the image (e.g., upper third with half-size blocks) to increase the resolution. An adaptive block size changing in time or depending on the dimensions of the blobs is also conceivable.

Moreover, completely different data might be stored within a block or center (e.g., textons [10], HOGs [11] etc.). The set of rules for merging of centers and their respective areas can be based on these other data or parameters.

Since in the pre-processing stage the information of a block in the previous frame is reset, information about the last frame is lost. In future work it will be looked at how to use this information in order to improve the algorithm.

VII. CONCLUSION

An efficient block-based segmentation algorithm has been presented being capable of separating partly occluding objects and detecting shadows. It has been proven to perform in real time with a maximum duration of 47.48 ms per frame (for 8x8 blocks on a 720x576 image) with a true positive rate of 89.2%. The flexible structure of the algorithm enables adaptations and improvements with little effort. Most of the parameters correspond to relative differences between quantities extracted from the image and should therefore not depend on scene and lighting conditions. A minimal amount of parameter tuning is required, which makes the configuration simple.

The characteristics of the proposed algorithm are indispensable for privacy enhancing features in video surveillance applications. Exploiting intrinsic shadow cancellation leads to a significant improvement in privacy protection of innocent persons without sacrificing performance or security, as seen in Section 4.

ACKNOWLEDGMENT

Discussions with Prof. O. Martikainen, Department of Information Processing Science. Oulu University, Finland, are highly acknowledged. This work was supported by Austria Economic Service „Austrian Wirtschaftsservice" www.awsg.at, Austrian Research Promotion Agency "Österreichische Forschungsförderungsgesellschaft" www.ffg.at and the Academic Business Incubator INiTS www.inits.at. This work will be partially included in a Ph.D thesis at the University of Oulu.

REFERENCES

[1] L. Shapiro and G. Stockman: "Computer Vision", New Jersey, Prentice-Hall, ISBN 0-13-030796-3, 2001.

[2] S. Osher and N. Paragios: "Geometric Level Set Methods in Imaging Vision and Graphics", Springer Verlag, ISBN 0387954880, 2003.

[3] J. Shi and J. Malik: "Normalized Cuts and Image Segmentation", IEEE Conference on Computer Vision and Pattern Recognition, pp 731-737, 1997.

[4] S. Beucher and F. Meyer "The morphological approach to segmentation: The watershed transformation". In: Dougherty ER, ed. Mathematical morphology in image processing.New York: Marcel Dekker, 1993.

[5] Y Cheng: "Mean Shift, Mode Seeking, and Clustering", IEEE Transactions on Pattern Analysis and Machine, 1995.

[6] C. Beleznai, B. Fruhstuck and H. Bischof: "Human detection in groups using a fast mean shift procedure", Image Processing, ICIP '04: International Conference, 2004.

[7] A. Cavallaro: "Adding Privacy Constraints to Video-Based Applications". In Proceedings of the European Workshop on the Integration of Knowledge, Semantics and Digital Media Technology, page 8, 2004.

[8] T. Horprasert, D. Harwood and L. Davis: "A statistical approach for Real-time Robust Background Subtraction and Shadow Detection", EEE ICCV'99 Frame-Rate Workshop, 1999.

[9] PETS 2006, Ninth IEEE International Workshop on Performance Evaluation of Tracking and Surveillance: http://www.pets2006.net, 2006, last accessed: 2010/10/08

[10] T Leung and J Malik: "Representing and Recognizing the Visual Appearance of Materials using Three-dimensional Textons", International Journal of Computer Vision, 2001.

[1] N Dalai, B Triggs, I Rhone-Alps and F Montbonnot: "Histograms of oriented gradients for human detection", Computer Vision and Pattern Recognition, 2005.

 ISBN: 978-1-61208-103-8

Performance Enhancement: An Advanced Nearly Indestructible Video Surveillance System

Stephan Sutor
KiwiSecurity Software GmbH
1050 Vienna, Austria

Abstract— This paper presents a novel architecture for high performance, scalable video surveillance systems. The goal of this architecture concept is the creation of an advanced, virtually indestructible video surveillance system with the highest possible performance level, even under worst-case circumstances. The security management process used to build up such a system is introduced and analyzed, with the focus on how to achieve the highest possible security performance level. The presented system was subject to several test methods and phase levels to evaluate system performance. From the end user's point of view, all the achieved results verified the high performance and nearly indestructible characteristics of the system.

Keywords— Video Surveillance; Performance; Availability; Reliability; Security Management Process; Physical Security; System Architecture; Quality of Service.

I. INTRODUCTION

During the last few years, the world was confronted with several security tragedies, bank robberies and art crimes. Accordingly, the advance in research and development of security systems, especially video surveillance systems, has reached a remarkable progress worldwide. However, this progress has in turn urged organized crime and underground organizations to develop corresponding advanced technologies. Recently masked robbers brandishing handguns succeeded to steal four 19th century masterpieces by van Gogh & Monet from a Zurich museum in broad daylight [1]. This incident demonstrates the lack of security management processes for many systems.

II. AVAILABLE STATE-OF-THE-ART VIDEO SURVEILLANCE SYSTEMS

Video surveillance (hereafter VS) products currently available on the market can be segmented into four categories:

- Household surveillance products
- Commercial small scale products
- Commercial large scale systems
- High Performance Surveillance Systems/Solutions (HPSS)

The HPSS category is providing different levels of availability, reliability, integrity and performance measures in a scalable configuration.

This paper presents a new architecture of video surveillance systems, with extremely high level of robustness and performance; subsequently the configuration and technological aspects are presented. Accordingly a new category is presented: "Advanced Nearly Indestructible Video Surveillance System: NIVSS"

III. NIVSS SECURITY MANAGEMENT PROCESS

In order to design and develop a large-scale VS system, which should measure up to the NIVSS standard, security has to be part of the planning stage. ISO 27001 [2] provides a general blueprint for such a security management process. This section gives a first idea of how this process needs to be customized to be specifically valid for NIVSS systems. The topics to fill this blueprint with can be taken from ISO 17799 [3], which is considered the best practice collection of activities for security management. Fig.1 shows the selected topics from ISO 17799, which will be relevant for achieving an NIVSS standard in the future.

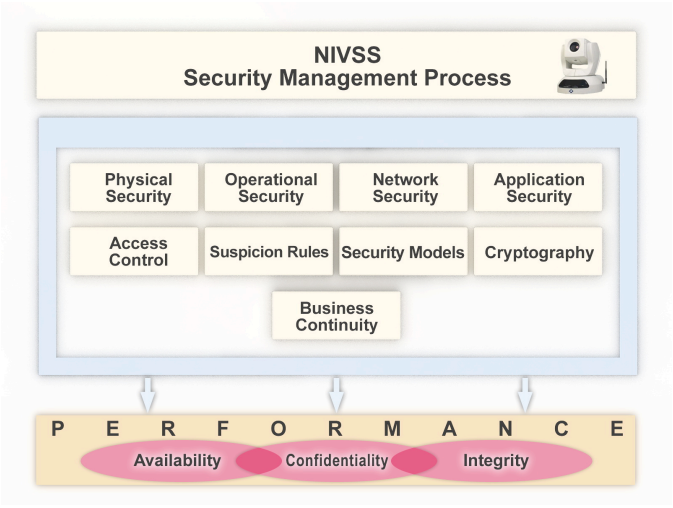


Figure 1: Security Management Process Components, relevant for the NIVSS

According to a recent study [4] the building blocks of the VS security management process are the following:

Access Control takes care of all situations, where assets need to be accessed and can possibly be manipulated (modified or deleted). Here identification methods and technologies for authentication and authorization have to be evaluated for aptness; policies for the implementation have to be developed. The scope ranges from user management aspects for the users accessing the database to the privileges under which an application process can access the operating

 ISBN: 978-1-61208-103-8

system's kernel processes.

Modern advanced security models deal with the formalized description of how information flows within a system in a secure way. This includes the states in which a secure system should be. In this respect the selected security model will also have a distinct influence on access control [5].

Physical security in general is concerned with any kind of physical influence on the system. For the proposed architecture this will be a major concern, as it will be physically distributed and the locations of the cameras mainly a given fact. So the planning of measures against natural disasters as well as intentional assault will be a challenge. Access control of the data center on a physical level, of the server farms hosting the applications and of the transmission network is also a topic warranting utmost attention.

Network security is concerned with the design of secure data transmission systems. This encompasses the development of a network architecture, which allows the separation of different types of networks (data, operations, security management, etc.), that provide secure areas (for the database) and demilitarized zones for data access of the various applications. It also deals with the lock down of the various network elements so that they do not provide any more functionality than absolutely necessary and only to those users (human or processes) who are allowed to access this functionality [6].

Cryptography deals with privacy aspects of transported or stored information. This is especially important as not only security critical data, but also personal data is sent via networks (possibly passing insecure channels). On the one hand the data must not be intercepted, on the other hand data must not be tampered with. Methods of how to achieve this and which cryptographic algorithms to choose are also a major concern of the VS security management process (e.g., if video surveillance is used across geographically distributed systems).

Application Security is dealing with the questions pertaining to the secure development installation and maintenance of application SW. Already in the planning phase detailed policies need to be worked out, which describe the way the applications have to be implemented.

Operation Security is concerned with all details, which concern the actual setup, execution and decommissioning of the system. The management process here needs to ensure that all processes are well described, operational security measures broken down to procedures, and operational descriptions for all parts of the system are present.

Business continuity is a topic, which is concerned with the high availability of the system. In the case of a surveillance system this is a number of policies for the implementation of software (e.g., backup & restore system), of redundant hardware (e.g., a complete offsite backup facility) and procedures to follow (e.g., definition of first response team and its actions) in case of disruptive events. Evidently business continuity will also play an important role in network security and application security, where secure and continuous operations of networks and applications are being planned.

IV. THREATS AGAINST THE NIVSS SYSTEM

In order to define the NIVSS and its protection mechanisms, all types of threats against VS systems have to be investigated. These threats are divided into two main categories: The first category is the threat of content manipulation, i.e., forging or removing of data, or adding useless or misleading data. The second category is based on system disruption, which ranges from simple attacks like destruction of cameras and destroying cables to more sophisticated ones like network or software based attacks.

These two types of threats are discussed for each main building block of the NIVSS security management process as follows:

A. Content manipulation

Content manipulation is defined as tampering and manipulation of data in the system, ranging from cameras to the storage of the video streams. In this section the threat of content manipulation is discussed for each aspect of the security management process.

- As tampering with data is categorized as part of network security, content manipulation on a physical level is not possible.
- Social engineering, using interactions to obtain confidential information, is considered as the greatest operational security threat for a VS system.
- Network Security: An attack on data transferred between one of the cameras and the final stage of processing, including attacks on analog signals and digital TCP/IP streams. When the attacker gets in a position to observe and intercept data streaming, this is called man-in-the-middle attack. This would allow the forging of video data by playing pre-recorded video streams or by editing the actual video delivered by the camera.
- Application Security: The goal of application security is to gain control over the VS system processing the pre-processed information or the machine storing the data via standard attacks like Trojans, worms, buffer overflows or exploiting backdoors, which causes a denial of service [7].

B. System disruption

The four categories of threats in this case are:

 ISBN: 978-1-61208-103-8

- Physical Security: An attack against any hardware component of the VC system e.g., camera, processing unit, network element or power supply. This type is considered to be the most severe attack on the VC.

- Operational Security: During updates wrong configuration data could be introduced into the system and could cause disruption.

- Network Security: System disruption attacks on network security are most likely in the periphery and user interface blocks of the NIVSS.

- Application Security: In this attack hackers try to get control over machines to manipulate content in order to achieve a system disruption. Furthermore, the traditional method of flooding the network can be used to achieve DoS.

- The architecture of the current NIVSS system was designed to withstand any of the above-mentioned attacks.

V. NIVSS: SYSTEM ARCHITECTURE

To achieve maximum availability, optimal reliability and best performance, the NIVSS was designed according to the following boundary conditions [8]:

- Redundant components (e.g., hardware): There must not be any part of the network without another component that can take over its function automatically. That means that the whole system has no single point of failure.

- Redundant networks (including wired and wireless networks as well as network components such as switches and routers): Each part of the system has to be reached through an alternative route. Backup networks have to be in place to take over networks which are down or overloaded.

- Tampering protection (software mechanisms to prevent tampering of cameras, network components and server hardware)

- Scalability (both in terms of video analytics as well as network):
  The system has to be scalable up to 100.000 cameras and more, interconnecting multiple video surveillance sites.

The system is divided into four hierarchical zones:

- Periphery
- Data Processing Center
- Demilitarized Zone (DMZ) and
- User Interfaces

Each of the four zones is subject to high security conditions in addition to the overall security concept of the system.

Periphery

As shown in Fig. 2, the periphery zone includes:

- Smart cameras:
  These cameras are analyzing the video image on their internal hardware and send results of this analysis to the data processing center. They track persons and vehicles, classify those and perform other tasks which are computationally intensive and which need to be done on the video images. This way the system stays scalable without sacrificing analysis quality. In addition to video analysis the smart cameras also encrypt and sign video streams to prevent unauthorized video streams from being injected.

- Cameras with image processing units:
  These are cameras (IP or analog), which are connected to an image-processing unit, which fulfills the same tasks as smart camera. With these units, existing cameras can be used in the system.

- Miscellaneous sensors:
  Various sensors (such as audio) help to identify events and tampering of cameras. Using multiple kinds of sensors, in addition to video, increases the security of the system considerably. If an intruder manages to tamper with a camera, injecting a video feed, simultaneous injection of audio material is more unlikely. Thus, if the system detects that sound does not correspond to video anymore, a tampering alarm can be triggered.

- Wired and wireless networks:
  Video and audio sensors are connected to the network with CAT 6 cables and wireless LAN, thus creating a redundant network. If one fails, the other one is still delivering data.

The NIVVS architecture is illustrated in Figure 2.

 ISBN: 978-1-61208-103-8

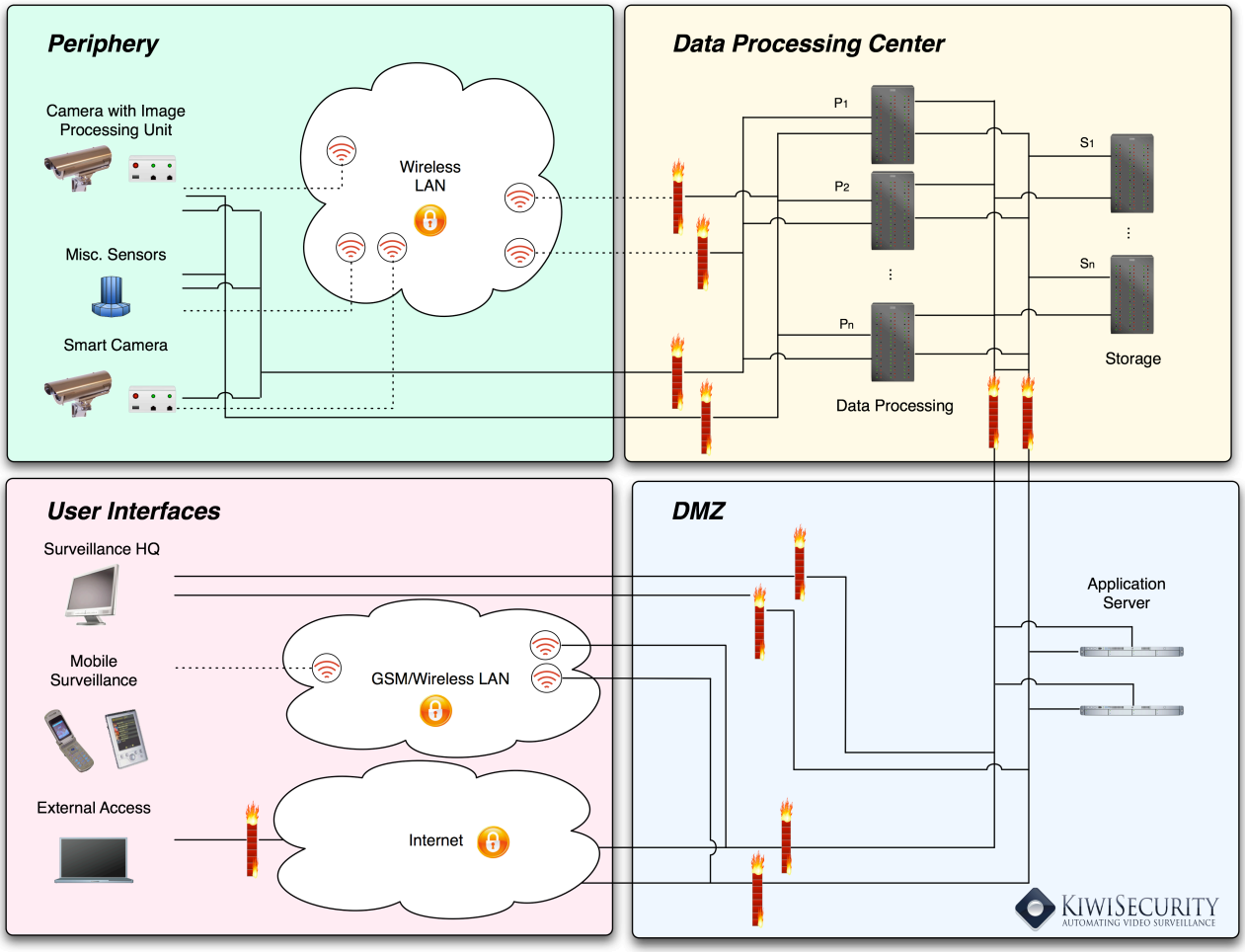


Figure 2: System architecture of the NIVSS system

Data processing center

Video information, including meta-information from image processing, is entering the "core zone" of the security system, the server network, through redundant firewalls with state full advanced package inspection and special intrusion detection systems. Just encrypted and correctly signed packages are allowed to pass. Both processing and the storage on the respective server hardware happen in parallel in consistency-checked self-healing clusters.

Demilitarized Zone (DMZ)

The collected and processed information is provided to the security guards through application servers. These servers are positioned in an internal Demilitarized Zone (DMZ) just beside each other because there is no need for them to be clustered. To protect the inner and core server nodes, solely reading rights on selected data are granted for these application servers. The network outside the DMZ is a virtual network with no machines connected directly to it; even though it should be physically built of multiple firewalls for security and reliability reasons. The only function of this network is to be the endpoint of the secured tunnel connections from the client applications and application networks.

User Interface

Relevant information from the NIVSS is presented to the user on various interfaces, including terminals, PCs and mobile devices (such as PDAs), which are used by security guards on patrol to get live alarms instantaneously to react to security breaches. In addition to the interfaces for the end user, operation and maintenance personnel have the right to define suspicion rules and conditions to be detected by the system.

Redundancy in the system

At the camera level power and the network connections have to be redundant. To guarantee this with analog cameras is anything but trivial. To reach this in digital systems based on computer network protocols is more simple. For power redundancy, power over Ethernet (PoE) can be used aside local power and buffer batteries. Parallel networks can be based on different media from copper wires over different optical to radio communication. Especially for a camera the cooperation of radio networks and wired connections, that can be attacked only in completely different ways, is essential. Wireless communication is another advantage of digital systems. There, protection against tapping is a must on any

 ISBN: 978-1-61208-103-8

network leaving a single secure room: Here, multi-tier encryption (application and network level) and signing provide protection must be implemented. The same applies to any communication channel [9-10] connecting nodes that are not servers in the same secure server room and physical redundant networks within this room.

Firewalls

The use of firewalls and separate networks for such a system is essential. However, firewalls should not be the only protection in the communication for clients. Furthermore, a security system must not only ever be connected to highly insecure networks like the internet or 'normal' corporate networks, but also an application server in an internal DMZ is required. This server has the most limited reading rights and connection rights in the primary server network. The processing servers and the DBMS servers have to be clustered as a consistency checked self-healing system. A well designed, maintained backup system is necessary, as well as an indestructible power supply including UPS, emergency power, over voltage and short circuit protection.

VI. VANDALISM, SABOTAGE AND NATURAL CATASTROPHIC SCENARIOS

A system level failure is defined as the status of having a single non observed point or no alarm upon detecting a suspected object/person/behavior. Following scenarios of attack, destruction or vandalism against the system are discussed taking into account the system structure.

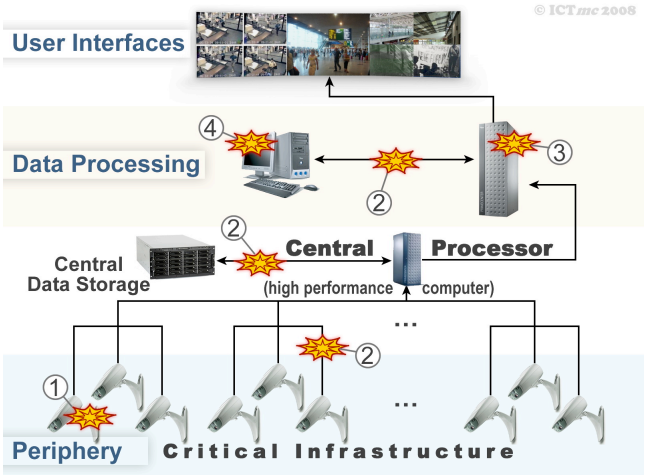


Figure 3: Attacks on a Single Site NIVSS

The periphery zone involves all the equipment outside the secured processing center. This includes cameras, image processing units attached directly to cameras and all networking connections and devices. Fig. 3 shows the infrastructure of a single site surveillance system, where all cameras and equipments are in the same VS centre, e.g., the main railway station, with the VS operation center directly connected to the central security authority or police station. Similarly, Fig. 4 illustrates attack attempts on a multiple site VS system, e.g., a large-scale airport, with several terminals and scattered infrastructure. Each of them is equipped with a local VS system, which is interconnected to the central VS operation & control room.

Four attack points or points of failure were identified. In attack-point 1 one of the cameras is destroyed or damaged. The system automatically detects such an event and causes an alert to be triggered and the task of the damaged camera is immediately covered either by the redundant camera or by one in the vicinity. In attack-point 2 the network connection is interrupted: The camera now uses full wireless transmission, and an alert describing the network failure is being sent.

In attack-point 3, the wireless network is jammed. However, the system can still rely on the cable connections and an alarm will be set off. If the attacker starts with the wired network, the system will switch into wireless mode and trigger an alarm immediately; hence the attacker will be stopped.

Attack-point 4 resembles an attack on one of the data processing servers or a storage server. Both these events will not affect the entire system, because each component is redundant, and the system will automatically distribute the load of the failed unit to the others. Measures of the required redundancy will be discussed in future work.

Even in case of disaster or natural catastrophes, where the whole surveillance site is under attack, an alert would be reported immediately to the authorities.

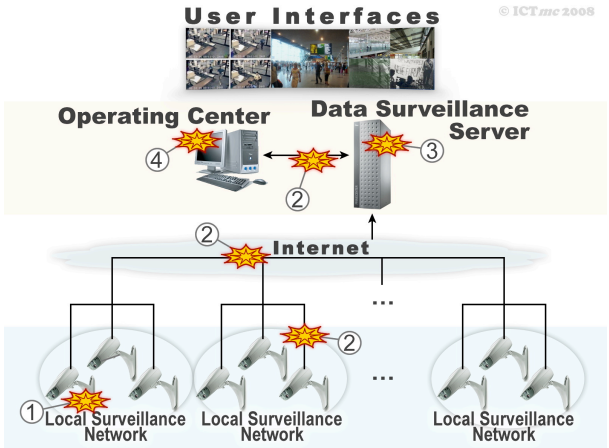


Figure 4: Attacks on a Multiple Site NIVSS Infrastructure.

VII. CONCLUSION

This paper presented a new security management process used to build a nearly indestructible, high-performance video surveillance system. Subsequently, a novel architecture for such a system was presented and analyzed. Accordingly, this video surveillance system was tested, validated and verified while running in real time at the highest possible performance level, even under worst-case circumstances, identifying different types of threats and corresponding counter-measures.

 ISBN: 978-1-61208-103-8

A patent application was filed for the NIVSS system architecture.

ACKNOWLEDGMENT

Discussions with Prof. O. Martikainen, Department of Information Processing Science at the University of Oulu, Finland, are highly acknowledged.

This work was further supported by the Austrian Research Promotion Agency "Österreichische Forschungsförderungs-gesellschaft" www.ffg.at and the Academic Business Incubator INiTS, www.inits.at. This work will be partially included in a Ph.D. thesis at the University of Oulu, Finland.

REFERENCES

[1] L. Moran, www.mainstreet.com, "Challenges, Dealing with being a Crime Victim, 12.02.2008

[2] "ISO/IEC FDIS 27001 "Information technology — Security techniques Information security management systems — Requirements" Final Draft 200

[3] ISO/IEC 17799:2005, Information technology — Security techniques — Code of practice for information security management.

[4] K. Kraus, Security Management Process in Distributed, Large Scale High Performance Systems", Proceedings of the World Congress on Power and Energy Engineering, WCPEE'10, Alexandria, Egypt, October 3-7, 2010

[5] B. Howard, O. J. Paridaens and B. J. Gamm. "Information security: threats and protection mechanisms", Alcatel Telecommunications Review, pp. 117-121, 2001

[6] J. Gonzalez, V. Paxson, and N. Weaver, Shunting, "A Hardware/ Software Architecture for Flexible, High-Performance Network Intrusion Prevention", Proceedings of ACM CCS, October 2007

[7] J. Bellardo and S. Savage. "802.11 Denial-of-Service Attacks: Real Vulnerabilities and Practical Solutions", Proceedings of the 12th conference on USENIX Security Symposium, pp 2-2, 2003

[8] F. Matusek, S. Sutor, F. Kruse, K. Kraus and R. Reda, "Large-Scale Video Surveillance Systems: New Performance Parameters and Metrics", The Third International Conference on Internet Monitoring and Protection, Bucharest June 29 - July 5, 2008

[9] R. Reda and N. Jordan, "Signpots for the Future of Mobile Communication" , e&I, elektronik und informationstechnik, heft 9.2006,published by Springer Wien New York, ISSN: 0932-383X EIEIEE 123(9) 361-408, a1-a44(2006)

[10] R. Reda and H. Volopich, Siemens AG Austria, (2003), e-Business Evolution, Market Trends, and Business Opportunities: Keynotes presented at the International Conference on e-Commerce, Hong Kong 2002

 ISBN: 978-1-61208-103-8

Past-based Search Function in Pastry

Wang-Cheol Song
Department of Computer Engineering, RIAT
Jeju National University
Jeju, South Korea
philo@jejunu.ac.kr

Seung-Chan Lee
Department of Computer Engineering
Jeju National University
Jeju, South Korea
this.dreamzinn@gmail.com

***Abstract*—The P2P technology is very popular to share several kinds of contents such as mp3 and video in the Internet. Recently the structured P2P has been studied widely, and there are many P2P algorithms such as Pastry, CAN, Chord, Tapestry and so on based on the Distributed Hash Table (DHT). As these structured P2P provides the context-aware routing, content can be shared among user-peers without help of any servers. However, the general users are used to utilize the keyword-based search engine in the Internet for searching content, and the structured P2P supports only the exact matching. Hence, a keyword-based searching algorithm is required. In this paper we propose a search function in Pastry – one of the structured P2Ps. The search function is designed using PAST which is the file storage system of Pastry.**

Keywords-Search Function; Structured P2P; Pastry; PAST; DHT

I. INTRODUCTION

The P2P technology is very popular to share several kinds of contents such as mp3 and video in the Internet. From year 2000 various structured P2P technologies have been announced to access the contents without the server [1-3]. Most of them are based on the Distributed Hash Table (DHT), so that if the name of the contents is known, the location of the contents can be found without help of any servers.

When we want to use services in the Internet, we are used to find something using a couple of keywords through search engine in the web. Because of that, as you know, there are many search engines in the Internet nowadays. So, some people may guess the structured P2P does not need the search function because if we only know the name of contents, we can be routed to a location of the service that we want. That is one of the advantages of the structured P2P over the unstructured P2P. But, the general users are not familiar with such way. That is, users generally do not know the exact name of contents they want. They ordinarily know a couple of keywords or remember only some parts of the name.

As the structured P2P operates with no servers, the user-peers cannot ask to search something to a search engine. With the exact name of the content, they just trust the structured P2P to take users to where the content is without searching operation. In other words, if you do not know the exact name, you never can go to the node having the content. If we want to provide the search function to users, the existing server based search algorithms to provide location information of content cannot be applied because the structured P2P assumes no server. Therefore, as the nodes of a peer-to-peer network cannot rely on a central server coordinating the exchange of content location, they are required to actively participate by independently and unilaterally performing tasks such as searching for other nodes, locating content.

There have been recent works related with the search function in P2P networks. [11] introduces several structured P2P technologies and describes required functions and issues for them. [4] outlines a research agenda for building complex query facilities on top of the DHT-based P2P systems. Lundgren et al. [5] propose a search engine called SCAN on Pastry. [6, 7] says P2P keyword searching based on DHT. Also, there are other approaches to design new DHTs for peer to be routed to the content without searching function [12] [13]. Although these works verify need of the search function in the structured P2P network, no one has proposed the keyword-based search function without the server.

Pastry [2] used in this paper is a structured P2P overlay network as the location and routing substrate that is efficient, scalable, fault resilient, and self organizing. It represents a second generation of peer-to-peer routing and location schemes along with Tapestry [8], Chord [1] and CAN [3]. Pastry assigns the unique identifier from a circular 128-bit namespace to every node and every object as a nodeId and key, respectively. When a message and a key are given, the message can be efficiently routed to the node with the nodeId numerically closest to the key. It guarantees a definite answer to a query in a bounded number of network hops. Also, Pastry assumes an existing infrastructure at the network layer, and the emphasis is on self-organization and the integration of content location and routing. In the viewpoint of scalability, it can be noticed that Pastry nodes only use local information. The global information exchange in the routing algorithms limits the scalability, necessitating a hierarchical routing architecture like the one used in the Internet. In addition, Pastry achieves network locality so that the entries in the routing table of each Pastry node are chosen to be close to the present node according to the proximity metric. With these merits, we have chosen Pastry as the platform to develop P2P based applications in the further research.

PAST [9] is an application of Pastry as the P2P storage system. Replicas of an object are stored at the k nodes whose

 ISBN: 978-1-61208-103-8

nodeIds are the numerically closest to the object's key as in Figure 1. PAST also maintains the invariant that the object is replicated on k nodes, regardless of node addition or failure. Since nodeId assignment is random, these *k* nodes are unlikely to suffer correlated failures.

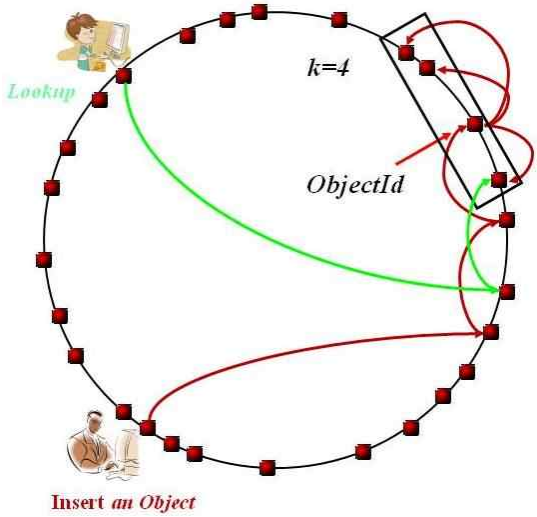


Figure 1. PAST operation in the Pastry ring

In order to develop the search function in Pastry, first we need to summarize the characteristics of the DHT in the aspect of search function.

- When the id of an object is known, we can access the node having the object directly and get it. That is, Pastry is content-addressable. It is because Pastry uses a name space for both nodes and objects and the object is stored in a node with nodeId identical to the object id.
- The id is uniquely generated by the hash function. As a similar keyword does not generate a similar hash code, an object' id is independent from id of other objects with the similar name.
- Pastry supports DHT based exact matched searching, but provides no facilities to search objects or nodes with related keywords.

We can easily imagine that when a user uses network based applications, he or she may want to begin with searching some contents by related keywords. As the exact content names are not usually known in most of cases and some people may want to know a list of more things than the exact content as what they want, keyword-based search function must be supported, although Pastry provides the content-addressability.

The rest of this paper is organized as follows. Section 2 describes the related work. Section 3 presents the design of search function in Pastry and Section 4 briefly explain it with a scenario. Section 5 shows performance evaluation and we conclude in Section6.

II. RELATED WORK

There have been recent works for search function in P2P networks. A system to support complex queries over structured peer-to-peer systems is proposed in [4]. This approach relies on the underlying peer-to-peer system for both indexing and routing, and implements a parallel query processing layer on top of it.

[5] proposes a search engine called SCAN to effectively perform distributed user lookup based on Pastry. This paper has the same approach as the keyword-based search engine, but it encodes Content meta-data e.g., keywords using ASCII table, into a set of Pastry keys for NodeIds that are inserted into the network. [6, 7] says peer-to-peer keyword searching and are based on DHT. But they propose the mechanisms for collaboration among peers as the search engine servers sharing indices.

[13] is the searching algorithm for a structured P2P - CAN in [3]. It proposes a searching algorithm named Recursive Partitioning Search (RPS) and develops the DHT properly modified for the searching function. This approach is intended to build the modified DHT and perform the blind searching. [12] proposes a peer-to-peer information retrieval system to support content- and semantic-based full-text searches. It also modifies the DHT of CAN, and presents resources and queries as vectors so that documents in the network are organized around their vector representations using a ranking algorithm. [14] is a extension from [12] to use a two-phase distributed semantic indexing method. [15] also proposes to modify the DHT, but it takes the Ontology approach.

Compared to the above works, we intend to design a keyword-based searching function. Usually such a search function is supposed to need the server operation, but we are sure that if we utilize PAST - the P2P storage of Pastry we can develop an efficient and scalable searching function specified to Pastry.

III. DESIGN OF SEARCH FUNCTION IN PASTRY

The structured P2P uses the DHT to find the requested contents by using the content's exact name without servers such as DNS. It supports only the exact matching. It is because as the DHT is based on the hash function, similar keywords produce entirely different results. However, users usually want to search what they want by using a couple of interesting keywords and get the list of the available contents.

We have designed the search function in the Pastry. We assume that every Pastry object has one or more keywords and the owner of the object register the keywords when he joins the Pastry ring. A couple of points should be considered for design of the search function as follows: As the node id and the object id are hashed in the Pastry, every id is independent each other. So, object ids hashed with keywords related to the object must be independent from an object id hashed with object's name. But the general users want to search objects by using one or more related keywords. Secondly, we should find where the owner of the object can store the keywords. This question arises because we do not want to lean on any server storage.

In this paper we have resolved them by using the P2P storage – the PAST. When an object joins the Pastry ring, it stores every keyword in the PAST as follows: *n* hashed keywords and the object's id are stored as *n* pairs of the

 ISBN: 978-1-61208-103-8

content's names and the content itself like (h_1, N), (h_2, N) … (h_n, N) through a PAST command – **Insert** as in Figure 2. Then, when a peer wants to find the object by one or more hashed keywords, he can try to find the object's id through a PAST command like **Lookup**(h_3) and get the object's id, N. As we can expect, the more popular the keyword is, the more objects' ids the **Lookup** function returns. Then, this search function returns a list of objects and the most overlapped ids in the **Lookup** results are ahead listed up.

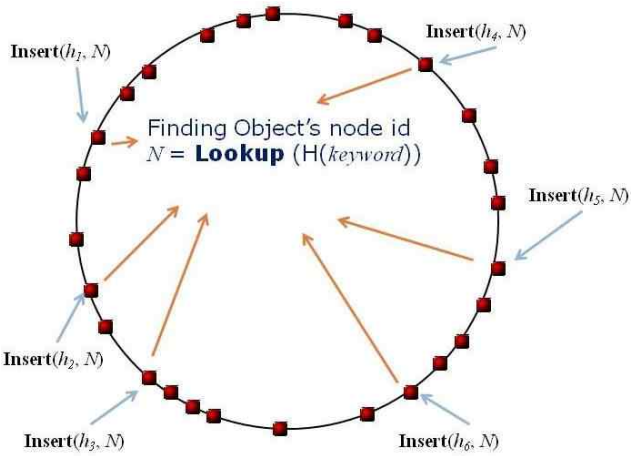


Figure 2. Storing and Searching keywords in PAST

PAST itself provides the robust P2P storage system by storing the replicas in the leaf set nodes. And, the objects stored in PAST are maintained as replicated on k nodes, regardless of node addition or failure. Therefore, the proposed search function based on the PAST is expected to be robust, too.

IV. A SIMPLE SCENARIO

If we think a *rose* as a Pastry content, it may have some keywords such as *flower*, *red* and *thorn*. And, if codes of the keywords – *flower*, *red* and *thorn* can be hashed as h_1, h_2 and h_3, the pairs of (h_1, *rose*), (h_2, *rose*) and (h_3, *rose*) can be stored using PAST command – **Insert**. Also, for keywords - *star*, *red*, *energy* and *center* hashed as h_4, h_2, h_5, h_6 in case of *sun*, the pairs of (h_4, *sun*), (h_2, *sun*), (h_5, *sun*) and (h_6, *sun*) can be stored in PAST.

Then, when a user want search an object with two keywords – *flower* and *red*, by using the hashed id h_1 of *flower* and the hashed id h_2 of *red* we can search the object like **Lookup**(h_1) and **Lookup**(h_2). Two lookup commands may return results as (*rose*) and (*rose*, *sun*) respectively. From these results the user can get a list as the search result in a way that the most overlapped object is first displayed. So, the user gets *rose* as the most related object.

V. PERFORMANCE EVALUATION

As the proposed search function is based on PAST, we think it could be robust and scalable like the characteristics of PAST. In order to evaluate the performance, we have simulated it on the open source code of the FreePastry [10].

For the simulation environment, we have run the FreePastry version 2.1 with JAVA JDK 6 on a Microsoft Window XP machine. The FreePastry is implementation of the Pastry as well as PAST. It could operate on the real network, but we have done the simulation on the Network Simulator of the FreePastry. We have selected the EuclideanNetwork as the network topology in the simulator. As we use the virtual clock in this simulation environment, we assume there is neither error nor delay in the network.

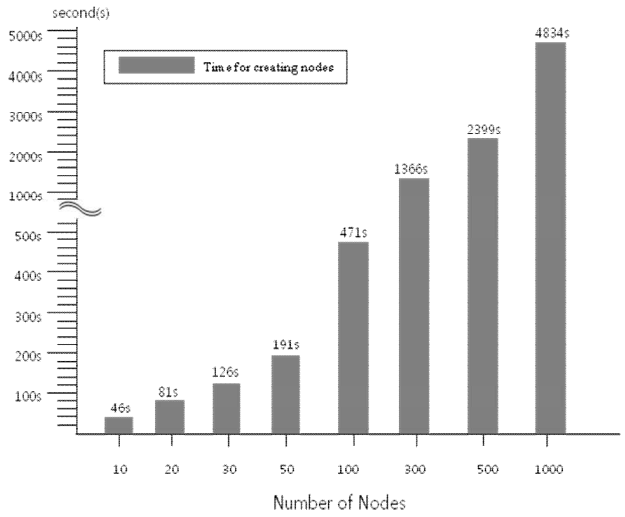


Figure 3. Time for Creating nodes in Pastry ring

As the starting point, we have measured time for creating nodes in the Pastry ring as in Figure3. We have measured time for the following number of nodes: 10, 20, 30, 50, 100, 300, 500 and 1000. Although measured time increases according to the number of nodes, we can know the averaged time per a node is almost constant near 4 seconds.

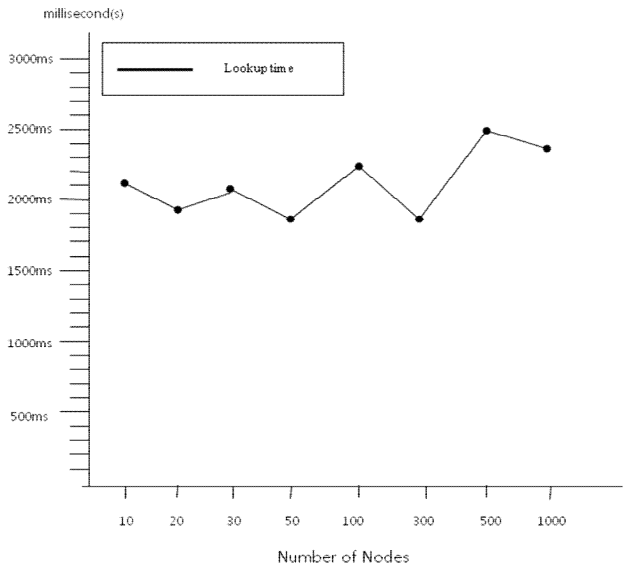


Figure 4. Time to retrieve five keywords

We have measured time to retrieve keywords as the Lookup time. We have assumed each node has five keywords. Every node simultaneously puts the **Lookup** command to retrieve a keyword five times, and measures time for a node to retrieve all of five keywords in cases of the following number of nodes: 10, 20, 30, 50, 100, 300, 500 and 1000. Figure 4 shows the measured time for the Lookup. The Lookup time for querying five keywords is almost flat near 2 seconds regardless of the number of nodes. From this figure we can say this search function is scalable.

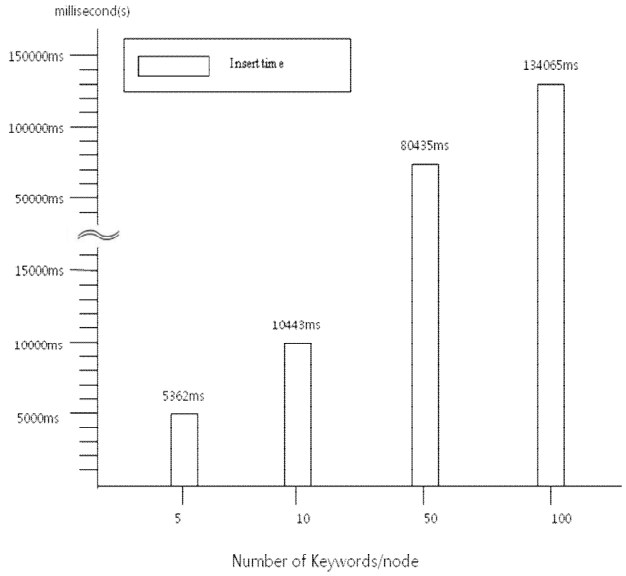


Figure 5. Insert time in 100 nodes

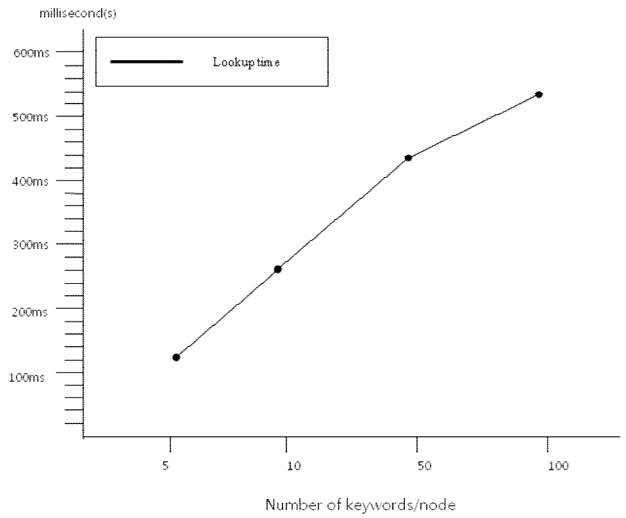


Figure 6. Lookup time in 100 nodes

Figure 5 and Figure 6 show the Insert time and the Lookup time when varying number of keywords for 100 nodes. In Figure 5 every node simultaneously issues the **Insert** commands of the number of keywords to store a keyword, and measures time for a node to store all keywords in cases of the following number of keywords: 5, 10, 50 and 100. The Insert time is shown to increases according to the number of keywords. However, the measured time per a keyword is 1.07, 1.04, 1.61 and 1.34 seconds in each case. These values look rather flat.

Figure 6 is the same case as the Figure 5, except the **Insert** command is replaced with **Lookup**. The Lookup time increases according to the number of keywords. But, the Lookup time per keyword is just 24, 26, 8.4 and 5.4 milliseconds, and it looks decreasing. We think it is due to parallel processing of retrieving keywords. When the number of keywords is small, its effect is a little shown, but as the number of keywords increases the parallel processing gets the effect. Therefore, we think the Lookup time get no effect from the number of keywords.

With the results in the above, we can conclude our search function is scalable. Therefore, we can say it could be valuably utilized in the Pastry based applications.

VI. CONCLUSION

Since many structured P2P algorithms are proposed and developed recently, there are so many tries to use one of them to develop various P2P systems. Our team has also tried to develop a system in Pastry, but we realize we need a search function. We wanted to use some keywords to discover content object, but we could not. As we have described, the structured P2P provides only the exact matching.

We have proposed a search function and evaluate the performance to be scalable. We can see it gets little effects from the number of nodes as well as the number of keywords. Pastry is a popular structured P2P platform and PAST is provided as an open source as the well matched system. If the DHT should be developed in other way for the search function, applying the Pastry to a system development is difficult. But, as we just use the original Pastry DHT with PAST, we think we can keep the advantages of Pastry DHT to apply Pastry to develop a system. PAST itself is already proven robust and scalable. Therefore, we think our system can be scalable. Also we expect it could be robust. We expect the proposed search function could be applied to various developments in the near future.

ACKNOWLEDGMENT

"This research was supported by the MKE(The Ministry of Knowledge Economy), Korea, under the ITRC(Information Technology Research Center) support program supervised by the NIPA(National IT Industry Promotion Agency)" (NIPA-2010-(C1090-1011-0009))

REFERENCES

[1] Ion Stoica, Robert Morris, David Karger, M. Frans Kaashoek, and Hari Balakrishnan, “Chord: A Scalable Peer-to-Peer Lookup Service for Internet Applications,”Proc. of the 2001 ACM SIGCOMM Conf., pp. 149-160, 2001.

[2] A. Rowstron and P. Druschel, “Pastry: Scalable, Decentralized Object Location and Routing for Largescale Peer-to-Peer Systems,” in IFIP/ACM Int’l Conf. on Distr. Systems Platforms (Middleware), 2001, pp.329-350.

 ISBN: 978-1-61208-103-8

[3] S. Ratnasamy, P. Francis, M. Handley, R. Karp, and S. Shenker, "A Scalable Content-Addressable Network," Proc. of ACM SIGCOMM, pp. 161-172, 2001.

[4] M. Harren, J. Hellerstein, R. Huebsch, B. Loo, S. Shenker, and I. Stoica, "Complex queries in DHT-based peer-to-peer networks," LNCS Vol. 2429, pp. 242 – 259, In IPTPS 2002.

[5] H. Lundgren, R. Gold, E. Nordström, M. Wiggberg, "A Distributed Instant Messaging Architecture based on the Pastry Peer-To-Peer Routing Substrate," In Proc. of Swedish National Computer Networking Workshop, Stockholm, Sept. 2003.

[6] Hanhua Chen, Hai Jin, Jiliang Wang, Lei Chen, Yunhao Liu, Lionel M. Ni, " Efficient multi-keyword search over p2p web", In WWW pp. 989-998, 2008.

[7] Patrick Reynolds and Amin Vahdat, "Efficient Peer-to-Peer Keyword Searching", In Middleware, pp. 21-40, 2003.

[8] B. Y. Zhao, J. D. Kubiatowicz, and A. D. Joseph, "Tapestry: An infrastructure for faultresilient wide-area location and routing," Technical Report UCB//CSD-01-1141, U. C. Berkeley, April 2001.

[9] Peter Druschel and Antony Rowstron, "PAST: A large-scale, persistent peer-to-peer storage utility," In Proc. HotOS VIII, Schloss Elmau, Germany, pp. 75-80, May 2001.

[10] http://www.freepastry.org/FreePastry/, May 2010.

[11] Androutsellis-Theotokis, S. and Spinellis, "A survey of peer-to-peer content distribution technologies," *ACM Comput. Surv.* 36, 4, pp. 335-371, Dec. 2004.

[12] Tang, C., Xu, Z., and Mahalingam, M., "pSearch: information retrieval in structured overlays," SIGCOMM Comput. Commun. Rev. 33, 1, pp. 89-94, Jan 2003.

[13] Vishnevsky, V., Safonov, A., Yakimov, M., Shim, E., and Gelman, A. D., "Scalable Blind Search and Broadcasting in Peer-to-Peer Networks," In Proceedings of the Sixth IEEE international Conference on Peer-To-Peer Computing P2P. IEEE Computer Society, Washington, DC, pp. 259-266, 2006.

[14] Y.Chen, Z. Xu, C. Zhai, "A Scalable Semantic Indexing Framework for Peer-to-Peer Information Retrieval," in ACM SIGIR 05 Workshop on Heterogeneous and Distributed Information Retrieval, 2005.

[15] C. Sangpachatanaruk, T. Znati, "A P2P Overlay Architecture for Personalized Resource Discovery, Access, and Sharing over the Internet," in CCNC'05, pp. 24-29, 2005.

 ISBN: 978-1-61208-103-8

A Novel TOPSIS-based Chunk Scheduling Approach for Layered P2P Streaming

Wei Chen, Sen Su, Fangchun Yang, Kai Shuang and Xinchao Zhao
State Key Laboratory of Networking and Switching Technology
Beijing University of Posts and Telecommunications
Beijing, China
chenwei348@gmail.com, {susen, fcyang, shuangk}@bupt.edu.cn, xcmmrc@gmail.com

***Abstract*—Although layered P2P streaming is perfectly adapted to heterogeneous network environments and heterogeneous user requirements, it suffers from bad delay performance like single-layer P2P streaming. In this paper, we analyze new characteristics of Pull-based P2P chunk scheduling problem caused by layered coding, and propose a Pull-based scheduling problem model aimed at enhancing the delay performance under the guarantee of video quality for layered P2P streaming. And then we put forward a heuristic chunk scheduling algorithm with aperiodic scheduling interval, where technique for order preference by similarity to ideal solution (TOPSIS) is utilized to solve several multiple-attribute decision making problems. Finally, we develop a new metric comprehensively evaluating video playback quality and delay performance, and simulations illustrate that our algorithm can greatly outperform the existing classical related work by a small increase in control overhead.**

Keywords-layered P2P streaming; heterogeneous peers; chunk scheduling; delay performance; MADM-TOPSIS

I. INTRODUCTION

P2P Streaming wins a big success in the large scale streaming systems in recent years due to low deployment cost and high scalability. And the development of Mobile access technologies (like 3G and LTE) and widespread adoption of broadband residential access make it possible that computers, mobile phones and set-top boxes share video streaming in Peer-to-Peer in the emerging scenario, i.e., heterogeneous network resources (especially bandwidth resources) and heterogeneous user requirements (especially subject to user terminals' limit, like display capacities). Layered coding is a promising solution adapted to the emerging scenario [3]. Accordingly, layered P2P streaming has drawn great interest in recent years [6][7][8][9][12].

Mesh-Pull P2P streaming has been proved with higher practicability, scalability, capability of coping with peer churn and bandwidth utilization than other P2P solutions by the success deployments of many commercial softwares, like PPStream [1], PPLive [2] and so on. However, it suffers from bad delay performance [4][5], and this problem would be much worse in layered P2P streaming. The delay origins from two aspects: overlay construction and chunk scheduling. On one hand, there are many related work [6][7] to focus on the overlay construction to meet QoS requirements. On the other hand, much work about chunk scheduling for layered P2P streaming aims to maximize the system's throughput [6][8]. However, to the best of our knowledge, there is no work about Pull-based chunk scheduling to optimize its delay performance under the guarantee of video quality for layered P2P streaming.

In this paper, we focus on the Pull-based chunk scheduling problem of layered P2P streaming to enhance the delay performance under the guarantee of video quality in the emerging scenario of heterogeneous upload and download bandwidth, heterogeneous and dynamic propagation delays, and heterogeneous requirements of video quality, where computers, mobile phones and set-top boxes share video streaming. First, we analyze new challenges brought by layered coding to Mesh-Pull P2P streaming, and then propose a chunk scheduling problem model to minimize the delivery time in a scheduling interval. Secondly, we put forward a TOPSIS-based, variable scheduling interval and heuristic Pull-based scheduling algorithm. Simulations illustrate that our algorithm can outperform the existing classical solutions.

The rest of the paper is organized as follows. Section II presents the related work. Section III describes new challenges and chunk scheduling problem model of layered P2P streaming. Section IV introduces our scheduling algorithm. We show the simulation results in Section V. Finally, Section VI presents our conclusion.

II. RELATED WORK

We briefly review the main chunk scheduling solutions for layered P2P streaming.

PALS [9] focused on the Mesh-Pull chunk scheduling problem, adopted a diagonal chunk priority order method and employed a Round-Robin method to request chunk. [10] aimed at optimizing the streaming transmission performance in mobile ad-hoc network. LION [11] utilized alternative paths and network coding to improve throughput. [12] studied the video-on-demand media distribution problem and put forward a peer-to-peer streaming solution which focused on how to optimally allocate the desired layers among multiple senders. In brief, these works aimed at maximizing the system throughput or delivering the satisfying number of layers according to available bandwidths. However, they ignored an important user experience, i.e., startup delay, and they did not simultaneously optimize delay performance and throughput so as to have longer startup delay. Therefore, we expect to decrease the startup delay caused by chunk scheduling under the guarantee of video play quality.

 ISBN: 978-1-61208-103-8

III. CHUNK SCHEDULING PROBLEM MODEL

In this section, we will analyze new characteristics of layered P2P streaming, and then propose a problem model for layered P2P streaming.

A. New Characteristics of Layered P2P Streaming

Layered coding [3] can be adapted to the emerging scenario where mobile terminals, personal computers and set-top boxes share the video streaming in Peer-to-Peer way. And compared with traditional P2P streaming, layered P2P streaming has new characteristics as follows:

- Layer characteristic of a chunk: layer characteristic should be considered because decoding of chunks in upper layer depend on that chunks with the same time identifier have been obtained in all lower layers.
- Heterogeneous video quality requirements: peers have diverse video quality (i.e., number of layers) requirements because of limits of different download bandwidths, different terminals' screen sizes and so on.
- Congestion in download bandwidths of peers: study on traditional P2P streaming in Internet usually assumes that download links of peers are not the bottleneck link [5] so that the number of neighbor peers delivering chunks is generally not limited. However, in the emerging scenario vast diversity among peers' download and upload bandwidths results in that download links of peers may be the bottleneck links. Therefore, a peer can only receive chunks simultaneously from limited number of neighbor peers in a scheduling interval so as not to the congestion in its download link.

These new characteristics must be considered in the model of chunk scheduling problem.

B. Problem Model for Layered P2P Streaming

In Pull-based chunk scheduling approach, each peer generally requests absent chunks in its request window from its neighbor peers autonomously and periodically. We pay close attention to how to get an optimal chunk assignment for a peer in a scheduling interval in which absent chunks as more as possible can be obtained in the shortest time, in order to improve the delay performance under the guarantee of video quality. For the scheduling problem of layered P2P streaming, we will consider the following factors:

- Absent chunks of a peer and their availabilities in its neighbor peers.
- Chunks' time and layer characteristics.
- Heterogeneous video quality requirements of peers.
- Heterogeneous and constant upload and download bandwidth of peers (It can be extended to variable bandwidth scenario by utilizing bandwidth measurement or prediction approaches [19], which is not what we cares about.)
- Congestion avoidance by limiting the number of neighbor peers delivering chunks.
- Heterogeneous and dynamic propagation delays among peers.

Suppose decision variable $\theta_{iljk} \in \{0,1\}$, "$\theta_{iljk}$ =1" means that scheduling peer P_r (any peer in the system) assigns the chunk C_{jk} to the neighbor P_i, and P_i will send C_{jk} as the l th of assigned chunks to P_r; otherwise "θ_{iljk} =0".

According to above consideration, we model Pull-based scheduling problem for layered P2P streaming in merging scenario as an integer programming problem to minimize the delivery time of absent chunks (i.e., the objective function) under the condition of ensuring the number of delivered chunks as more as possible (Constraint k)):

$$\min\{T\} \tag{1}$$

s.t.

$$a) \quad T = \max\{T_1, T_2, ..., T_{N-1}, T_N\}$$

$$b) \quad T_i = d_{ri} + d_{ir} + XNum_i / B_{ir}$$

$$c) \quad Num_i = \sum_{l=1}^{|S_r|} \sum_{j=p_r}^{p_r+W_B+W_C} \sum_{k=1}^{K} \theta_{iljk}$$

$$d) \quad \theta_{iljk} \le H_{ijk}$$

$$e) \quad t_{ijk} = \sum_{l_i=1}^{|S_r|} \sum_{s=1}^{N} (d_{ri} + d_{ir} + Xl_i / B_{ir}) \theta_{sl_i jk}$$

$$f) \quad t_{ijk} < Deadline(C_{jk})$$

$$g) \quad \sum_{i=1}^{N} \sum_{l=1}^{|S_r|} \theta_{iljk} \le 1$$

$$h) \quad \sum_{j=p_r}^{p_r+W_B+W_C} \sum_{k=1}^{K} \theta_{iljk} \le 1$$

$$i) \quad \theta_{iljk} \le \theta_{i(l-1)jk}$$

$$j) \quad \theta_{iljk} \le \theta_{ilj(k-1)}$$

$$k) \quad \sum_{i=1}^{N} \sum_{l=1}^{|S_r|} \sum_{j=p_r}^{p_r+W_B+W_C} \sum_{k=1}^{K} Priority(C_{jk}) \theta_{iljk} \ge \alpha \sum_{j=p_r}^{p_r+W_B+W_C} \sum_{k=1}^{K} Priority(C_{jk}) M_{jk}$$

$$l) \quad \sum_{i=1}^{N} B_{ir} g(Num_i) \le D_r$$

$$m) \quad g(Num_i) = \begin{cases} 1, & \text{if } Num_i \ge 1 \\ 0, & \text{if } Num_i = 0 \end{cases}$$

The symbols are defined in TABLE I. The delivery time in this scheduling interval is the maximum of delivery times of N neighbor peers, as shown in Constraint a). Constraint b) introduces that the propagation delay of request message, transmission delay of chunks and propagation delay of chunks should be considered when evaluating the delivery time of chunks assigned to P_i. Constraint c) shows the number of chunks assigned to neighbor P_i. Constraint d) requires chunk C_{jk} can be assigned to neighbor P_i only if the neighbor P_i has the chunk C_{jk}. Constraint e) introduces how to quantify the delivery time of a chunk C_{jk}, like the quantification of T_i. Constraint f) requires that chunk C_{jk} must be obtained in its deadline. Constraint g) requires a chunk can be assigned to one neighbor peer at most. Constraint h) requires a neighbor P_i delivers one chunk at most to P_r in l th sequence. Constraint i) requires a neighbor

 ISBN: 978-1-61208-103-8

TABLE I. NOTATIONS

Symbols	**Description** (all the symbols are confined to a scheduling interval; delivery time is the absolute time; time unit, i.e., K*X/Rate, is the absolute time of K chunks with the same time identifier due to the layered partition of video.)
T	Delivery time of chunks from requested to obtained by scheduling peer P_r
T_i	Delivery time of chunks assigned to neighbor P_i from requested to obtained by scheduling peer P_r
d_{ri}	Propagation delay from scheduling peer P_r to neighbor P_i
d_{ir}	Propagation delay from neighbor P_i to scheduling peer P_r
Num_i	Number of chunks assigned to neighbor P_i
C_{jk}	Identifier of chunk representing jth time unit and kth layer (K is the maximum number of video layers; J is the maximum number of time units of video)
$H_{ijk} \in \{0,1\}$	"H_{ijk} =1" denotes neighbor P_i has the chunk C_{jk}; otherwise, "H_{ijk} =0"
t_{ijk}	Delivery time of chunk C_{jk} by neighbor P_i
X	Size of a chunk
Rate	Bit rate of video playback with K layers
B_{ir}	Available upload bandwidth from neighbor P_i to scheduling peer P_r
D_r	Download bandwidth of scheduling peer P_r
Deadline(C_{jk})	Playback deadline (absolute time) of chunk C_{jk}
Priority(C_{jk})	Priority value of chunk C_{jk}
$p_r \in \{1,2,\ldots,J\}$	Number of time units of chunk being about to be played by scheduling peer P_r
W_A	Number of time units held by A area
W_B	Number of time units held by B (i.e., urgent) area
W_C	Number of time units held by C (i.e., loose) area
S_r	Set of absent chunks in request window of scheduling peer P_r (S_r ={ C_{jk} | H_{ijk} =0, k⩽K, p_r⩽j⩽p_r + W_B + W_C })
$M_{jk} \in \{0,1\}$	"M_{jk} =1" denotes chunk C_{jk} can be provided by some neighbor peer (i.e., $\sum^{N}_{i=1} H_{ijk} \geqslant 1$); otherwise, "$M_{jk}$ =0"

P_i can deliver a chunk in higher sequence to P_r only if there are assigned chunks in all the lower sequences of P_i in order to reduce the delivery time of chunk. Constraint j) requires a chunk can be assigned to some neighbor only if all the chunks in the lower layers have been assigned, in order to decrease the number of chunks which cannot be decoded. Constraint k) requires that sum of all the assigned chunks' priorities should be bigger than α times of sum of all the chunks' priorities, which can be provided by neighbors, which ensures the video play quality; Constraint l) requires sum of upload bandwidths of neighbors delivering chunks to P_r is less than or equal to the download bandwidth of P_r so as not to the congestion in its download link. Constraint m) shows "$g(Num_i)$ =1" represents neighbor P_i will deliver chunks to P_r, otherwise "$g(Num_i)$ =0".

Formula (1) is called distributed and local delay-optimum chunk scheduling problem model under the guarantee of video quality for layered P2P streaming. The integer programming problem is a NP-hard problem with $N^{|Sr|}$ combination solutions [13], where $|S_r|$ is always a large number. Therefore, we propose a distributed and heuristic scheduling algorithm to solve the problem.

IV. TOPSIS-BASED SCHEDULING ALGORITHM

According to the above problem model, we propose a heuristic and TOPSIS-based chunk scheduling algorithm for layered P2P streaming. First, we'd like to emphasize its four important problems:

1) Priority ordering problem: Considering the characteristics of chunks in layer and time (i.e., some chunks must be obtained as soon as possible, and other chunks can be obtained later), scheduling peer P_r should order its absent chunks by their importance to P_r and assign the chunk to some neighbor peer one by one following the chunk priority sequence;

2) Candidate neighbor selection problem: Due to the limit of download bandwidth, P_r should select candidate neighbor peers used to deliver chunks from its neighbor peers in order to avoid the congestion;

3) Chunk assignment problem: P_r should assign a chunk to a neighbor which deliver the chunk and chunks assigned in the shortest time;

4) Aperiodic scheduling interval: too long or too short scheduling interval would result in the longer delay or more repeated chunks, therefore we develop an aperiodic scheduling interval based on delivery time of absent chunks in each scheduling interval to reduce the delay or number of repeated chunks.

TOPSIS [14] is utilized in the following design to solve the MADM problems. Due to the limit of space, we do not introduce the TOPSIS in detail. As described in [14], we can quantify and order the alternatives according to performance data for n alternatives, attributes and their weights.

B. Chunks' Priorities Ordering

For layered coding, the video stream is encoded into several layers, and each layer is partitioned into chunks. Thus each chunk (C_{jk}) has a layer ID (i.e., k) and time unit ID (i.e., j). Buffer, shown in Figure.1, is divided into three parts: A area, B area (also called urgent area or playback window) and C area (also called loose area). A area stores video chunks just played. Urgent area is close to playback point, and each new peer cannot start playing the video until obtaining the large part of or all the chunks in this area. Request window is composed of B and C area and a peer

 ISBN: 978-1-61208-103-8

hopes to request absent chunks in request window from its neighbor peers in each scheduling interval.

Quantifying and ordering the absent chunks is a MADM problem and we adopt TOPSIS to solve the problem. According to characteristics of layered coding, four attributes should be considered as follows:

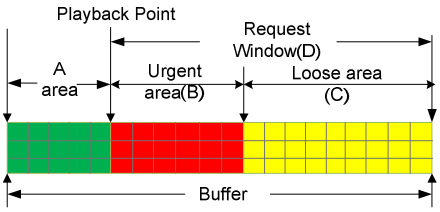


Figure 1. Design of Buffer

- Layer ID of chunk C_{jk} (denoted as $Property_1(C_{jk})$): the priority of C_{jk} increases with the decrease of its layer ID.
  $Property_1(C_{jk})= k$ (2)
- Number of chunks which cannot be decoded due to the absence of chunk C_{jk} (denoted as $Property_2(C_{jk})$): the priority of the chunk increase as the decrease of its $Property_2(C_{jk})$.
  $Property_2(C_{jk})= \sum^{K}_{s=k+1} H_{rjs}$ (3)
- Number of neighbors having chunk C_{jk} (denoted as $Property_3(C_{jk})$): the priority of C_{jk} increases with the decrease of its $Property_3(C_{jk})$, because this strategy can achieve good performance [15][16].
  $Property_3(C_{jk})= \sum^{N}_{s=0} H_{sjk}$ (4)
- Number of time units of the playback deadline of chunk C_{jk} (denoted as $Property_4(C_{jk})$): the priority of C_{jk} increases with the decrease of its $Property_4(C_{jk})$.
  $Property_4(C_{jk})=j-p_r$ (5)

Considering the characteristics of urgent and loose areas, we design the chunk priority algorithm as follows:

1) The priorities of chunks in urgent area are higher than ones of chunks in loose area, because chunks in urgent area are closer to playback point.

2) Scheduling objective of chunks in urgent area is to reduce the number of chunks which cannot be decoded, and to obtain video chunks with more layers (i.e., higher video quality). Therefore, we consider three attributes: $Property_1$、$Property_2$ and $Property_4$. And the corresponding weights are ω_{B1}、ω_{B2} and ω_{B4}, where $\omega_{B2} > \omega_{B4} > \omega_{B1}$ ($\omega_{B2}+\omega_{B1}+\omega_{B4}=1$). ω_{B2} is the largest value in order to decrease the number of chunks which cannot be decoded to improve the video play quality. The set of absent chunks in this area is S_{rB} ($=\{C_{jk} \mid H_{rjk}=0, k\leqslant K, p_r\leqslant j\leqslant p_r + W_B, M_{jk}=1\}$). According to TOPSIS, performance data a_{nm} ($n=|S_{rB}|$, $m=3$) of chunks obtained by Formula (2,3,5), monotonicity and weight of each attribute are put together to quantify the chunks' priorities, denoted as $Priority_{rB}(C_{jk})$, and order the chunks.

3) Scheduling objective of chunks in this area is to obtain more chunks so that these chunks with a high video quality can enter the urgent area. Therefore, we consider three attributes: $Property_1$、$Property_3$ and $Property_4$. And the corresponding weights are ω_{C1}、ω_{C3} and ω_{C4}, where $\omega_{C3} > \omega_{C1}$ and $\omega_{C3} > \omega_{C4}$ ($\omega_{C1}+\omega_{C3}+\omega_{C4}=1$). ω_{C3} is the largest value to increase the number of assigned chunks. The larger ω_{C1} leads to higher video playback quality while the larger ω_{C4} brings higher playback continuity. Therefore, ω_{C1} and ω_{C4} are a tradeoff between playback quality and playback continuity. The set of absent chunks is S_{rC} ($=\{C_{jk} \mid H_{rjk}=0, k \leqslant K, p_r+W_B\leqslant j\leqslant p_r+W_B+W_C, M_{jk}=1\}$). According to TOPSIS, performance data a_{nm} ($n=|S_{rC}|$, $m=3$) of chunks obtained by Formula (2,4,5), monotonicity and weight of each property are put together to quantify the chunks' priorities, denoted as $Priority_{rC}(C_{jk})$, and order the chunks.

Utilizing the above chunk priority algorithm, the priorities of chunks in S_r ($=S_{rB}\cup S_{rC}$) can be quantified, and we can obtain a chunk sequence $(C_1,C_2,\ldots,C_S)$ with $S=|S_r|$.

C. Candidate Neighbor Selection

Scheduling peer should choose as more neighbor peers with higher performances as possible as candidate neighbor peers whose upload bandwidths' sum is less than or equal to its download bandwidth to avoid the congestion.

We adopt TOPSIS method to quantify neighbors' importance index I_{ri} (i.e., the importance degree of neighbor peer P_i to scheduling peer P_r). Two factors affecting I_{ri} are the chunks owned by neighbor peer and the capability of neighbor peer delivering chunks.

First, we adopt the sum of priorities of chunks needed by P_r and owned by P_i to evaluate the first factor, because different chunks have different priorities or importance for P_r. Considering the design of urgent area and loose area, we take sum of priorities of chunks owned by P_i in urgent area and sum of priorities of chunks owned by P_i in loose area as two attributes (PR_{Bi} and PR_{Ci}):

$$PR_{Bi} = \sum_{C_{jk}\in S_{rB}} H_{ijk}\,Priority_{rB}(C_{jk}) \quad (6)$$

$$PR_{Ci} = \sum_{C_{jk}\in S_{rC}} H_{ijk}\,Priority_{rC}(C_{jk}) \quad (7)$$

The corresponding weights of PR_{Bi} and PR_{Ci} are ω_B and ω_C, where $\omega_B > \omega_C$. That is because chunks in urgent area are closer to playback point and more important to video play quality. With the rise of ω_B or ω_C, I_{ri} increases.

Secondly, the capability of neighbor peer P_i delivering a chunk (denoted as $Performance_i$) is the third attribute, which is evaluated by the propagation delay of request message, the propagation delay of a chunk and transmission delay of P_i:

$Performance_i=X/B_{ir}+d_{ri}+d_{ir}$ (8)

The I_{ri} of neighbor peer decreases with the increase of $Performance_i$. The corresponding weight of this attribute is ω_p. Candidate neighbor peer selection is greatly related to the current scheduling situation and the number of chunks owned by candidate neighbor peers affects the number of chunks which can be delivered, so $\omega_B+\omega_C > \omega_p$ ($\omega_B+\omega_C+\omega_p=1$).

According to TOPSIS, performance data a_{nm} of neighbor peers ($n=N$, $m=3$), monotonicity and weight of each attribute

 ISBN: 978-1-61208-103-8

are put together to quantify the I_{ri}.

By I_{ri}, we can order the neighbor peers and obtain the decreasing sequence $(P_1,P_2,\ldots,P_N)$. The candidate neighbor selection algorithm is to choose the first N_r neighbor peers so that sum of these N_r neighbors' upload bandwidths is equal to scheduling peer's download bandwidth. Maybe, the upload bandwidth of P_{Nr} cannot be fully utilized due to the limit of scheduling peer's download bandwidth. Therefore, the set $\{P_1,P_2,\ldots,P_{Nr}\}$ is the candidate neighbor peers set.

D. Delay-Optimum Chunk Assignment

For the chunk sequence $(C_1,C_2,\ldots,C_{|Sr|})$ and candidate neighbor peers set $\{P_1,P_2,\ldots,P_{Nr}\}$, scheduling peer choose a neighbor peer in $\{P_1,P_2,\ldots,P_{Nr}\}$ for each chunk one by one following the chunk sequence. Due to the same chunk size, the delivery time of a chunk only depends on the transmission bandwidth of a neighbor peer and propagation delay between the neighbor and scheduling peer. Therefore, delay-optimum chunk assignment algorithm is that scheduling peer greedily chooses the neighbor peer which meets the formula (9) (i.e., choosing the neighbor peer which delivers the chunk in the shortest time) for each chunk according to the chunk sequence.

$$\min\{t_{ijk}\} = \min\{\sum_{l=1}^{|S_r|}\sum_{s=1}^{N}(d_{ri}+d_{ir}+Xl_i/B_{ir})\theta_{sljk}\} \quad (9)$$

E. Aperiodic Scheduling

Pull-based chunk scheduling algorithms [9][15] usually adopt constant periodical scheduling interval. That's because they cannot quantify the delivery time of each scheduling. If the interval is too long, there exists a span when upload bandwidths of neighbor peers and download bandwidth of the scheduling peer are idle so as to decrease the utilization of bandwidths and increase the video play delay. And if the interval is too short, the scheduling peer may request chunks which have been requested in the last scheduling interval so as to waste the bandwidths of neighbor peers and scheduling peer and also increase the play delay.

In our algorithm, we adopt the aperiodic scheduling interval, and take the delivery time (i.e., $T=\max\{T_1,T_2,\ldots,T_{Nr}\}$) of chunks in a scheduling as the interval between this scheduling and next scheduling. This will reduce the idle time of upload and download bandwidths and the number of repeatedly requested chunks to further improve the delay performance.

V. Simulation and Evaluation

To validate and evaluate our scheduling algorithm, we have conducted extensive simulations based on PeerSim [17]. We only focus on the chunk scheduling algorithm, and so adopt the same overlay construction approach [7]. In the beginning, 100 peers join in the system and make up a mesh with eight neighbor peers. The video with 100s duration is divided into chunks with the same size 1250byte, close to a maximum of MTU, and encoded ten layers with 100kbps of each layer. The buffer has 10s duration with request window of 8s duration. A new peer joins in the system every two seconds and an online peer quits the system every three seconds. Propagation delays among peers are randomly assigned from the delay matrix (2500*2500) in the Internet measurement [18] and reassigned every five seconds. Due to the unpredictability of TCP retransmission delay, UDP is adopted. And we assume packet loss ratio is 2%. The streaming server's upload bandwidth is 2Mbps and the bandwidth distribution and desired video qualities for three kinds of peers are shown in TABLE II.

TABLE II. Bandwidth Distribution and Desired Video Quality

	Mobile terminals	**PCs**	**Set-top boxes**
Upload(kbps)	300	600	1000
Download(kbps)	2000	4000	8000
Video quality	2 layers	6 layers	10 layers
Ratio	30%	40%	30%

A. Metrics

For layered streaming, we adopt the following metrics:

- Layer delivery ratio: Ratio of the number of a peer's video playback layers to one of its desired layers. This metric reflects the peer's video playback quality.
- Useless chunks ratio: Ratio of chunks unable to be decoded for a peer. This metric should be kept low.
- Number of control messages: Number of a peer' control messages, like chunk availability messages, request messages, maintenance messages for peers' departure. This metric reflects the control overhead.
- (λ_1, λ_2)-Startup Delay: A peer usually starts to playback the video after obtaining all or the large part of chunks in playback window. With the rise of playback window, a peer has more time to request the absent chunks so as to have a better video quality, however, have a longer startup delay; otherwise, a peer has a worse video playback quality and a shorter startup delay. This metric represents the minimal startup delay for a peer when it enjoys the video quality with λ_1 layer delivery ratio and λ_2 useless chunk ratio. This metric comprehensively reflects a peer's video quality and delay performance.

B. Effects of algorithm parameters

In our chunk scheduling algorithm, there are several important parameters: the weights [ω_{B2} ω_{B4} ω_{B1}], the weights [ω_{C3} ω_{C4} ω_{C1}] and the weights [ω_B ω_C ω_P].

Figure.2 shows the cumulative distribution function of peers' (0.999,0.001)-Startup Delay with different [ω_{B2} ω_{B4} ω_{B1}], and the performance decreases by the configuration 2, 7, 4, 1, 3, 5, 6. For different configurations, peers have better delay performance with larger ratio of ω_{B2} or ω_{B4}, like configuration 1,2,3,4 and 7; peers have worse delay performance with larger ratio of ω_{B1}. That is because for chunks in urgent area ω_{B2} affects useless chunk ratio, and lower useless chunk ratio may result in higher utilization of download bandwidth, and then higher video playback quality; the larger ω_{B4} makes scheduling peer prioritily request chunks closer to playback point, which improves the video playback quality; the larger ω_{B1} makes scheduling peer prioritily request chunks in lower layers and yet ignore the

 ISBN: 978-1-61208-103-8

absent chunks to be played soon, and then results in a worse video play video, like configuration 5 and 6. Therefore, $\omega_{B2} > \omega_{B1} > \omega_{B4}$ can achieve better delay performance under the guarantee of high video quality and the configuration [0.5 0.3 0.2] is adopted in the following simulation.

Figure. 3 shows the cumulative distribution function of peers' (0.999, 0.001) –Startup delay with different [ω_{C3} ω_{C4} ω_{C1}], and the performance decreases by the configuration 2, 1, 6, 3, 4, 5. When ω_{C3} is bigger, peers have better delay performance, like configuration 1, 2 and 6; otherwise, even if ω_{C3} or ω_{C1} is bigger, peers cannot achieve better delay performance, like configuration 3, 4 and 5. That is because the bigger ω_{C3} is illustrated in [15][16] to achieve better delivery performance. Therefore, $\omega_{C3} > \omega_{C1}$ and $\omega_{C3} > \omega_{C4}$ can achieve better delay performance under the guarantee of high video quality and the configuration [0.8 0.1 0.1] is adopted in the following simulation.

Figure. 4 shows the cumulative distribution function of peers' (0.999, 0.001) –Startup delay with different [ω_B ω_C ω_P], and the performance decreases by the configuration 7, 6, 5, 4, 3, 1, 2. Three weights are important factors of

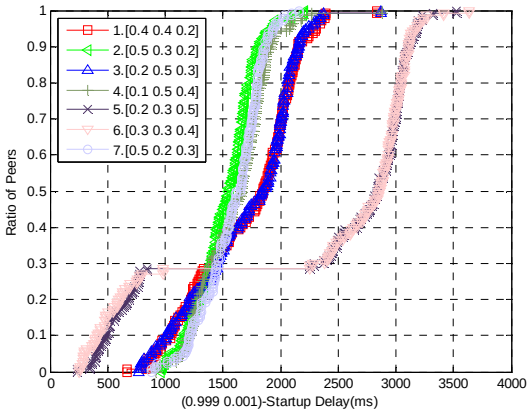


Figure 2. CDF of peers' (0.999,0.001)-Startup Delay with different [ω_{B2} ω_{B4} ω_{B1}]

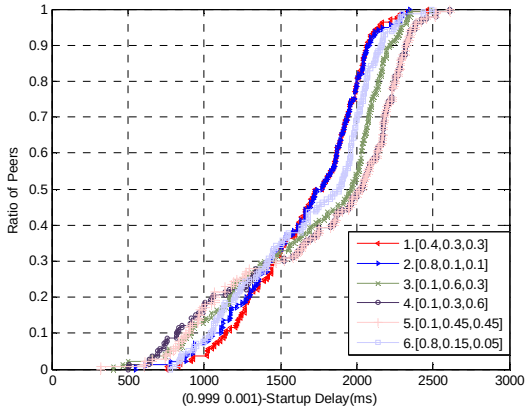


Figure 3. CDF of peers' (0.999,0.001)-Startup Delay with different [ω_{C1} ω_{C3} ω_{C4}]

importance index and when each of them is smaller, peers have the worse delay performances, e.g., configuration 5, 6, 7 are better than configuration 1, 2, 3, 4. That is because ω_B, ω_C and ω_P are respectively responsible for playback quality of video to be played soon, the number of chunks in loose area affecting the video quality of entering urgent area, and the delivery time of chunks. Besides, compared with the other weights, ω_C is a little more important, e.g., configuration 7 is better than 5 and 6. That is because the larger ω_C is to obtain more chunks in this scheduling interval and ensure high video quality entering into urgent area. Therefore, following that ω_C is a little larger than ω_B and ω_P can achieve better delay performance under the guarantee of high video quality. And the configuration [0.3 0.4 0.3] is adopted in the following simulation.

C. *Comparisons*

We compare our algorithm with two classical related work PALS [9] and Random Scheduling [15] to verify our algorithm's performance.

Figure. 5 shows the cumulative distribution function of peers' (0.995, 0.005) –Startup delay with different

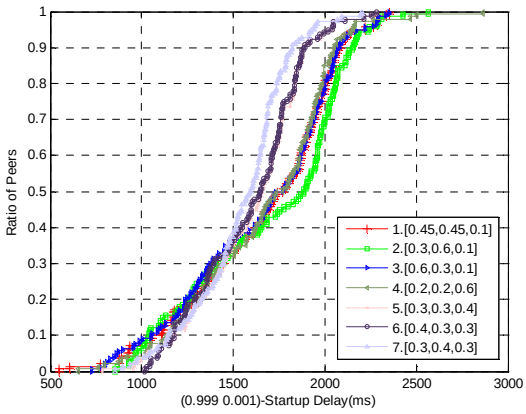


Figure 4. CDF of peers' (0.999,0.001)-Startup Delay with different [ω_B ω_C ω_P]

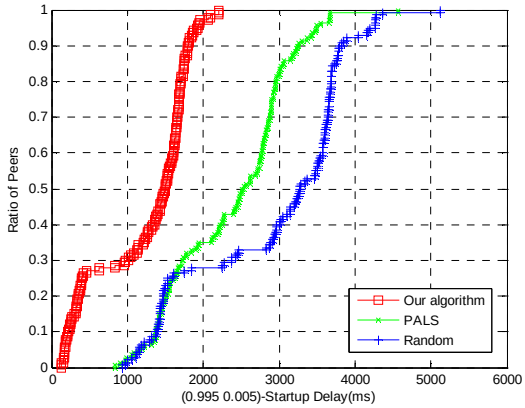


Figure 5. CDF of peers' (0.995,0.005)-Startup Delay with different Scheduling algorithm

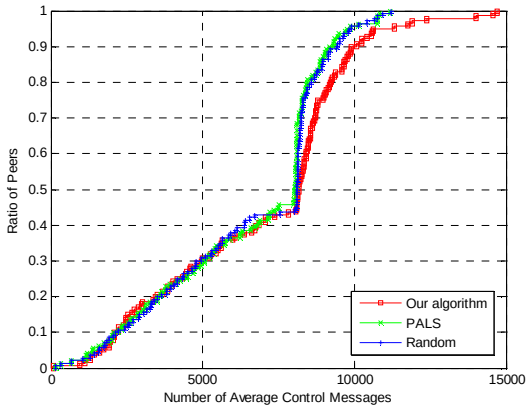


Figure 6. CDF of peers' number of control messages with different scheduling algorithms

scheduling algorithms and our algorithm reduces the startup delay respectively by 48.6% and 57.9% than PALS and random scheduling. PALS adopts Round-Robin scheduling algorithm to ignore the differences among capabilities of neighbor peers delivering chunks; Random scheduling algorithm ignores not only the differences among capabilities of neighbor peers but also the differences among priorities of absent chunks. These result in not achieving higher layer delivery ratio in shorter time for PALS and Random scheduling. Our algorithm proposes a priorities' quantification method based on the characteristics of urgent area and loose area, and prioritily requests chunks with higher priorities from neighbor peers which deliver them in the shortest time. Therefore, our algorithm can achieve higher layer delivery ratio in the shortest time.

Figure. 6 shows the cumulative distribution function of peers' number of control messages when achieving (0.995, 0.005) –Startup delay for different scheduling algorithms. And our algorithm increases respectively by 6% and 5.5% than PALS and Random scheduling. The rise of number of control messages for our algorithm results from the rise of number of request messages which are used to obtain the lost chunk due to link loss, which improves the layer delivery ratio. However, PALS and Random scheduling do not have enough time to request the lost chunks due to the worse performance of their scheduling algorithm.

In summary, our algorithm can greatly (respectively 48.6% and 57.9%) outperform two classical approaches PALS and Random scheduling by control overhead's slight increase (respectively 6% and 5.5%).

I. CONCLUSION

In this paper, we proposed a Pull-based chunk scheduling problem model and TOPSIS-based chunk scheduling algorithm for layered P2P streaming in the emerging scenario to deal with the problem of long startup delay. And simulations illustrated our algorithm could greatly enhance delay performance under the guarantee of high video quality than the existing classical related work by a slight increase of control overhead.

ACKNOWLEDGMENT

This work was supported by National Key Basic Research Program of China (973 Program) (2009CB320504), National High Technology Research and Development Program of China ("863"Program) (2008AA01A317), the Foundation for Innovative Research Groups of the National Natural Science Foundation of China (Grant No. 60821001), Research Fund for the Doctoral Program of Higher Education of China (20090005120012), National Key Basic Research Program of China (973 Program) (2009CB320406).

REFERENCES

[1] PPStream, http://www.ppstream.com/ , 07.03.2010

[2] PPLive, http://www.pplive.com/, 07.03.2010

[3] B. Li, J. Liu, "Multirate video multicast over the internet: an overview", IEEE Network, vol. 17, no. 1, pp. 24-29, 2003.

[4] S. Agarwal, J. P. Singh, A. Mavlankar, P. Baccichet, and B. Girod,"Performance and Quality-of-Service Analysis of a Live P2P Video Multicast Session on the Internet," IEEE IwQoS'08, Enschede, NL, pp. 11-19, June 2008.

[5] H. Xiaojun, L. Chao, L. Jian, L. Yong, and K. W. Ross, "A Measurement Study of a Large-Scale P2P IPTV System," Multimedia, IEEE Transactions on, vol. 9, pp. 1672-1687, 2007.

[6] L. Dai, Y. Cui, and Y. Xue, "Maximizing Throughput in Layered Peer-to-peer Streaming", in Proc. IEEE ICC, Glasgow, Scotland, pp. 1734-1739, 2007.

[7] X. Xiao, Y.C. Shi, B.P. Zhang and Y. Gao, "OCals: A Novel Overlay Construction Approach for Layered Streaming", in Proc. IEEE ICC, Beijing, China, pp. 1807-1812, 2008.

[8] Y. Okada, M. Oguro, et al., "A New Approach for the Construction of ALM Trees using Layered Video Coding", in Proc. P2PMMS, Hilton, Singapore, pp.12-12, 2005.

[9] R. Rejaie, A. Ortega, "PALS: Peer-to-Peer Adaptive Layered Streaming", In Proc. ACM NOSSDAV, Monterey, California, pp. 153-161, 2003.

[10] M. Qin, R. Zimmermann, "Improving Mobile Adhoc Streaming Performance through Adaptive Layer Selection With Scalable Video Coding",in Proc. ACM Multimedia, Augsburg, Germany, pp. 717 - 726, 2007.

[11] J. Zhao, F. Yang, et al, "On Improving the Throughput of Media Delivery Applications in Heterogenous Overlay Network", in Proc. IEEE Globecom, San Francisco, USA, pp. 1-6, 2006.

[12] Y. Cui and K. Nahrstedt, "Layered Peer-to-Peer Streaming", in Proc. ACM NOSSDAV, Monterey, USA, pp.162 -171,2003.

[13] M.R. Garey and D.S. Johnson, Computers and Intractability - A Guide to the Theory of NP-completeness, Freeman, 1979.

[14] 4. K. Yoon and C.L. Hwang, Multiple Attribute Decision Making: An Introduction, Sage, Thousand Oaks, CA,(1995).

[15] V. Pai, K. Kumar, et al., "Chainsaw: Eliminating trees from overlay multicast", in Proc. IEEE INFOCOM 2005, Miami, USA, pp. 127-140, 2005.

[16] A. R. Bharambe, C. Herley, and V. N. Padmanabhan, "Analyzing and improving a bittorrent networks performance mechanisms," in IEEE INFOCOM 2006, Barcelona, Spain, pp. 1-12, Apr. 2006.

[17] Jelasity, M., Montresor, A., Jesi, G.P.: Peersim: Peer-to-Peer simulator (2004), http://peersim.sourceforge.net

[18] Meridian node to node latency matrix (2500×2500):http://www.cs.cornell.edu/People/egs/meridian/data.php. Meridian Project, 2005.

[19] S. Floyd et al., "Equation-based Congestion Control for Unicast Applications", In Proc. ACM SIGCOMM 2000, Stockholm, Sweden, pp. 43-56, 2000

 ISBN: 978-1-61208-103-8

Optimal State Surveillance under Budget Constraints

Praveen Bommannavar
Management Science and Engineering
Stanford University
bommanna@stanford.edu

Nicholas Bambos
Electrical Engineering and
Management Science and Engineering
Stanford University
bambos@stanford.edu

***Abstract*—In this paper we consider the problem of monitoring an intruder in a setting where the number of opportunities to conduct surveillance is budgeted. Specifically, we consider a problem in which we model the state of an intruder in our system with a Markov chain of finite state space. These problems are considered in a setting in which we have a hard limit on the number of times we may view the state. Such a constraint is natural when considering surveillance where mobile devices are involved and battery power is at a premium. We consider the Markov chain together with an associated metric that measures the distance between any two states. We develop a policy to optimally (with respect to the specified metric) keep track of the state of the chain at each time step over a finite horizon when we may only observe the chain a limited number of times. The tradeoff captured is the budget for surveillance versus having a more accurate estimate of the state; the decision at each time step is whether or not to use an opportunity to observe the process.**

***Keywords*-monitoring; surveillance; budget; resource allocation; dynamic programming**

I. Introduction

The importance of monitoring technologies in today's world can hardly be overstated. Indeed, there are volumes dedicated to this field [1] [2]. In recent years, the need for effective security measures has become especially evident. Indeed, at present, Microsoft announces almost one hundred new vulnerabilities *each week* [3]. Perhaps more alarming is the fact that government agencies routinely must manage defenses for network security and are hardly equipped to do so. This is evidenced by the fact that 10 agencies accounting for 98% of the Federal budget have been attacked with as high of a success rate as 64% [4].

This paper is concerned with a mathematical treatment of these important problems. Specifically, we consider a scenario in which we model the activities of an intruder as a state in a Markov chain. We develop the problem of monitoring the state in a finite-horizon discrete-time setting where we are only able to make observations a limited number of times. Such a budget arises naturally in wireless settings, for example. We present an algorithm for deciding when to use opportunities to view the process in order to minimize the surveillance error. This error is accrued at each time step according to a metric indicating how far from the true state the estimate was.

A growing literature addresses security from a mathematical perspective, with a range of theoretical tools being employed for managing threats. In [5], a network dynamically allocates defenses to make the system secure in the appropriate areas as time progresses. Parallels between the security problem and queuing theory are drawn upon, where vulnerabilities are treated as jobs in a backlog. The model of [6] uses ideas from game theory for intrusion detection where an attacker and the network administrator are playing a non-cooperative game. A related problem is addressed in [7] as well.

More generally, theoretical work in signal estimation has also been greatly developed [8]. Related works have considered aspects of decision making with limitations on the available information. In [9], an estimation problem is considered in which the received signal may or may not contain information. Similar issues are studied in [10] but in a control theoretic context in which the actuator has a non-zero probability of dropping estimation and control packets.

The unique aspect of our formulation is the nature of the power limitation. This non-standard constraint was introduced in [11] and developed in other works such as [12]. All of these problems consider finite horizon frameworks in which decisions are usage limited and hence the ability to make actions is a resource which must be appropriately allocated.

In Section II, we begin by introducing the monitoring problem mathematically. We continue with a derivation of the optimal policy using dynamic programming and then present the implementation of the optimal policy. In Section III, we demonstrate performance using numerical results and finally, in Section IV, we conclude the paper and offer directions for future work in this vein.

II. Monitoring

Let us now examine the monitoring/surveillance problem in greater detail. In what follows, we shall consider the states of a Markov chain as an abstraction for the position of an intruder in our system. Such a model is able to capture several scenarios. In one, we may wish to spatially monitor the location of an adversary using equipment that has usage constraints. Another situation is that we can consider the state of the intruder to be a location in a

 ISBN: 978-1-61208-103-8

data network. Although many interpretations are possible, our goal is to be able to track this state with as little error as possible. We begin by presenting the model in a mathematical state estimation framework, and then present the solution structure.

A. Model

Consider a Markov chain $\mathcal{M}$ with finite state space S, transition matrix P and an associated measure $d : S \times S \rightarrow \mathbf{R}$. The metric gives a sense of how close states are so that we can measure the effectiveness of an estimate of the true state. We assume that the process is known to start at initial state x_0 and we are interested in having an accurate estimate of the process over a finite horizon $k = 1, ..., N-1$. The decision space is simply $u \in \{0, 1\}$ where 0 corresponds to no observation being made and 1 corresponds to an observation being made. When an observation is made, the state x_k of $\mathcal{M}$ is perfectly known. Without an observation, on the other hand, we must form an estimate $\hat{x}_k$ for the state given all observed information thus far. The number of times observations may be made is limited to $M < N$.

The cost of making estimate $\hat{x}_k$ at time k when the true state is actually x_k is $d(x_k, \hat{x}_k)$. If d is a metric, we have the important properties

1. $d(x, y) \geq 0 \quad \forall x, y \in S$
2. $d(x, x) = 0 \quad \forall x \in S$
3. $d(x, y) = d(y, x) \quad \forall x, y \in S$
4. $d(x, z) \leq d(x, y) + d(x, z) \quad \forall x, y, z \in S$

At each time k, the state of our system can be represented by $\{(r, s, t); x_{N-t-r}; x_{N-t}\}$ where r is the number of time slots that have passed since the last observation, s is the number of opportunities remaining to make an observation, t is the number of time slots remaining in the problem, x_{N-t-r} is the last observed state of $\mathcal{M}$ and x_{N-t} is the current state. We seek a policy $\pi = \{\mu_k\}_{k=1}^{N-1}$ such that the actions $u_k = \mu_k((r, s, t), x_{N-t-r}) \in \{0, 1\}$ are chosen to minimize the cumulative estimation error. The policy π is admissible if it abides by the additional constraint that the number of times observations are made is no greater than M. Denote the class of admissible policies by Π.

We want to find a policy $\pi^* \in \Pi$ to minimize

$$\mathbf{E}\left\{\sum_{k=1}^{N-1} d(x_k, \hat{x}_k)\right\}$$

It should be noted that the estimate $\hat{x}_k$ depends on the action u_k because if $u_k = 1$ then $\hat{x}_k = x_k$ and there is no estimation error, while if $u_k = 0$ then we must make the best guess of the state that is possible with the known information.

Deciding on the distance metric is an issue of modeling and may be specific to the application at hand. We consider a few alternatives here:

1) Probability of Error: To recover a cost structure that results in the same penalty regardless of which state is chosen in error (probability of error criterion), we simply set the distance metric as

$$d(x, y) = \begin{cases} 0 & \text{if } x = y \\ 1 & \text{otherwise} \end{cases}$$

Such a choice maximizes the likelihood of estimating the correct state.

2) Euclidean distance: We may suppose that states correspond to physical locations - in this case, we may choose to let the distance $d(\cdot, \cdot)$ correspond to the Euclidean distance between states so that best estimates minimize the error as measured spatially.

B. Dynamic Programming

We use a dynamic programming approach to obtain an optimal policy [13]. Before presenting our algorithm for determining π^*, however, we first develop some important notation. In order to proceed, we must begin by determining several quantities offline. Let $\mathbf{d}(w)$ be the vector of distances of each state from w. Then we proceed by cataloging the quantities

$$\begin{aligned} w_r^*(x) &= arg\min_{w \in S}\left\{\sum_{y \in S} \mathbf{P}[x_r = y | x_0 = x] d(y, w)\right\} \\ &= arg\min_{w \in S}\{(P^r \mathbf{d}(w))(x)\} \\ e_r^*(x) &= (P^r \mathbf{d}(w_r^*(x)))(x) \end{aligned}$$

for $r = 1, ..., N$. The values $w_r^*(x)$ and $e_r^*(x)$ correspond to the optimal estimate and estimation error, respectively, when we must determine the current state given that r time steps ago we observed that the state was x. There may, in some cases, be an efficient way to determine these quantities, but in general we must do this by simply cataloging these quantities offline through brute force. This may be done with relative ease if the state space is of tractable size or if the specific application displays certain sparsity (if our intruder is moving at a bounded rate then we may narrow down his location to a sparse set of states).

Now we proceed to construct the solution using backwards induction. We begin with $t = 1$, which corresponds to one unit of time remaining in the problem, and then continue for $t = 2, 3, ...$ until we are able to determine a recursion. As we build backwards in time (and forward in t), we let s vary and keep track of the cost $J_{r,s,t}(x)$ where x is a state of the Markov chain.

For $t = 1$, we can either have $s = 0$ or $s = 1$. These costs, respectively, are (in vector form)

$$\begin{aligned} J_{(r,0,1)} &= e_r^* \\ J_{(r,1,1)} &= 0 \end{aligned}$$

 ISBN: 978-1-61208-103-8

since not having an observation means we need to make a best estimate, and having an observation leads to zero cost.

Moving on to $t = 2$, the values of s can range from $s = 0$, $s = 1$ or $s = 2$. For $s = 0$ we have

$$J_{(r,0,2)} = e_r^* + e_{r+1}^*$$

since we would need to make an optimal estimate with no further information for the next two time slots. When $s = 1$, there are two choices: use an opportunity to make an observation so that $u = 1$ or do not observe, in which case $u = 0$. These choices can be denoted with superscripts above the cost function for each stage:

$$J^{(0)}_{(r,1,2)}(x) = e_r^*(x) + J_{(r+1,1,1)}(x) = e_r^*(x)$$
$$J^{(1)}_{(r,1,2)}(x) = 0 + \sum_{y\in S} P[x_{N-2} = y|x_{N-2-r} = x]e_1^*(y)$$

For $u = 0$, we accrue error for the current time slot and no error afterwards. When an observation is made, no error is accrued for the current time slot $N-2$, but there is error in the next time slot which depends on the current observation. In vector form, we may write

$$J^{(0)}_{(r,1,2)} = e_r^* + J_{(r+1,1,1)} = e_r^*$$
$$J^{(1)}_{(r,1,2)} = P^r e_1^*$$

We now introduce some new notation:

$$\Delta_{(r,1,2)} = J^{(0)}_{(r,1,2)} - J^{(1)}_{(r,1,2)}$$
$$= e_r^* - P^r e_1^*$$

so that if $\Delta_{(r,1,2)}(x) \leq 0$, then we should not make an observation, whereas we should make an observation if $\Delta_{(r,1,2)}(x) > 0$. We proceed now by defining sets $\tau_{(r,1,2)}$ and $\tau^c_{(r,1,2)}$ such that

$$x \in \tau^c_{(r,1,2)} \Leftrightarrow \Delta_{(r,1,2)}(x) \leq 0$$
$$x \in \tau_{(r,1,2)} \Leftrightarrow \Delta_{(r,1,2)}(x) > 0$$

and we also define an associated vector $\mathbf{1}_{(r,1,2)} \in \{0,1\}^S$

$$\mathbf{1}_{(r,1,2)}(x) = \begin{cases} 1 & \text{if } x \in \tau^c_{(r,1,2)} \\ 0 & \text{otherwise} \end{cases}$$

Moving on to $s = 2$, we have $J_{(r,2,2)} = 0$, since there are as many opportunities to observe the process as there are remaining time slots. We continue with $t = 3$:

$$J_{(r,0,3)} = e_r^* + e_{r+1}^* + e_{r+2}^*$$

since there are three time slots to make estimates for with no new information arriving. For $s = 1$, we again have a choice of $u = 0$ and $u = 1$. For $u = 0$, we accrue a cost for the current stage, and then count the future cost depending on the current state:

$$J^{(0)}_{(r,1,3)}(x) = e_r^*(x) + \mathbf{1}_{(r+1,1,2)}(x)J^{(0)}_{(r+1,1,2)}(x)$$
$$+ (1 - \mathbf{1}_{(r+1,1,2)}(x))J^{(1)}_{(r+1,1,2)}(x)$$

and combining terms gives us

$$J^{(0)}_{(r,1,3)}(x) = e_r^*(x) + J^{(1)}_{(r+1,1,2)}(x)$$
$$+ \mathbf{1}_{(r+1,1,2)}(x)\Delta_{(r+1,1,2)}(x)$$

which after substituting the value of $J^{(1)}_{(r+1,1,2)}(x)$ and putting things in vector form gives us:

$$J^{(0)}_{(r,1,3)} = e_r^* + P^{r+1}e_1^* + diag(\mathbf{1}_{(r+1,1,2)})\Delta_{(r+1,1,2)}$$

Now we consider the $u = 1$ case:

$$J^{(1)}_{(r,1,3)}(x) = 0 + \sum_{y\in S} P[x_{N-3} = y|x_{N-3-r} = x]J_{(1,0,2)}(y)$$
$$= \sum_{y\in S} P[x_{N-3} = y|x_{N-3-r} = x](e_1^*(y) + e_2^*(y))$$

which can be put in vector form:

$$J^{(1)}_{(r,1,3)} = P^r(e_1^* + e_2^*)$$

We now write the expression for $\Delta_{(r,1,3)} = J^{(0)}_{(r,1,3)} - J^{(1)}_{(r,1,3)}$:

$$\Delta_{(r,1,3)} = e_r^* + P^{r+1}e_1^* + diag(\mathbf{1}_{(r+1,1,2)})\Delta_{(r+1,1,2)}$$
$$- P^r(e_1^* + e_2^*)$$

Continuing with $s = 2$,

$$J^{(0)}_{(r,2,3)}(x) = e_r^*(x) + 0$$

whereas for $u = 1$,

$$J^{(1)}{}_{(r,2,3)}(x) = 0 + \sum_{y\in S} P[x_{N-3} = y|x_{N-3-r} = x]$$
$$\Big(\mathbf{1}_{(1,1,2)}(y)J^{(0)}_{(1,1,2)}(y) + (1 - \mathbf{1}_{(1,1,2)}(y))J^{(1)}_{(1,1,2)}(y)\Big)$$

where we have accounted for the cost stage by stage: in the current stage, no error is accrued since an observation is made but future costs depend on the observation that is made. That is, future costs depend on whether the current state x_{N-3} is observed to be in the set $\tau_{(1,1,2)}$. Averaging over these, we obtain the expression above. Combining like terms as above, we arrive at:

$$J^{(1)}_{(r,2,3)}(x) = 0 + \sum_{y\in S} P[x_{N-3} = y|x_{N-3-r} = x]$$
$$\Big(J^{(1)}_{(1,1,2)}(y) + \mathbf{1}_{(1,1,2)}(y)\Delta_{(1,1,2)}(y)\Big)$$

Substituting the expression for $J^{(1)}_{(1,1,2)}(y)$, we get

$$J^{(1)}_{(r,2,3)}(x) = \sum_{y\in S} P[x_{N-3} = y|x_{N-3-r} = x]$$
$$\Bigg(\sum_{z\in S} P[x_{N-2} = z|x_{N-3} = y]e_1^*(z)$$
$$+\mathbf{1}_{(1,1,2)}(y)\Delta_{(1,1,2)}(y)\Bigg)$$

We simplify the expression by bringing the first summation in the parentheses. Then we apply the Kolmogorov-Chapman equation to get

$$J^{(1)}{}_{(r,2,3)}(x) = \sum_{z\in S} P[x_{N-2} = z | x_{N-3-r} = x] e_1^*(z) + \sum_{y\in\tau^c_{(1,1,2)}} P[x_{N-3} = y | x_{N-3-r} = x]\Delta_{(1,1,2)}(y)$$

Putting this into vector form, we have the expression:

$$J^{(1)}{}_{(r,2,3)} = P^{r+1}e_1^* + P^r \mathbf{1}_{(1,1,2)}\Delta_{(1,1,2)}$$

We use these expressions to get $\Delta_{(r,2,3)}$.

$$\Delta_{(r,2,3)} = e_r^* - P^{r+1}e_1^* - P^r\mathbf{1}_{(1,1,2)}\Delta_{(1,1,2)}$$

Finally, letting $s = 3$, we get

$$J_{(r,3,3)}(x) = 0$$

This process can be continued for $t = 4, 5, ...$ For each stage (r, s, t), we may determine $J^{(0)}_{(r,s,t)}$ and $J^{(1)}_{(r,s,t)}$. These costs then allow us to determine when we should make an observation in the process and when we should not. The implementation of this policy is detailed in the following subsection.

C. Solution

We now present a method for constructing an optimal policy. We do this by storing for each (r, s, t) a subset of S, denoted by $\tau^c_{(r,s,t)}$, which is the set of last observed states for which we do not use an opportunity to view the process when we are at stage (r, s, t). That is, if the last observed state x was seen r time slots ago, it is in the set $\tau^c_{(r,s,t)}$, there are s opportunities remaining to make observations and there are t time slots remaining in the horizon then we should not make an observation at this time and simply make an estimate $w_r^*(x)$. On the other hand, if $x \in \tau_{(r,s,t)}$ then we should make an observation at stage (r, s, t) and accrue zero cost for that stage.

More precisely, an optimal policy π^* is given by

$$u_{(r,s,t)}(x) = \begin{cases} 0 & \text{if } x \in \tau^c_{(r,s,t)} \\ 1 & \text{otherwise} \end{cases}$$

Let us introduce three vector valued functions: $F_{(r,s,t)}, \Delta_{(r,s,t)} \in \mathbf{R}^S$ and $\mathbf{1}_{(r,s,t)} \in \{0,1\}^S$. We fill in values for these functions by using the following recursions:

$$F_{(r,s,t)} = F_{(r+1,s-1,t-1)} + P^r\mathbf{1}_{(1,s-1,t-1)}\Delta_{(1,s-1,t-1)}$$
$$\Delta_{(r,s,t)} = e_r^* + F_{(r+1,s,t-1)} - F_{(r,s,t)} + diag(\mathbf{1}_{(r+1,s,t-1)})\Delta_{(r+1,s,t-1)}$$
$$\mathbf{1}_{(r,s,t)}(x) = \begin{cases} 0 & \text{if } \Delta_{(r,s,t)}(x) > 0 \\ 1 & \text{otherwise} \end{cases}$$

for $1 < s < t < N$ and $1 \le r \le N - t + 1$. We also have the boundary conditions

$$F_{(r,t,t)} = 0, \quad F_{(r,1,t)} = P^r \sum_{j=1}^{t-1} e_j^*, \quad \Delta_{(r,t,t)} = e_r^*$$

These recursions allow us to determine the sets $\tau^c_{(r,s,t)}$ for s, t, r in the bounds specified, which in turn defines our optimal policy. Specifically, we assign

$$x \in \tau^c_{(r,s,t)} \Leftrightarrow \Delta_{(r,s,t)}(x) \le 0$$

We conclude by giving expressions for the cost-to-go from any particular state when a particular action $u \in \{0, 1\}$ is taken. The superscripts denote whether or not an observation will be made in the current stage.

$$J^{(0)}_{(r,s,t)} = e_r^* + F_{(r+1,s,t-1)} + diag(\mathbf{1}_{(r+1,s,t-1)})\Delta_{(r+1,s,t-1)}$$
$$J^{(1)}_{(r,s,t)} = F_{(r,s,t)}$$

Observe that $\Delta_{(r,s,t)}$ is the difference between these two quantities. Hence, $\Delta_{(r,s,t)}$ functions as a method of determining whether or not to make an observation in the current time step.

We note that although the curse of dimensionality can make the operations required for the solution to be intractable for large scale problems, the structure of specific problems may allow us to generate good approximations to the solution. For medium sized problems, we see that with the given algorithms we do not need to conduct any sort of value iteration to converge at the optimum, but rather the dynamic programming has been reduced to matrix multiplications. Hence, the algorithm provided here outperforms conventional Dynamic Programming tools such as Dynamic Programming via Linear Programming or value iteration because this algorithm has been tailored to our specific problem. In the following section we apply our results to small example problems.

III. Numerical Results

Let us now examine the performance of our algorithm. We shall fix a horizon length and plot the cost that the prescribed algorithm accrues versus the number of opportunities to make observations. Let us consider Markov chains of the type $\mathcal{M}_{n\times n}$ in Fig. 1, which is an n-by-n grid of states where the transition probabilities are given in the figure. Such a construction is simple enough for quick simulation but can capture the inherent variations which our algorithm is able to leverage.

A. Surveillance

Suppose we would like to track the position of an intruder in an environment modeled by the Markov chain $\mathcal{M}_{3\times3}$ over a discrete-time horizon of 30 time slots. However, updating the location of the intruder requires battery power of a mobile device due to communications with a satellite and

 ISBN: 978-1-61208-103-8

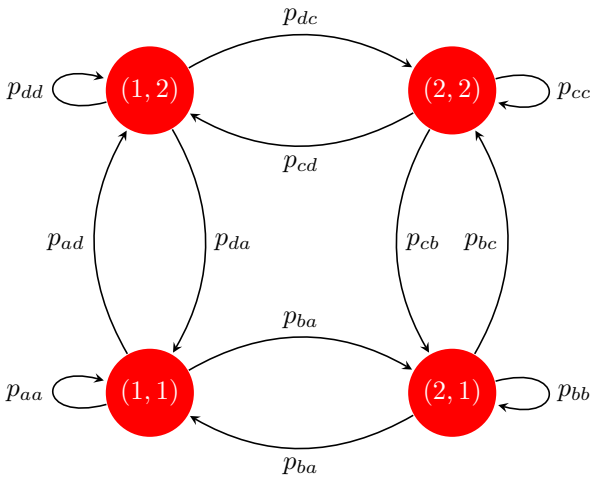


Figure 1. Markov chain $\mathcal{M}_{2\times 2}$

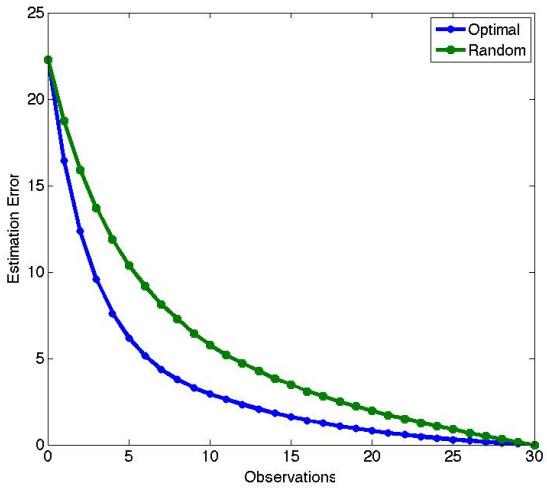


Figure 2. Plot of Optimal Error vs. Number of Observation Opportunities

hence we are not able to request the position of the intruder at every time. Fixing the initial position of the device to be $(2,1)$, let us vary the number of opportunities to retrieve the true location from 0 to 30. The distance metric we take is the standard Euclidian norm, which may be represented in matrix form as:

$$D = \begin{bmatrix} 0 & 1 & 2 & 1 & \sqrt{2} & \sqrt{5} & 2 & \sqrt{5} & \sqrt{8} \\ 1 & 0 & 1 & \sqrt{2} & 1 & \sqrt{2} & \sqrt{5} & 2 & \sqrt{5} \\ 2 & 1 & 0 & \sqrt{5} & \sqrt{2} & 1 & \sqrt{8} & \sqrt{5} & 2 \\ 1 & \sqrt{2} & \sqrt{5} & 0 & 1 & 2 & 1 & \sqrt{2} & \sqrt{5} \\ \sqrt{2} & 1 & \sqrt{2} & 1 & 0 & 1 & \sqrt{2} & 1 & \sqrt{2} \\ \sqrt{5} & \sqrt{2} & 1 & 2 & 1 & 0 & \sqrt{5} & \sqrt{2} & 1 \\ 2 & \sqrt{5} & \sqrt{8} & 1 & \sqrt{2} & \sqrt{5} & 0 & 1 & 2 \\ \sqrt{5} & 2 & \sqrt{5} & \sqrt{2} & 1 & \sqrt{2} & 1 & 0 & 1 \\ \sqrt{8} & \sqrt{5} & 2 & \sqrt{5} & \sqrt{2} & 1 & 2 & 1 & 0 \end{bmatrix}$$

and we choose the transition matrix to be

$$P = \begin{bmatrix} 0 & 0.1 & 0 & 0.9 & 0 & 0 & 0 & 0 & 0 \\ 0.1 & 0 & 0.9 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.1 & 0.8 & 0 & 0 & 0.1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.2 & 0 & 0.8 & 0 & 0 \\ 0 & 0.7 & 0 & 0.15 & 0 & 0.15 & 0 & 0 & 0 \\ 0 & 0 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0.9 & 0 & 0 & 0 & 0.1 & 0 \\ 0 & 0 & 0 & 0 & 0.8 & 0 & 0 & 0 & 0.2 \\ 0 & 0 & 0 & 0 & 0 & 0.5 & 0 & 0.4 & 0.1 \end{bmatrix}$$

where we have ordered the states by the first index and then the second (that is in order $(1,1),(1,2),(1,3),(2,1),..$).

We expect the estimation error to monotonically decrease with the number of opportunities to learn the true state. In Fig. 2, we see that this indeed the case, and also compare it to a benchmark strategy of randomly distributing observations.

B. Analysis of Performance

We now note several properties of our curve in Fig. 2. First, the endpoints are fixed no matter what policy is used - this is because when there are zero opportunities to make observations or there are 30 chances to view the process, there is no way to come up with policies that result in different decisions. There is only one way to allocate opportunities to observe the process. Next, we note that our algorithm outperforms a benchmark strategy of randomly placing observations over the 30 time slots. We see that the greatest "savings" occurs when we have a sparsity of opportunities to make observations. This is the case in most practical situations.

Finally, we observe the convexity of the curve. This is interpreted to mean that as opportunities to observe the process are more readily available, there is a law of diminishing returns and these opportunities become less valuable. The degree of convexity depends greatly on the transition matrix P of the Markov chain. For example, if the grid $\mathcal{M}_{n\times n}$ has transitions that are all equal, the benchmark and our algorithm both produce a straight line. This is because there is no variation in the Markov chain to exploit.

IV. Conclusion and Future Work

In this paper, we have developed a problem in monitoring over a finite horizon when there are a limited number of opportunities to conduct surveillance. We mathematically model this as a problem of state estimation. In the estimation problem we hope to minimize the distortion from estimating the state of a Markov chain when the number of time the process may be viewed is limited to a few times over the total horizon. The distortion is measured using a specified metric $d(x,y)$ which tells us how "far apart" states x and y are.

 ISBN: 978-1-61208-103-8

In our optimal policy, a set of recursive equations with boundary conditions give a practical method for determining an optimal policy. Although the policy could have been determined using standard methods in dynamic programming, such as value iteration, the algorithm given here relies only on the ability to store data and conduct matrix multiplications. Hence, larger problems can be handled before intractability results due to state space complexity.

There are many further problems to consider in future work. If the state space complexity becomes unmanageable, we must develop policies that are near optimal or find some other way around the complexity using approximation schemes. Also, we may consider problems with a variable horizon length. That is, we might consider problems in which the Markov chain dictates a random stopping time for the process during which we may only make observations a limited number of times. Additionally, there are practical scenarios in which one does not have complete information about the transition matrix. In this case, we may be interested in coupling parameter estimation with efficient budget allocation. Finally, we can generalize the model so that observations not only are limited in number but also carry a cost per usage.

ACKNOWLEDGMENT

The first author is supported by the Department of Defense (DoD) through the National Defense Science & Engineering Graduate Fellowship (NDSEG) Program.

REFERENCES

[1] E. Wilson, *Network Monitoring and Analysis: A Protocol Approach to Troubleshooting*. Prentice Hall, 2000.

[2] D. Josephsen, *Building a Monitoring Infrastructure with Nagios*. 1st ed., Prentice Hall, 2007.

[3] Microsoft Security Center. Retrieved from http://technet.microsoft.com/en-us/security. May, 2010.

[4] General Accounting Office. Information Security: Computer Attacks at Department of Defense Pose Increasing Risks. GAO/AIMD-96-84, May, 1996.

[5] R. A. Miura-Ko and N. Bambos, "Dynamic risk mitigation in computing infrastructures," in *Third International Symposium on Information Assurance and Security*. IEEE, 2007, pp. 325 - 328.

[6] K. C. Nguyen, T. Alpcan, and T. Basar, "Fictitious play with imperfect observations for network intrusion detection," *13th Intl. Symp. Dynamic Games and Applications (ISDGA)* Wroclaw, Poland, June 2008.

[7] T. Alpcan and X. Liu, "A game theoretic recommendation system for security alert dissemination," in *Proc. of IEEE/IFIP Intl. Conf. on Network and Service Security (N2S 2009)*, Paris, France, June 2009.

[8] H.V. Poor, *An Introduction to Signal Detection and Estimation*. 2nd ed., Springer-Verlag, 1994.

[9] N. E. Nahi, "Optimal recursive estimation with uncertain observation," in *IEEE Transactions on Information Theory*, vol. 15, no. 4, pp. 457 - 462, July 1969.

[10] B. Sinopoli, L. Schenato, M. Franceschetti, K. Poolla, M. I. Jordan, and S. S. Sastry, "Kalman filtering with intermittent observations," *IEEE Transactions on Automatic Control*, vol. 49, no. 9 , pp. 1453 - 1464, September 2004.

[11] O.C. Imer. Optimal Estimation and Control under Communication Network Constraints. Ph.D. Dissertation, UIUC, 2005.

[12] P. Bommannavar and N. Bambos, Patch Scheduling for Risk Exposure Mitigation Under Service Disruption Constraints. Technical Report, Stanford University, 2010.

[13] D. P. Bertsekas, *Dynamic Programming and Optimal Control*. Belmont, MA: Athena Scientic, 1995.

 ISBN: 978-1-61208-103-8

An Enhanced RED-Based Weighted Fair Priority Queuing Algorithm for IEEE 802.16 Subscriber Station Scheduler

Serda Kasaci, Sema Oktug
Department of Computer Engineering
Istanbul Technical University
Istanbul, Turkey
e-mail: {kasaci, oktug}@itu.edu.tr

Abstract— **IEEE 802.16, which is called as Worldwide Interoperability for Microwave Access (WiMAX), is an air interface for Fixed Broadband Wireless Access Systems. In the 802.16 IEEE standard, different service types are introduced, such as Unsolicited Grant Service (UGS), Real-time Polling Service (rtPS), Non-Real-time Polling Service (nrtPS) and Best Effort (BE). Each service type is associated with a set of Quality of Service (QoS) parameters; however WiMAX does not specify how to schedule the granted bandwidth efficiently between these service classes. In this paper, we propose an Enhanced RED-based Weighted Fair Priority Queuing algorithm for Subscriber Stations. The weights are calculated according to the traffic load of the rtPS and nrtPS service classes. Simulation results show that, both rtPS and nrtPS throughputs are improved without starving lower priority service classes.**

Keywords- wimax; scheduling; uplink; GPSS; QoS

I. INTRODUCTION

The IEEE 802.16 standard, widely known as WiMAX, has been developed for Broadband Wireless Access (BWA). The advantages of this standard are easy and low-cost deployment, high speed data rate, last mile wireless access, and QoS support for multimedia applications [1]. The standard defines two possible network topologies, such as *Point-to-Multipoint (PMP)* and *Mesh* Networks. In the PMP networks, communication between Subscriber Stations (SSs) is possible only through a Base Station (BS). In the mesh mode, SSs can communicate with each other directly. In this paper, we employ PMP topology.

The standard IEEE 802.16 defines the physical layer and the *MAC (Medium Access Control)* layer. The main purpose of the MAC protocol is to share radio channel resources among multiple accesses of different users. As the MAC protocol is connection-oriented, all data transmission takes place in connections, even for connectionless packets. The MAC layer contains three sublayers such as: Convergence Sublayer (CS), Common Part Sublayer (CPS), and the Security Sublayer. CS accepts Protocol Data Units (PDUs) from higher layers. The MAC SDUs are classified and mapped into appropriate Connection IDentifiers (CIDs) and they are transmitted to CPS by CS. CPS is responsible for fragmentation and segmentation of each MAC SDU into MAC PDUs, system access, bandwidth allocation, connection maintenance, QoS control, and scheduling transmission. The Security Sublayer is responsible for security, authentication, and encryption.

The PHY Layer establishes the physical connection between uplink and downlink directions. This layer is responsible for transmission of the bit sequences. There are two dublexing techniques for PHY layer of downlink and uplink such as; Frequency Division Duplex (FDD) and Time Division Duplex (TDD). FDD requires two distinct channels to transmit downlink sub-frame and uplink sub-frame at the same time slot. In TDD, downlink (DL) and uplink (UL) subframes share the same frequency; but they take place at different times. DL Subframe has DL-MAP, UL-MAP and DL PHY PDUs. The DL-MAP message defines the usage of the downlink intervals. The UL-MAP defines the uplink usage in terms of the offset of the burst relative to the Allocation Start Time [2]. UL Subframe contains contention slot for initial ranging, contention slot for bandwidth requests and UL PHY PDUs from SSs. Via Initial Ranging IE, BS provides an interval for new stations to join to the network. Ranging Request (RNG-REQ) packets are used in this interval. Via Request IE, BS specifies an uplink interval which can be used by SS to send a bandwidth requests using contention slots.

There are four service types defined in IEEE.802.16-2004; Unsolicited Grant Service (UGS), real-Time Polling Service (rtPS), non-real-time Polling Service (nrtPS), and Best Effort (BE). In the 802.16e standard [3], a new service type, called extended real time Polling Service (ertPS), has been added. However, it is out of the scope of this paper.

UGS supports Constant Bit Rate (CBR) flows for real-time applications; such as VoIP without silence suppression or E1/T1 data streams. The BS allocates fixed sized data grants at periodic intervals based on the Maximum Sustained Traffic Rate of the service flow. The overhead and latency of SS requests are eliminated for UGS connections. However, UGS is more expensive than other service types.

Variable Bit Rate (VBR) flows, which have variable packet length and periodic packet intervals, such as Moving Pictures Expert Group video, are supported by rtPS. BS provides unicast request opportunities to SSs periodically.

Variable-sized packets, which are delay-tolerant data streams, such as File Transfer Protocol (FTP), are supported by nrtPS. Therefore the minimum data rate is required for this service. BS provides unicast request opportunities periodically as in the rtPS service, so this will guarantee data

 ISBN: 978-1-61208-103-8

granting during network congestion. In addition to this, SSs can use contention request mechanism.

Best Effort is designed for best-effort traffic such as HTTP, and this service does not have any minimum service guarantee. SS can use contention request opportunities to send any bandwidth request.

A bandwidth request may be a standalone BW request header or it may come as a Piggyback Request. Some policies are used to send the BW such as unicast or multicast polling, using contention request slots or setting Poll-Me bit. There are two types of BW Requests: incremental and aggregate. BW is requested on a CID basis, but bandwidth grants are allocated on an SS basis. IEEE 802.16 MAC accommodates two modes of SS, differentiated by their ability to accept bandwidth grants simply for a connection or for the SS as a whole. In Grant Per Connection (GPC) mode, bandwidth is granted to a connection, so SS can use this grant for this connection only. In the Grant Per Subscriber (GPSS) mode, the BS grants bandwidth to an SS as an aggregate of grants in response to per connection requests from the SS. Then the SS distributes bandwidth among its connections, with respect to their QoS requirements. Therefore, the GPSS mode is more complex than the GPC mode.

A scheduling algorithm has to determine the allocation of the bandwidth among the users and their transmission order. QoS requirements of the users need to be satisfied while utilizing the available bandwidth efficiently [4]. There are two types of schedulers: the SS Scheduler and the BS scheduler. The SS Scheduler is more complicated in the GPSS mode, as the algorithm which works in the SS scheduler distributes the granted bandwidth between its connections [5]. As the WiMAX standard does not specify how to efficiently schedule traffic to fulfill QoS requirements, a lot of research has been done on this topic. Several works have introduced algorithms for the schedulers in the Base Station (BS) and the Subscriber Station (SS).

In this paper, we focus on the GPSS type of SS scheduler and their performance. Strict Priority (SP), Weighted Fair Priority Queuing (WFPQ), and RED-based Deficit Fair Priority Queuing (DFPQ) are investigated. We propose an enhanced RED-based WFPQ algorithm to increase both rtPS and nrtPS throughput. This algorithm is called as Enhanced RED-based Weighted Fair Priority Queuing and has a dynamic structure while granting bandwidth between service classes. This algorithm takes the packet size information of rtPS and nrtPS, and then calculates the weights of the service flows based on the RED technique.

The rest of the paper is organized as follows. Section II presents previously introduced scheduling algorithms for SS Schedulers in PMP WiMAX networks. The proposed scheduling algorithm is described in Section III. Simulation results are shown and discussed in Section IV. Section V concludes the paper by giving future directions.

II. ALGORITHMS FOR SS SCHEDULER IN 802.16 NETWORKS

Several scheduling algorithms have been proposed for SS schedulers in PMP WiMAX networks to improve the performance of the system; such as Strict Priority, Weighted Fair Priority Queuing, and RED-based Deficit Fair Priority Queuing.

A. Strict Priority:

Bandwidth is allocated for rtPS service flows first, then bandwidth is allocated for nrtPS service flows, and finally the remaining bandwidth is allocated for BE service flows. Consequently, under heavy rtPS traffic load, nrtPS and BE service flows may starve. Strict Priority scheduling does not guarantee the QoS requirements of the traffic that comes from lower priority service classes.

B. Weighted Fair Priority Queuing

WFPQ scheduling is a generalization of Fair Queuing. WFPQ allows different sessions to have different service shares. A link data rate (R), is serviced for the active data flows (N). The data rate of session j is calculated as follows:

$$R_j = \frac{R \times w_j}{\sum_{i=1}^{N} w_i} \quad (1)$$

where w_j represents the weight assigned to session j.

According to (1), the available bandwidth is shared between the service types in the SS Scheduler. Therefore, we need to define the weights for service types efficiently. For example, as the priority of rtPS is higher than nrtPS, rtPS needs to be given a higher weight than nrtPS.

C. RED-based Deficit Fair Priority Queuing

Chen et al. proposed the Deficit Fair Priority Queuing based scheduler for bandwidth allocation among the service classes of WiMAX networks [6]. It uses Deficit Counters (DCs) for rtPS, nrtPS, and BE. In Fig. 1, the DC for rtPS service class is adaptively calculated according to RED technique. If the current packet length of the rtPS queue ($QL_{current}$) is less than $QL_{threshold1}$, the DC value will be equal to DC_{min}. If the $QL_{current}$ is between $QL_{threshold1}$ and $QL_{threshold2}$, DC will be equal to $DC_{dynamic}$. The $DC_{dynamic}$ is calculated using (2). If the $QL_{current}$ is more than $QL_{threshold2}$, DC equals to DC_{max}.

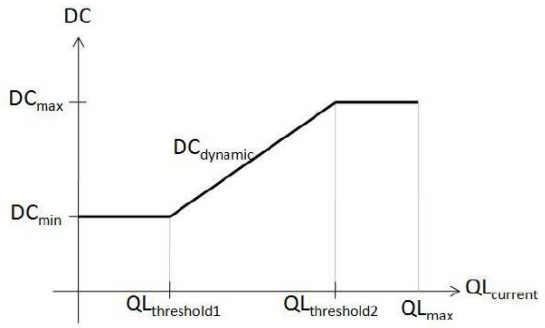


Figure 1. RED-based Deficit Fair Priority Queuing [6]

$$DC_{\min} = Q_{rtPS}$$
$$DC_{dynamic} = Q_{rtPS} + \frac{QL_{current} - QL_{threshold\ 1}}{QL_{threshold\ 2} - QL_{threshold\ 1}} \times Q_{rtPS} \quad (2)$$
$$DC_{\max} = 2 \times Q_{rtPS}$$

Following the transmission of rtPS packets, nrtPS packets will be transmitted. If there is no rtPS or nrtPS packet left, scheduler transmits BE packets.

III. PROPOSED ALGORITHM

In this paper, we propose a new SS Scheduler algorithm which is called Enhanced RED-based Weighted Fair Priority Queuing and derived from RED-based Weighted Fair Priority Queuing (WFPQ). RED-based WFPQ is simple than RED-based DFPQ, as we do not deal with deficit counters; we only determine weights for the service types.

A. RED-Based Weighted Fair Priority Queuing

When the rtPS queue length is lower than $QL\_Th_{rtps\_min}$, W_{rtps_min} is assigned for the weight of the rtPS service flow. When the rtPS queue length is higher than $QL\_Th_{rtps\_max}$, W_{rtps_max} is assigned to rtPS. When the rtPS queue length is between $QL\_Th_{rtps\_min}$ and $QL\_Th_{rtps\_max}$, the rtPS weight changes dynamically according to the rtPS queue length. The slope of W_{rtps} is calculated according to (3). Equation (5) represents for the weight assignment of rtPS.

$$m_{rtps} = \frac{W_{rtps\_max} - W_{rtps\_min}}{QL\_Th_{rtps\_max} - QL\_Th_{rtps\_min}} \quad (3)$$

According to Fig. 2, W_{rtps} is calculated at the beginning of every frame by using the diagram. The rest of the available weights are distributed between nrtPS and BE flows according to their weights (W_{nrtps} and W_{BE}).

B. Enhanced RED-Based Weighted Fair Priority Queuing

In Enhanced RED-based WFPQ algorithm, we do not define static weight for nrtPS. We apply the dynamic weight assignment of RED-based WFPQ algorithm to nrtPS service types. As the dynamic weight assignment is used for both rtPS and nrtPS, we call this algorithm "Enhanced RED-based WFPQ" algorithm.

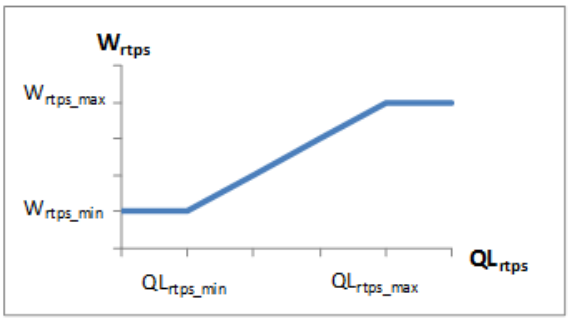


Figure 2. RED-based weights for rtPS service class

TABLE I. RTPS QUEUE LENGTH CONDITIONS

Condition-1	[$QL_{rtps} < QL\_Th_{rtps\_min}$]
Condition-2	[$QL\_Th_{rtps\_min} < QL_{rtps} < QL\_Th_{rtps\_max}$]
Condition-3	[$QL_{rtps} > QL\_Th_{rtps\_max}$]

The weight assignment of the rtPS service type is the same as in RED-based WFPQ. As the nrtPS weight depends on the variation of the rtPS weight (total weight is distributed between all service types), we need to consider three conditions while determining the nrtPS weight. The conditions are given in Table I. rtPS values are determined based on the conditions; therefore nrtPS weight characteristics are dynamic and depend on rtPS weight.

Condition-1

In Fig. 2, when the rtPS queue length (QL_{rtps}) is lower than $QL\_Th_{rtps\_min}$, the rtPS weight is set to the predefined W_{rtps_min} value. Weight assignment of nrtPS is represented in (6). When nrtPS queue length (QL_{nrtps}) is lower than the minimum threshold of nrtPS, W_{nrtps_min} is assigned as the weight of nrtPS. When QL_{nrtps} is between minimum and maximum threshold values, the nrtPS weight varies dynamically. When QL_{nrtps} is higher than the maximum threshold of nrtPS, W_{nrtps_max} is assigned as the nrtPS weight. Fig. 3 represents for the RED-based weights for nrtPS service class. In each condition, weights for BE service types are calculated according to (4).

$$W_{BE} = W_{total} - W_{rtps} - W_{nrtps} \quad (4)$$

$$W_{rtps} = \begin{cases} W_{rtps\_min} & if\ (QL_{rtps} \le QL\_Th_{rtps\_min}) \\ W_{rtps\_min} + m_{rtps} \times (QL_{rtps} - QL\_Th_{rtps\_min}) & if\ (QL\_Th_{rtps\_min} < QL_{rtps} < QL\_Th_{rtps\_max}) \\ W_{rtps\_max} & if\ (QL_{rtps} \ge QL\_Th_{rtps\_max}) \end{cases} \quad (5)$$

 ISBN: 978-1-61208-103-8

$$W_{nrtps} = \begin{cases} W_{nrtps\_min} & if\left(QL_{nrtps} \le QL\_Th_{nrtps\_min}\right) \\ W_{n\,rtps\_min} + m_{n\,rtps} \times \left(QL_{n\,rtps} - QL\_Th_{n\,rtps\_min}\right) & if\left(QL\_Th_{nrtps\_min} < QL_{nrtps} < QL\_Th_{nrtps\_max}\right) \\ W_{nrtps\_max} & if\left(QL_{nrtps} \ge QL\_Th_{nrtps\_max}\right) \end{cases} \quad (6)$$

$$W_{nrtps} = \begin{cases} W_{nrtps\_min} = \left(W_{Total} - W_{BE\_min} - \left(QL\_Th_{nrtps\_max} - QL\_Th_{nrtps\_min}\right) \times m_{n\,rtps} - W_{rtps}\right) & if\ \left(QL_{nrtps} \le QL\_Th_{nrtps\_min}\right) \\ W_{n\,rtps\_min} + m_{n\,rtps} \times \left(QL_{n\,rtps} - QL\_Th_{n\,rtps\_min}\right) & if\ \ \left(QL\_Th_{nrtps\_min} < QL_{nrtps} < QL\_Th_{nrtps\_max}\right) \\ W_{nrtps\_max} = W_{Total} - W_{rtps} - W_{BE\_min} & if\left(QL_{nrtps} \ge QL\_Th_{nrtps\_max}\right) \end{cases} \quad (7)$$

Condition-2

In Fig. 2, when the rtPS queue length (QL_{rtps}) is lower than $QL\_Th_{rtps\_max}$ and higher than $QL\_Th_{rtps\_min}$, the rtPS weight is set dynamically according to queue length by using (5). Consequently, the available weight, that remains for nrtPS and BE service type, is ($W_{total} - W_{rtps}$). Fig. 3 displays the variation of the nrtPS weight is RED-based, and details are given in (7).

In each condition, weights for BE service types are calculated according to (4). We reserve a little bandwidth for BE flow (W_{BE_min}) to prevent from starving in a congested network.

Condition-3

In Condition-3, rtPS queue length (QL_{rtps}) is higher than $QL\_Th_{rtps\_max}$. Consequently, the rtPS weight is set to W_{rtps_max}. In this condition, nrtPS and BE weights are statically assigned. The possible maximum value for nrtPS (W_{nrtps_max}) is assigned to W_{nrtps}. The weight of BE is calculated according to (4).

In all three conditions, the maximum values of nrtPS are not the same, as the weight intervals depend on the weight rtPS.

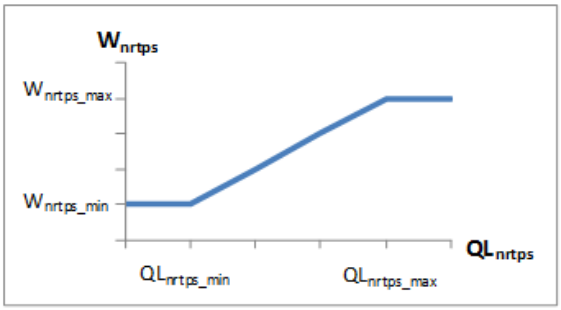


Figure 3. RED-based weights for nrtPS service class

IV. SIMULATION RESULTS

In this paper, the simulations are performed by using IEEE 802.16 WiMAX NIST [7] module which has been developed on NS-2 version 2.29 [8]. We used the WIMAX QoS patch which is designed for NIST WIMAX module [9, 10]. We added nrtPS service class to the patch. The fundamental simulation parameters are shown in Table II. The existing QoS-included WiMAX Patch supports only one connection per subscriber, so we modified the patch to support GPSS mode. We run simulation for throughput analysis 5 times to achieve results with 95% confidence interval.

TABLE II. SIMULATION PARAMETERS

PHY specification	WirelessMAN-OFDM
Frequency Band	5MHz
Antenna Model	Omni Antenna
Antenna Height	1.5 m
Propagation Model	TwoRayGround
Transmit Antenna Gain	1
Transmit Power	0.25 W
Frame Duration	20 ms
Cyclic Prefix	0.025 s
Simulation Duration	100 s
Packet Length	1000 bytes
Frame Structure	TDD

 ISBN: 978-1-61208-103-8

TABLE III. RED-BASED WFPQ SCHEDULER PARAMETERS

Variables	Selected Values
$W_{rtps\ min}$	0.5
$W_{rtps\ max}$	0.7
$QL\_Th_{rtps\ min}$	10 packets
$QL\_Th_{rtps\_max}$	30 packets
$QL_{rtps\ max}$	50 packets
W_{nrtps}	W_{nrtps}
W_{BE}	(2/3) W_{nrtps}
m_{rtps}	0.01
W_{Total}	1

TABLE IV. ENHANCED RED-BASED WFPQ SCHEDULER PARAMETERS IN CONDITION-1

Variables	Selected Values
W_{nrtps_min}	0.35
W_{nrtps_max}	0.45
$QL\_Th_{nrtps\_min}$	10 packets
$QL\_Th_{nrtps\_max}$	30 packets
QL_{nrtps}	50 packets
m_{nrtPS}	0.05
W_{rtps_min}	0.5

The scheduler parameters used throughout the simulations are given in Table III. Fig. 4 shows the behavior of RED-based WFPQ when we use the values in Table III. As we use the same values for rtPS parameters in RED-based WFPQ and Enhanced RED-based WFPQ, their weight graphs are the same. The parameters of Enhanced RED-based WFPQ in Condition-1, which are used for nrtPS weight determination, are chosen as in Table IV.

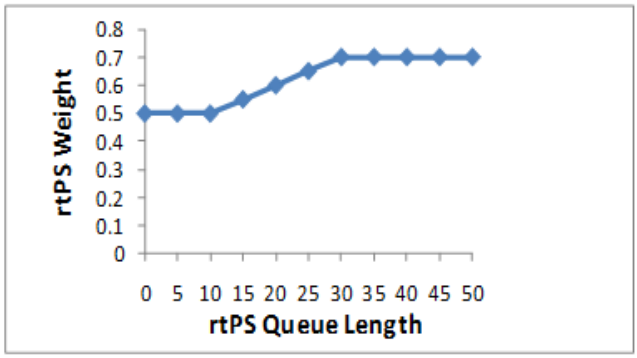


Figure 4. RED-based weights for rtPS service flow

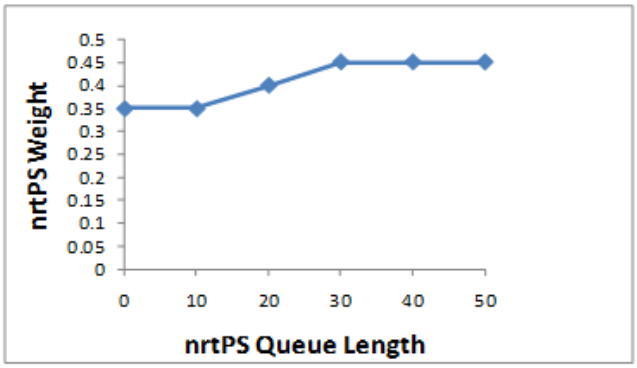


Figure 5. nrtPS weights of Enhanced RED-based WFPQ in Condition-1

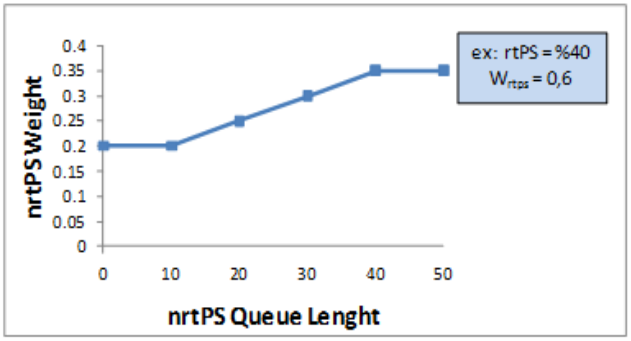


Figure 6. nrtPS weights of Enhanced RED-based WFPQ in Condition-2

Fig. 5 shows the behavior of the algorithm when we use the parameter values in Table V.

Parameters of nrtPS and rtPS, which are used in Condition-2 for Enhanced RED-based WFPQ, are given in Table V. Fig. 6 shows the behavior of the algorithm when we use the parameter values nrtPS as in Table V. In this condition, the diagram depends on rtPS weights. Therefore, to draw a diagram for nrtPS weights, we need to select an rtPS weight value. In our example, we take 0.6 as an example for rtPS.

TABLE V. ENHANCED RED-BASED WFPQ SCHEDULER PARAMETERS IN CONDITION-2

Variables	Selected Values
W_{nrtps_min}	0.2
W_{nrtps_max}	0.35
$QL\_Th_{nrtps\_min}$	10 packets
$QL\_Th_{nrtps\_max}$	40 packets
QL_{nrtps}	50 packets
m_{nrtPS}	0.05
rtPS weight example	0.6
rtPS QL percentage	40%
W_{BE_min}	0.05

 ISBN: 978-1-61208-103-8

TABLE VI. ENHANCED RED-BASED WFPQ SCHEDULER PARAMETERS IN CONDITION-3

Variables	Selected Values
$W_{rtps} = W_{rtps\_max}$	0.7
$W_{nrtps} = W_{nrtps\_max}$	0.25
$W_{BE} = W_{Total} - W_{rtps\_max} - W_{nrtps\_max}$	0.05

Parameters of service flows for Condition-3, are given in Table VI. W_{BE} is calculated according to (4). We need to subtract W_{rtps} and W_{nrtps} from W_{Total}.

We measured throughput of rtPS, nrtPS and BE service class flows. We also calculated the queuing delay, dropped packet percentage and fairness index. We compared the following schedulers with each other:

- Strict Priority Scheduling
- WFPQ Scheduling
- RED-Based WFPQ Scheduling
- Enhanced RED-Based WFPQ Scheduling

In Fig. 7, we consider the rtPS throughput versus increasing rtPS traffic load. Strict Priority scheduling has the maximum throughput level as the algorithm always grants bandwidth for rtPS first, if there is no packet in the rtPS queue and there is available bandwidth left for the SS, then the bandwidth is allocated for the nrtPS service flow. If there are no packets in rtPS and nrtPS queues and there is available bandwidth left for the SS, then the bandwidth is allocated to the BE service flow. In the WFPQ algorithm, as the weights are chosen statically, we cannot increase the throughput of rtPS significantly while increasing rtPS load. RED-based WFPQ and Enhanced RED-based WFPQ have higher rtPS throughput as they can dynamically change the weights of rtPS according to the queue length of the rtPS flow. Initial weights are the same in WFPQ and RED algorithms. However, as rtPS load submission increases, due to higher rtPS traffic, the queue length will also increase. Consequently, RED-based WFPQ and Enhanced RED-based WFPQ yield better performance than WFPQ.

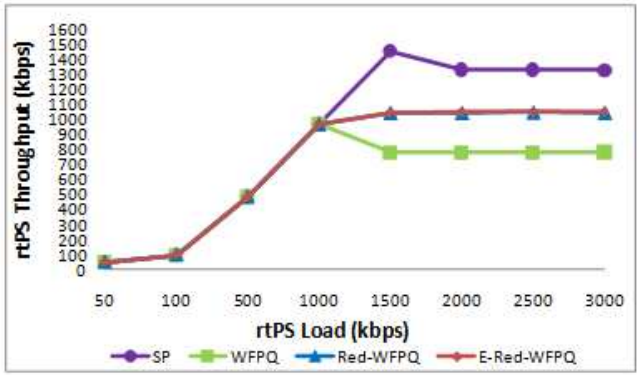


Figure 7. Comparison of rtPS Throughput

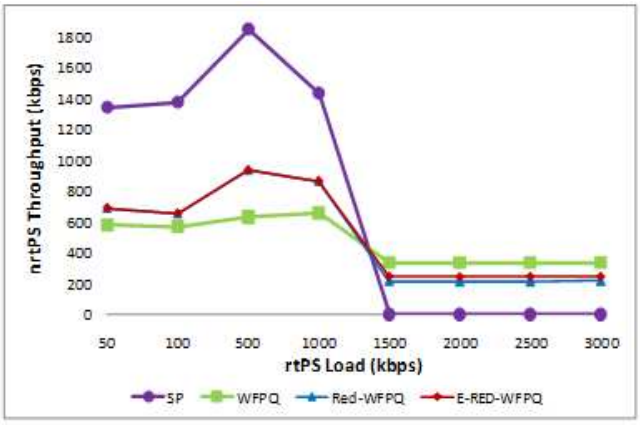


Figure 8. Comparison of nrtPS Throughput

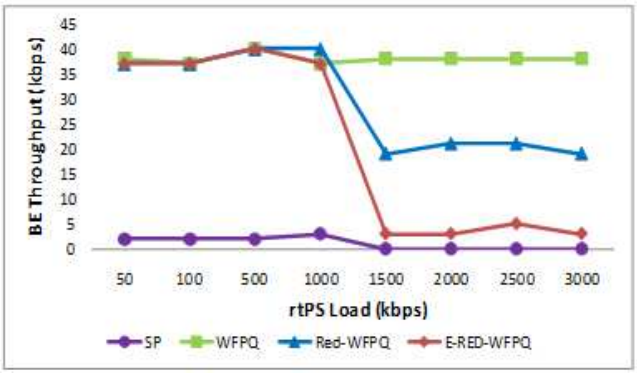


Figure 9. Comparison of BE Throughput

In Fig. 8, for Strict Priority, when rtPS traffic increases significantly, there will be no resource left for nrtPS and BE flows. Therefore, their throughputs will drop to zero. As a result, nrtPS and BE flows may starve under high rtPS traffic. Under high rtPS traffic, WFPQ has the highest nrtPS throughput as it has the maximum nrtPS weights than the others. The main reason is that, WFPQ has lower rtPS load than the others, so it can grant more weight for the nrtPS flow. Enhanced RED-based WFPQ has higher nrtPS throughput than RED-based WFPQ. When the rtPS load is 3000 kbps, Enhanced RED-based WFPQ yields 244 kbps and RED-based WFPQ yields 214 kbps throughput. This means that Enhanced RED-based WFPQ increases the nrtPS throughput as much as 14% over RED-based WFPQ.

In Fig. 9, the WFPQ algorithm yields the best throughput. This is because among all the algorithms, WFPQ assigns the highest weight to BE. However, we do not need to grant bandwidth for BE flows, as they do not have significant QoS requirements. As long as we prevent the starvation of the BE flows in a congested network, we have an acceptable QoS-based system. Therefore, in RED-based WFPQ and Enhanced RED-based WFPQ, we allocate very little weight to BE, so that we can continue serving them. In Strict Priority, the BE users have no chance of being served if the network is congested. Consequently, as rtPS load increases, BE flows cannot transmit their packets, and their throughputs drops to zero beyond 1500 kbps rtPS load.

 ISBN: 978-1-61208-103-8

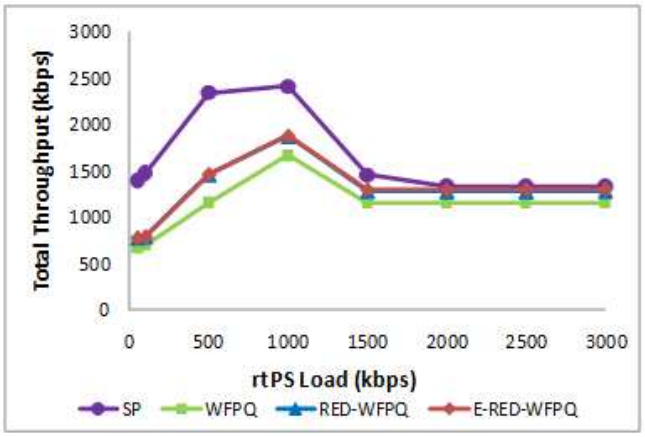


Figure 10. Comparison of Total Throughput

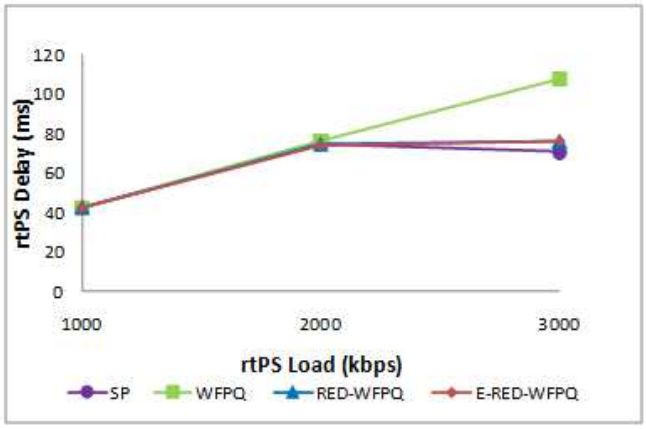


Figure 11. Comparison of rtPS Delay

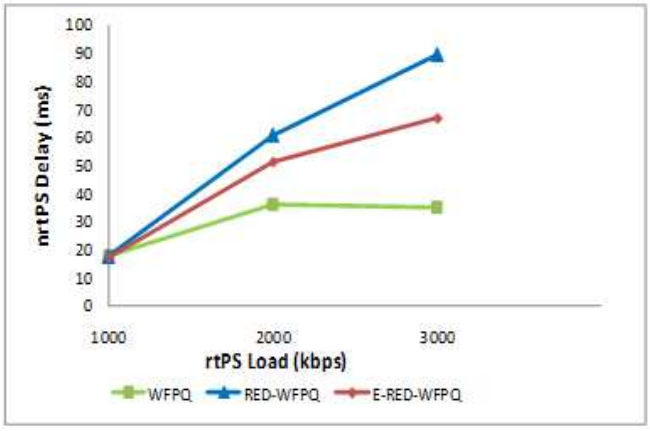


Figure 12. Comparison of nrtPS Delay

According to Fig. 10, Strict Priority scheduling has the maximum total throughput; however, it is not fair and acceptable for a QoS-based system. RED-based WFPQ and Enhanced RED-based WFPQ algorithms yield better throughput than WFPQ. In addition, RED-based WFPQ and Enhanced RED-based WFPQ algorithms yield approximately the same total throughput. The reason is that in our algorithm, Enhanced RED-based WFPQ, we decrease the BE throughput and so increase the nrtPS throughput.

In Fig. 11, as we increase the rtPS load, we observe that Strict Priority yields the lowest delay. This is because Strict Priority allocates higher bandwidth for rtPS flows than the others. WFPQ has the highest rtPS delay as its throughput is lower than the others. RED-based WFPQ and Enhanced RED-based WFPQ decrease the delay of rtPS, as they control the weight of rtPS according to the queue length of rtPS.

Fig. 12 shows the variation of nrtPS delay with increasing rtPS load. In this graph, we do not show the results of Strict Priority. The reason is that beyond 1500 kbps the nrtPS flow cannot transmit any packets, and the delay increases extremely. Consequently, the result for SP is not comparable with the other algorithms. The Enhanced RED-based WFPQ algorithm succeeds in decreasing the delay of nrtPS flows over RED-based WFPQ and WFPQ increasing the throughput. Among all three algorithms, WFPQ allocates the highest weight for nrtPS, thus, it has the lowest nrtPS delay.

In Fig. 13, we do not show the results for SP. The reason is that beyond 1500 kbps, the BE flow cannot transmit any packets, and the delay increases extremely. Consequently, the results are not comparable with the other algorithms. Enhanced RED-based WFPQ algorithm increases the delay of BE flows, as the algorithm increases the throughput of nrtPS. Therefore, RED-based WFPQ has lower BE delay than Enhanced RED-based WFPQ. In that point, we provide to transmit BE flows but as BE flow do not have QoS requirement, we increase nrtPS throughput in order to BE's. Among the three algorithms, WFPQ allocates the highest weight to BE, so it yields the lowest delay.

In Fig. 14, we observe that SP yields the lowest dropped packet percentage. WFPQ yields the highest dropped packet percentage, as RED-based WFPQ and Enhanced RED-based WFPQ algorithms increase the throughput of rtPS over that of WFPQ.

According to Fig. 15, WFPQ has the highest number of rtPS packets dropped. As Strict Priority yields the highest throughput for rtPS, its number of dropped packets is the lowest. RED-based WFPQ and Enhanced RED-based WFPQ have the same number of dropped packets. The reason is that their granting mechanism for rtPS flows is the same.

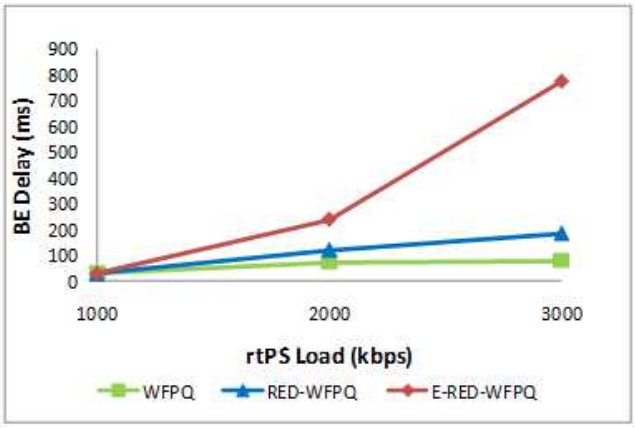


Figure 13. Comparison of BE Delay

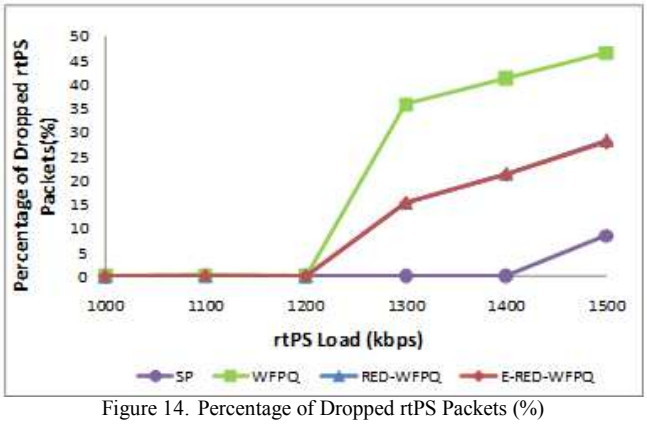


Figure 14. Percentage of Dropped rtPS Packets (%)

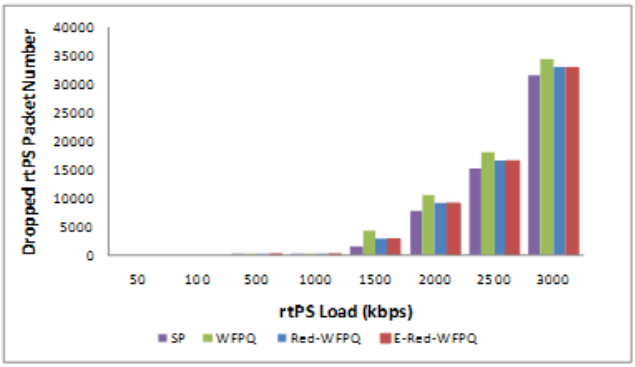


Figure 15. Dropped rtPS Packet Number

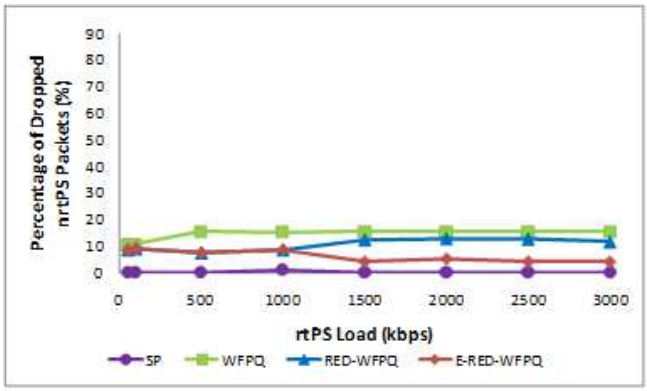


Figure 16. Percentage of Dropped nrtPS Packets (%)

According to Fig. 16, Strict Priority sends nrtPS packets until 1500 kbps rtPS load is reached. Beyond that, due to TCP congestion, nrtPS flows reduce their transmission rate to zero. Since there are no nrtPS packets submitted, the percentage of dropped nrtPS packets equals zero. WFPQ has the highest percentage of dropped nrtPS packets, as RED-based WFPQ and Enhanced RED-based WFPQ increase the nrtPS throughput over that of WFPQ. Also, beyond 1000 kbps rtPS load, Enhanced RED-based WFPQ has a lower percentage of dropped nrtPS packets than RED-based WFPQ. This is to be expected, as Enhanced RED-based WFPQ increases the nrtPS throughput.

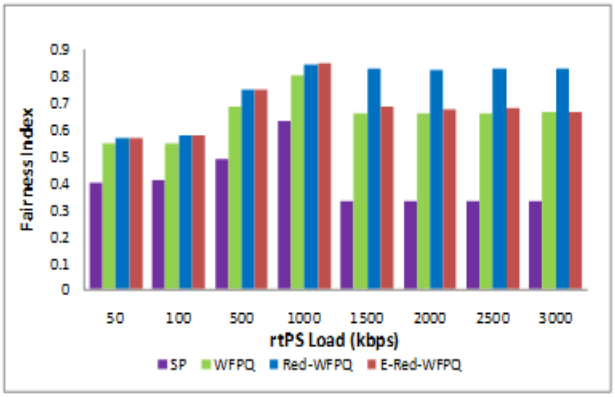


Figure 17. Fairness Index

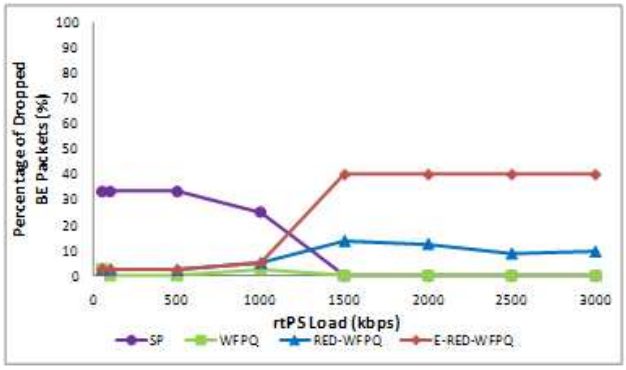


Figure 18. Percentage of Dropped BE Packets (%)

According to Fig. 17, Strict Priority scheduler sends BE packets until 1500 kbps rtPS load is reached. Beyond that, due to TCP congestion, BE flows can not be allocated any bandwidth. Since there are no BE packets submitted, the percentage of dropped packets equals zero. WFPQ has the lowest percentage of dropped BE packets, as RED-based WFPQ and Enhanced RED-based WFPQ decrease the BE throughput. Also, beyond 1000 kbps rtPS load, Enhanced RED-based WFPQ has a higher percentage of dropped BE packets than RED-based WFPQ. This is to be expected, as Enhanced RED-based WFPQ decreases the BE throughput.

Fairness Index is calculated according to [11]. Therefore, we normalized rtPS flow with 64 kbps, nrtPS flow with 45 kbps, and BE flow with 1 kbps. According to Fig. 18, Strict Priority scheduling has the lowest Fairness Index, as it is an unfair algorithm. Enhanced RED-based WFPQ is slightly fairer than WFPQ. As Enhanced RED-based WFPQ allocates sufficient bandwidth for rtPS flows, it achieves higher fairness over WFPQ. RED-based WFPQ has the highest Fairness Index because the algorithm provides more allocation for BE flows. Consequently, normalized values of rtPS, nrtPS, and BE are closer to each other, resulting in a higher Fairness Index. In a QoS-based system, we do not need to provide strong fairness if we increase the QoS of the system. But, we still evaluate if fairness is provided at an acceptable level.

 ISBN: 978-1-61208-103-8

V. CONCLUSION

In this paper, an Enhanced RED-based WFPQ algorithm for SS uplink scheduler is proposed and the throughput of nrtPS is increased while keeping rtPS throughput at high levels. The algorithm is compared with Strict Priority, WFPQ, and RED-Based WFPQ with the respect to throughput, delay, packet loss rate, and fairness. It is observed that the proposed algorithm gives promising results while keeping fairness at reasonable levels among the different QoS classes. The details of the work are available in [12].

Simulation results showed that RED-based WFPQ and Enhanced RED-based WFPQ increase the rtPS throughput, and they follow the same approach while allocating rtPS bandwidth. The rtPS throughput of Strict Priority is the highest and the throughput of WFPQ is the lowest. The nrtPS throughput of Enhanced RED-based WFPQ is higher than that of RED-based WFPQ. The BE throughput of Enhanced RED-based WFPQ is lower than that of RED-based WFPQ. Enhanced RED-based WFPQ increases nrtPS throughput, but it decreases the throughput of BE flows. However, the starvation of BE flows in congested network is prevented.

RED-based WFPQ and Enhanced RED-based WFPQ significantly decrease the delay of rtPS. Strict Priority has the lowest delay, and WFPQ has the highest delay. The delays experienced depend on the throughput of the flows. The nrtPS delay of Enhanced RED-based WFPQ algorithm is lower than that of RED-based WFPQ. The BE delay of RED-based WFPQ algorithm is lower than that of Enhanced RED-based WFPQ.

The number of dropped rtPS packets is directly proportional to the delay; therefore, Strict Priority exhibits the lowest number of dropped rtPS packets, while WFPQ exhibits the highest number of dropped rtPS packets. The number of dropped rtPS packets for RED-based WFPQ and Enhanced RED-based WFPQ are the same.

Performance of the studied and proposed schedulers is given in terms of Fairness Index also. It is observed that, SP scheduling is unfair and its Fairness Index is the lowest. RED-based WFPQ and Enhanced RED-based WFPQ have higher Fairness Index than WFPQ. The reason is the allocation of the bandwidth depend on the queue length and shows dynamic characteristic.

Currently we are working on dynamically changing weight thresholds. Here, the thresholds could adapt to the state of the service flow queues.

ACKNOWLEDGMENT

This project is partially supported by Istanbul Chamber of Industry and TUBITAK.

REFERENCES

[1] Borin, J. F. and Fonseca, N. L. S. da, “Simulator for WiMAX networks”, Elsevier Simulation Modeling Practice and Theory 16, 1, pp. 817-833, 2008.

[2] IEEE Standard 802.16-2004, “IEEE standard for Local Metropolitan Area Networks, Part16: Air interface For Fixed Broadband Wireless Access Systems ”, June 2004.

[3] IEEE Standard 802.16e-2005, “IEEE standard for Local Metropolitan Area Networks, Part16: Air Interface for Fixed Broadband Wireless Access Systems - Physical and Medium Access Control Layers for Combined Fixed and Mobile Operation in Licensed Bands”, December 2005.

[4] Pratik, D. et al., “A Performance Study of Uplink Scheduling Algorithms in Point-to-Multipoint WiMAX Networks”, Elsevier Computer Communications 32 511-521, 2009.

[5] Masri et al., “An Uplink Bandwidth Management Framework for IEEE 802.16 with QoS Guarantees”, IFIP Networking, Aachen, Germany, 2009.

[6] Po-Chun Ting, Chia-Yu Yu, and Naveen Chilamkurti., “A Proposed RED-Based Scheduling Scheme for QoS in WiMAX Networks”, IEEE International Symposium on Wireless Pervasive Computing (ISWPC), February 2009.

[7] <http://w3.antd.nist.gov/seamlessandsecure/>, last accessed in May 2010.

[8] <http://www.isi.edu/nsnam/ns/>, last accessed in July 2010.

[9] NIST, Seamless and secure mobility, NS-2 NIST WIMAX Module, http://www.antd.nist.gov/seamlessandsecure, May 2010.

[10] A. Belghith and L. Nuaymi, "Design and Implementation of a QoS-included WiMAX Module for NS-2 Simulator", Simutools, ACM, 2008

[11] Pratik, D., 2007. A Performance Study of Uplink Scheduling Algorithms in Point to Multipoint WiMAX Networks: MSc Thesis, *Queen’s University, Canada*.

[12] Kasaci, S., 2010, An Enhanced RED-Based Weighted Fair Priority Queuing Algorithm for IEEE802.16 SS Scheduler, MS Thesis, Istanbul Technical University, Institute of Science and Technology.

 ISBN: 978-1-61208-103-8

On the Link Layer Performance of Narrowband Body Area Networks

Jean-Michel Dricot[1], Gianluigi Ferrari[2], Stéphane Van Roy[1], François Horlin[1], and Philippe De Doncker[1]
1. Université Libre de Bruxelles
OPERA – Wireless Communications Group
E-mail: {jdricot, svroy, fhorlin, pdedonck}@ulb.ac.be
2. University of Parma, Italy
Dept. of Information Engineering
E-mail: gianluigi.ferrari@unipr.it

***Abstract*—Personal area networks and, more specifically, body area networks (BANs) are key building blocks of the future generation networks and the Internet of Things as well. In the last years, research has focused on the channel modeling and the definition of efficient medium access control (MAC) mechanisms. Less attention was paid to network-level performance. Thereby, this paper presents a novel analytical model for network performance analysis with centralized and mesh topologies. This model takes into account the channel statistics (i.e., the large-scale fading) and delivers several insights on the BAN implementation.**

***Index Terms*—body area networks, wireless networks, fading, performance analysis.**

I. Introduction

Recent advances in ultra-low power sensors have fostered the research in the field of body-centric networks, also referred to as *body area networks* (BANs). In these networks, a set of nodes (called *sensors*) is deployed on the human body. They aim at monitoring and reporting several physiological values, such as blood pressure, breath rate, skin temperature, or heart beating rate. A pictoral example of a BAN is shown in Fig. 1, where two illustrative topologies are presented: (i) a centralized topology, where a special node (denoted as "HUB") acts as a sink for all communications initiated by the sensor and (ii) a mesh topology (or "multi-hop topology"), where several intermediary nodes relay the information from the source node to the destination (e.g., when data fusion or sensor cooperation is required).

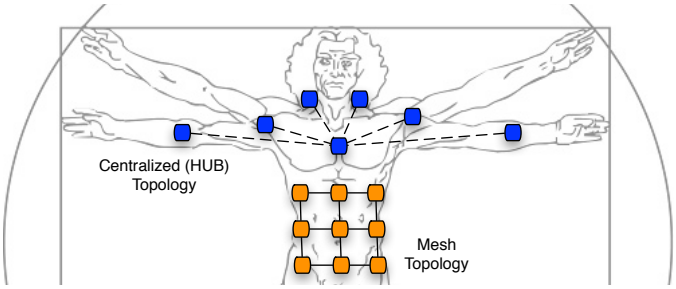


Fig. 1: Body Area Network.

Most of the time, sensing is performed at low rates but, in case of emergency, the network load may increase in seconds. Therefore, an in-depth analysis of the network outage, throughput, and achievable transmission rate can give insights on the maximum supported reporting rate and the corresponding performance.

The focus of this paper is on multi-hop communications and the impact of the specificities of the propagation channel. The modeling of the BAN channel has recently been thoroughly investigated [1]–[5]. The main findings on the body radio propagation channel can be summarized as follows. First, the average value of the power decreases as an exponential function of the distance. However, unlike classical propagation models, where the received power P is a decreasing function of the distance of the form $d^{-\alpha}$, in [6] the authors prove that a law of the form $10^{-\gamma d}$ characterizes more accurately body radio propagation. Second, the propagation channel is subject to large-scale fading (that is, shadowing). This variation follows a zero-mean Gaussian distribution in the dB scale or a Log-normal distribution in the linear scale.

This paper addresses the evaluation of the *throughput* for BANs, being this metric a traditional measure of how much traffic can be delivered by the network [7], [8]. Therefore, our analysis is expedient to understand the level of information which could be collected and processed in body-related applications (e.g., health or fitness monitoring). We consider slotted and asynchronous communications such that, in every time slot, each node transmits independently with a probability p. Indeed, in a generic scenario, the traffic distribution in a sensor network can be considered as spatially and temporally bursty, that is, reporting periods alternate temporally and spatially with periods and areas with little or no traffic (or even with a scheduled sleep of the nodes). It may therefore be impractical to employ reservation-based MAC schemes, such as those based on time/frequency division multiple access (TDMA/FDMA), that require a substantial amount of coordination traffic and cannot be implemented efficiently in energy- and computation-constrained sensor nodes.

The rest of the paper is organized as follows. In Section II, the models, definitions, and notations are introduced. Then, in Section III, the conditional success probability of a transmission for a node given the transmitter-receiver and interference-receiver distances is derived. Section IV investigates the average link throughput and achievable transmission rate for centralized and mesh topologies. Section V concludes the paper.

 ISBN: 978-1-61208-103-8

II. MODELS, NOTATION, AND DEFINITIONS

A. Stochastic Channel Model

Defining as $P_i^{(\mathrm{r})}$ the received power from the i-th node at distance d gives

$$P_i(d) = P_i \; L(d_i)\mathbf{X}_i$$

where P_i is the emitted power, $L(d)$ is the loss at distance d, it accounts for the antenna gains and carrier frequency, and $\mathbf{X}_i$ is a random variable (RV) which depends on the channel characteristics. It is shown in [9] that $\mathbf{X}_i$ has a log-normal distribution[1] with parameters μ and σ, i.e., its cumulative distribution function (cdf) is

$$F_{\mathbf{X}_i}(x;\mu,\sigma) = \frac{1}{2} - \frac{1}{2}\mathrm{erf}\left(\frac{\mu_{\mathrm{dB}} - 10\log_{10} x}{\sigma_{\mathrm{dB}}\sqrt{2}}\right)$$

where σ_{dB} typically ranges from 4 dB to 10 dB, μ_{dB} is the average path loss on the link (dimension: [dB]). Since the loss is accounted for by the term $L(d)$, it follows that $\mu_{\mathrm{dB}} = 0$ and the cdf of $\mathbf{X}_i$ reduces to the following:

$$F_{\mathbf{X}_i}(x;0,\sigma) = \frac{1}{2} - \frac{1}{2}\mathrm{erf}\left(\frac{-10\log_{10} x}{\sigma\sqrt{2}}\right)$$

and its corresponding pdf is

$$f_{\mathbf{X}_i}(x;0,\sigma) = \frac{10}{(\ln 10)\; x\; \sqrt{2\pi}\sigma} \exp\left\{-\frac{(10\log_{10} x)^2}{2\sigma^2}\right\}. \quad (1)$$

In [6], [10], it is shown, on the basis of an extensive campaign of measurements, that the path loss (in dB scale) is a linearly increasing function of the distance, i.e.:

$$P_i - P_i(d_i) = L_{\mathrm{ref}} + 10\gamma(d_i - d_{\mathrm{ref}}) \qquad d_i > d_{\mathrm{ref}}$$

where d_{ref} is a reference distance, L_{ref} is the loss at that distance, and γ a suitable constant. For instance, in [10] the authors found $d_{\mathrm{ref}} = 8$ cm, $L_{\mathrm{ref}} = 55.18$ dB, and $\gamma = 1.26$dB/cm. The path loss, in linear scale, can then be expressed as follows:

$$\begin{aligned} L(d_i) &= 10^{(10\gamma d_{\mathrm{ref}} - L_{\mathrm{ref}})/10} .10^{-\gamma d_i} \\ &\triangleq L_0\, 10^{-\gamma d_i}. \end{aligned} \quad (2)$$

B. Traffic Model

The transmission state of the i-th node at time[2] t is represented by the following RV:

$$\mathbf{\Lambda}_i(t) = \begin{cases} 1 & \text{if the } i\text{-th node transmits at time } t \\ 0 & \text{if the } i\text{-th node is silent at time } t. \end{cases}$$

A simple random access scheme is such that, at each time slot, a node transmits with probability p [11, p. 278]. Therefore, $\{\mathbf{\Lambda}_i(t)\}_{t=1}^{\infty}$ is a sequence of Bernoulli RVs with $\forall t$: $\mathbb{P}\{\Lambda_i(t) = 1\} = p$.

[1]Note that we use the $\log_{10}$ variant of the log-normal since the widely-used shadowing model uses an additive Gaussian variation expressed in dB.

[2]For the sake of simplicity, we assume that t can take integer values, i.e., we refer to a slotted communication system.

C. Total Interference Power

A transmission in a given link is successful if and only if the signal-to-noise and interference ratio (SINR) is above a certain threshold θ. This threshold value depends on the receiver characteristics, the modulation format, and the coding scheme, among others. The SINR at the receiving node of the link is given by

$$\mathrm{SINR} \triangleq \frac{P_0(d_0)}{W + P_{\mathrm{int}}} \quad (3)$$

where $P_0(d_0)$ is the received power from the link source located at distance d_0, W is ambient the noise power, and P_{int} is the total interference power at the link receiver, that is, the sum of the received power from all the undesired transmitters:

$$P_{\mathrm{int}} \triangleq W + \sum_{i=1}^{N} P_i(d_i)\; \mathbf{\Lambda}_i = W + \sum_{i=1}^{N} P_i L(d_i)\; \mathbf{X}_i\mathbf{\Lambda}_i.$$

D. Link Throughput and Link Transport Capacity

A *transmission is successful* if the channel is not in an outage, i.e., if the (instantaneous) SINR exceeds a certain threshold θ, that is, $\mathcal{P}_{\mathrm{s}} = \mathbb{P}\{\mathrm{SINR} > \theta\}$.

The *probabilistic link throughput* (adimensional) is defined to be the success probability multiplied by the probability that the transmitter actually transmits (in full-duplex operation) and, in addition in half-duplex operation, the receiver actually listens. So it is the unconditioned reception probability. This is the throughput achievable with a simple ARQ scheme (with error-free feedback) [12]. For the slotted transmission scheme we consider, the half-duplex probabilistic throughput is $\tau^{\mathrm{half}} \triangleq p(1-p)\mathcal{P}_{\mathrm{s}}$ and for full-duplex it is $\tau^{\mathrm{full}} \triangleq p\,\mathcal{P}_{\mathrm{s}}$.

Finally, the *link achievable transmission rate* (dimension: [bit/s/Hz]) is defined as the product of the probabilistic throughput and the link capacity, i.e., $T = \tau\, \log_2(1 + \mathrm{SINR})$.

III. SUCCESS PROBABILITY OF A TRANSMISSION

A. Derivation

The link probability of success for a required threshold SINR value equal to θ is

$$\begin{aligned} \mathcal{P}_{\mathrm{s}} &= \mathbb{P}\{\mathrm{SINR} > \theta\} \\ &= \mathbb{E}\left[\mathbb{P}\left\{\frac{P_0\; L(d_0)\mathbf{X}_0}{P_{\mathrm{int}}} > \theta \;\middle|\; P_{\mathrm{int}}\right\}\right] \\ &= \mathbb{E}_{X,\lambda}\left[1 - \mathbb{P}\left\{\mathbf{X}_0 \le \theta\frac{W + \sum_{i=1}^{N} P_i\, L(d_i)\mathbf{X}_i\mathbf{\Lambda}_i}{P_0\, L(d_0)}\right\}\right] \\ &= \mathbb{E}_{X,\lambda}\left[\frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{-10\log_{10}\theta\boldsymbol{\eta}}{\sigma\sqrt{2}}\right)\right] \end{aligned} \quad (4)$$

where

$$\boldsymbol{\eta} \triangleq \frac{W + \sum_{i=1}^{N} P_i\, L(d_i)\mathbf{X}_i\mathbf{\Lambda}_i}{P_0\, L(d_0)} = \frac{W}{P_0\, L(d_0)} + \sum_{i=1}^{N} \eta_i \mathbf{X}_i\mathbf{\Lambda}_i. \quad (5)$$

 ISBN: 978-1-61208-103-8

$$\mathcal{P}_{\rm s} = \sum_{j=1}^{n} c_j \underbrace{\exp\left(\frac{-a_j \theta W}{P_0\, L_0\, 10^{-\gamma d_0}}\right)}_{\text{Background noise}} \prod_{i=1}^{N} \left[p \underbrace{\int_0^{\infty} \exp\left(-a_j \theta \frac{P_i}{P_0}\, 10^{\gamma(d_0-d_i)} x_i\right) f_{\mathbf{X}}(x_i)\, {\rm d}x_i}_{\text{Interference}} + (1-p) \right] \tag{6}$$

In the Appendix, it is shown that

$$\zeta(z;\sigma) = \frac{1}{2} + \frac{1}{2}{\rm erf}\left(\frac{-10\log_{10} z}{\sigma\sqrt{2}}\right) \approx \sum_{j}^{n} c_j \exp(-a_j z)$$

where $\{c_j\}_{j=1\ldots n}$, $\{-a_j\}_{j=1\ldots n}$ are suitable coefficients, so that

$$\begin{aligned}\zeta(\theta\boldsymbol{\eta}) &= \sum_{j=1}^{n} c_j \exp\left(\frac{-a_j\theta W}{P_0\, L(d_0)}\right) \exp\left(-a_j\,\theta \sum_{i=1}^{N} \eta_i \mathbf{X}_i \boldsymbol{\Lambda}_i\right) \\ &= \sum_{j=1}^{n} \underbrace{c_j \exp\left(\frac{-a_j\theta W}{P_0\, L(d_0)}\right)}_{c'_j} \prod_{i=1}^{N} \exp\left(-a_j \theta \eta_i \mathbf{X}_i \boldsymbol{\Lambda}_i\right)\end{aligned}$$

and

$$\begin{aligned}\mathbb{E}_X\left[\frac{1}{2} + \frac{1}{2}{\rm erf}\left(\frac{-10\log_{10}\theta\boldsymbol{\eta}}{\sigma\sqrt{2}}\right)\right] &= \mathbb{E}_X\left[\zeta(\theta\boldsymbol{\eta})\right] \\ &= \underbrace{\int_0^{\infty} \cdots \int_0^{\infty} \int_0^{\infty}}_{N \text{ times}} \zeta(\theta\boldsymbol{\eta}) \times \prod_{i=1}^{N} f_{\mathbf{X}}(x_i)\, {\rm d}x_i \\ &= \sum_{j=1}^{n} c'_j \prod_{i=1}^{N} \int_0^{\infty} \exp\left(-a_j \theta \eta_i x_i \boldsymbol{\Lambda}_i\right) f_{\mathbf{X}}(x_i)\, {\rm d}x_i\end{aligned}$$

Finally, (4) becomes

$$\begin{aligned}&\mathcal{P}_{\rm s} = \mathbb{E}_{\Lambda}\, \mathbb{E}_X\left[\zeta(\theta\boldsymbol{\eta})\right] \\ &= \mathbb{E}_{\Lambda}\left[\sum_{j=1}^{n} c'_j \prod_{i=1}^{N} \int_0^{\infty} \exp\left(-a_j \theta \eta_i x_i \boldsymbol{\Lambda}_i\right) f_{\mathbf{X}}(x_i)\, {\rm d}x_i\right] \\ &= \sum_{j=1}^{n} c'_j \prod_{i=1}^{N} \mathbb{E}_{\Lambda}\left[\int_0^{\infty} \exp\left(-a_j \theta \eta_i x_i \boldsymbol{\Lambda}_i\right) f_{\mathbf{X}}(x_i)\, {\rm d}x_i\right] \\ &= \sum_{j=1}^{n} c'_j \prod_{i=1}^{N} \left[\mathbb{P}\left\{\lambda_i = 1\right\} \int_0^{\infty} \exp\left(-a_j \theta \eta_i x_i\right) f_{\mathbf{X}}(x_i)\, {\rm d}x_i \right. \\ &\qquad \left. + \mathbb{P}\left\{\lambda_i = 0\right\} \int_0^{\infty} f_{\mathbf{X}}(x_i)\, {\rm d}x_i\right] \\ &= (6)\end{aligned}$$

which can be numerically computed.

B. Narrowband Communications

The first term of (6) defines the link probability of success in a noise-limited regime, i.e., even if no interferers are present. Starting from (4) with $P_{\rm int} = 0$ gives:

$$\mathcal{P}_{\rm s} = \zeta\left(\frac{\theta W}{P_0\, L_0\, 10^{-\gamma d_0}}\right).$$

If a threshold link probability of success equal to $\mathcal{P}_{\rm s}^{\rm th}$ is required and W is the thermal noise, a transmission is possible if and only if

$$P_0 \geq \frac{\theta\, kTB}{L_0\, \zeta^{-1}(\mathcal{P}_{\rm s}^{\rm th})}\, 10^{\gamma d_0} \qquad 0 < \mathcal{P}_{\rm s}^{\rm th} < 1 \tag{7}$$

where $T = 300$ K is the room temperature, k is the Boltzmann's constant, and B is the transmission bandwidth. For instance, in Fig. 2 the minimum transmission power P_0 for a ZigBee equipment ($B = 5$ MHz), operating at $T = 30$°C with a SINR $\theta = 10$ dB and $\sigma = 4$ dB, is shown as a function of the distance, considering various values of the required link probability of success of $\mathcal{P}_{\rm s}^{\rm th}$.

It can be seen that (i) the value of $\mathcal{P}_{\rm s}^{\rm th}$ plays a limited role on the transmission power[3] and (ii) if the transmission power is constrained by energy concerns, only short-range communications (some tenths of centimeters) will be possible. A *multi-hop* network architecture is therefore preferred.

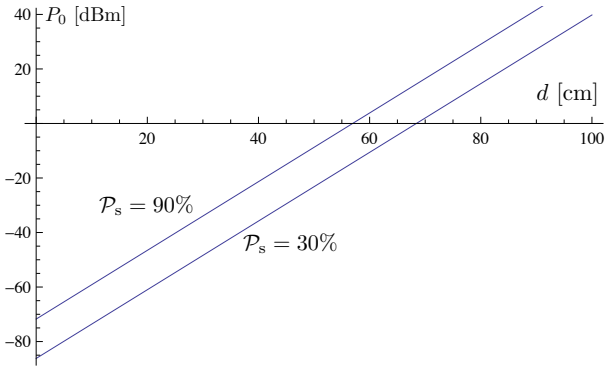


Fig. 2: Minimum transmission power as a function of the distance.

In the following, we will consider only interfence-limited networks, i.e., scenarios where condition (7) is satisfied. Formally, this situation is equivalent to letting $W = 0$ in (6).

IV. Performance Analysis

In this section, we investigate networks with two topologies: (i) a centralized (hub) architecture in which all nodes connect to a central sink and (ii) a mesh architecture where every node has the same number of nearest neighbors and the same distance to all nearest neighbors. Without any loss of generality, we assume that all nodes are equivalent and, therefore, transmit at equal power levels, i.e., $\forall i, j \in \{0, N\} : P_i = P_j$.

 ISBN: 978-1-61208-103-8

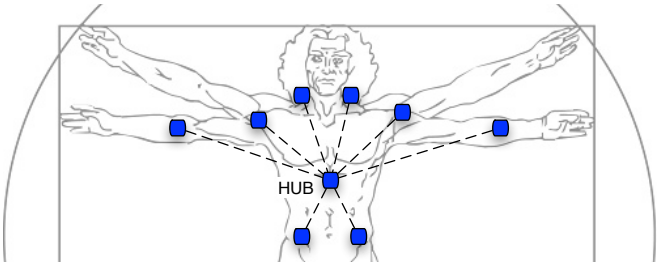


Fig. 3: Central hub surrounded by several nodes.

A. Hub Topologies

Fig. 3 presents a centralized architecture, where a central node (called *hub* or *sink*) is surrounded by several nodes. These nodes can be leaves of the routing tree or, in a star topology, the coordinators of each star. In this architecture, all distances between the nodes and the hub are approximately the same (i.e., $\forall i \in \{1, \ldots, N\} : d_i \approx d_0$) and the link probability of success at the hub becomes

$$\mathcal{P}_{\mathrm{s}} = N \sum_{j=1}^{n} c_j \left[p \int_0^{\infty} \exp\left(-a_j \theta x_i\right) f_{\mathbf{X}}(x_i)\, \mathrm{d}x_i + (1-p) \right].$$

The link achievable transmission rate is shown in Fig. 4 for full- and half-duplex systems. It can be seen that both transmission strategies exhibit same performance when (i) the number of nodes N to serve increases and (ii) the data collection (through p) remains limited.

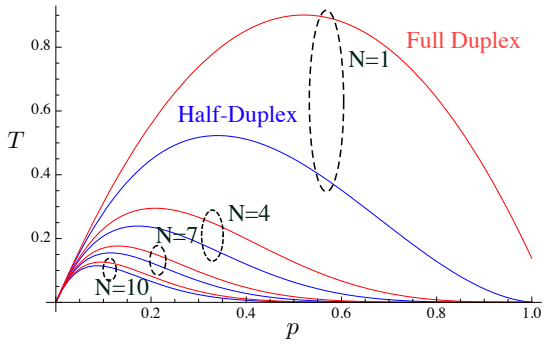


Fig. 4: Channel achievable transmission rate T in bits/s/Hz as a function of the number of nodes N connected to the hub. $\theta = 10$dB, $\sigma = 4dB$, links SINR=10dB.

B. Mesh Topologies

Referring back to Fig. 1, it can be seen that a BAN build using a mesh topology can be approximated as a cylindrical surface of radius r and height $2h$ deployed on a human torso. Without any loss of generality, the center of the reference axes can be located on the receiver node. Therefore, as shown in Fig. 5, the body area network can be modeled as a finite, rectangular area of width $2\pi r$ and height $2h$. For a regular deployment of N nodes on the surface $A = 2\pi r.2h$, the inter-nodes distance is $\delta = \sqrt{2\pi r.2h/N}$. For instance, let us fix: $2\pi r = 1$ m and $2h = 0.5$ m, so that $\delta = 1/\sqrt{2N}$.

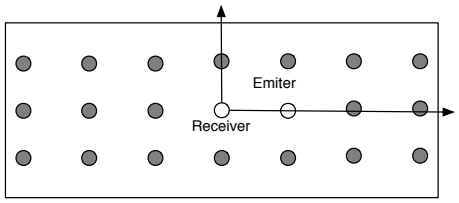


Fig. 5: Schematic representation of a regular deployment of sensors.

It is interesting to note that the term $10^{\gamma(d_0-d_i)}$ is virtually insensitive to an increase in the amount of nodes deployed on the body surface. Indeed, if the amount of nodes is multiplied by a factor α, the transmission distances are divided accordingly, so that the term becomes $10^{\gamma(d_0-d_i)/\alpha}$. Since γ is large (authors in [10] report $\gamma = 126$dB/m) and distances are limited, the value of α has a small impact on the exponent expression. This can also be interpreted as the fact that only the direct neighbors do interfere on communications: these are limited by interference from *dominant nodes*.

In Fig. 6, the throughput is plotted in a scenario similar to Fig. 5 and with respect to the transmission distance. This distance is expressed in terms of nodes hopped over, i.e., $d_0 = n\delta$. More precisely, the transmission goes from $n = 1$ (direct transmission to the closest neighbor) to $n = 4$ (transmission four hops away).

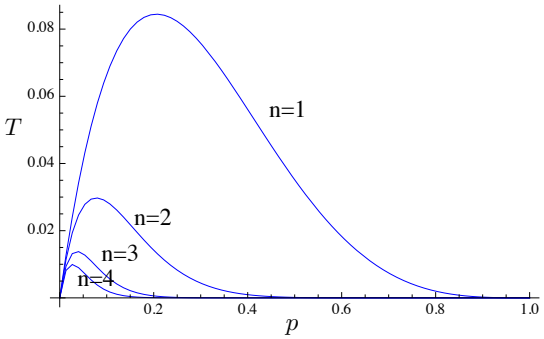


Fig. 6: Channel achievable transmission rate T in bits/s/Hz as a function of the hop distance (expressed in number of intermediary hop nodes n). $\theta = 10$dB, $\sigma = 4$dB, links SINR=10dB.

It can be seen that, when the sensing rate is low ($p \ll 1$), long-range transmissions are to be preferred (for delay and energy considerations). On the other hand, when $p \geq 0.05$, short-range communications and multi-hopping are more suitable.

C. Sustainable Number of Hops

In the context of multi-hop communications, each transmission takes place on a route in which a certain amount of intermediary nodes act as relays. The maximum sustainable number of nodes (that is, the route length) is a critical parameter since it quantifies the effective distance that can

[3] Indeed, $\forall x \in \mathbb{R} : |x| < 1 \Rightarrow |\operatorname{erf}(x)| \leq 2$. Therefore, since $10\log_{10}\zeta_{\mathrm{dB}}^{-1}(x) = -\sigma\sqrt{2}\,\operatorname{erf}^{-1}(2x-1)$, one has $|10\log_{10}\zeta_{\mathrm{dB}}^{-1}(x)| \leq 2\sigma\sqrt{2}$. For instance, with $\sigma = 4$ dBm, this bound is $2\sigma\sqrt{2} = 11.3$ dBm.

 ISBN: 978-1-61208-103-8

be covered and is an indicator of the connectivity of whole the network [13]. It depends on the acceptable packet loss, link interference, and topology among others.

In a conservative scenario, each node stores and forwards every packet it receives, without any consideration for re-transmissions. This case, though pessimistic, corresponds to UDP-style connections and allows to derive the lower of bound of the network performance. Since all links are considered equals, the *route* probability of success is

$$\mathcal{P}_{\mathrm{s}}^{\mathrm{route}} = \prod_{i=1}^{n_{\mathrm{hops}}} \mathcal{P}_{\mathrm{s}}^{(i)} = \left(\mathcal{P}_{\mathrm{s}}\right)^{n_{\mathrm{hops}}}.$$

Therefore, the maximum sustainable amount of hops for an acceptable final transmission success equal to $\mathcal{P}_{\mathrm{s}}^{\mathrm{th}}$ is

$$n_{\mathrm{hops}} = \left\lfloor \frac{\log\left(\mathcal{P}_{\mathrm{s}}^{\mathrm{th}}\right)}{\log\left(\mathcal{P}_{\mathrm{s}}\right)} \right\rfloor$$

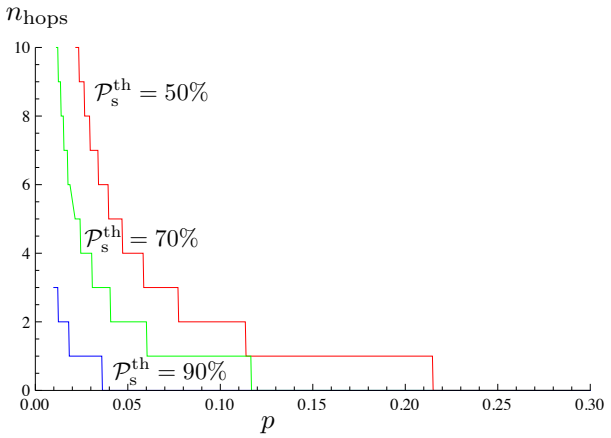


Fig. 7: Maximum sustainable number of hops with regular topology and for different values of the acceptable route probability of sucess. $\theta = 10$dB, $\sigma = 4dB$.

In Fig. 7, the maximum number of sustainable hops is presented for various values of the minimal route probability of success $\mathcal{P}_{\mathrm{s}}^{\mathrm{th}}$. It can be observed that (without any form of transmission control), highly-reliable routes can only be achieved at very low transmission rates. Also, most routes exhibit a limited amount of possible hops but these values should be sufficient enough to achieve full connectivity in the specific context of body area networks.

V. Conclusions

In the presence of a centralized network topology, it has been shown that the duplex capability of the nodes does not play a critical role, especially in the presence of limited sensing rates. It has been shown that a maximum achievable transmission rate exists and depends on the amount of deployed nodes. Beyond this maximum, the interference makes the achievable transmission rate decrease.

For BANs with a mesh topology, the transmission strategy depends on the traffic profile. More precisely: when the transmission probability of each node is limited (passive monitoring of a patient or deep sleep of the nodes), long-range transmissions can be used in order to save energy and avoid multiple relays. On the other hand, when the sensors have a substantial amount activity, the performance decreases exponentially with the number of hops. Shortest possible hop strategy should be used but it comes at the cost of numerous relaying. However, even though the maximum sustainable number of hops is small, it is still suitable for BAN-based applications, though it is extremely difficult to reach a security level of 90% without a transmission control protocol.

Finally, the main contribution of this paper consists in the derivation of a closed-form expression for the link probability of success in interference-limited BANs subject to large-scale fading.

VI. Acknowledgement

The support from the Belgian National Fund for Scientific Research (FRS-FNRS) is gratefully acknowledged.

Appendix

The modeling of slow-scale fading as a Log-normal distribution (that is, a zero-mean Gaussian in dB scale) raises mathematical difficulties, as shown in (4). The complementary cdf. of the zero-mean Log-normal distribution is

$$\zeta(z;\sigma) = \frac{1}{2} + \frac{1}{2}\mathrm{erf}\left(\frac{-10\log_{10} z}{\sigma\sqrt{2}}\right) \tag{8}$$

This function is plotted on Fig. 8 for $\sigma \in [4, 16]$. It can be observed that (i) $\zeta(z)$ saturates for $z \to \infty$ and (ii) it has the shape of a decreasing exponential function.

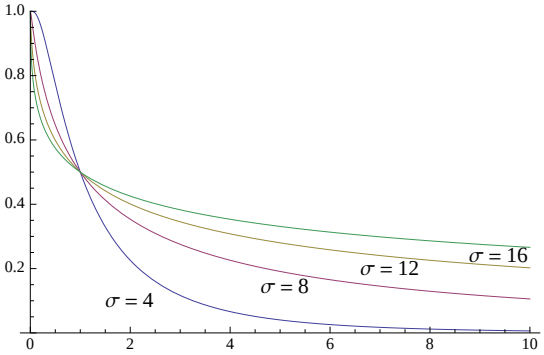


Fig. 8: $The function \zeta(z)$ for $\sigma = 4 - 16$dB

Indeed, the erf(.) function is known to have the following approximations:

$$\begin{aligned}
\mathrm{erf}(x) &\approx 1 - \frac{e^{-x}}{x\sqrt{\pi}} \\
&\approx \frac{1}{x\sqrt{2\pi}}\left(1 - \frac{1}{x^2}\right)e^{-x^2} \\
&\approx \sqrt{1 - \exp\left(-\frac{2x}{\sqrt{\pi}}\right)}\ , \text{ etc.}
\end{aligned}$$

 ISBN: 978-1-61208-103-8

The $\zeta(.)$ function can therefore be approximated with a linear combination of negative exponential function, as in Prony's approximation [14]:

$$\zeta(z) = \sum_{j}^{\infty} c_j \exp(-a_j z) \approx \sum_{j}^{n} c_j \exp(-a_j z)$$

where the coefficients $\{c_j\}_{j=1...n}$, $\{-a_j\}_{j=1...n}$ are determined in a least square sense by means of $q \geq 2n$ known points of the $\zeta(.)$ function. A Levenberg-Marquardt algorithm [15], [16] was used to determine the coefficients $\{c_j\}$ and $\{a_j\}$ for different values of σ and $q = 10000$ points over the interval $z \in [0, 1000]$. The corresponding values are reported in Table I along with the corresponding residual sum of squares.

TABLE I: Prony coefficients for the approximation of $\zeta(.)$

	c_1	a_1	c_2	a_2	c_3	a_3	residual
$\sigma = 4$	0.49	0.75	0.49	0.75	0.03	0.16	$4.68\,10^{-5}$
$\sigma = 6$	0.38	0.31	0.56	1.21	0.06	0.07	$4.23\,10^{-6}$
$\sigma = 8$	0.59	1.32	0.34	0.18	0.06	0.02	$1.04\,10^{-4}$
$\sigma = 10$	0.29	0.09	0.65	1.17	0.05	0.01	$7.53\,10^{-4}$
$\sigma = 12$	0.04	0	0.24	0.04	0.70	0.93	$3.52\,10^{-3}$
$\sigma = 14$	0.20	0.01	0.03	0	0.72	0.64	$1.03\,10^{-2}$
$\sigma = 16$	0.18	0.01	0.70	0.49	0.04	0	$1.67\,10^{-2}$

REFERENCES

[1] E. Reusens, W. Joseph, G. Vermeeren, and L. Martens, "On-body measurements and characterization of wireless communication channel for arm and torso of human," *4th international workshop on wearable and implantable body sensor networks (BSN 2007)*, vol. 13, pp. 264–269, 2007.

[2] J. M. Choi, H.-J. Kang, and Y.-S. Choi, "A study on the wireless body area network applications and channel models," in *FGCN '08: Proceedings of the 2008 Second International Conference on Future Generation Communication and Networking*. Washington, DC, USA: IEEE Computer Society, 2008, pp. 263–266.

[3] N. Katayama, K. Takizawa, T. Aoyagi, J.-I. Takada, H.-B. Li, and R. Kohno, "Channel model on various frequency bands for wearable body area network," *IEICE Transactions on Communications*, vol. E92.B, no. 2, pp. 418–424, 2009.

[4] A. Fort, C. Desset, P. De Doncker, P. Wambacq, and L. Van Biesen, "An ultra-wideband body area propagation channel model-from statistics to implementation," *IEEE Transactions on Microwave Theory and Techniques*, vol. 54, no. 4, pp. 1820– 1826, 2006.

[5] H. Sawada, J. Takada, S. Choi, K. Yazdandoost, and R. Kohno, "Review of body area network channel model," in *Proc. IEICE Gen. Conf.*, 2007.

[6] J. Ryckaert, P. D. Doncker, R. Meys, A. de Le Hoye, and S. Donnay, "Channel model for wireless communication around human body," *Electron. Lett.,*, vol. 40, no. 9, pp. 543–544, 2004.

[7] P. Gupta and P. R. Kumar, "The capacity of wireless networks," vol. 46, no. 2, pp. 388–404, Mar. 2000.

[8] X. Liu and M. Haenggi, "Throughput analysis of fading sensor networks with regular and random topologies," *EURASIP J. Wirel. Commun. Netw.*, vol. 2005, no. 4, pp. 554–564, 2005.

[9] J. ichi Takada, T. Aoyagi, K. Takizawa, N. Katayama, H. Sawada, T. Kobayashi, K. Y. Yazdandoost, H. bang Li, , and R. Kohno, "Static propagation and channel models in body area,," in *COST 2100 6th Management Committee Meeting, TD(08)639*, 2008.

[10] S. van Roy, C. Oestges, F. Horlin, and P. De Doncker, "Propagation modeling for uwb body area networks: Power decay and multi-sensor correlations," in *IEEE 10th International Symposium on Spread Spectrum Techniques and Application*, 2008, pp. 649–653.

[11] D. Bertsekas and R. Gallager, *Data Networks*, Prentice-Hall, Ed., 1991.

[12] S. Ahmed and M. S. Alam, "Performance evaluation of important ad hoc network protocols," *EURASIP Journal on Wireless Communications and Networking*, vol. 2006, pp. Article ID 78 645, 11 pages, 2006.

[13] O. K. Tonguz and G. Ferrari, *Ad Hoc Wireless Networks: A Communication-Theoretic Perspective*. Chichester, UK: John Wiley and Sons, March 2006.

[14] Baron G. de Prony, "Essai expérimental et analytique sur les lois de la dilatabilité des fluides élastique et sur celles de la force expansive de la vapeur de l'eau et de la vapeur de l'alkool, à différentes températures," *Journal de l'École Polytechnique*, vol. 1, no. 2, pp. 24–76, 1795.

[15] K. Levenberg, "A method for the solution of certain non-linear problems in least squares," *The Quarterly of Applied Mathematics*, vol. 2, pp. 164–168, 1944.

[16] D. Marquardt, "An algorithm for least-squares estimation of nonlinear parameters," *SIAM Journal on Applied Mathematics*, vol. 11, pp. 431–44, 1963.

 ISBN: 978-1-61208-103-8

Implementation of the Wireless Autonomous Spanning Tree Protocol on Mote-Class Devices

Kamini Garg, Daniele Puccinelli, and Silvia Giordano
Networking Lab
University of Applied Sciences of Southern Switzerland
CH-6928 Manno
Email:{kamini.garg,daniele.puccinelli,silvia.giordano}@supsi.ch

***Abstract*—The Wireless Autonomous Spanning Tree Protocol combines medium access control and routing to streamline energy-efficient communication in Wireless Body Area Networks. As these networks generally contain low-end resource-constrained devices, a thorough evaluation on real hardware is essential to the validation of any protocol. We present our implementation of the Wireless Autonomous Spanning Tree Protocol on mote-class devices, and highlight the challenges of taming the vagaries of low-power wireless that do not arise in simulation-based evaluations. We study the performance of the protocol in several dimensions, with a special emphasis on energy-efficiency. Our comprehensive set of experimental results indicates that the protocol can achieve low duty cycles if used jointly with a low-power link layer.**

***Keywords*-Wireless Body Area Networks; Wireless Autonomous Spanning Tree Protocol; Low Power Listening, Medium Access; Routing**

I. Introduction

A Wireless Body Area Network (WBAN) is a sensor network whose nodes are either attached or implanted into the human body. As for sensor networks in general, energy-efficiency is of paramount importance for WBANs. Because the radio notoriously accounts for the lion's share of the overall energy consumption, WBAN protocols must streamline communication. Due to their energy constraints, WBAN nodes are forced to employ low-power radios whose limited transmit power often precludes the option of forming a simple star network topology and require multihop communication as a matter of course.

In this context, channel access and routing decisions cannot be made in a vacuum and must account for the conditions of the wireless medium. While most existing solutions address the MAC and the network layer separately, the Wireless Autonomous Spanning Tree Protocol (WASP) [1] offers a unified framework for the coordination of medium access and multihop routing tailored to the relatively small network size of WBANs. With the WASP, WBAN nodes self-organize in a tree topology where parents set up a medium access schedule for children. In the original WASP work, the existence of a static tree topology is taken for granted, and the protocol is mainly evaluated through simulation. In practice, however, static connectivity cannot be expected in low-power wireless networks, because time-varying link dynamics affect even networks of stationary nodes. In a WBAN, nodes are generally stationary with respect to one another, but the network as a whole moves with the person and its link dynamics are affected accordingly. In this paper we tackle the challenging task of taming the vagaries of low-power wireless to implement the WASP on mote-class devices and evaluate it experimentally in various dimensions. Because the paramount goal of the WASP is energy-efficiency, we use the duty-cycle of our nodes as our primary figure of merit, and investigate the interplay of the WASP with the standard link-layer duty-cycling technique known as Low Power Listening (LPL) [2].

II. Related Work

In the sensor network literature, communication protocols can be mainly categorized as contention-based protocols, such as B-MAC [2], Wise-MAC [3], and X-MAC [4], and slotted protocols, such as S-MAC [5]. B-MAC is a CSMA-based technique that leverages asynchronous LPL, the standard technique for link-layer duty-cycling that enables nodes to periodically put their radios into sleep mode while maintaining the illusion of an always-on link. LPL dampens the idle listening problem by shifting the energy cost of communication from the receiver to the transmitter, which needs to match the preamble of its outgoing packets to the sleep interval of the receiver (long preamble). Wise-MAC gets transmitters to learn the wake-up schedule of their intended receivers and shortens LPL's long preamble through synchronization, while X-MAC sticks to asynchronous LPL but reduces the energy impact of the preamble. S-MAC uses a periodic listen/sleep cycle and ensures the schedule synchronization of neighboring nodes. Tree-based collection routing is a basic primitive for sensor networks employed by several protocols, such as MintRoute [6], the Collection Tree Protocol (CTP) [7], and Arbutus [8]. Cross-layer protocols that merge medium access and routing have been proposed to meet the specific needs of WBANs, whose network size is typically rather small (less than 20 nodes). The Wireless Autonomous Spanning Tree Protocol (WASP) [1] combines slotted medium access control and tree-based col-

 ISBN: 978-1-61208-103-8

lection routing to streamline energy-efficient communication in WBANs.

III. Protocol Description

A. Overview of the WASP Protocol

As described in [1], the WASP protocol is a slotted cross-layer protocol that uses a multi-level spanning tree for the coordination of medium access and multihop routing. The WASP employs a slotted notion of time, which is viewed as a succession of WASP-cycles (sets of time-slots). Moreover, the WASP presupposes the pre-existence of a spanning tree rooted at the sink. It further assumes that every node in the tree has exactly one parent as well as a complete knowledge of its neighborhood (*i.e.*, its parent, its siblings, and its children), and that the sink has a complete knowledge of the whole tree. Each node provides channel access information to its children by broadcasting a node-specific message called WASP-scheme. The WASP-scheme serves to regulate medium access from parent to children: a parent uses a WASP-scheme to tell its children when to access the medium and when to sleep. The sink initiates the process by broadcasting its own WASP-scheme. Upon reception of the WASP-scheme from the sink, its children broadcast their own WASP-scheme, which they derive from the sink‘s WASP-scheme. This process continues until all nodes in the tree have learned the correct timing for channel access using WASP-schemes. A node whose many children cannot be accommodated within the medium access time-slots allocated through its parent's WASP-scheme can use its own WASP-scheme to request additional time-slots for its children to use in future WASP-cycles.

Table I
Format of the WASP-Scheme for the Sink

SinkID	ChildIDs	SP	TFS	CS

As shown in Table I, the format of the sink's WASP-scheme comprises the SinkID, the ChildIDs, the Silent Period (SP), the Total number of Forwarding Slots (TFS), and the Contention Slot (CS). The SinkID and ChildIDs are the addresses assigned, respectively, to the sink and its child nodes. The WASP-scheme of the sink indicates the timeline of one WASP-cycle that includes a slot for the sink to broadcast its WASP-scheme, a slot for each of its children to send out their own WASP-schemes, a radio sleep mode period (whose slot count is the SP), a period during which the sink receives data forwarded by its children (the forwarding slots, whose total count is the TFS), and a special slot, the CS, in which new nodes may join the network using CSMA-CA. Let S denote the sink and Λ_m denote the m^{th} level in the spanning tree, with $\Lambda_0 \triangleq \{S\}$. Also, let T_i denote the set of nodes in the subtree of node i. The SP of the sink can be computed as

$$\mathrm{SP}_S = \max_{i \in \Lambda_1} |T_i|. \quad (1)$$

The TFS needed by the sink is given by the sum of all forwarding slots required by each node in Λ_1. During the forwarding slots, the nodes in Λ_1 forward their received data to the sink, and the number of forwarding slots required by each node in Λ_1 is equal to the total number of nodes in its sub-tree.

Table II
Format of the WASP-Scheme for a Node in Λ_m ($m \geq 1$)

NodeID	SP	ChildIDs	TFS	CS	DATA

Every node derives its own WASP-scheme based on the parent's WASP-scheme and therefore learns about its role in each time-slot of the WASP-cycle. While a sink's WASP-scheme is a dedicated control packet, the WASP-schemes of all other nodes may include data. As shown in Table II, the format of the WASP-scheme for nodes other than the sink contains the NodeID (the address of the node), the SP, the ChildIDs, the TFS, the CS, and the data itself. For any node other than the sink, the SP is used to tell its children in which time-slots they can go into sleep mode. As explained in [1], for a node $j \in \Lambda_1$, the length of the SP is equal to the slot number of the start of the SP of the sink S minus the slot number of its first occurrence in the WASP-scheme of S, minus 1. For a node $i \in \Lambda_m$ ($m > 1$), the length of the SP is equal to the slot number of the CS minus the slot number of its first occurrence in its parent's WASP-scheme. The number of forwarding slots for each node can be computed based on the requirements of the children and may vary accordingly over different WASP-cycles. For any node in any particular WASP-cycle, the TFS is given by the sum of the data packets received from its children. The length of a WASP-cycle depends on the length of the sink's WASP-scheme. The total number of WASP-cycles needed to send data from all the nodes to the sink depends on the depth the spanning tree. Data up to Λ_2 can be sent only in one WASP-cycle while for the further level nodes more WASP-cycles are required.

The original work on WASP uses analysis and simulation to evaluate the protocol, and takes for granted the tree formation process as well as the distribution of connectivity information across the tree. In practice, however, tree topologies [7][8] are never static and are continuously subjected to real-life link dynamics: parents, siblings, and children may and do change. To enable an implementation of WASP on Berkeley motes, we approximate the static tree that WASP presupposes by constructing a stable tree, *i.e.*, a tree that is solely constituted by links with high noise margins, or, equivalently, a high Received Signal Strength (RSS). We

 ISBN: 978-1-61208-103-8

will refer to links whose RSS exceeds a cutoff threshold Θ as highly reliable links. As long as the noise margins of its links are sufficiently high to withstand the link dynamics, our empirical evidence indicates that a stable tree can be expected to remain static with high probability.

B. Generation of a Stable Tree

We obtain a robust connectivity graph by blacklisting all links other than the highly reliable ones. A stable tree is obtained by using a randomized subset of the links in the robust connectivity graph. The stable tree formation process initiates from the sink and continues across the network until all the nodes learn about their local connectivity. In accordance with the WASP's requirements, every node has to have a highly reliable link to its single parent, and the nodes that share a parent are also required to have a highly reliable link to each other. Let $R_{i,j}$ denote the RSS measured at j when i transmits. The connectivity matrix between nodes i and j can be defined as

$$\mathrm{C}_{ij} = \mathbf{1}_{R_{i,j} \geq \Theta}. \quad (2)$$

By way of a sink-initiated connectivity discovery sweep, each node in Λ_m ($m > 0$) selects a unique parent and implicitly assigns itself to a given level Λ_m based on the availability of a highly reliable link to a unique parent and highly reliable links to one or more siblings (nodes that share the same parent). The basic condition for the assignment of a node j (with k as its parent) to level m is

$$\mathrm{C}_{jk}\mathrm{C}_{kj} = \mathbf{1}, k \in \Lambda_{m-1} \quad (3)$$

If multiple nodes satisfy (3), then we examine the cross-links between the nodes to find all pairs (i, j) such that

$$\mathrm{C}_{ji}\mathrm{C}_{ij} = \mathbf{1}, i \in \Lambda_m, j \in \Lambda_m \quad (4)$$

After arbitrarily selecting one pair (i, j) that satisfies (4), we check whether other nodes that satisfy (3) also have highly reliable links to both i and j (according to (4)). All such nodes are added to Λ_m. If no pair (i, j) exists that satisfies (4), one single node that satisfies (3) is selected. Once the stable tree has been set up, every node learns the structure of its subtree and sends it to its parent. Subsequently this structure is propagated until the sink receives it. At the end of this process the sink learns the structure of the whole tree and determines the number of slots in its SP in its own WASP-scheme accordingly.

Depending on the connectivity properties of the network, it may not be possible for the stable tree to span all the nodes for a given value of the cutoff Θ. Network partitioning is a well-known side effect of blacklisting [9], but it is not likely to occur at the levels of node density that are typical of WBANs.

C. Example

Let us illustrate the tree formation process with a sample network of ten nodes. Nodes are labeled with numbers ranging from 0 to 9, and the sink is node 0. The connectivity matrix of our sample network is displayed in Table III.

In the matrix highly reliable links are represented with a 1 and all other links with a 0. For our sample network the sink has highly reliable links to nodes 1, 2, 3, 4, 5 and 9, as shown in Figure 1. All these nodes satisfy (3) and are eligible to become children of the sink. We further check the cross-links and find the pairs of nodes satisfying equation (4). For our sample network such pairs are (1,3), (1,4), (1,9), (2,4), (2,5), (3,9), (4,5), (4,9) and (5,9). We arbitrarily choose pair (1,3) and further check for the remaining nodes satisfying (3) and (4). If we consider node 4, we find that it satisfies (3) with the sink as its parent; node 4 also satisfies (4) with node 1 but does not satisfy (4) with node 3, and therefore it is not elected to join the pair (1,3) as a sibling. Likewise, we check for all the eligible nodes and, as an outcome of this process, we find that node 9 is the only remaining node that satisfies both (3) and (4). Therefore nodes 1, 3, and 9 will belong to Λ_1. For our sample network the structure of the stable tree up to level 1 is given in Figure 2.

Table III
CONNECTIVITY MATRIX OF HIGH CONNECTIVITY NODES

NodeID	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
0	1	1	1	1	1	1	0	0	0	1
1	1	1	0	1	1	0	1	1	1	1
2	1	0	1	0	1	1	0	0	1	0
3	1	1	0	1	0	0	1	0	1	1
4	1	1	1	0	1	1	1	0	0	1
5	0	0	1	0	1	1	0	0	0	1
6	0	1	0	1	0	0	1	0	0	1
7	0	1	0	0	0	0	0	1	1	1
8	1	1	1	1	0	0	1	1	1	1
9	1	1	0	1	1	1	0	1	1	1

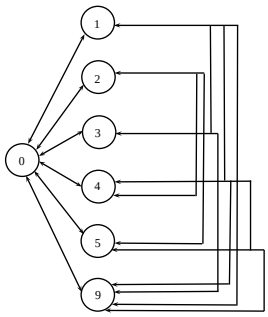


Figure 1 : Connectivity Graph of the Sink's Neighborhood

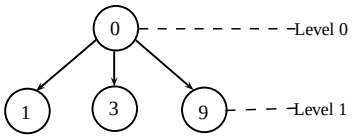


Figure 2: The Stable Tree Formation Up to Level 1

Nodes at Λ_1 start identifying their own children by using the same procedure as the sink. Let us assume that node 1 begins the child formation, followed by node 3 and 9. Highly reliable links to node 1 are 0, 3, 4, 6, 7, 8 and 9. Node 0 is the parent of 2 while nodes 3 and 9 are the siblings. Therefore the nodes that are eligible to become the children of node 1 are 4, 6, 7 and 8. Further we find that only pair (7,8) satisfies (4) and therefore belong to Λ_2. The procedure continues at nodes 3 and 9 and the outcome is shown in Figure 3.

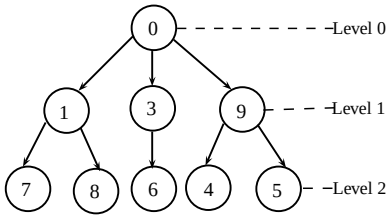


Figure 3 : The Stable Tree Formation Up to Level 2

Afterwards, Λ_2 nodes will also perform the same procedure for the formation of Λ_3. Node 2 will be the only node in Λ_3. For our sample network the final stable tree (comprising of three levels) is displayed in Figure 4.

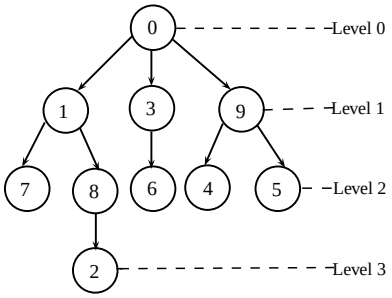


Figure 4 : The Final Stable Tree of Level 3

As soon as the complete tree is built, the sink learns the structure of the whole tree. After the tree formation is complete, the sink constructs its WASP-scheme and broadcasts it. To generate its WASP-scheme, the sink first calculates the SP and the TFS. In our example, $\Lambda_1 = \{1, 3, 9\}$, and from (1) $\mathrm{SP}_S = 4$. The TFS of the sink is given by the sum of all the forwarding slots needed by each child in Λ_1 and is therefore equal to 6. Table IV shows the SP and TFS for each node in two WASP-cycles.

Table IV
SP AND TFS VALUES IN TWO WASP-CYCLES

	WASP-cycle 1		WASP-cycle 2	
NodeID	**SP**	**TFS**	**SP**	**TFS**
0	4	6	4	6
1	2	0	2	1
2	1	0	1	0
3	1	0	1	0
4	2	0	2	0
5	1	0	1	0
6	1	0	1	0
7	2	0	3	0
8	1	0	2	0
9	0	0	0	0

IV. PROTOCOL IMPLEMENTATION AND EXPERIMENTAL RESULTS

We have implemented the WASP using MICAz motes and the TinyOS operating system. MICAz is a widely used research platform built around the CC2420 transceiver, which employs the 802.15.4 physical layer. We evaluated our implementation on an indoor testbed of ten MICAz motes. The testbed consists of a sink acting as the coordinator and nine other nodes that inject their data into the network so that it can be delivered to the sink. The sink in turn forwards everything to a base station node connected to a Crossbow MIB600 gateway that exports the data for offline processing. The 5ms slot length used as a simulation parameter in [1] is simply not workable with our hardware. To simplify the implementation we relax the slot length to one second, thereby eliminating the need for a tight time synchronization technique. Our own experimental results suggest that the RSS threshold $\Theta = -60$dBm is more than sufficient for the link dynamics that are typical of WBANs (in general, it is a rather conservative calibration).

Since the ultimate goal of WASP is energy efficiency, a paramount figure of merit is the duty-cycle, which we measure with online software estimation [10]. In our implementation, we compare WASP's own duty-cycling and the joint action of WASP's duty-cycling with LPL [2]. We use the standard TinyOS implementation of LPL, integrated in a link layer called BoX-MAC [11].

Each of our experiments was run for the duration of one hour. We explored two different LPL settings and ran experiments with sleep intervals of 100ms and 150ms for each node. In general, the LPL sleep interval must be significantly smaller than the duration of a WASP slot. Figure 5(a), 5(b) and 5(c) show the stable tree obtained by applying our tree formation algorithm in three different experiments (respectively with no LPL, with LPL at 100ms, and with LPL at 150ms). Although we employed a similar network setup in all experiments, the tree formation process (selection of a subset of the highly reliable links in the robust communication graph) is randomized, leaving us with

 ISBN: 978-1-61208-103-8

no control over the specific tree layout for each individual experiment. In general, because of the requirements that the WASP imposes on the tree topology, routes may be set up using more hops than needed based on connectivity; sacrificing a low hop count is certainly a drawback of the WASP approach [12].

In our experiments we measure the duty cycle and the Packet Delivery Ratio (PDR) of each node in the network. The PDR for a given node other than the sink is defined as the total number of data packets delivered to the sink over the total number of data packets injected by the node into the network. The PDR for the sink is the end-to-end delivery ratio computed as the total number of delivered packets over the total number of injected packets by all nodes. We group nodes based on their Λ_m ($m \geq 0$) membership and compare the duty cycle and the PDR of the nodes residing in the same level. We also measure the Control Overhead for the network, defined as the ratio of the number of dedicated control packets over the total number of transmitted packet (control and data packets). Tables V, VI, and VII show the duty cycle and PDR for all the experiments.

In Tables V, VI, and VII, we observe that by imposing LPL on the WASP protocol we obtain a sharp drop in the duty cycle of each node as compared to WASP's own application-based duty cycling. The sharp decrease in the duty cycle allows a drastic reduction of the overall energy consumption needed for network communication and boosts the energy-efficiency of the WASP protocol. We observed from Tables V, VI, and VII that the duty cycle for the sink without using LPL is 61.35%, which drops sharply to 9.48% with a 100ms LPL and 7.51% with a 150ms LPL. The drop in the duty cycle for the sink is around 52% with either LPL setting. Our results show that WASP can peacefully coexist with a low-power link layer and greatly benefit from it, because its baseline duty-cycling is not sufficiently aggressive to ensure significant energy savings. Nodes in Λ_1 are responsible for forwarding the data of the lower levels, and therefore they consume the largest amount of energy as they have to keep their radio on for the longest time. Our results show that both nodes that are very active (like Λ_1 nodes) and nodes that are see relatively little action (like the leaf nodes) can greatly benefit from the joint action of WASP and LPL.

LPL only takes a small toll on the reliability of the protocol. For our experiments the end-to-end PDR is 100% without LPL, 99.83% with 100ms LPL, and 99.82% with 150ms LPL. We obtain a better PDR performance compared to [1] where packet loss rate is 30%, but this is largely a byproduct of our relaxed time-slotting as well as the overly pessimistic channel model employed in the WASP simulations in [1].

We also measure the Control Overhead for the network. In our implementation the control packets are the packets broadcast for the network set up and the control WASP-scheme broadcast by the sink node. The Control Overhead is about 6% in all of our setups, suggesting that neither the tree layout nor the presence/absence of LPL has a considerable impact on the overhead. We display the total number of control and data packets transmitted by the WASP protocol (without LPL) over time in Figure 6.

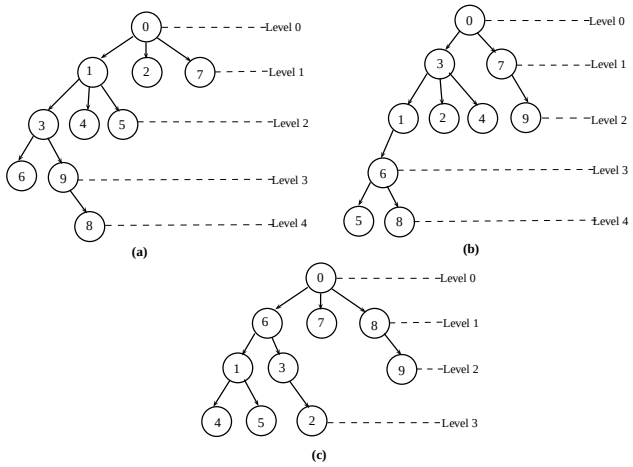


Figure 5: (a) Tree obtained with WASP-No LPL (b) Tree obtained with WASP-LPL-100ms (c) Tree obtained with WASP-LPL-150ms

Table V
DUTY CYCLE AND PDR VALUES FOR THE WASP-NO LPL

Level Λ_m **($m \geq 0$)**	**NodeID**	**Duty Cycle (in %)**	**PDR (in %)**
Λ_0	0	61.35	100
Λ_1	1	91.71	100
	2	22.76	100
	7	28.29	100
Λ_2	3	72.15	100
	4	31.04	100
	5	36.54	100
Λ_3	6	42.03	99.49
	9	61.25	99.49
Λ_4	8	86.08	99.49

Table VI
DUTY CYCLE AND PDR VALUES FOR THE WASP-LPL AT 100MS

Level Λ_m **($m \geq 0$)**	**NodeID**	**Duty Cycle (in %)**	**PDR (in %)**
Λ_0	0	9.48	99.83
Λ_1	3	15.64	100
	7	14.46	100
Λ_2	1	11.29	100
	2	3.63	100
	4	4.10	100
	9	4.11	99.49
Λ_3	6	10.33	99.66
Λ_4	5	7.91	99.49
Λ_4	8	9.91	99.49

 ISBN: 978-1-61208-103-8

Table VII
DUTY CYCLE AND PDR VALUES FOR THE WASP-LPL AT 150MS

Level Λ_m ($m \geq 0$)	NodeID	Duty Cycle (in %)	PDR (in %)
Λ_0	0	7.51	99.82
Λ_1	6	14.05	100
	7	2.71	100
	8	12.13	100
Λ_2	1	8.80	100
	3	7.38	100
	9	4.49	100
Λ_3	4	4.73	99.47
	5	5.49	100
	2	5.83	99.47

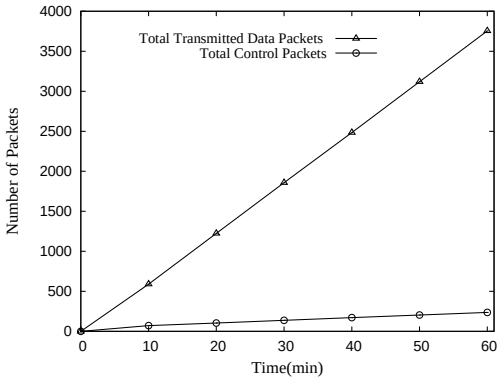


Figure 6: Total Control and Transmitted Data Packets for the WASP over Time

V. CONCLUSIONS

We successfully implemented the Wireless Autonomous Spanning Tree Protocol (WASP) on mote-class devices and evaluated its performance on 10-node testbed. We chose the WASP because of its promising cross-layer design approach that combines and coordinates medium access and multihop routing. The WASP presupposes the existence of a static tree structure to be superimposed to the network, but in practice low-power wireless connectivity is highly dynamic, even for networks of (almost) stationary nodes like WBANs. To obtain a stable tree topology out of inherently unstable low-power links, we adopted a blacklisting-based approach to build a tree of highly reliable links. While this approach is indispensable to approximate the static tree topology that the WASP presupposes, in practice it may lead to routes consisting of too many short hops, with the consequent loss of the benefits of long-hop routing [12]. The approximation of static connectivity also has scalability issues, and the lack of overlap between logical and physical (broadcast) connectivity may thwart WASP's slotted medium access with omnipresent hidden node effects. From the standpoint of energy-efficiency, while WASP's baseline duty-cycling is insufficient to meet the lifetime demands of WBANs, our results show that the WASP can be effectively complemented by a standard link-layer duty-cycling technique, which reduces the mean node duty-cycle from over 50% to about 8%, and the standard deviation from 25% to 4%.

REFERENCES

[1] B. Braem, B. Latré, I. Moerman, C. Blondia, and P. Demeester. The Wireless Autonomous Spanning tree Protocol for multi hop wireless body area networks. In *The 3rd Ann. Int. Conf. on Mobile and Ubiquitous Systems: Networks and Services (MOBIQUITOUS'06)*, San Jose, CA, USA, July 2006.

[2] J. Polastre, J. Hill, and D. Culler. Versatile Low Power Media Access for Wireless Sensor Networks. In *2nd ACM Conference on Embedded Networked Sensor Systems (SenSys'04)*, Baltimore, MD, USA, November 2004.

[3] A. El-Hoiydi and J.-D. Decotignie. Wisemac, An ultra low power MAC protocol for multi-hop wireless sensor networks. In *First Int. Workshop on Algorithmic Aspects of Wireless Sensor Networks (ALGOSENSORS'04)*, Turku, Finland, July 2004.

[4] M. Buettner, G. Yee, E. Anderson, and R. Han. X-MAC: a short preamble MAC protocol for duty-cycled wireless sensor networks. In *4th ACM Conference on Embedded Networked Sensor Systems (SenSys'06)*, Boulder, CO, USA, November 2006.

[5] W. Ye, J. Heidemann, and D. Estrin. An energy-efficient MAC protocol for wireless sensor networks. In *21th Annual Joint Conference of the IEEE Computer and Communications Societies (INFOCOM'02)*, pages 1567–1576, New York, NY, USA, June 2002.

[6] A. Woo, T. Tong, and D. Culler. Taming the Underlying Challenges of Reliable Multihop Routing in Sensor Networks. In *1st ACM Conference on Embedded Networked Sensor Systems (SenSys'03)*, Los Angeles, CA, November 2003.

[7] O. Gnawali, R. Fonseca, K. Jamieson, D. Moss, and P. Levis. Collection Tree Protocol. In *7th ACM Conference on Embedded Networked Sensor Systems (SenSys'09)*, Berkeley, CA, November 2009.

[8] D. Puccinelli and M. Haenggi. Reliable Data Delivery in Large-Scale Low-Power Sensor Networks. *ACM Transactions on Sensor Networks*, July. 2010.

[9] O. Gnawali, M. Yarvis, J. Heidemann, and R. Govindan. Interaction of Retransmission, Blacklisting, and Routing Metrics for Reliability in Sensor Network Routing. In *1st IEEE Conf. on Sensor and Ad Hoc Comm. and Networks (SECON'04)*, Santa Clara, CA, USA, October 2004.

[10] A. Dunkels, F. Osterlind, N. Tsiftes, and Z. He. Software-based on-line energy estimation for sensor nodes. In *4th Workshop on Emb. Networked Sensors (Emnets IV)*, Cork, Ireland, June 2007.

[11] D. Moss and P. Levis. BoX-MACs: Exploiting Physical and Link Layer. Technical Report 08-00, Stanford University, 2008.

[12] M. Haenggi and D. Puccinelli. Routing in Ad Hoc Networks: A Case for Long Hops. *IEEE Communications Magazine*, 43:93–101, October 2005.

 ISBN: 978-1-61208-103-8

Minimizing Energy Consumption Through Mobility in Wireless Video Sensor Networks

Moufida Maimour*, Khadidja Fellah†, Bouabdellah Kechar†, Congduc Pham ‡, Hafid Haffaf †
*CRAN laboratory, Nancy University, CNRS, France
Moufida.Maimour@cran.uhp-nancy.fr
†Oran Es-Sénia University, Oran, Algeria
fellahkhadidja@yahoo.fr, kechar.bouabdellah@univ-oran.dz,haffaf_hafid@yahoo.fr
‡LIUPPA laboratory, Pau University
congduc.pham@univ-pau.fr, France

***Abstract*—Energy is an important issue in designing wireless sensor networks. Coverage and connectivity are not of less important since they are necessary for the network to be operational. In this work, we consider the case of wireless video sensor networks where some sensors have visual capabilities. We study the benefit of having some mobile nodes able to move in the network field so coverage and connectivity constraints are satisfied while saving energy. We formulated this problem using integer linear programming. We performed several experiments using the CPLEX solver, to get some insight into the contribution of mobility in the context of video streaming. We mainly show that, even if mobility cost is much higher than communication, the latter tends to be predominant in the overall consumed energy as the video session duration increases. Thus, justifying the mobility cost.**

***Keywords*-Video; mobility; optimization; energy saving; topology control;**

I. INTRODUCTION

Recent advances in micro-electronics and wireless communications allow the emergence of wireless sensor networks (WSN) which are currently a hot research area [1]. Research effort in WSN mainly focused on scalar ones with large number of sensors able to sense environmental data (temperature, pressure, location of objects ...), perform specific processing on them and collaborate to achieve applications' requirements. More recently, the availability of low-cost CMOS cameras and microphones, Wireless Multimedia Sensor Networks (WMSN) [2] gained more interest and research effort. In a WMSN, the scalar WSN is strengthened by introducing the ability of retrieving richer information content through image and video/audio sensors [3][4][5][6]. This can significantly enhance a wide range of applications like object detection, surveillance, recognition, localization, and tracking.

While sensors are small devices mostly running on batteries, the network should operate autonomously for long periods of time in most applications. This is why energy is an important issue in WSN and becomes an important research topic. Other issues such as coverage and connectivity in a WSN are not of less importance. When some areas of the field become uncovered, the mission of the entire network may be affected especially when the uncovered area is security critical. Connectivity, for its part, allows the different sensors to be able to reach each other as well as the sink (central controller or a gateway). Lack of connectivity could create unconnected sets in the network leading to some sensors to be unable to reach the sink.

Due to connectivity and coverage issues, nodes have to be placed carefully when deployed in the network field according to the target application. Good coverage and strong connectivity can be achieved through careful planning of node densities and fields of view so the network topology can be defined before startup [7][8]. However, a sensor network is dynamic by nature since sensors stop working when they exhaust their on-board energy supply. In a dynamic, hostile or hard-to-access environment, there is a need to be able to dynamically redeploy the network such that the application's requirements in terms of coverage and connectivity continue to be met while saving energy. This is what we call On-demand repositioning. In [9] for instance, sensor's ability to move is used to distribute them as evenly as possible in the region so coverage is achieved within the shortest time duration and with minimal overhead. A survey on node placement in WSN can be found in [10].

In this work, we consider a wireless video sensor network (WVSN) where a subset of the nodes are equipped with cameras. We explore the possibility of having locomotion capabilities at some sensors so they are able to move [11]. The aim of this work is to save the overall communication energy in a video session by allowing mobile nodes to move. Even if mobility cost may be higher than communication, moves can be justified by preserving coverage and connectivity in the network. Moreover, moves are generally performed only once, at the beginning of a session, so video applications characterized by their large amount of data can have a small mobility cost as the video duration increases.

Our approach is based on linear programming where we

 ISBN: 978-1-61208-103-8

extended the work of [12] so both coverage and connectivity are considered. Additionally, our formulation fits the case of heterogeneous networks where video and scalar nodes coexist. Nodes may have different types of energy supplies (traditional batteries, solar or wind energy, etc.). Energy levels at nodes can be considered in the model so the network lifetime is increased. The paper is organized as follows. Section II summarizes the related work and Section III presents our network model with the different parameters and assumptions made in this work. Our problem formulation is presented in Section IV. Some numerical results are given in Section V. Finally, Section VI concludes and gives some future directions.

II. Related Work

The closest work to ours is the one of Kadayif et al. [12]. An integer linear program is proposed to minimizes energy consumption with the presence of mobile nodes. Our work is an extended version of the linear program they proposed so both coverage and connectivity are satisfied in a heterogeneous WSN. In such a network, nodes may have different type of energy supplies and can have different sensing roles (video/scalar) and capabilities.

Assuming mobile sinks, [13] considers the case of multiple sink nodes positioning so the network lifetime is maximized. The problem is formulated using a linear programming model. Bredin et al. [14] studied the problem of placing nodes at the network setup time where K-Connectivity is achieved. K-Connectivity implies having K independent paths among every pair of nodes. They formulated the problem as an optimization model where the number of additional nodes required by the K-Connectivity is minimized.

One important concern in nodes placement is field coverage. In [15] the problem of maximum lifetime sensor deployment with coverage constraints is considered and an energy-efficient INformation Gathering (SPRING) algorithm is proposed. Cardel et al. [16] addressed the coverage problem in WSN with adjustable sensing range. Based on the assumption that longer sensing ranges consume more energy, the aim is to give each sensor a coverage radius so the overall consumed energy is minimized while assuring the entire field coverage.

Jaggi et al. [17] considered the problem of maximizing WSN lifetime through activating a minimal set of sensor nodes at any given time while both coverage and connectivity constraints are satisfied. A linear program is formulated and a distributed algorithm for practical use in sensor networks is developed. The WSN considered is composed of static sensors. In this work, however, we aim to find optimal moves of mobile sensors so overall energy in the network is minimized.

III. Network Model

We consider a wireless sensor network of N sensor nodes among which some are video sensors located in strategic positions of a two-dimensional grid ($N_1 \times N_2$). All sensor nodes positions are assumed to be known and are given by a boolean matrix P:

$$p_{i,j} = \begin{cases} 1 & \text{if there is a sensor at position } (i,j) \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

where $0 \leq i \leq N_1 - 1$ and $0 \leq j \leq N_2 - 1$. This assumes that we consider a WSN with location awareness. Even if only a few nodes have known positions by equipping them with GPS or placing them deterministically, the remainder of nodes positions can be computed from knowledge about communication links [18].

The network is heterogeneous since it contains video and scalar sensors with different energies and processing powers. $c_{i,j,i',j'}$ is the amount of energy needed to transmit a 1-bit message by a sensor located at (i,j) and to be received by the one located at (i',j') and can be estimated using [19]:

$$c_{i,j,i',j'} = \alpha_{i,j}(2 \times E_{elec} + \epsilon_{amp} \times d^2_{i,j,i',j'}) \quad (2)$$

where $d_{i,j,i',j'}$ is the distance between the two sensors located at (i,j) and (i',j') positions, E_{elec} is the dissipated energy by the radio to run the transmitter or the receiver circuitry and ϵ_{amp} is the required energy by the transmit amplifier. We introduced a parameter $\alpha_{i,j}$, $0 \leq \alpha_{i,j} \leq 1$, defined on a per sensor basis in order to individually consider the energy capacities of each sensor node. For instance, a mobile node with solar cells can be assigned an $\alpha_{i,j}$ close to 0 and a node with a low energy level at a given time (possibly with ubiquitous energy) can be assigned an $\alpha_{i,j}$ close to 1. Sensors in the network can have different energy capacities. They can operate on batteries or even use energy extracted from the environment, such as solar energy or vibrations. This does not mean that the energy could become infinite [20] since harvesting energy can not be possible all the time and could be insufficient to provide sensors mobility for instance.

In our network model, some nodes have locomotion capabilities [11] so they are able to move. Their positions can be known thanks to the mobility matrix $B(N_1 \times N_2)$:

$$b_{i,j} = \begin{cases} 1 & \text{if the sensor at location } (i,j) \text{ is mobile} \\ 0 & \text{otherwise} \end{cases} \quad (3)$$

To move from point (i,j) to (i',j') in the sensor field, the required energy is noted $m_{i,j,i',j'}$ and assumed to drain much more energy compared to communication cost per bit for the same distances, that is,

$$\forall i,j,i',j' : m_{i,j,i',j'}/c_{i,j,i',j'} = \rho > 1$$

 ISBN: 978-1-61208-103-8

In order to cover a given region or to avoid obstacles, a video sensor with locomotion facility may move mainly as a response to a sink request. However, a video sensor is assumed to stay at its location for the whole session when it begins capturing/transmitting images. Since there is a big amount of data to be transmitted in a video session and assuming that the transportation path is provided from the network layer, a relatively long schedule of messages send/receive can be obtained. We note by L, the number of messages to be transmitted. S and R are the transmission and reception matrices respectively before move where $s_{i,j,l} = 1$ if node at position (i, j) (before moving) sends the l^{th} message to another node and $r_{i,j,l} = 1$ if node at position (i, j) (before moving) receives the l^{th} message from another node. Each sensor node has a radio communication range r_c which is fixed and can not be varied during the video session.

Finally, we assume that each sensor node is able to sense within a disk of constant radius r_s and introduce the notion of *coverage degree*. Noted d_c, it is the number of redundant sensors that cover a given area. For video sensors, we aim to obtain a *soft* video coverage as opposed to *hard* coverage. a video sensor is able to move when there is another node to replace it even if it is not a video sensor and can not insure the same service degree (rich video capture). Nevertheless, it can contribute in covering the sensor field by sensing other physical (scalar) phenomenon such as movement detection. In a hard video coverage however, a video sensor moves only if there is another video sensor that it is able to replace it in the coverage of a given zone.

Notations and different parameters and variables used in this paper are listed in tables I and II.

TABLE I
NOTATIONS: PARAMETERS

N	number of sensor nodes.
$N_1 \times N_2$	sensor field dimensions.
P	matrix position: $p_{i,j} = 1$ if there is a node at (i, j).
$d_{i,j,i',j'}$	the distance between sensors located at (i, j) and (i', j').
B	mobility matrix: $b_{i,j} = 1$ if node at (i, j) is able to move.
k	number of bits per message.
L	number of messages to send.
S	transmission matrix before move, $s_{i,j,l} = 1$ if node at (i, j) (before moving) sends the l^{th} message, $1 \leq l \leq L$.
R	reception matrix before move: $r_{i,j,l} = 1$ if node at (i, j) (before moving) receives the l^{th} message, $1 \leq l \leq L$.
$\alpha_{i,j}$	weight given to node located at (i, j).
ρ	ratio of mobility to communication per bit cost: $\rho > 1$
C	communication energy matrix: $c_{i,j,i',j'}$ is the required energy to send a 1-bit message by a sensor located at (i, j) and to be received by the another one located at (i', j').
M	mobility energy matrix: $m_{i,j,i',j'}$ is the required energy to move from point (i, j) to (i', j').
r_c	communication radio range of the different sensors.
r_s	sensing (coverage) radius of each sensor.
d_c	required degree of coverage.

TABLE II
NOTATIONS: VARIABLES

$\dot{S}$	sending matrix after move: $\dot{s}_{i,j,l} = 1$ if node at (i, j) (after a move) sends the l^{th} message to any other node, $(1 \leq l \leq L)$.
$\dot{R}$	reception matrix after move: $\dot{rp}_{i,j,l} = 1$ if node at (i, j) (after a move) receives the l^{th} message from any other node, $(1 \leq l \leq L)$.
Δ	movement matrix: $\delta_{i,j,i',j'} = 1$ if node at (i, j) moves to optimal location (i', j').
SR	send/receive matrix after move: $sr_{i,j,i',j',l} = 1$ if (after move) node (i, j) takes part in the communication of message number l and sends it to a node located at (i', j'), $1 \leq l \leq L$.

IV. PROBLEM FORMULATION

In this section, we present our formulation to the problem of minimizing energy through mobility while preserving connectivity and coverage in our relatively heterogeneous network as described in the previous section. The problem can be formulated as an integer linear program (ILP) as follows:

Minimize

$$E = \sum_{i=0}^{N_1-1}\sum_{j=0}^{N_2-1}\sum_{i'=0}^{N_1-1}\sum_{j'=0}^{N_2-1} \delta_{i,j,i',j'} \times m_{i,j,i',j'} + \sum_{i=0}^{N_1-1}\sum_{j=0}^{N_2-1}\sum_{i'=0}^{N_1-1}\sum_{j'=0}^{N_2-1}\sum_{l=1}^{L} sr_{i,j,i',j',l} \times k \times c_{i,j,i',j'} \quad (4)$$

subject to

$$\forall i \in 0..N_1-1, \forall j \in 0..N_2-1, \sum_{i'=0}^{N_1-1}\sum_{j'=0}^{N_2-1} \delta_{i,j,i',j'} = p_{i,j} \quad (5)$$

$$\forall i' \in 0..N_1-1, \forall j' \in 0..N_2-1, \sum_{i=0}^{N_1-1}\sum_{j=0}^{N_2-1} \delta_{i,j,i',j'} \leq 1 \quad (6)$$

$$\forall i' \in 0..N_1-1, \forall j' \in 0..N_2-1, \forall l \in 1..L, \quad \dot{r}_{i',j',l} = \sum_{i=0}^{N_1-1}\sum_{j=0}^{N_2-1} \delta_{i,j,i',j'} \times r_{i,j,l} \quad (7)$$

$$\forall i' \in 0..N_1-1, \forall j' \in 0..N_2-1, \forall l \in 1..L, \quad \dot{s}_{i',j',l} = \sum_{i=0}^{N_1-1}\sum_{j=0}^{N_2-1} \delta_{i,j,i',j'} \times s_{i,j,l} \quad (8)$$

$$\forall i \in 0..N_1-1, \forall j \in 0..N_2-1, \\ (p_{i,j}=1) \wedge (b_{i,j}=0) \Rightarrow \delta_{i,j,i,j}=1 \quad (9)$$

$$\forall i \in 0..N_1-1, \forall j \in 0..N_2-1, \forall l \in 1..L, \\ \sum_{i'=0}^{N_1-1}\sum_{j'=0}^{N_2-1} sr_{i,j,i',j',l} = \dot{s}_{i,j,l} \text{ with } d_{i,j,i',j'} \leq r_c \quad (10)$$

$$\forall i \in 0..N_1-1, \forall j \in 0..N_2-1, \forall l \in 1..L, \\ \sum_{i'=0}^{N_1-1}\sum_{j'=0}^{N_2-1} sr_{i',j',i,j,l} = \dot{r}_{i,j,l} \text{ with } d_{i,j,i',j'} \leq r_c \quad (11)$$

$$\forall i \in 0..N_1-1, \forall j \in 0..N_2-1, \\ \sum_{i''=0}^{N_1-1}\sum_{j''=0}^{N_2-1}\sum_{i'=0}^{N_1-1}\sum_{j'=0}^{N_2-1} \delta_{i'',j'',i',j'} \geq d_c \quad (12)$$

with $(i' \geq i-r_s) \wedge (i' \leq i+r_s) \wedge (j' \geq j-r_s) \wedge (j' \leq j+r_s) \wedge ((i \neq x) \vee (j \neq y))$ where (x,y) is the sink coordinates.

where E is the overall consumed energy including both communication and movement cost and k is the number of bits per transmitted packet. The different joined constraints are explained below:

(5): a sensor node can move to any non-occupied place and a move can only take place from an occupied position in the network [12].

(6): any move to a non-occupied position is performed by only one node ; otherwise this latter stays in its initial position [12].

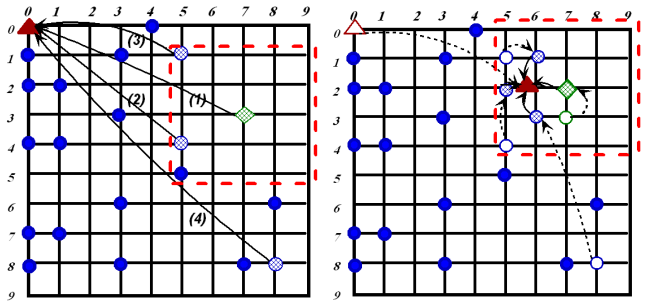
(7) and (8) give expressions of $\dot{S}$ and $\dot{R}$, the emission and reception matrices respectively after move.

(9): a non mobile node located at (i,j) (i.e. $b_{i,j}=0$) stays at its initial position.

(10) and (11) are the connectivity constraints. A message m is sent by one node and received by only one node (unicast communication). Moreover two nodes can not communicate unless they are in each other radio range. The distance between the two nodes (after move) is less or equal to the communication radio range [12].

(12) is the coverage constraint. Each position in the field is covered by at least d_c nodes to satisfy the required coverage degree. A node moves to position (i',j') from another one (i'',j'') or it stays at its initial position i.e. $i'=i''$ and $j'=j''$. Position (i,j) must be in the zone covered by the sensor located at (i',j').

Illustrative Example: we consider a field 10×10 where 20 sensor nodes are deployed as depicted by Figure 1(a) with 4 sources (at $(3,7)$, $(4,5)$, $(1,5)$ and $(8,8)$ willing to transmit one message each to the sink. Taking $r_s=2$, each sensor node covers in addition to its own position, the 24 neighboring ones: the node located at $(3,7)$ covers the square area within the dotted boundary as shown in Figure 1. In this sensor field, positions $(0,8)$ and $(0,9)$ are not covered. We assume that the communication radio range $r_s=4$ and that communications are only possible in single hop (there is no underlying routing protocol). All sources can not reach the sink in one hop.



(a) (b)

Figure 1. Illustrative Example: coverage and connectivity constraints

After applying our optimization program, all source nodes as well as the sink move as shown by dashed arrows in Figure 1(b). In this way, the required connectivity is satisfied so all the sources can achieve the sink in one hop. Additionally, the node located at $(3,7)$ moves to position $(2,7)$ so the problem of coverage is solved (points $(0,8)$ and $(0,9)$ become covered). The consumed energy is also reduced (for one message with 1024 bits, $401mJ$ is consumed instead of $403mJ$).

V. Numerical Results

In order to get some insight into the benefit of mobility to save energy in a WVSN, our formulated problem was coded using AMPL (A Mathematical Programming Language) [21] and solved using the CPLEX solver [22] on NEOS server [23].

We consider the case of a grid of dimension 10×10 where 40 nodes among which a given ratio is assumed to be mobile, are randomly placed. The sink is located at position $(0,0)$ and depending on the experiment, one to seven sources are randomly chosen in the field. Paths from each source to the sink are generated using MFR (Most Forward within Radius) [24]. Each source is assumed to capture and transmit a 10-second video sequence (*Hall Monitor* [25]). Data packets are assumed to have 1024 bits of payload. Information about paths, amount of data to be transmitted and the size of packets allow us to generate the corresponding communication schedule required as an input of our ILP. For the energy model, we put in equation (2), $E_{elec}=50nJ/bit$ and $\epsilon_{amp}=0.1nJ/bit/m^2$.

Figure 2 plots the mobility to the overall consumed energy ratio as a function of video duration for different values of ρ. In this scenario, 20% nodes have locomotion facilities and only one source is transmitting. The overall

 ISBN: 978-1-61208-103-8

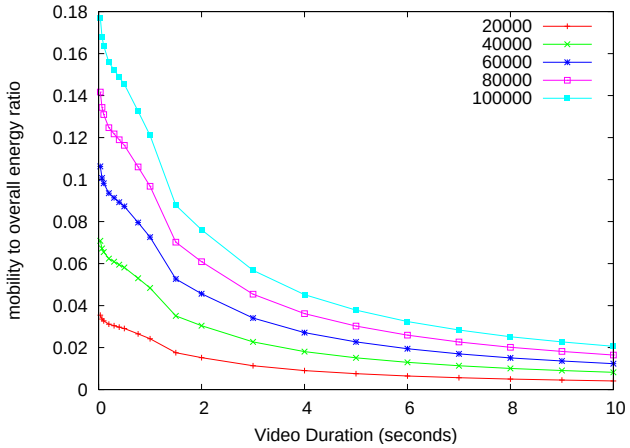


Figure 2. Mobility to overall energy ratio for different values of ρ. One source, 20% of nodes are mobile

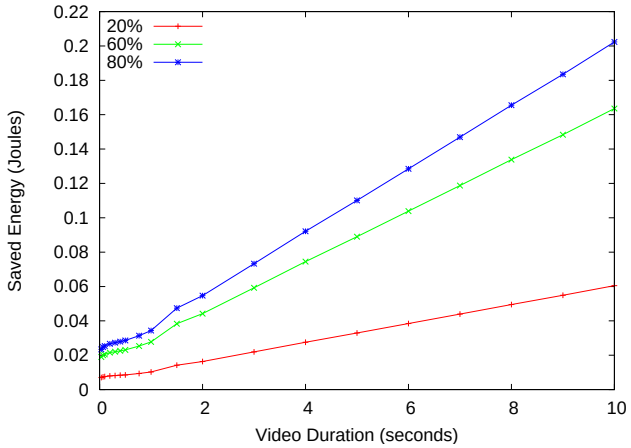


Figure 3. Saved energy as a function of video duration for different densities. $\rho = 1000$

consumed energy includes energy required by nodes to move to their optimal positions and the consumed energy due to transmitting/receiving data packets. We can see that if we increase ρ (till 100,000) so the mobility cost is much higher than the communication one and even for a small video duration (0.1 second for instance), mobility cost is at most about 18% of the overall consumed energy. It is also to notice that the share of mobility in overall energy consumption decreases with session duration. In fact, the longer the video session, the larger the amount of data to deliver. As a result, the communication cost increases compared the mobility one where moves are performed only once at the beginning of a session. Consequently, mobility is well justified in the context of WMSN characterized by their big amount of data.

In order to assess the gain obtained thanks to nodes mobility, we plot curves of Figure 3 showing the amount of energy (in Joules) saved when applying our optimization problem as a function of video duration for different densities of mobile nodes in the field. We can see that the amount of saved energy is higher for bigger number of mobile nodes. Furthermore, when the video session duration increases, saved energy is also increased. This confirms results obtained and observed in Figure 2. The amount of energy saved allows for augmenting the lifetime of the entire network.

Finally, we varied the number of transmitting sources from 1 to 7 and reported the amount of saved energy for different video streaming durations ranging from 1 to 5 minutes. Figure 4 plots this saved energy and shows, once again, that when increasing the video duration, the saved energy increases. When augmenting the number of sources until 5 sources, we save more energy. However when the number of sources reaches 6, we get less energy saving. This is due to the fact that when increasing the number of sources, some

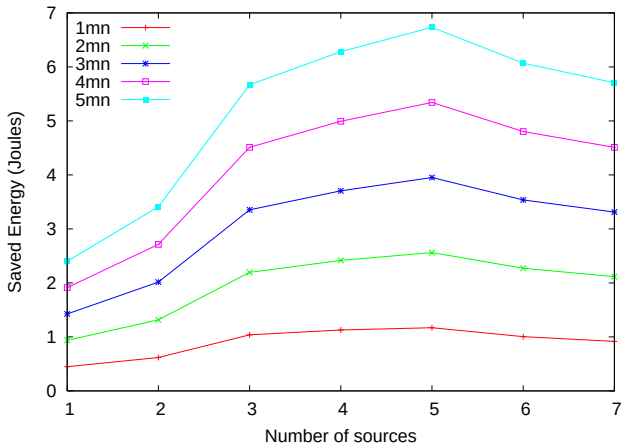


Figure 4. Saved energy as a function of the number of sources. $\rho = 1000$

nodes are likely to belong simultaneously to more than one path.

VI. Conclusion and Future Work

In this work, we formulated the problem of optimal node placement in a WVSN so energy consumption is minimized under coverage and connectivity constraints using ILP. We performed several experiments to get some insight into the benefit of having some mobile nodes in the network in the context of video streaming. We mainly showed that even if mobility cost is much higher than communication, the latter tends to be predominant in the overall consumed energy as the video session duration increases.

Our problem formulation is $O(N_1^2 \times N_2^2 \times L)$ and it is difficult to scale to large sensor networks. We suggest to execute it at the sink for relatively small networks (at the beginning of a video session) and off-line for larger ones for optimal initial deployment. This study allowed us to assess the contribution of mobility in saving energy. Our

 ISBN: 978-1-61208-103-8

optimal solution can be used to derive and evaluate the effectiveness of localized and less complex algorithms based on heuristics. Actually, we are developing and evaluating distributed heuristics that approximate the optimal solution.

ACKNOWLEDGEMENTS

This work was supported in part by the French National Research Agency (ANR TCAP project, Nb. 06-JCJC-0072) and the Tassili project (CMEP 09mdu784 and EGIDE 20281UB).

REFERENCES

[1] I. Akylidiz, Y. S. W. Su, and E. Cayirci, "Wireless sensor networks: A survey," *Computer Networks*, vol. 38, no. 4, pp. 393–422, December 2002.

[2] I. F. Akyildiz, T. Melodia, and K. R. Chowdhury, "A survey on wireless multimedia sensor networks," *Computer Networks*, vol. 51, no. 4, pp. 921–960, March 2007. [Online]. Available: http://dx.doi.org/10.1016/j.comnet.2006.10.002

[3] M. H. Rahimi, R. Baer, O. I. Iroezi, J. C. Garcia, J. Warrior, D. Estrin, and M. B. Srivastava, "Cyclops: in situ image sensing and interpretation in wireless sensor networks," in *SenSys*, 2005, pp. 192–204.

[4] W. chi Feng, B. Code, E. Kaiser, M. Shea, and W. chang Feng, "Panoptes: A scalable architecture for video sensor networking applications," *ACM Transactions on Multimedia Computing, Communications and Applications*, January 2005.

[5] P. Kulkarni, D. Ganesan, P. Shenoy, and Q. Lu, "Senseye: A multi-tier camera sensor network," in *ACM Multimedia*, Singapore, 2005.

[6] C. M., R. J.E., and Z. F., "Distributed attention for large video sensor networks," in *Intelligent Distributed Surveillance Systems seminar*, London, UK, 2004.

[7] M. Ishizuka and M. Aida, "Performance study of node placement in sensor networks," in *24th International Conference on Distributed Computing Systems Workshops - W7: EC (Icdcsw'04)*, Hachioji, Tokyo, Japan, March 2004.

[8] A. Boukerche, X. Fei, and R. B. Araujo, "A coverage preserving and fault tolerant based scheme for irregular sensing range in wireless sensor networks," in *49th Annual IEEE Global Communication Conference (Globecom'06)*, San Francisco, CA, November 2006.

[9] G. Wang, G. Cao, and T. F. La Porta, "Movement-assisted sensor deployment," *IEEE Transactions on Mobile Computing*, vol. 5, no. 6, pp. 640–652, 2006.

[10] M. Younis and K. Akkaya, "Strategies and techniques for node placement in wireless sensor networks: A survey," *Ad Hoc Netw.*, vol. 6, no. 4, pp. 621–655, 2008.

[11] G. T. Sibley, M. H. Rahimi, and G. S. Sukhatme, "Robomote: A tiny mobile robot platform for large-scale ad-hoc sensor networks," Washington, DC, USA, pp. 1143–1148, May 2002.

[12] I. Kadayif, M. T. Kandemir, N. Vijaykrishnan, and M. J. Irwin, "An integer linear programming-based tool for wireless sensor networks," *J. Parallel Distrib. Comput.*, vol. 65, no. 3, pp. 247–260, 2005.

[13] H. Kim, Y. Seok, N. Choi, Y. Choi, and T. Kwon, "Optimal multi-sink positioning and energy-efficient routing in wireless sensor networks," in *ICOIN*, 2005, pp. 264–274.

[14] J. L. Bredin, E. D. Demaine, M. Hajiaghayi, and D. Rus, "Deploying sensor networks with guaranteed capacity and fault tolerance," 2005, pp. 309–319.

[15] K. Dasgupta, M. Kukreja, and K. Kalpakis, "Topology-aware placement and role assignment for energy-efficient information gathering in sensor networks," in *ISCC '03: Proceedings of the Eighth IEEE International Symposium on Computers and Communications*. Washington, DC, USA: IEEE Computer Society, 2003, p. 341.

[16] M. Cardei, J. Wu, and M. Lu, "Improving network lifetime using sensors with adjustable sensing ranges," *IJSNet*, vol. 1, no. 1/2, pp. 41–49, 2006.

[17] N. Jaggi and A. A. Abouzeid, "Energy-efficient connected coverage in wireless sensor networks," Kolkota, 01/2006 2006, pp. 77–86.

[18] L. Doherty, K. S. J. Pister, and El, "Convex position estimation in wireless sensor networks," in *INFOCOM 2001. Twentieth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings. IEEE*, vol. 3, 2001, pp. 1655–1663 vol.3.

[19] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan, "Energy-efficient communication protocol for wireless microsensor networks," in *Proceedings of the 33rd Hawaii International Conference on System Sciences (HICSS'00)*, January 2000.

[20] D. E. Tiliute, "Battery management in wireless sensor networks," *Electronics and Electrical Engineering*, vol. 4, no. 76, pp. 9–12, 2007.

[21] K. B. W. Fourer R, Gay D M, *AMPL: A Modelling language for Mathematical Programming*. Duxury, 2002.

[22] "Cplex," www.cplex.com [accessed, August 2nd, 2011].

[23] "Neos," http://neos.mcs.anl.gov/neos/solvers/milp:scip/AMPL.html [accessed, August 2nd, 2011].

[24] X. Lin and I. Stojmenovic, "Location-based localized alternate, disjoint and multi-path routing algorithms for wireless networks," *J. Parallel Distrib. Comput.*, vol. 63, no. 1, pp. 22–32, 2003.

[25] "Evalvid," http://www.tkn.tu-berlin.de/research/evalvid/cif.html [accessed, August 2nd, 2011].

 ISBN: 978-1-61208-103-8

A Nontraditional Approach for a Highly Interactive Collective-Adaptive System

Emergent Lightweight Learning and Collective Adaptation

Dávid Lányi
Department of Telecommunications
Budapest University of Technology and Economics
Budapest, Hungary
dlanyi@hit.bme.hu

Borbála Katalin Benkő
Department of Telecommunications
Budapest University of Technology and Economics
Budapest, Hungary
bbenko@hit.bme.hu

***Abstract*—In this paper, we describe a nontraditional approach for realizing a highly interactive adaptive system with a minimal level of intrinsic knowledge; which opens up the way for a more open and active adaptation to the rules and dynamic changes of the environment. First, we discuss a lightweight, incremental adaptation model where knowledge and strategy emerge during (and throughout) the system's normal operation. Then, this model is extended with a collective mechanism: a society of learners makes use of each other's findings in order to converge faster. Finally, principles are evaluated with simulation in a theoretical showcase ('Connect-5') world.**

Keywords — collective-adaptive systems, open learning

I. INTRODUCTION

Emergence is one of nature's strongest weapons for handling complex, dynamic and large scale problems. The application of surprisingly simple algorithms, when done in multitudes, may lead to a different quality of global results. Certain fields of emergent behavior have already been utilized in computer science (such as self-organization), while other possible application areas are yet untouched. In this paper we propose to face the challenge of open adaptive systems with an emergent solution.

Adaptiveness–in our context–means that the system is able to dynamically respond to the changes of the environment. A traditional approach for achieving adaptivity is to maintain a self- and/or environment model, detect when changes occur, and explicitly start an adjustment process within the feedback loop. [1] Powerful tools like reasoning, semantics and ontology guarantee that the adaptation is effecttive and convergent.[3] However, the success of this approach largely relies on the accuracy of the system's *explicit knowledge*: on the completeness of the world *model* and on the efficiency of built-in *adjustment mechanisms*. As long as environmental changes are in line with it, the system is guaranteed to adapt efficiently; but it is theoretically impossible to react to changes that are not included in the model, or to choose adjustment strategies that were not encoded previously.[4,5,6] Our research focuses on the question: how much can we avoid this burdening explicit knowledge in an adaptive system, thus, unbind the learning process, and let the system openly find its way for achieving adaptation (instead of following what scientist taught it to do)? An extremity may be a system without any pre-injected model or strategy, an idea which opens up dimensions yet unknown in dynamic adaptation. In this paper we propose a knowledge-poor adaptation model (the amount of explicit, pre-injected model is kept very low) that we prove to be open and effective for a complex dynamic problem case.

Adaptivity is often required in distributed situations, where autonomous building blocks of the system need to find their optimum locally, without central help or control. Collective adaptive systems make use of the connectedness of individual blocks–through communication–in order to help converging faster or globally better. The presence of an explicit world model facilitates the collective behavior in this case, as members of the society share a common and well-defined understanding of the world and of possible strategies. The question we asked was: is it possible to establish meaningful cooperation between independent adaptive systems without sharing an explicit world model; is it possible to share ones expertise and help others to adapt better even if their–independently developed–knowledge and strategy is highly different? We describe a collective self-evaluation and expertise sharing mechanism for the society of knowledge-poor adaptive systems, and discuss the effects.

The structure of the paper is the following. Section II defines the problem space: a theoretical world with an adaptation challenge and easy measurability. Section III introduces the knowledge-poor learning and adaptation model. Section IV generalizes the model from a collective perspective, where independent learners share knowledge with each other. Section V discusses further aspects, consequences and limitations of the pro-posed models and algorithms. Section VI discloses evaluation results about the goodness of the individual and collective algorithms. Section VII concludes.

II. PROBLEM STATEMENT

The problem space tackled with in this paper is a world with actors who observe their environment and make actions from time to time. Certain states of the world mean reward to the actor who reaches it, while other states result in penalty. The world is only roughly modeled by the actors–the set of ob-servable factors is limited–and actors don't have pre-injected knowledge or assumptions about rules, requirements or strategies of the world. We focus on scenarios which can be described in means of discrete time-steps, one or more actors take part, and their actions modify to the world's state.

 ISBN: 978-1-61208-103-8

The environment is not only dynamic because of the presence of other actors (who also act), but also in means of general requirements or rules which may change from time to time. Actors get known with the actual requirements implicitly, through reward or penalty provided by the environment (reinforcement learning).

While numerous cases can be imagined satisfying the above specification; we choose a showcase in which the basic abilities of the system can be demonstrated and efficiently evaluated. This is actually a generalization of a simple two-player, deterministic, fully observable, zero-sum board game, often known as connect-5, gomoku, or amoeba.

A. Starting point: a basic Connect-5 World

Connect-5 is a simple board game, an extension of tic-tac-toe for bigger sized boards and longer combinations. There are two players in the game, one with mark X and the other one with mark O. The game starts with a board of tabular arranged empty square cells. Players make steps intermittently; each one places their mark onto an empty cell. A player wins the game, if five of their signs is placed consecutively in one row, column, or in a diagonal line on the board. When a player wins, the other one looses. Tie is reached, if the board has no more empty cells, but no one has won.

More formally, the *state* of the game is represented by the board itself (cells and their contents). The transition between states is the *action* of an actor, and the game is basically a time series of game states. Only one actor is allowed to perform action in each particular state. The action is mandatory, so if there is an empty cell on the board, the upcoming actor must act. After each action, the new state is evaluated by the environment, and if winner or tie state is reached, actors receive *feedback*.

Actors are able to observe a rough model of the actual environment at any time; and are also able to receive feedback about winning/losing/tie state. They also may accumulate a local knowledge base from these observations.

The goal of the actor is to perform actions that lead to a winning state, while the environment is dynamically modified by the opponent from time to time.

B. Generalized Connect-5 World

While the basic connect-5 world incorporates several important properties of our problem space, we decided to generalize it in order to include further aspects of dynamism and collectiveness.

- ***Collectiveness***. Instead of a single actor-opponent pair, the world consists of a society of players, engaged in multitudes of parallel games. The members of the society are still autonomous actors–with individual experience, strategy and decision ability –, but they also possess the ability of communicating with each other. Actors may share their expertise with other actors, or learn from others' shared knowledge. Please note that it's not guaranteed that the members of the society face the same problem instance, e.g., same opponent style or the same rules –; nor do we say that the knowledge of any individual agent is guaranteed to be of help for others. However, the pure ability of sharing one's dynamically built knowledge is an important attribute for a collective system, also from the theoretical point of view. Pair-wise knowledge sharing may also–but not necessarily–lead to the emergence of society level "common knowledge".
- ***Dynamic opponent style and strength***. The strategy, goodness and consistency of the opponent may change over time, resulting in dynamically changing environmental requirements from the actor's point of view. Extremities may be a random opponent (just picking random steps), and an analytically optimal opponent (using a mathematically optimal strategy).
- ***Changing game rules***. We also allow the game rules to be changed dynamically during the actor's life cycle. For example, the competitive aspect may be removed, so, the player gets rewarded for their own 5-long series regardless of the other player's moves. Another way of changing the rules is to modify the length of the required series: e.g., 4 or 10 items long series may mean victory.

The generalized model keeps the following attributes of the basic world (a) The states of the world form a time series. (b) Actors are able to observe the world's state and perform actions. They may receive feedback from the environment in certain states. (c) The world changes because of the actor's action or because of factors that are outside of the actor's control (e.g., opponent's action). Besides that, we also made the assumption that the actor has no pre-injected knowledge of the rules of the world or about the goal to reach–it has to reach (positive) game states by 'learning by doing'. [9] This may be a selfish requirement under static conditions (where explicit world mode-ling could result in optimal behavior from the startup), but our goal here is to ensure the openness of the system and its dynamic adaptation ability for immensely new requirements.

The state space in the showcase example–supposing a limited board size–is finite. However, we don't see that as a hard limitation from the theoretic side, because the observation ability of an agent is finite by definition (so any board size larger than the player's vision would work as infinite), and the mathematical model we use for learning can be easily extended for non-binary (multi-value or continuous interval) cases.

Summarizing the above, the agent's job is to learn the dynamically changing rules of the world through a feedback mechanism in order to select actions that lead to success; plus, to do this on-the-fly, without having had any pre-injected knowledge or preliminary training session; and possibly in a collective manner (by knowledge sharing).

III. BASIC LEARNING MODEL

This section describes the basic learning and adaptation model used within the open autonomous agents. The model combines known basic models (Markov Decision Process and Temporal Difference Learning) with specific extensions.

A. Markov Decision Process

The inspiration of our model comes from *reinforcement learning* (RL), a general approach that tackles with problems very similar to our problem statement. RL is not one specific

 ISBN: 978-1-61208-103-8

mechanism but a dynamically improving domain of machine learning models. The common approach [7] focuses on finding actions in an actual world state, in order to achieve a goal desired in that context. It is assumed, that an agent completing this task is able to sense the environment to some extent, is able to perform actions which influence that state, and is able to receive feedback from the environment about its success.

A widely used mathematical model describing reinforcement learning problems is the *Markov Decision Process* (MDP). It is defined as a five-tuple (S, A, R, P, γ), where S is the set of world states, A is a set of actions, P is a state transition probability function $P : S \times A \times S \mapsto [0, 1]$, (where $P(s', a, s)$ tells the probability of reaching state s' after performing action a in state s), R is the reward function $R \mapsto \mathbb{R}$, and γ is a discount factor from interval (0, 1]. It's important to note that the observation capacity of the agent is limited; hence, the state observed may significantly differ from the real world state. At this stage, S refers to the real world state (later, it will be replaced by the agent's perception).

RL models are often equipped with value functions and policies to help making automatic decisions. We follow the terminology and notation of [10]. There is a *value function* defined for states which is able to better describe the real utility of a state than the reward function itself (which would only tell whether state is terminal). The value function is associated with a *policy* π, which is a mapping from states to actions, $\pi : S \mapsto A$. A value function for a given policy, $V^{\pi} : S \mapsto \mathbb{R}$ is defined as the expected discounted sum of rewards received when starting (in t=0) from state s, and following policy π: $V^{\pi}(s) = E\left[\sum_{t=0}^{\infty} \gamma^t R(s_t) | s_0 = s, \pi\right]$. It can be shown [2] that this value function satisfies Bellman's equation, and may be expressed in the following way:

$$V^{\pi}(s) = R(s) + \gamma \sum_{s' \in S} P(s', a^{\pi}, s) V^{\pi}(s')$$

If the reward function R and transition probability function P are both known, this formula can be analytically solved as a linear system. However, in most RL settings–including our case–the probability function P is not known; the agent only has access to a subset of state transitions (own experience) and to the feedback coming from the reward function. [8]

To limit the size of $|S|$ –to keep computations on an easy-to-handle level –, often, estimations or approximations are used instead of exact models or values.

We also introduced a feature extraction step between the raw perceived state and the state principle used within the model. Feature extraction helps highlighting important properties of a state, by preprocessing the raw observation before learning. The exact realization of this feature extraction step is an important attribute of the agent, as the extracted information deeply influences the learning process. In the basic model, the feature extraction mechanism is wired-in (and this is all the knowledge–even though being implicit, we call it knowledge– the agent gets at startup). In general versions of the model, features may be introduced or removed on the fly.

In means of terminology, a feature based linear approximator [2] for the value function is defined as:

$$V^{\pi}(s) \approx w^T \phi(s)$$

where $\boldsymbol{\Phi} \in \mathbb{R}^k$ is a feature vector based abstraction belonging to the state s, and $w \in \mathbb{R}^k$ is a parameter vector.

While the usage of feature vectors was originally suggested in order to keep the computational complexity under control and to be able to deal with large or even infinite state spaces, we use it for two other purposes, respectively: (a) to facilitate convergence with the selection and usage of relevant and meaningful features, and (b) to bring openness into the model through the possibility of dynamically adding and removing features–hence refreshing the implicit world model of the agent.

With these approximations we lose the applicability of Bellman's equation, but there are other efficient ways for finding the solution, such as the LSTD.

B. Least-Squares Temporal Difference Method

The Least-Squares Temporal Difference (LSTD) algorithm provides way for finding a parameter vector w that approximately satisfies Bellman's equation. Without the full deduction of the method discussed in [10] and recalled in [2] we denote the main formulae, and review it from the aspect of our setting. LSTD attempts to find a fixed point of the approximation

$$w = \tilde{f}(w) = \underset{u \in \Re^k}{\operatorname{argmin}} \left\| \Phi u - (\tilde{R} + \gamma \Phi' w) \right\|^2$$

in which Φ and Φ' are matrices containing m samples of observed state transitions from s to s' in their rows, represented with $\Phi(s)^T$ and $\Phi(s')^T$ in each row; $\tilde{R}$ is a vector containing the obtained reward r_i for each of the m transitions. Because the term to be minimized contains Euclidian norms only, the optimal fixed point can be analytically determined by solving a linear system $\tilde{A}^{-1}\tilde{b}$, where

$$\tilde{A} = \sum_{i=1}^{m} \phi(s_i)(\phi(s_i) - \gamma\, \phi(s_i'))^T \qquad \tilde{b} = \sum_{i=1}^{m} \phi(s_i)\, r_i$$

In other words, the only knowledge required by the agent for selecting the desirable next state is only a vector (b) and a matrix (A).

- ***Vector b*** gives a picture about the perceived goodness of each state, based on the total (positive or negative) reward experienced there.
- ***Matrix A*** describes the experienced state transition pairs. Transitions model the effect of the agent's action along with the effect of the opponent's action, in one unit. The discount factor helps in distinguishing between states immediately preceding an end state and states that are far away. Please note that this abstraction does not include any preconception about the number or nature of opponents, so the model is also applicable for n > 2 players.

From our point of view, the most important property of vector b and matrix A is that they can be constructed iteratively; each new experience means a minor addition to them.

The agent may use two different approaches for translating the knowledge (matrix A, vector b) into an action:

(1) Calculate the expected effect of each possible action, and select the most desirable one based on the value of the result state. (If the number of actions is too large–or infinite–, a sampled subset of the possible actions may be used, with hierarchical refinement.)
(2) Analytically identify the most desirable result state and then search for an action that leads to it. This approach is more problematic because (a) it's not guaranteed that the state space is connected enough so an arbitrary end state can be reached from the current state, and (b) transitions are not deterministic because of the effect of the opponent, so we may end up in a very different end state than desired.

We used the approach (1) in the model.

IV. COLLECTIVE LEARNING MODEL

In case of the collective learning model autonomous agents share their locally developed knowledge with each other. Knowledge sharing is realized as multitude of pair-wise shares, in a self-organizing manner, without central control and without stashing common knowledge centrally.

This requires the followings: (a) Knowledge import model: a mechanism to integrate external knowledge into the one's own knowledge base, (b) Self-evaluation mechanism: a metric for the agent to evaluate the goodness of its knowledge in the actual environment, and (c) Sharing and acceptance mechanism: a mechanism that initiates and controls knowledge sharing / acceptance. Participants of the share–donor and receptor–are autonomous elements, so it's their free decision what, when and with whom to share or accept.

A. Knowledge Import Model

The import model uses the previously described matrix A and vector b as the manifestation of the knowledge; this is what the donor shares with the receptor. As shown previously, A and b are built up iteratively, thus, they're additive.

We defined the knowledge import mechanisms the following way: the new knowledge of the receptor is a weighted combination of its old and the donor's shared knowledge. Weights are denoted by c_1 and c_2.

$$A_{new} = c_1 A_{original} + c_2 A_{import} \qquad b_{new} = c_1 b_{original} + c_2 b_{import}$$

Weights define the influence of the imported elements. An extreme case is when the original knowledge gets zero weight meaning that the imported knowledge replaces the receptor's own knowledge (suppressive import). In this case the receptor becomes the donor's equal copy or clone. This may be desirable if the imported knowledge is guaranteed to be of high value while the knowledge of the receptor is clearly non-performing. However, suppressive import may easily lead to a drastic drop in the population's diversity which may become dangerous when the environment changes.

Non-suppressive import of good knowledge may perform somewhat weaker on the short term, but it keeps the population diverse which is a useful property on the long term. Combinatory import may also accumulate a more general knowledge than suppressive one, because it tends to store information about uncommon parts of the state space (which may become handy when usual strategies stop working).

B. Self-Evaluation Mechanism

The self-evaluation mechanism of the agents is based on a sliding window memory about the outcome of the last few games. Contents of the sliding window are summarized: a game won counts as +1, a lost game counts as -1 and tie counts as zero. (This is just a despotic choice, more complex models could also consider trends, the goodness of the opponent etc.)

C. Sharing and Acceptance Decisions

We didn't define explicit triggers for knowledge sharing because we believe it's not possible to say that a certain knowledge instance is guaranteed to be helpful or unhelpful for others. Instead, a sharing protocol was defined: when a donor is ready to share, it contacts another agent who–if decides so–becomes the receptor. We examined the following sharing patterns:

- Random sharing model. Agents initiate/accept the transaction with a given probability.
- Self-confidence based model. Agents with high self-evaluation values offer their knowledge, and agents with low self-evaluation values accept it.
- Knowledge density (completeness) based model. The goodness of the knowledge is measured by its completeness (e.g., number of filled cells in A).
- Opponent-based model. The receptor prefers knowledge that contains data about its current opponent.

Our model extends the state of the art in the followings: (a) applies temporal difference (TD) learning for the problem of adaptive systems where requirements change dynamically over time (b) introduces the possibility of on-the-fly, automatic feature injection (c) brings TD learning into a collective dimension.

V. DISCUSSION

This section discusses consequences, generalization directions and limitations of the pervious models.

The descriptive properties of the model were partially covered in Section III: the model is suitable for problems where the state of the environment changes from time to time and the actor is able to perform actions picked from a finite or infinite set of possible actions. Opponents and environmental rules are not directly modeled within the agent's knowledge, so the model is generally applicable for multi-actor situations. The learning process is knowledge-poor, so, except for the initial features, the agent does not need pre-injected knowledge.

Learning happens naturally, during the agent's normal activity; which is in contrast with today's popular adaptive system approaches, where the learning phase precedes the phase of normal operation.

The most important factors influencing the learnability of a problem are (a) what the agent perceives from the world, hence feature extraction, and (b) how well the actual problem case is presented, hence, the behavior of the opponent.

Opponents may significantly influence the convergence of the learning process, especially for upexperienced agents. When playing against a dummy (e.g., random) opponent, the

 ISBN: 978-1-61208-103-8

agent easily spends significant amount of time in irrelevant sections of the problem space–as none of the players knows how to become successful. In case of a strong opponent the agent learns fast what to avoid and, probably, also what to do to win (see Section VI). It's an interesting question whether the strongest opponent is the best, or it's more optimal to learn against consequent but imperfect players. The advantage of an imperfect opponent is that it leaves space for the agent to learn how to correct errors and how to make use of the other's mistakes. We believe, and tests indicate, that the variety of opponent styles leads to the best kind of knowledge for static and dynamic cases, respectively.

Feature extraction is the heart of the agent's learning model. A novelty in our model is that features are not bound to be static: they may be introduced or removed dynamically, during runtime.

- Features may be introduced (a) at knowledge import, where the donor agent does not only transfer its knowledge but also the mechanism how the new feature can be calculated, or (b) through systematic generation where the agent uses data mining based techniques to generate new features or to derive them by combining existing ones. (The exact mechanism of data mining based feature generation is outside of the scope of this paper.)
- Irrelevant features may be removed in order to minimize the problem space, thus, facilitate convergence. The trigger of that may be (a) knowledge import where the agent may decide to remove those features that are missing from some/ any of the knowledge bases (b) in an explicitly executed feature optimization step which may be a mathematical matrix minimization method (e.g., principal component analysis or singular value decomposition) or the society may use a genetic algorithm based feature selection. (Experiments confirmed both directions, but the details are again outside of the scope of this paper.)

Too fast convergence in the knowledge may be dangerous because it develops over-specialized strategies that work well against the current opponent but may not help if the environment changes. To avoid overspecialization, the agent may choose prevention strategies, such as picking second-best directions. Such a strategy leads to a better coverage of the problem space, which may be suboptimal in the current game, but could help against future opponents with yet unknown strategies.

The challenge of the agent is not just to learn adapting to (playing well within) the current environment; but to *adjust to the dynamic changes of the environment*. Changes may range from mild (slightly different opponent style) to drastic (essential rule change in the game). Agents may face this challenge alone or as a society.

- Standalone adaptation strategies include: (a) for slight changes: prevention of over fitting by better problem space coverage, and (b) for drastic changes: self-evaluation triggered knowledge deprecation –when the agent feels a significant performance drop it devalues or completely clears up its existing knowledge.
- Collective adaptation strategies include: (a) when the problem space is locally homogeneous: learning from neighbors, (b) knowledge generalization through sharing and combination and (c) for drastic changes: fast, population-wide propagation of the up-to-date knowledge.

Collective mechanisms may be biased by distortions of the self-evaluation metrics. Self-evaluation is intrinsically subjective; the agent possesses empirical information only which means that the metrics is biased by its experience–opponents–by definition. Agents facing weak opponents may overvalue themselves, while agents with strong opponents may do the opposite. The evaluation bias may gain attention when it comes to sharing the knowledge: a receptor in a hard environment, so with low self-confidence, may overvalue the weak-environmental donor's knowledge just because of its false self-confidence. However, it would be heedless to say that receiving such knowledge–unless suppressive–is guaranteed to be unhelpful. When treating it (choosing c_1 and c_2) with caution, even that kind of import may prove to be useful, because it covers a different, yet un-known sub-domain of the problem space (see Section VI).

Altogether, we think that the descriptive power and generalizability of the model is high. We can also imagine a possible convergence between this knowledge-poor approach and today's knowledge intensive directions, where the two may strengthen each other (e.g., in feature generation).

VI. EVALUATION

Models were evaluated through simulation. This section includes the most important results about the standalone learning, adaptation ability, and the collective dimension.

The *standalone model* was evaluated along two lines. First we wanted to see, how efficient is the on-line learning ability in the connect-5 world, with no initial experience, trained against opponents with different strengths. In this setting, we also measured, how the self-evaluation mecha-nism performs compared to an objective evaluator. After this we examined how trained agents react to drastic environmental changes.

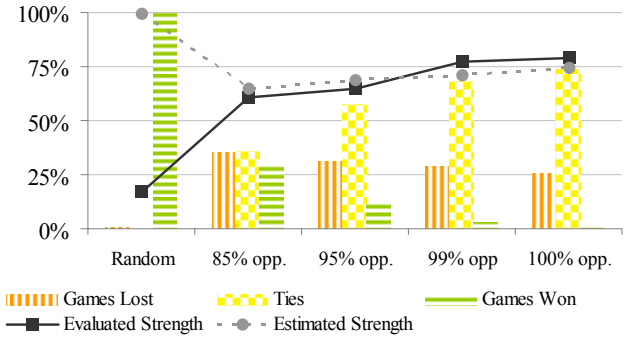


Figure 1. Learning characteristics and self-evaluation vs. opponent style.

On-line learning with no initial knowledge. The experiment consisted of an adaptation phase and an objective evaluation phase. (1) First, the untrained agent plays 350 games against a fixed-algorithm opponent, and uses its learning mechanism to adapt. Win/lost/tie statistics were also collected here. We evaluated five identical agents, each playing with a different opponent, namely: one random player; and the four opponents with mathematically optimal strategies but with

 ISBN: 978-1-61208-103-8

some–15%, 10%, 1% and 0%–chance of making an error (failing to choose the perfect action). (2) In the evaluation phase, learning was switched off in order to get an unbiased picture about the knowledge of each agent. Agents were allowed to use their existing knowledge, and were evaluated by playing 100 games against the "perfect" (0% failure rate) opponent, as an absolute measure. Their preliminary self-evaluation (based on the training phase) was compared to the actual measured strength. (Strength is defined as the percentage of non-lost games.)

Columns in Figure 1 visualize the outcome of the adaptation phase, while the curves refer to the self-evaluated and objectively measured strength. Training results show that the number of games won by the agent falls as opponents get stronger. Surprisingly, the number of lost games does not increase with stronger opponents; instead, games tend to end more often with a tie. Evaluation results indicate that the real gameplay strength is higher for agents trained against stronger opponents. Please note, that although the agent trained with the random player holds the lowest strength, it could also fray out a tie in 17 percent of the games against the strongest opponent. The difference between the self-estimated strength and the actual strength is unexpectedly small, expect for the divergent (random) training environemnt.

Adaptivity. The second experiment examines the level of adaptivity to world changes. We used a trained agent, which had a training session of 50 connect-4 games (a game with the same rules as connect-5, except that the combination of four is enough for the victory). Then, the agent had to play connect-5 against the strongest opponent, without any notification or adjustment regarding the rule change. Figure 2 shows that in the first 25 games the agent had serious problems using the experience gathered earlier, resulting in a defeat rate of almost 96 percent. Although, after 50 games it could defend with 50 percent accuracy, and after 125 games, it reached almost the same strength level, as in the previous on-line learning test. Tests with other rule change schemes (C-5 to C-4, no competitiveness) brought similar results.

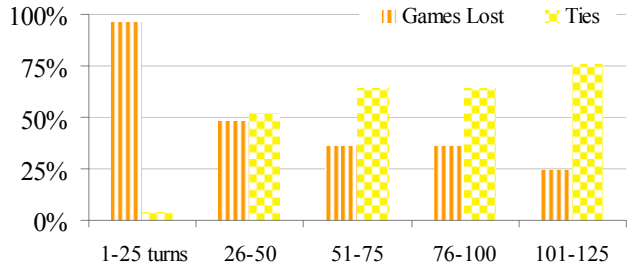


Figure 2. Adaptation to changing rules (Connect-4 to Connect-5).

The ***collective dimension*** was evaluated along various aspects, here we present one. Differently trained agents got knowledge injections, and then, got evaluated off-line. Listed combinations are: empty (untrained) receptor + best trained donor (trained with the strongest opponent), randomly trained receptor + best trained donor, Connect-4 trained receptor + best trained donor, and best trained receptor + random donor. Figure 3 shows that knowledge injection had positive effects in all cases. This is not surpising in the first three cases when the injected knowledge was clrealy more accurate than the agent's own. In the last case, a good agent received "worse" knowledge, and still, this helped it to gain winning which was out of ques-tion beforehand (however, general strength dropped slightly). This effect can be explained with the nonlinearity of the problem space.

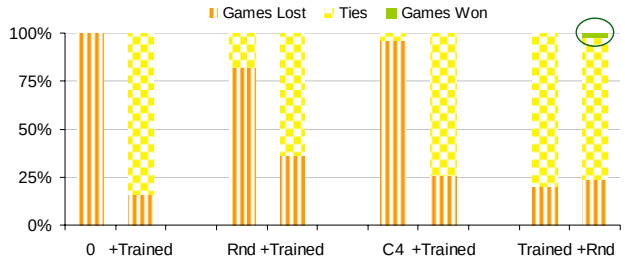


Figure 3. Collective effects (evaluated against the 100% opponent).

VII. SUMMARY

We described an emergent, knowledge-poor and open approach for adaptive systems where adaptation emerges from simple steps during the system's normal operation, even amongst drastically changing environmental rules. Dynamic features and the lack of mandatory and too explicit semantics bring real openness. The cooperative knowledge sharing mechanism brings the adaptation of knowledge-poor learners to a collective level.

REFERENCES

[1] Brun Y. et al., "Engineering self-adaptive systems through feedback loops," LNCS 5525/2009. Springer, Heidelberg, 2009.

[2] Kolter J.Z. and Ng A.Y., "Regularization and feature selection in least-squares temporal difference learning," Proc. of the 26th Annual International Conference on Machine Learning (ICML 09), vol. 382, 2009, pp. 521-528.

[3] Goldsby H.J. et al., "Goal-based modeling of dynamically adaptive system requirements," Proc. of 15th Annual IEEE International Conference on the Engineering of Computer Based Systems (ECBS), 2008.

[4] Benkő B. K., Brgulja N., Höfig E., and Kusber R., "Adaptive services in a distributed environment," Proc. of 8th International Workshop on Applications and Services in Wireless Networks, Kassel, 2008.

[5] Nelly Bencomo N. et al. "Genie: supporting the model driven development of reflective, component-based adaptive systems," Proc. of the 30th International Conference on Software Engineering, Leipzig, 2008, pp 811-814.

[6] Harel D and Marelly R., "Come, let's play: scenario-based programming using LSCs and the play-engine. Springer, Heidelberg, 2005.

[7] Sutton R.S. and Barto A.G., "Reinforcement learning: an introduction," The MIT Press, 1998.

[8] Sutton R. S., "Learning to predict by the methods of temporal differences," Machine Learning 1988/3, 1988, pp. 9-44.

[9] Baxter J. and Weaver A., "TDLeaf(lambda): Combining temporal difference learning with game-tree search," In Proceedings of the 9th Australian Conference on Neural Networks, 1998, pp. 39-43.

[10] Bradtke S.J. and Barto A.G., "Linear least-squares algorithms for temporal difference learning," Machine Learning 1996/22, 1996, pp. 33-57.

 ISBN: 978-1-61208-103-8

MTM Parameters Optimization for 64-FFT Cognitive Radio Spectrum Sensing using Monte Carlo Simulation

Owayed A. Alghamdi, Mosa A. Abu-Rgheff and Mohammed Z. Ahmed
Mobile Communications Research Group
University of Plymouth
Plymouth PL4 8AA, UK
{owayed.alghamdi, mosa, m.ahmed}@plymouth.ac.uk

***Abstract*—This paper presents parameter optimization of the multi taper spectrum estimation method (MTM) for 64-FFT based cognitive radio (CR) spectrum sensing. The design problem is formulated to determine the MTM parameters pair; the half time bandwidth product, and the number of tapers that maximize the performance at a fixed number of data samples. Maximum performance is defined by the highest probability of detection at a fixed false alarm probability. A Monte Carlo simulation is implemented to find the optimal parameters. The binary hypothesis test is developed to insure that the effect of choosing optimum MTM parameters is based upon performance evaluation. The whole band under sensing is divided into subbands, some contain primary user signal (PR), and the other does not. Consequentially, in addition to the variance of the estimate, the spectral leakage outside the PR subband is included in the performance evaluation. We found that the half time bandwidth product of 4 and 5 tapers gives the highest performance. We examined both MTM and periodogram (i.e., energy detector) methods in Gaussian (AWGN) and Rayleigh flat fading environments. The CR system performance using the MTM technique outperforms the performance of the same system that uses periodogram in all the cases we examined.**

Keywords-cognitive radio; spectrum sensing; multitaper spectrum estimation.

I. INTRODUCTION

High data rate applications in the emerging wireless technologies are faced with the problem of the ever-increasing scarcity of spectrum, coupled with the underutilization of the current licensed spectrum. Cognitive radio's basic idea is the opportunistic use of the unused spectrum of a licensed PR user. Consequently, CR technology is expected to become an increasingly popular part of future wireless networking technologies.

Cognitive radio, proposed by Mitola in 1999 [1], addresses the problem of secondary usage of underutilized spectrum using techniques of accurate spectrum sensing. It intelligently interacts with its operational environment to dynamically and autonomously adjust the radio operating parameters accordingly, to avoid interference with PR transmission.

The key enabling functionality for practical CR concept is a reliable spectrum sensing scheme to avoid harmful interference to licensed users. The classical spectrum sensing techniques such as the matched filtering and the cyclostationary detector have high performance for CR applications [2-4]. But such techniques require prior information about PR's signaling. The periodogram (i.e., energy detector) is a simple method at the expense of performance. Large variance, and bad biasing of the power spectrum estimates are main drawbacks of periodogram [5].

Thomson proposed the 'Multitaper spectral estimation Method (MTM)' to produce single spectrum estimate by multiplying the sampled data by several leakage resistant tapers [6]. Haykin , on the other hand, suggested the use of MTM as an efficient method for spectrum sensing in cognitive radio systems [7].

The CR systems using the Orthogonal Frequency Division Multiplexing (OFDM) technology have the ability to dynamically fill the spectrum holes by activating the available OFDM subcarriers, and deactivating the remaining subcarriers. The FFT operation in the OFDM demodulation process can be used for the analysis of the spectral activity of the licensed users [8].

Practically, using the MTM in the OFDM-based CR systems will be supported by the already available IFFT/FFT processors to perform the spectrum estimations. MTM has to be optimized for implementation in OFDM-based CR systems. Half time bandwidth product (*NW*) and the number of tapers (*K*) play key functions in the MTM process. In [9], and in [10], the recommended range of *NW* is recommended to be between 4 and 10, and *K* between 10 and 16.

Based on these recommendations, it is clear that such parameters are still an open issue, and have to be optimized towards achieving high performance and low complexity by determining a specific number of tapers.

In this paper, we consider issues of optimizing the MTM parameters for 64-FFT CR systems. In order to determine the optimal *NW*, and *K* we examine the performance using the

 ISBN: 978-1-61208-103-8

binary hypothesis. The objective is to include the spectral leakage effect and the large variance as performance metric parameters in the evaluation to determine the *NW*, and *K* that maximize the performance. A Monte Carlo simulation is implemented for the formulated problem. A comparison to the periodogram is presented in term of performance and complexity.

The rest of the paper is organized as follows: Section II defines the model for the system under consideration and reviews MTM technique. Section III considers the optimization of the MTM parameters used in the system model. Section IV presents the results and Section V concludes the paper.

II. SYSTEM MODEL

Our system model consists of a single PR transmit/receive node, transmitting QPSK-OFDM signal in the sub-band between f_a and f_b as shown in Fig. 1, and an OFDM-based CR sensor (node) that detects the PR user's signal and decides whether the PR's signal is present or absent in the searched frequency band.

A family of orthonormal tapers is generated using Discrete Prolate Slepian Sequences (DPSS) [10], of length N to concentrate the received PR energy in the frequency interval Δf between$(-W, W)$. The total number of sequences (tapers) produced, is $N_{tapers} = 2NW = N\,\Delta f$, and K is the number of tapers used in the estimation. The associated eigenvalues of the K tapers, are $1 > \lambda_0(N,W) > \lambda_1(N,W) > \lambda_2(N,W) > \cdots > \lambda_{K-1}(N,W) > 0$. The k^{th} taper is represented by $v_t^{(k)}(N,W)$, where $t = 0,1,\dots,N-1$, is a time index.

The received PR signal at a CR sensor (node) is sampled to generate a finite discrete time samples series $\{x_t; t = 0,1,\dots,N-1\}$ that is 'dot multiplied' with different tapers. The product is applied to Fourier Transform to compute the energy concentrated in the bandwidth $(-W, W)$ centred at frequency f. For K orthonormal tapers, there will be K different eigenspectrums produced and defined as :

$$Y_k(f_i) = \sum_{t=0}^{N-1} v_t^{(k)}(N,W) x_t e^{-j2\pi f_i t} \qquad (1)$$

where $f_i = 0, 1, 2, \dots, N-1$ are the normalized frequency bins.

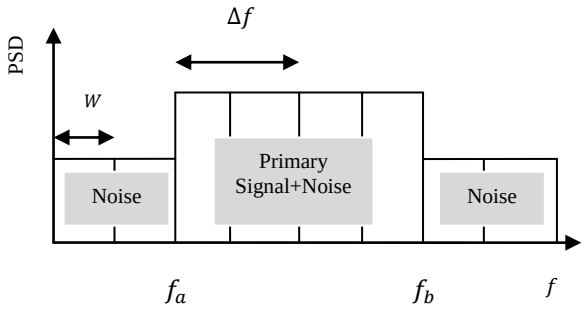


Fig. 1. System Model.

The spectrum estimate given by Thomson's theoretical work is defined as:

$$S_{MTM}(f_i) = \frac{\sum_{k=0}^{K-1} \lambda_k(N,W)|Y_k(f_i)|^2}{\sum_{k=0}^{K-1} \lambda_k(N,W)} \qquad (2)$$

On the other hand, the periodogram method, when the samples are taken at uniform time spacing, gives the power spectrum density estimation as:

$$S_{PE}(f_i) = \frac{1}{N}\left|\sum_{t=0}^{N-1} x_t e^{-j2\pi f_i t}\right|^2 \qquad (3)$$

III. OPTIMAZIATION OF MTM PARAMETERS

Maximum-likelihood methods provide an optimal estimate of the power spectrum. MTM technique is an approximation to the Maximum-likelihood power spectral estimates but at reduced computation [11], [12]. The main motivation of the work presented in this paper is to fnd the MTM parameters that optimize the performance of this technique.

In this section, we investigate the optimization of MTM parameters in OFDM-based CR systems. The CR transceiver carries out 64-IFFT/FFT digital processing for both transmission and receiving operations. Consequently, the MTM processing in the spectrum sensing will not add additional hardware at the receiver except for taper sequences generation, multiplication and adding operations.

MTM tolerates the classical problems which occurred in spectrum estimation by averaging over a number of orthonormal tapers/windows. The tapering sequences concentrate the energy within a bandwidth $2W$, where $0 < W < 1/2$. The half time bandwidth product NW determines the bandwidth resolution for fixed length$N$. As the half time-bandwidth product decreases, the half bandwidth W decreases resulting in higher resolution in the spectrum sensing and vice versa. The main spectrum lobe of each taper/window is $2NW$ frequency bins (where the FFT- frequency bin spacing is $1/N$) [13]. Thus in OFDM-based CR applications with 64-FFT, the main band under sensing can be divided into a number of subbands based on the half time bandwidth product. For example using NW=2, means that there will be 16 subbands with $2W$ width each, and then the main lobe is 4 frequency bins out of the 64. Therefore in such applications the useful half time bandwidth products should be 0.5,1,2,4,8, or 16, and 32 to concentrate the energy in one band, which is the whole band under sensing; consequently, the higher edge of the half time bandwidth is 16.

Furthermore, the number of the tapers in the higher resolution sensing is smaller than that in the lower resolution since the total number of tapers is $N_{tapers} = 2NW$.

The eignevalues $\lambda_k(N,W)$ of the first few tapers for the higher bandwidth resolution is much smaller than eignvalues in the lower resolution which implies that lower bandwidth resolution sequences have more energy concentration than sequences in the higher bandwidth resolution.

 ISBN: 978-1-61208-103-8

Furthermore, the first few eigenvalues of a specific time half bandwidth product are close to one. As the number of taper sequences increases, the eigenvalues decrease indicating bad bias properties, and as the number of tapers decreases, the eigenvalues increase towards 1 indicating good bias properties.

Our work for choosing appropriate values of *NW* and *K* for MTM estimator uses two approaches: in the first approach, we compute the power spectral density using (2) to show that random choice of these values may generate a lot of leakage causing an increase in false alarm probability during the estimation process. These results are presented in Fig. 2, and Fig. 3. The second approach is Monte Carlo simulation to estimate the probabilities of detection and false alarm for various values of *NW*, *K* and SNR, and then to find the optimal (*NW*, *K*).

Fig. 2 shows the PR's power spectral density (PSD) computed using (2) with *NW*=4, and 16 where the number of tapers used is *K*=5, and 25 respectively at AWGN channel with SNR=−15dB and number of averaged samples is 2500. PR transmits OFDM-QPSK signal from normalized frequency $f_a = 16$ to $f_b = 48$ with power normalized to one over the whole band. Both PR and CR use 64-IFFT/FFT signal processing. CR receiver implements MTM to estimate the PR's PSD using different values for *NW* and *K* parameters. The ideal curve represents the levels of noise, and noise plus signal. We can clearly note how much power spectral leakage outside the PR's signal band when using *NW*=16 and *K*=25. Such leakage of power will affect the decision outside the PR's band by introducing more false alarms. At the same time we can see how such leakage is reduced when using *NW*=4 and *K*=5.

Fig.3 shows the PSD for the same system with *NW*=8 computed using (2). This figure clearly shows that using a small number of tapers *K*=2 introduces large variance in the estimate due to the averaging over small number of tapers. At the same time using a large number of tapers *K*=14 improves the variance but at the expense of spectral leakage which is noticeable in the figure. Using *K*=5 produces leakage that is between the previous two cases.

We may conclude from these two figures that an unwise choice of *NW* and *K* within the range, suggested by Haykin, may have catastrophic results on false alarm of the MTM estimator.

Fig. 4 shows a representative diagram of the MTM parameters' optimization problem in a 64-FFT based CR system that is used in the simulation. In our case *NW* axes values are *NW*=0.5, 1, 2, 4, 8, and 16. On the *K* axes values are *K*=1, 2, 3,…, 32. In regions (R_1), and (R_3), the half time bandwidth *NW* has higher resolution than the other two regions (R_2), and (R_4). At the same time R_1, and R_2 have a small number of tapers with good bias properties at the expense of higher variance when used in computing the PSD using (2). R_3, and R_4 regions have large number of tapers that improve the variance of the spectrum estimate, but at the expense of larger spectral leakage. The recommended values of *NW*, and *K* ranges in the literature are shown in the figure. Although these ranges are useful in the CR spectrum sensing, they still need to be optimized to get the highest performance for 64-FFT CR systems. In addition to maximizing the performance, optimum parameters will contribute to reducing the MTM estimator complexity.

The mathematical derivation of the optimal MTM parameters is intractable. Therefore, a Monte Carlo simulation program has been used in this paper to examine the effect of the different values of *NW* and *K* on the spectral leakage outside the PR's subband and the MTM estimator decision statistic. Consequently the binary hypotheses at each frequency bin will be used to evaluate MTM estimator performance.

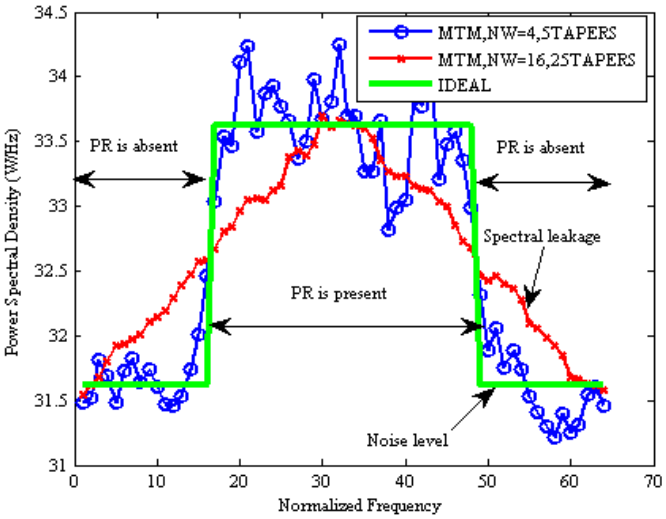


Fig. 2. Power spectral density (PSD) using MTM computed with *NW*=4, and 16 and different values of *K* at AWGN with SNR=−15dB.

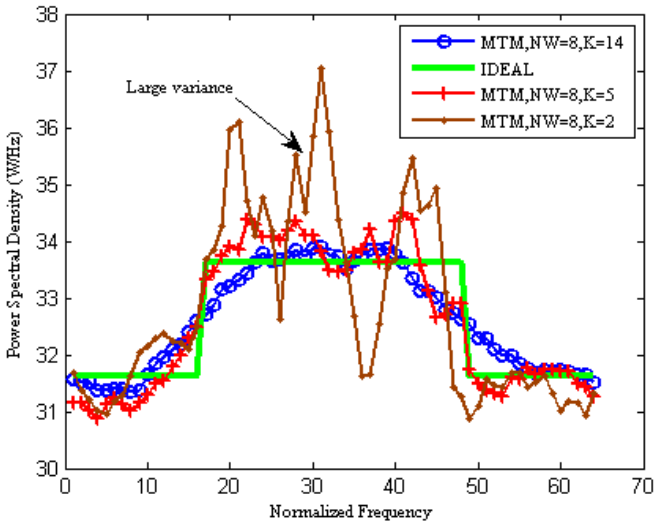


Fig. 3. Power spectral density (PSD) using MTM with *NW*=8, and different values of *K* at AWGN with SNR=−15dB.

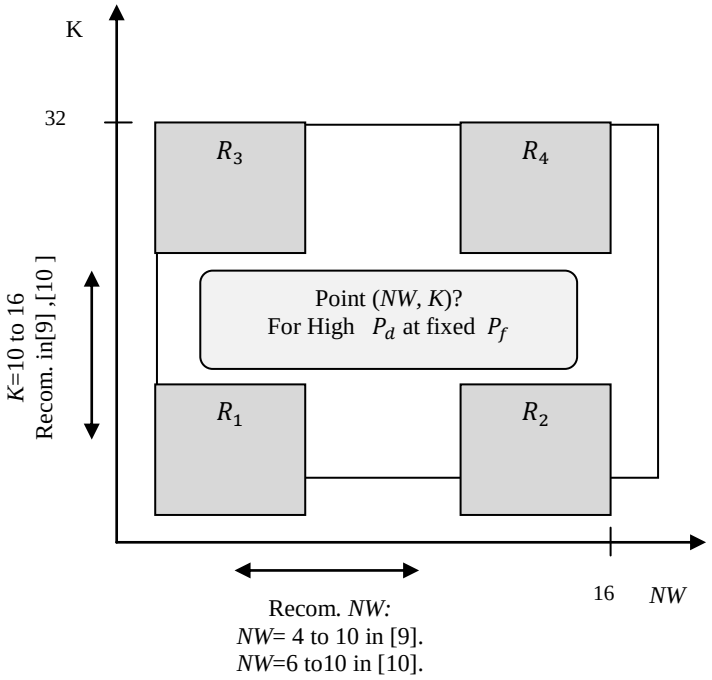


Fig. 4. Representative diagram of the MTM parameters optimization problem for 64-FFT based CR systems.

Clearly, optimizing the MTM estimator performance requires maximizing probability of PR signal detection P_d for a predefined probability of false alarm P_f. Here P_d is the probability the MTM estimator decides correctly the presence of the PR's signal, and P_f is the probability that the MTM estimator decides the PR's signal is present when it is absent.

The binary hypothesis test for MTM spectrum sensing at the l^{th} time is given by:

$$\begin{aligned} \mathcal{H}_0&: \quad x(l) = w(l) \\ \mathcal{H}_1&: \quad x(l) = hs(l) + w(l) \end{aligned} \tag{4}$$

where l = 0,1,…,*L*-1 is OFDM block's index, $x(l)$, $w(l)$ and $s(l)$ denote the CR received, noise, and PR transmitted samples. The transmitted PR signal is distorted by the zero mean additive white Gaussian noise $w(l) \sim \mathcal{CN}(0, \sigma^2_{noise})$. Additionally the channel between PR transmitter and CR receiver is subjected to flat fading. The channel gain h is assumed to be constant during the sensing time. The time instant l comes from the samples over different OFDM blocks; and time instant t comes from the samples from the same OFDM block (i.e., IFFT/FFT samples).

Decision *DEC* over time interval L, and at a specific frequency bin using the MTM can be formulated as:

$$DEC_{MTM}(f_i) = \frac{1}{L}\sum_{l=0}^{L-1} S^l_{MTM}(f_i) \tag{5}$$

Thus, we can reformulate the eigenspectrum in (1), using (4) at the l^{th} time when the binary hypothesis $\mathcal{H}_1$ is valid to be as follows:

$$Y_k(f_i) = \sum_{t=0}^{N-1} v_t^{(k)}(N,W)(hs_t(l) + w_t(l))e^{-j2\pi f_i t} \tag{6}$$

The decision at a specific frequency bin over the spectrum sensing time duration L can be rewritten using (6) to be as follows:

$$DEC_{MTM}(f_i) = \frac{1}{L}\sum_{l=0}^{L-1} \frac{\sum_{k=0}^{K-1}\lambda_k(N,W)\left|\sum_{t=0}^{N-1} v_t^{(k)}(N,W)(hs_t(l)+w_t(l))e^{-j2\pi f_i t}\right|^2}{\sum_{k=0}^{K-1}\lambda_k(N,W)} \tag{7}$$

When using the periodogram, the decision can be formulated as:

$$DEC_{PE}(f_i) = \frac{1}{L}\sum_{l=0}^{L-1} S^l_{PE}(f_i) \tag{8}$$

By rewriting (8) using (3) and (4), the decision at a specific frequency bin using the periodogram when the binary hypothesis $\mathcal{H}_1$ is valid, is as follow:

$$DEC_{PE}(f_i) = \frac{1}{LN}\sum_{l=0}^{L-1}\left|\sum_{t=0}^{N-1}(hs_t(l) + w_t(l))e^{-j2\pi f_i t}\right|^2 \tag{9}$$

The detection and false alarm probabilities at each frequency bin are defined as:

$$\begin{aligned} P_D(f_i) &= Pr\{DEC_{MTM}(f_i), DEC_{PE}(f_i) > \gamma | \mathcal{H}_1\} \\ P_F(f_i) &= Pr\{DEC_{MTM}(f_i), DEC_{PE}(f_i) > \gamma | \mathcal{H}_0\} \end{aligned} \tag{10}$$

The threshold γ, is defined according to the noise variance σ^2_{noise}. The decision statistics ($DEC_{MTM}(f_i)$, $DEC_{PE}(f_i)$) are calculated at each frequency bin using (4) to (9), and then the probabilities of detection and false alarm can be evaluated by comparing the decision statistic to the predefined threshold over a number of realizations using (10).

The binary hypothesis $\mathcal{H}_0$ will be examined through all frequency bins that don't contain PR' signal (i.e., $\{f_i \in ([0, f_a) \cup (f_b, 63])\}$). The binary hypothesis $\mathcal{H}_1$ will be examined through all frequency bins that contain the PR's (i.e., $\{f_i \in [f_1, f_2]\}$).

The probability of detection P_d over the band under sensing can be achieved from the averaged summation of the individual probability of detection $P_D(f_i)$ of the all the bins which lie within the subbands used by the PR user, and can be written as:

$$P_d = \frac{\sum_{f_i=f_a}^{f_i=f_b} P_D(f_i)}{32} \tag{11}$$

The probability of false alarm P_f over the band under sensing can be achieved from the averaged summation of the individual probability of false alarm $P_F(f_i)$ of the all the bins which lie within the subbands outside the PR's subband, and can be written as:

 ISBN: 978-1-61208-103-8

$$P_f = \frac{\sum_{f_i=0}^{f_i=f_a-1} P_F(f_i) + \sum_{f_i=f_b+1}^{63} P_F(f_i)}{32} \quad (12)$$

where 32 represents the total number of frequency bins of the hypotheses $\mathcal{H}_0$, and $\mathcal{H}_1$ of the model.

The optimization problem here can be written simply as:

$$find\ (NW^*, K^*)\ \ that\ maximizes\{\ P_d\}\ at\ P_f = \alpha \quad (13)$$

where α is a constant false alarm, and is assumed as 10% in this paper.

The complexity of MTM estimator for producing the spectrum estimate at a specific frequency bin f_i and *N*-FFT over *L* OFDM-Blocks, in terms of the number of mathematical operations (i.e., adding, and multiplication) is defined as:

$$com_{MTM} = L[K(3N-1) + 4K - 2] \quad (14)$$

Using the periodogram to produce spectrum estimate at a specific frequency bin f_i , the complexity can be defined as follows:

$$com_{PE} = L(2N) \quad (15)$$

IV. Simulation Results

The frequency band under study is divided into three non-overlapped subbands as shown in Fig. 1. The PR user is transmitting QPSK-OFDM signal using the subband between the frequencies $f_a = 16$ to $f_b = 48$, and with normalized averaged power of 1 over the whole band. The PR user's transmitter uses 64-IFFT with sampling frequency 20 MHz, where the symbol duration T_s =0.05µs. The CR's node uses 64-FFT with sampling frequency 20 MHz as well. The performance is evaluated using number of samples at the CR user' node as $N_{samples} = 20(N_{FFT}) = 1280$, which corresponds to sensing time of $64\mu s$, that is sensing process is carried out every $L = 20$ OFDM blocks. In all cases of simulations the results are averaged over 100000 simulation runs. The channels considered in the simulation are AWGN with zero mean and variance σ^2_{noise}, and Rayleigh flat fading.

The probabilities of detection for *NW*= 0.5, 1, and using different number of tapers are shown in table I. The wireless channel is assumed to be AWGN with SNR $= -5$ dB. The threshold γ that gives probability of false alarm 10% was estimated by Monte Carlo simulation using computer software platform. This threshold is then substitutes in (7) to (12) to find probability of detection and in (13) to find optimum *NW* and *K*. The highest probability of detection was found as P_d=98.8150%, which is achieved using *NW*=2 and *K*=3 tapers in the spectrum sensing.

Fig. 5 shows the probability of detection versus the number of tapers when the half time bandwidth product was as *NW*= 4, 8, and 16, at the same wireless environment applied before. We note that each curve has three different behaviors. It starts from a lower point that represents the minimum probability of detection which is achieved by the first taper. Then it increases sharply to a peak point, and starts finally to level off. The peak point for the different *NW* in this case is at *K*=5 tapers. Table II summarizes the probability of detection for *NW*=4, 8, and 16 for *K*=1, and 5 that obtained from Fig. 5.

The highest probability of detection is P_d =99.7138%, which is achieved using *N*W=4, and *K*=5. Generally, 5 tapers is a good compromise between the good bias properties, and improved variance. However, the maximum value of P_d may vary with the wireless environment conditions Additionally, *NW*=4 is the optimal resolution that gives the highest performance.

TABLE I. PROBABILITY OF DEDECTION FOR *NW*=0.5, 1,2 AND DIFFERENT *K* AT AWGN (SNR=−5dB) WHEN FALSE ALARM IS 10%.

NW	P_d (%)			
	K=1	*K*=2	*K*=3	*K*=4
0.5	83.233	-	-	-
1	81.2166	87.7376	-	-
2	75.9459	90.7959	98.8150	98.560

TABLE II. PROBABILITY OF DEDECTION FOR *NW*=4,8,16 AND DIFFERENT *K* AT AWGN (SNR=−5dB) WHEN FALSE ALARM IS 10%.

NW	P_d (%)	
	K=1	*K*=5
4	73.4531	99.7138
8	71.6678	99.2422
16	69.1575	98.6350

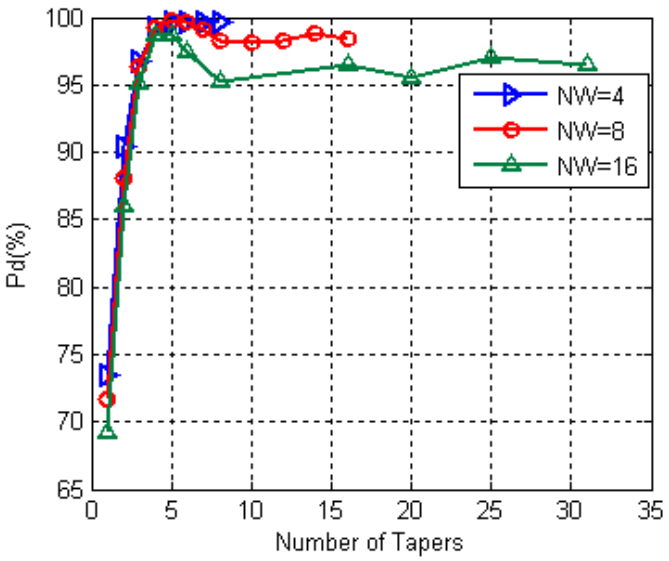


Fig. 5. Probability of detection versus number of tapers (*K*) using MTM with different half time bandwidth products (*NW*) where the probability of false alarm was 10% and at channel AWGN with SNR=−5dB.

 ISBN: 978-1-61208-103-8

Fig. 6 shows the probability of detection versus probability of false alarm for MTM with *NW*=4 using 5 tapers at AWGN with SNR= −5, and −10 dB. The results are compared with those obtained from the periodogram estimator to the same system. When the probability of false alarm is fixed at 10% and SNR = −5 dB, the probability of the detection using the periodogram is less than that using the MTM with *NW*=4 and 5 tapers by16 % .

When the SNR is decreased to −10 dB, the probability of detection of the MTM spectrum sensing with *NW*=4, and 5 tapers, is better than that for the periodogram by approximately 40% when the probability of false alarm is fixed at 10%. Consequently we can conclude that the MTM spectrum sensing performance is more robust than the periodogram at low SNR.

Fig. 7 shows the probability of detection versus probability of false alarm using periodogram and MTM with *NW*=4 and *K*=5 tapers schemes. The wireless channel is Rayleigh flat fading channel and SNR = −5 dB. We note that the probability of detection using the MTM, *NW*=4 and 5 tapers case is better than that using the periodogram by 8% when probability of false alarm $P_f = 10\%$.

Furthermore, comparing the results in Fig.7 with those in Fig.6, we conclude that the probability of detection is degraded in a flat fading channel compared to Gaussian channel for both schemes for the same probability of false alarm, *NW*, number of tapers, and SNR.

Table III shows the complexity of the MTM spectrum sensing based on (14) for different number of tapers *K* with length *N*=64 over one OFDM block(i.e., *L*=1).

It is clear that, in addition to the high performance achieved by *K*=5, it requires less mathematical operations for computation compared to $K > 5$ cases. The periodogram complexity is found as 128 operations at the same conditions of MTM using (15).

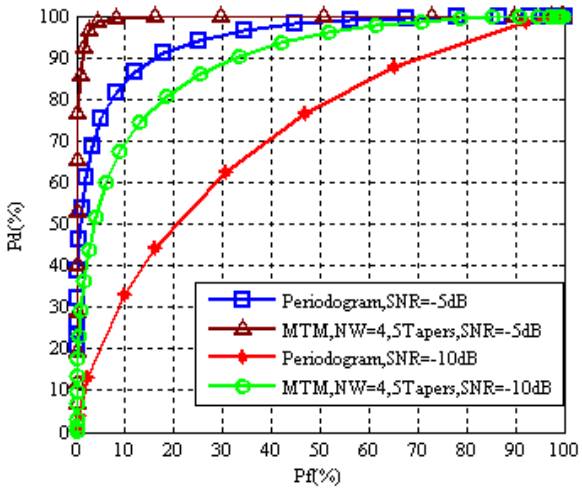


Fig. 6. Probability of detection versus probability of false alarm using MTM with *NW*=4, and *K*=5 compared to the periodogram at AWGN with SNR=−5 and− 10dB.

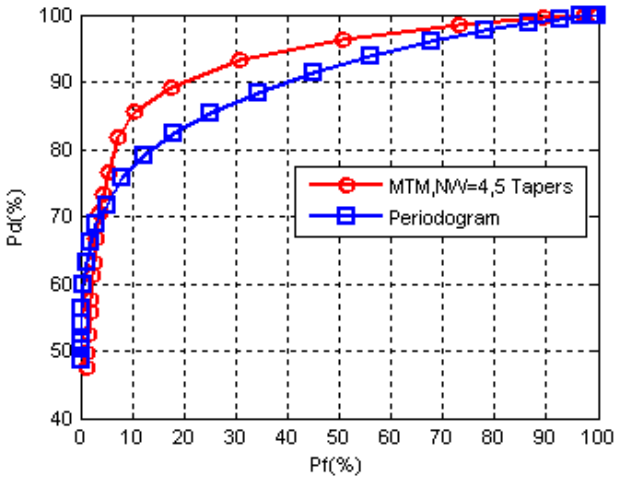


Fig. 7. Probability of detection versus probability of false alarm using MTM with *NW*=4 and *K*=5 tapers and periodogram using Rayleigh flat fading channel with SNR=− 5dB.

TABLE III. MTM COMPLEXITY EVALUATION FOR 64-FFT OVER L=1 USING DIFFERENT *K*.

MTM	*K*=1	*K*=5	*K*=10	*K*=20	*K*=31
Complexity	193	973	1948	3898	6043

V. CONCLUSION

In this paper, we have investigated the effects of the system parameters, within the range suggested in the literature, on the performance of the MTM spectrum sensor for opportunistic use by OFDM-based CR users. We have examined the MTM parameters to find their optimality to give higher probability of detection at lower probability of false alarm, and minimal complexity. The MTM technique has been analyzed and simulated in AWGN, and Rayleigh flat fading environments. Our primary and secondary users were communicating through OFDM-based systems with 64-IFFT/FFT.

Although the first few tapers (Slepian sequences) have the best spectral leakage properties, we found that they give the worst performance in terms of detection and false alarm probabilities. We found that unwise choice of *NW* and *K* from the range suggested in [9] produces catastrophic false alarms in the system. In our chosen 64-IFFT/FFT systems, the optimal number of tapers was 5 for the *NW*=4, 8, and 16 cases, and the optimal half time bandwidth product is given by *NW*=4 for 10% false alarm when system is operating in AWGN channel with SNR=− 5 dB. For cases where *NW*< 4, for example when *NW*=2, the bad bias properties of the tapers overcome the high resolution in this system. Generally, 5 tapers, and half time bandwidth *NW*=4 can be considered as optimal parameters for different FFT-sizes, since the change in FFT affects only resolution. We found that the performance of the system using MTM estimator is better than when using the periodogram estimator for any number of tapers except for one

 ISBN: 978-1-61208-103-8

taper when both systems are operating at the same channel conditions.

Both estimators suffer by the Rayleigh flat fading compared to AWGN environment. Furthermore, the MTM technique performance is more robust than the periodogram in the AWGN channel. Finally, the improvement in performance of the MTM estimator over the periodogram estimator comes at the cost of a slightly higher computational complexity. The additional complexity seems to be justifiable considering the advantages gained from using the MTM technique.

VI. REFERENCES

[1] J. Mitola and G. Q. Maguire, "Cognitive radio: making software radios more personal," *IEEE personal communications,* vol. 6, pp. 13-18, 1999.

[2] W.-Y. L. I. F. Akyildiz, M. C. Vuran, and S. Mohantly, "Next Generation/dynamic spectrum access /cognitive radiowireless network: A survey," *Elsevier Computer Networks,* vol. 50, pp. 2127-2159, Septemper 2006.

[3] M. Jun, G. Y. Li, and J. Biing Hwang, "Signal Processing in Cognitive Radio," *Proceedings of the IEEE,* vol. 97, pp. 805-823, 2009.

[4] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *Communications Surveys & Tutorials, IEEE,* vol. 11, pp. 116-130, 2009.

[5] D. B. Percival and A. T. Walden, *Spectral analysis for physical applications: multitaper and conventional univariate techniques*: Cambridge Univ Pr, 1993.

[6] D. J. Thomson, "Spectrum estimation and harmonic analysis," *Proceedings of the IEEE,* vol. 70, pp. 1055-1096, 1982.

[7] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE journal on selected areas in communications,* vol. 23, pp. 201-220, 2005.

[8] T. A. Weiss and F. K. Jondral, "Spectrum pooling: an innovative strategy for the enhancement of spectrum efficiency," *IEEE Communications Magazine,* vol. 42, pp. S8-14, 2004.

[9] S. Haykin, D. J. Thomson, and J. H. Reed, "Spectrum Sensing for Cognitive Radio," *Proceedings of the IEEE,* vol. 97, pp. 849-877, 2009.

[10] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty. V- The discrete case," Bell System Technical Journal, vol. 57, pp. 1371–1430, 1978.

[11] P. Stoica and T. Sundin, "On nonparametric spectral estimation," Circuits, Systems, and Signal Processing, vol. 18, pp. 169-181, 1999.

[12] D. J. Thomson, "Jackknifing multitaper spectrum estimates," IEEE Signal Processing Magazine, vol. 24, pp. 20-30, 2007.

[13] J. Park, C. R. Lindberg, and F. L. Vernon, "Multitaper spectral analysis of high-frequency seismograms," J. geophys. Res, vol. 92, pp. 675–12, 1987.

 ISBN: 978-1-61208-103-8

Probabilities of Detection and False Alarm in MTM- Based Spectrum Sensing for Cognitive Radio Systems

Owayed A. Alghamdi, Mosa A. Abu-Rgheff and Mohammed Z. Ahmed
Mobile Communications Research Group
University of Plymouth
Plymouth PL4 8AA, UK
{owayed.alghamdi, mosa, m.ahmed}@plymouth.ac.uk

***Abstract*— The paper presents closed-form expressions for the detection, and false alarm probabilities for spectrum sensing detection based on the Multitaper Spectrum Estimation Method (MTM) using Neyman-Pearson criterion. The MTM spectrum sensing is a powerful technique in Cognitive Radio (CR) systems. It tolerates problems related to bad biasing, and large variance of estimates, that are the main drawbacks in the periodogram (i.e., energy detector). The performance of the MTM spectrum sensing system is controlled by parameters, such as the chosen half time bandwidth product, Discrete Prolate Slepian Sequence (DPSS) (i.e., tapers), DPSS' eigenvalues, and the number of tapers used. These parameters determine the theoretical probabilities of detection and false alarm, which are used to evaluate the system performance. The paper shows a good match between the theoretical and numerical simulation results.**

Keywords*— *cognitive radio; spectrum sensing; multitier spectrum estimation.

I. INTRODUCTION

Cognitive radio is an innovative new technology in wireless communications, which was firstly proposed by Mitola in 1999 [1]. It allows secondary users (CRs), to opportunistically use the vacant spectrum subbands that are licensed already to primary users (PRs), at a specific time and geographical location. By full exploitation of the vacant spectrum subbands while keeping the PR users protected for harmful interference, CR technology provides efficient new spectral opportunities for next generations of wireless applications. It represents a new paradigm of spectrum allocation that helps reduce spectrum scarcity, and underutilization. Additionally, it can provide communications anywhere at any time [2].

A CR system should be capable of scanning through a given spectrum to find vacant bands to operate. The accurate CR system decision about the availability of vacant bands is totally dependent on the quality of the sensing techniques used. Clearly, CR technology can only be useful if an accurate sensing scheme is used.

Although the matched filtering and the cyclostationary feature detector have high performance as spectrum sensing techniques in CR, such techniques require prior information about the PR signaling [3-5].

On the other hand, the energy detector does not require prior information about the PR signaling and has low complexity. Such advantages come at the expense of moderate performance due to the use of single rectangular windows' tapering [6].

Multi taper spectrum estimation (MTM) [7], uses orthonormal tapers; known as the Discrete Prolate Slepian Sequence (DPSS) [8]. It produces a single spectrum estimate with minimum spectral leakage and good variance. MTM is an approximation of the optimal spectrum estimate; the Maximum-likelihood method but at reduced computation [9], [10]. Haykin, on the other hand, suggested the use of MTM as an efficient method for spectrum sensing in CR [2].

Using Neyman-Pearson criterion [11], theoretical derivations of probabilities of false alarm, and detection for the MTM spectrum sensing optimal detector are necessary to evaluate its performance. Furthermore, MTM detection system includes parameters, such as time bandwidth product, the DPSS, and their associated eigenvalues that control the quality of the spectrum estimate. Consequentially, a set of different parameters and thresholds can be chosen that maximizes the performance of the detection.

Although MTM was first studied by Thomson in 1982, statistics and probabilistic theoretical work are still an open research issue. In [12], the authors derived the probabilities of detection and false alarm formulae based on the spectrum estimate characteristic function (CHF) by formulating the MTM spectrum detector as a quadratic function of Gaussian vector.

In this paper, we present closed-form formulae for the probabilities of detection and false alarm for the MTM-based spectrum detector. The probability density function (PDF) of

 ISBN: 978-1-61208-103-8

the MTM spectrum estimate (decision statistic) is approximated to Gaussian. The mean and the variance of the PDF have been derived for both hypotheses, and used in the calculation of the probabilities.

Our theoretical work presented in this paper, includes two cases: firstly, the PR signal is known as a modulated signal, and secondly, the PR signal is unknown and assumed as Gaussian random variable.

The rest of the paper is organized as follows: Section II defines the model for the system under consideration and reviews MTM technique Section III presents the theoretical work of the MTM detector. Section IV presents the results and Section V concludes the paper.

II. SYSTEM MODEL

In our system model, we consider OFDM signaling scheme for the PR user. The PR transmitter with *N* subcarriers (*N*-IFFT/FFT) transmits OFDM-QPSK signal with energy E_s over each subcarrier. The CR transceiver is supported by (*N*-IFFT/FFT) processor as well so as to perform both tasks of communications, and sensing. Additionally, MTM spectrum detector is added to the CR receiver for spectrum sensing.

The received PR signal, at CR receiver, is sampled to generate a finite discrete time samples series $\{x_t; t = 0,1,\dots,N-1\}$, where t is time index. The discrete time samples are 'dot multiplied' with different tapers $v_{(t,k)}(N,W)$ (tapers are Discrete Prolate Slepian Sequences). The associated eigenvalues of the k^{th} taper is $\lambda_k(N,W)$. The product is applied to a Fourier Transform to compute the energy concentrated in the bandwidth $(-W,W)$ centered at frequency f. The half time bandwidth product is NW, and the total number of generated tapers is $2NW$. For K orthonormal tapers used in the MTM, there will be K different eigenspectrums produced and defined as [7]:

$$Y_k(f_i) = \sum_{t=0}^{N-1} v_{(t,k)}(N,W) x_t e^{-j2\pi f_i t} \qquad (1)$$

where, $f_i = 0, \frac{1}{N}, \frac{2}{N}, \dots, \frac{N-1}{N}$ are the normalized frequency bins. The spectrum estimate given by Thomson theoretical work is defined as [7]:

$$S_{MTM}(f_i) = \frac{\sum_{k=0}^{K-1} \lambda_k(N,W)|Y_k(f_i)|^2}{\sum_{k=0}^{K-1} \lambda_k(N,W)} \qquad (2)$$

On the other hand, the energy detector, when the samples are taken at uniform time spacing, gives the power spectrum density estimation as [6]:

$$S_{ED}(f_i) = \frac{1}{N}\left|\sum_{t=0}^{N-1} x_t e^{-j2\pi f_i t}\right|^2 \qquad (3)$$

In order to evaluate the performance of the MTM spectrum detector, we considered the probability of detection $P_d(f_i)$, the probability of false alarm $P_f(f_i)$, and the probability of miss detection $P_m(f_i)$ at each frequency bin f_i based on the Neyman-Pearson (NP) criterion. $P_d(f_i)$ is the probability that CR detector decides correctly the presence of the PR's signal, $P_f(f_i)$ is the probability that CR detector decides the PR's signal is present when it is absent, and $P_m(f_i)$ is the probability that CR fails to detect the PR's signal when it is present.

The binary hypothesis test for CR spectrum sensing at the lth time is given by:

$$\begin{aligned} \mathcal{H}_0: &\quad x_t(l) = w_t(l) \\ \mathcal{H}_1: &\quad x_t(l) = s_t(l) + w_t(l) \end{aligned} \qquad (4)$$

where l = 0,1,…,*L*-1 is OFDM block's index, $x_t(l)$, $w_t(l)$, and $s_t(l)$ denote the CR received, noise, and PR transmitted samples. The transmitted PR signal is distorted by the zero mean additive white Gaussian noise $w_t(l) \sim \mathcal{CN}(0, \sigma_w^2)$. The signal to noise ratio (SNR) is $SNR = \frac{E_s}{\sigma_w^2}$.

The time instant *l* comes from the samples over different OFDM blocks; and time instant *t* comes from the samples from the same OFDM block (i.e., IFFT/FFT samples). Thus, the spectrum sensing time in second is $(L)(N)(T_s)$, where T_s represents symbol duration, *L* represents the number of OFDM blocks that used in sensing, and *N* is the number of samples per OFDM block (i.e., FFT size). The decision statistic over *L* OFDM blocks using MTM is defined as follows:

$$DEC_{MTM}(f_i) = \sum_{l=0}^{L-1} \frac{\sum_{k=0}^{K-1} \lambda_k(N,W)\left|\sum_{t=0}^{N-1} v_{(t,k)}(N,W) x_t(l) e^{-j2\pi f_i t}\right|^2}{\sum_{k=0}^{K-1} \lambda_k(N,W)} \qquad (5)$$

III. DECISION STATISTIC PROBABILITY DENSITY FUNCTION

For large *L* at low SNR, we approximate the PDF of the eigenspectrum absolute square $|Y_k(f_i)|^2$ from chi-square to Gaussian, then the decision statistic ($DEC_{MTM}(f_i)$) is represented by the sum of *K* correlated Gaussian samples (i.e., eigenspectrum absolute square $|Y_k(f_i)|^2$).

Thus, the decision statistic using MTM detector is approximately normal Gaussian distributed as expected due to MTM linear processing.

We now consider the mean (E), and the variance (VAR) of the decision statistic $DEC_{MTM}(f_i)$ for both hypotheses. (i.e., $E(DEC_{MTM}(f_i)/\mathcal{H}_0$, $E(DEC_{MTM}(f_i)/\mathcal{H}_1)$, $VAR(DEC_{MTM}(f_i)/\mathcal{H}_0)$, and $VAR(DEC_{MTM}(f_i)/\mathcal{H}_1)$).

The probability of detection, and the probability of false alarm at frequency bin f_i $P_d^{MTM}(f_i)$, and $P_f^{MTM}(f_i)$, respectively, for the decision statistic with Gaussian distribution are defined as:

$$\begin{aligned} P_d^{MTM}(f_i) &= P(DEC_{MTM}(f_i) > \gamma / \mathcal{H}_1) \\ &= Q\left(\frac{\gamma - E(DEC_{MTM}(f_i)/\mathcal{H}_1)}{\sqrt{VAR(DEC_{MTM}(f_i)/\mathcal{H}_1)}}\right) \end{aligned} \qquad (6)$$

 ISBN: 978-1-61208-103-8

$$P_f^{MTM}(f_i) = P(DEC_{MTM}(f_i) > \gamma / \mathcal{H}_0)$$

$$= \mathcal{Q}(\frac{\gamma - E(DEC_{MTM}(f_i)/\mathcal{H}_0)}{\sqrt{VAR(DEC_{MTM}(f_i)/\mathcal{H}_0)}}) \quad (7)$$

The probability of miss detection can be defined as:

$$P_m^{MTM}(f_i) = P(DEC_{MTM}(f_i) < \gamma / \mathcal{H}_1)$$

$$= 1 - \mathcal{Q}(\frac{\gamma - E(DEC_{MTM}(f_i)/\mathcal{H}_1)}{\sqrt{VAR(DEC_{MTM}(f_i)/\mathcal{H}_1)}}) \quad (8)$$

the term $\mathcal{Q}(\xi)$ is given by the tails of the distribution, and γ represents the threshold. Note that γ can be controlled based on σ_w^2.

When only noise is present for $\mathcal{H}_0$ case at frequency bin f_i, and based on the linearity property of the FFT process, the mean of the decision statistic $E(DEC_{MTM}(f_i)/\mathcal{H}_0)$ can be defined for K Gaussian samples as:

$$E(DEC_{MTM}(f_i)/\mathcal{H}_0) =$$

$$C.\sum_{l=0}^{L-1}(\sum_{k=0}^{K-1}(\lambda_k(N,W).K\sum_{t=0}^{N-1}\sum_{t'=0}^{N-1}E(v_{(t,k)}(N,W) .v_{(t',k)}(N,W).w_t(l)w_{t'}(l)))) \quad (9)$$

where $C = E\left(\frac{1}{\sum_{k=0}^{K-1}\lambda_k(N,W)}\right) = \frac{1}{\sum_{k=0}^{K-1}\lambda_k(N,W)}$ (10)

It can be shown that (9) can be simplified as:

$$E(DEC_{MTM}(f_i)/\mathcal{H}_0) = \sum_{l=0}^{L-1}\sum_{t=0}^{N-1}v_{(t,k)}^2(N,W).K.E\big(w_t(l)w_{t'}(l)\big) \quad (11)$$

From the definition of the Discrete Prolate Slepian Sequence (DPSS), we have [8]:

$$\sum_{t=0}^{N-1}v_{(t,k)}(N,W).v_{(t,k')}(N,W) = \begin{cases}1, & k = k'\\ 0, & k \neq k'\end{cases} \quad (12)$$

The orthonormality of the sequences can be used to simplify (11), when $t = t'$as follows:

$$E(DEC_{MTM}(f_i)/\mathcal{H}_0) = \sum_{l=0}^{L-1}.K.E\big(w_t^2(l)\big) =$$

$$LK(\big(E\big(w_t(l)\big)\big)^2 + VAR\big(w_t(l)\big)) = LK(0 + \sigma_w^2) = LK\sigma_w^2 \quad (13)$$

When the PR signal is present for $\mathcal{H}_1$ case at frequency bin f_i, the mean of the decision statistic $E(DEC_{MTM}(f_i)/\mathcal{H}_1)$ can be defined following the same steps of $\mathcal{H}_0$ as:

$$E(DEC_{MTM}(f_i)/\mathcal{H}_1) = K\sum_{l=0}^{L-1}E\big(s_t^2(l) + 2s_t(l)w_t(l) + w_t^2(l)\big)$$

$$= LK(E_s + \sigma_w^2) \quad (14)$$

where $E(s_t^2(l)) = E(E_s) = E_s$, and $E(w_t^2(l)) = VAR(w_t(l)) + (E\big(w_t(l)\big))^2 = \sigma_w^2$, and $E\big(2s_t(l)w_t(l)\big) = 0$.

We now consider the variances of the hypotheses in the next stage of the derivation. In order to simplify our derivation, we redefine (5) using decision statistic coefficients α_k , $k = 0,1,2,\dots,K-1$, as follows:

$$DEC_{MTM}(f_i) = \sum_{l=0}^{L-1}\sum_{k=0}^{K-1}\alpha_k(f_i) \quad (15)$$

where coefficient α_k is defined as follows:

$$\alpha_k(f_i) = \frac{\lambda_k(N,W)|Y_k(f_i)|^2}{\sum_{k=0}^{K-1}\lambda_k(N,W)},\ k = 0,1,2,\dots,K-1 \quad (16)$$

Then, the variance of $\alpha_k(f_i)$ can be defined as follows:

$$VAR(\alpha_k(f_i)/\mathcal{H}_p) = \frac{\lambda_k^2(N,W).VAR(|Y_k(f_i)|^2/\mathcal{H}_p)}{(\sum_{k=0}^{K-1}\lambda_k(N,W))^2},\ p = 0,1 \quad (17)$$

The variance of the $\mathcal{H}_0$ hypothesis where the noise only is present for K correlated Gaussian samples (i.e., eigenspectrum absolute square $|Y_k(f_i)|^2$) $VAR(DEC_{MTM}(f_i)/\mathcal{H}_0)$, can be defined as follows:

$$VAR(DEC_{MTM}(f_i)/\mathcal{H}_0) = \sum_{l=0}^{L-1}(\sum_{k=0}^{K-1}VAR(\alpha_k(f_i)/\mathcal{H}_0) + 2\rho_{01}\sqrt{VAR(\alpha_0(f_i)/\mathcal{H}_0)}\sqrt{VAR(\alpha_1(f_i)/\mathcal{H}_0)} + 2\rho_{02}\sqrt{VAR(\alpha_0(f_i)/\mathcal{H}_0)}\sqrt{VAR(\alpha_2(f_i)/\mathcal{H}_0)} + \cdots + 2\rho_{0K-1}\sqrt{VAR(\alpha_0(f_i)/\mathcal{H}_0)}\sqrt{VAR(\alpha_{K-1}(f_i)/\mathcal{H}_0)} + 2\rho_{12}\sqrt{VAR(\alpha_1(f_i)/\mathcal{H}_0)}\sqrt{VAR(\alpha_2(f_i)/\mathcal{H}_0)} + 2\rho_{13}\sqrt{VAR(\alpha_1(f_i)/\mathcal{H}_0)}\sqrt{VAR(\alpha_3(f_i)/\mathcal{H}_0)} + \cdots + 2\rho_{1K-1}\sqrt{VAR(\alpha_1(f_i)/\mathcal{H}_0)}\sqrt{VAR(\alpha_{K-1}(f_i)/\mathcal{H}_0)} + \cdots + 2\rho_{K-2K-1}\sqrt{VAR(\alpha_{K-2}(f_i)/\mathcal{H}_0)}\sqrt{VAR(\alpha_{K-1}(f_i)/\mathcal{H}_0)}) \quad (18)$$

where $\rho_{ij} \approx 1$, is the correlation coefficient between α_i, and α_j, and since $VAR(|Y_k(f_i)|^2/\mathcal{H}_p) = VAR(|Y_k(f_i)|^2/\mathcal{H}_p)$, for $i,j = 0,1,2,\dots,K-1$ applying the orthonormality in (12). Then (18) can be rewritten using (10) and (17) as follows:

$$\begin{aligned}VAR(DEC_{MTM}(f_i)/\mathcal{H}_0) &= C^2.VAR(|Y_k(f_i)|^2/\mathcal{H}_0).\sum_{l=0}^{L-1}\Bigg(\sum_{k=0}^{K-1}\lambda_k^2(N,W) + 2\lambda_0(N,W)\lambda_1(N,W) + 2\lambda_0(N,W)\lambda_2(N,W) + \cdots + 2\lambda_0(N,W)\lambda_{K-1}(N,W) + 2\lambda_1(N,W)\lambda_2(N,W) + 2\lambda_1(N,W)\lambda_3(N,W) + \cdots + 2\lambda_1(N,W)\lambda_{K-1}(N,W) + \cdots + 2\lambda_{K-2}(N,W)\lambda_{K-1}(N,W)\Bigg)\end{aligned} \quad (19)$$

when $t = t'$, and since the variance of Gaussian random variable W, $VAR(W^2) = 2(VAR(W))^2$, then $VAR(|Y_k(f_i)|^2/\mathcal{H}_0)$, can be defined as follows:

$$\begin{aligned}VAR(|Y_k(f_i)|^2/\mathcal{H}_0) &= \left(\sum_{t=0}^{N-1}v_{(t,k)}^2(N,W)\right)^2.VAR(\big(w_t^2(l)\big)\\ &= \big(2\sigma_w^4\big)\end{aligned} \quad (20)$$

Finally, $VAR(DEC_{MTM}(f_i)/\mathcal{H}_0)$ over L can be rewritten using (19) and (20) as follows:

$$VAR(DEC_{MTM}(f_i)/\mathcal{H}_0) = 2C^2 L\lambda_\Sigma \sigma_w^4 \quad (21)$$

where λ_Σ, is defined as follows:

$$\lambda_\Sigma = \sum_{k=0}^{K-1} \lambda_k^2(N,W) + 2\lambda_0(N,W)\lambda_1(N,W) + 2\lambda_0(N,W)\lambda_2(N,W) + \cdots + 2\lambda_0(N,W)\lambda_{K-1}(N,W) + 2\lambda_1(N,W)\lambda_2(N,W) + 2\lambda_1(N,W)\lambda_3(N,W) + \cdots + 2\lambda_1(N,W)\lambda_{K-1}(N,W) + \cdots + 2\lambda_{K-2}(N,W)\lambda_{K-1}(N,W) \quad (22)$$

When the PR signal is present-for $\mathcal{H}_1$ case at frequency bin f_i, the variance of the decision statistic $VAR(DEC_{MTM}(f_i)/\mathcal{H}_1)$ is:

$$VAR(DEC_{MTM}(f_i)/\mathcal{H}_1) = C^2.L.\lambda_\Sigma.VAR(|Y_k(f_i)|^2/\mathcal{H}_1) \quad (23)$$

when $t = t'$, and since $VAR(s_t^2(l)) = 0$ in this case, and $VAR(2s_t(l)w_t(l)) = 4E_s\sigma_w^2$, then

$$VAR(|Y_k(f_i)|^2/\mathcal{H}_1) = VAR\,(s_t^2(l) + 2s_t(l)w_t(l) + w_t^2(l)) = (2\sigma_w^4 + 4E_s\sigma_w^2) \quad (24)$$

Finally, (23) can be written as follows:

$$VAR(DEC_{MTM}(f_i)/\mathcal{H}_1) = C^2\, L\lambda_\Sigma \left(2\sigma_w^4 + 4E_s\sigma_w^2\right) \quad (25)$$

The probabilities formulae in (6), (7), and (8) can now be rewritten as follow:

$$P_d^{MTM}(f_i) = Q\left(\frac{\gamma - LK(E_s+\sigma_w^2)}{\sqrt{2LC^2\lambda_\Sigma\sigma_w^2(\sigma_w^2+2E_s)}}\right) \quad (26)$$

$$P_f^{MTM}(f_i) = Q\left(\frac{\gamma - LK\sigma_w^2}{\sqrt{2LC^2\lambda_\Sigma\sigma_w^4}}\right) \quad (27)$$

$$P_m^{MTM}(f_i) = 1 - Q\left(\frac{\gamma - LK(E_s+\sigma_w^2)}{\sqrt{2LC^2\lambda_\Sigma\sigma_w^2(\sigma_w^2+2E_s)}}\right) \quad (28)$$

It is clear that, the main processing difference between the energy detector and the MTM detector is simply multiplying the signal by a number of orthonormal tapers; the *DPSS* to produce a single estimate, while the multiplication in the energy detector is by a single rectangular taper. Thus in order to see the effect of this difference, we use the probabilities formulae of the energy detector which can be defined for the same system conditions as follow [13], [14]:

$$P_d^{ED}(f_i) = Q\left(\frac{\gamma - L(E_s+\sigma_w^2)}{\sqrt{2L\sigma_w^2(\sigma_w^2+2E_s)}}\right) \quad (29)$$

$$P_f^{ED}(f_i) = Q\left(\frac{\gamma - L\sigma_w^2}{\sqrt{(2L\sigma_w^4)}}\right) \quad (30)$$

$$P_m^{ED}(f_i) = 1 - Q\left(\frac{\gamma - L(E_s+\sigma_w^2)}{\sqrt{2L\sigma_w^2(\sigma_w^2+2E_s)}}\right) \quad (31)$$

The number of OFDM blocks *L*, which is needed to achieve predefined probabilities of detection $P_d^{MTM}(f_i)$, and false alarm $P_f^{MTM}(f_i)$ in the MTM technique can be written using (26) and (27) to be as follows:

$$L = \left(\frac{\sqrt{2C^2\lambda_\Sigma\sigma_w^4}Q^{-1}\left(P_f^{MTM}(f_i)\right) - \sqrt{2C^2\lambda_\Sigma\sigma_w^2(\sigma_w^2+2E_s)}Q^{-1}\left(P_d^{MTM}(f_i)\right)}{KE_s}\right)^2 \quad (32)$$

which can be written in (dB) to be as follows:

$$L(dB) = 10log_{10}(L)$$

The probabilities formulae when the PR's signal is modeled as Gaussian random process are listed in the appendix.

In multipath fading environment, the binary hypothesis test in (4) can be redefined for $\mathcal{H}_1$ to be as follows:

$$x_t(l) = \sum_{m=0}^{M-1} h_m s_{t-m}(l) + w_t(l) \quad (33)$$

where the discrete channel impulse response between the PR's transmitter and CR's receiver is represented by h_m, $m = 0,1,\ldots,M-1$, and *M* is the total number of resolvable paths. The discrete frequency response of the channel is obtained by taking the *N* point FFT, with $N \geq M$ as follows [14]:

$$H(f_i) = \sum_{m=0}^{M-1} h_m e^{-j2\pi f_i m} \quad (34)$$

In this case, the formulae in (26), and (28) can be written as follow:

$$P_d^{MTM}(f_i) = Q\left(\frac{\gamma - LK(|H(f_i)|^2E_s+\sigma_w^2)}{\sqrt{2LC^2\lambda_\Sigma\sigma_w^2(\sigma_w^2+2|H(f_i)|^2E_s)}}\right) \quad (35)$$

$$P_m^{MTM}(f_i) = 1 - Q\left(\frac{\gamma - LK(|H(f_i)|^2E_s+\sigma_w^2)}{\sqrt{2LC^2\lambda_\Sigma\sigma_w^2(\sigma_w^2+2|H(f_i)|^2E_s)}}\right) \quad (36)$$

The SNR can be redefined here to be as follows:

$$SNR = \frac{|H(f_i)|^2E_s}{\sigma_w^2} \quad (37)$$

The same steps can be followed to rewrite the formulae (29), (30), and (31) for the energy detector case.

In this paper, we assume that the channel gain between the PR's transmitter and the CR's receiver is constant during the spectrum sensing duration, and $|H(f_i)|^2 = 1$. In practice, $|H(f_i)|^2$ can be estimated priori during the time that PR's transmitter occupies a specific band with specific power [14].

IV. SIMULATION RESULTS

We evaluate our theoretical work by running a simulation program where the PR's signal is QPSK with normalized energy equal to 1 over each subcarrier. Both-CR and PR users employ 64-IFFT/FFT digital signal processing in their communications with sampling frequency 20 MHz/ T_s=0.05μs, where T_s represents the symbol duration, the MTM parameters used are *NW*=4, and 5 tapers, and the results obtained over 1000000 realizations. Additionally, we compare the performance of MTM spectrum detector system to that of the energy detector under the same conditions. We

 ISBN: 978-1-61208-103-8

used theoretical and simulation results for a chosen frequency bin at the CR FFT to examine the hypotheses $\mathcal{H}_1$, and $\mathcal{H}_0$.

Fig. 1 shows the probability of detection P_d versus probability of false alarm P_f using MTM detector with *NW*=4 and 5 tapers (simulation and theory) and the energy detector at AWGN with SNR=−10dB and $L = 20$ OFDM blocks. Note that, the total number of samples used is $(L = 20) \times (N = 64) = 1280$, which approximately corresponds to sensing time $(L = 20) \times (N = 64) \times (T_s = 0.05\mu s) = 64\mu s$. By comparing the theoretical to the simulation in the MTM case, we note that the theoretical results match well the simulation one. At the same system conditions, the probability of detection P_d of MTM outperforms that for energy detector by 30%, when the probability of false alarm is P_f=10%, and the miss detection P_m in MTM is lower than that in energy detector case by 40%.

Fig. 2 shows the theoretical results of the number of OFDM blocks (L) required to achieve $P_d = 99\%$, and $P_f = 1\%$ at AWGN environment with different SNR using MTM with *NW*=4 and 5 tapers compared to the energy detector. It is clear that the number of OFDM blocks used in the sensing process in the MTM system is lower than that for the energy detector. For example, at SNR=−15dB, the L required by the MTM is 33dB, and the energy detector is 47dB. These two values correspond to 1995 and 5012 OFDM blocks for MTM and the energy detector, respectively, in the linear scale. Thus, the energy detector requires 2.5 times as many samples compared to MTM in order to achieve the same probabilities at the same SNR. Such a large number of the samples for sensing in CR system might hinder the opportunistic use of the vacant channels, as it is the main objective of the developing of CR systems.

Fig. 3 shows the probabilities of detection P_d that gives probabilities of false alarm $P_f = 5\%$, and 10% versus the SNR at AWGN using MTM with *NW*=4 and 5 tapers and $L = 50$. For both of predefined probabilities of false alarm the probabilities of detection are almost 100% for SNR=−7dB or higher with unnoticeable change for $P_f = 10\%$ curve, which is reasonable. Both probabilities of detection curves start to decrease with the decrease in the SNR with noticeable outperforming of the $P_f = 10\%$ curve. At SNR=−25dB, $P_d = 11\%$ for $P_f = 10\%$ curve, and $P_d = 6\%$ for $P_f = 5\%$ curve.

Fig. 4 shows the threshold versus probabilities of false alarm and detection using MTM with *NW*=4 and 5 tapers at AWGN with SNR=−7dB (i.e., $\sigma_w^2 = 5.0119$) and $L = 50$. Such a figure presents the range of the threshold that should be chosen in order to meet specific probability of false alarm and detection at defined SNR level and L used in the spectrum sensing. As an example, for threshold=5, the probability of false alarm and detection pair (P_f, P_d) is $(50\%, 100\%)$. By increasing the threshold level to 5.3, the pair becomes $(10\%, 100\%)$. This figure can be revaluated at different SNR and L conditions using (26) and (27).

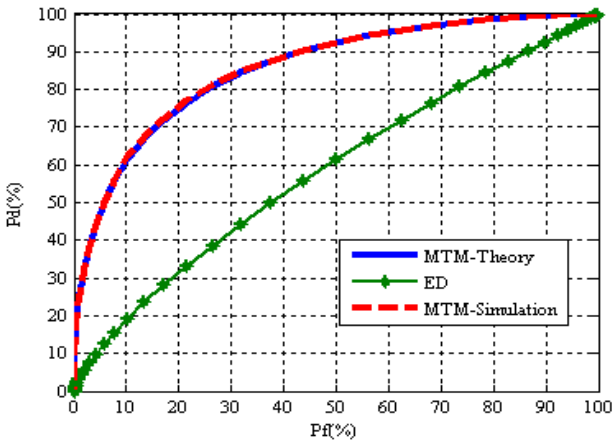


Fig. 1. Probability of detection versus probability of false alarm using MTM with NW=4 and 5 tapers (simulation and theory) and the periodogram (energy detector) at AWGN with SNR=−10dB and $L = 20$.

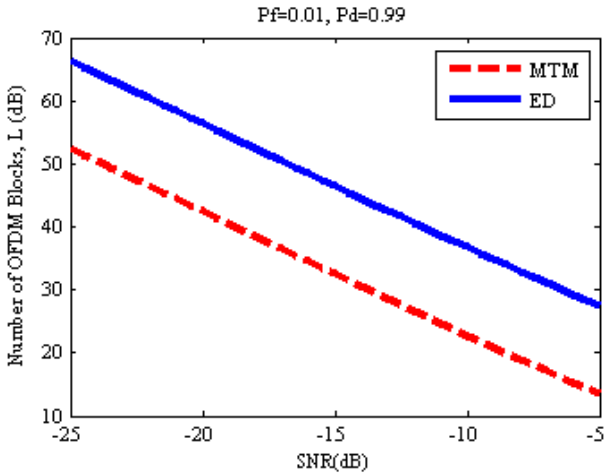


Fig. 2. Comparison between the number of OFDM blocks (L) required to achieve $P_d = 99\%$, and $P_f = 1\%$ at AWGN with different SNR using MTM with NW=4 and 5 tapers and the energy detector.

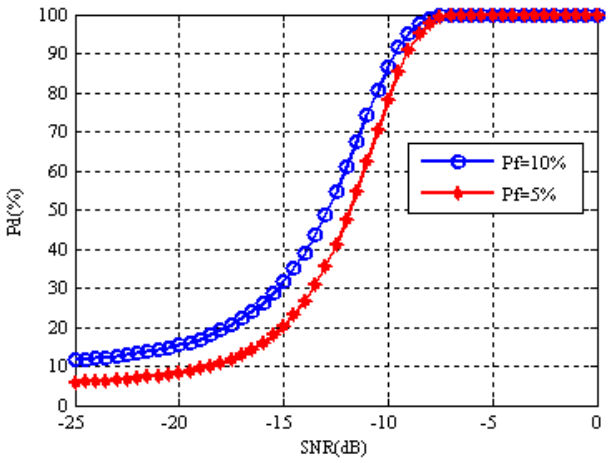


Fig. 3. Probability of detection that meets $P_f = 5$ and 10% versus the SNR at AWGN using MTM with *NW*=4 and 5 taper and $L = 50$ samples for spectrum sensing.

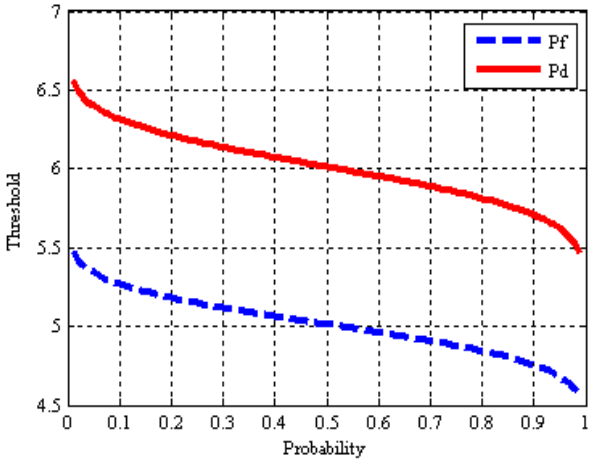


Fig. 4. Threshold versus probabilities of false alarm and detection using MTM with NW=4 and 5 tapers at AWGN with SNR=−7dB and $L = 50$ samples are used in the spectrum sensing.

V. CONCLUSION

In this paper, we have derived closed-form formulae for the probabilities of false alarm, detection, and miss detection as functions of the parameters of the MTM spectrum detector such as threshold, number of sensed blocks L, number of tapers, eigenvalues of the DPSS, PR signal power, and the noise power. These probabilities control the performance of the MTM-based spectrum sensing detector. Additionally, MTM probabilities can be used to choose the appropriate threshold that maximizes the probability of detection at fixed probability of false alarm.

In the process of the derivation, we defined the PDF of the MTM decision theory. Statistical parameters, such as the mean, and the variance of the distribution have been derived for different PR signals.

Comparing the performance of the MTM spectrum sensing detector to that for the energy detector, we found the MTM detector outperforms the performance of the energy detector by about 40% increase in the probability of detection at fixed probability of false alarm 10%. Furthermore the energy detector requires 2.5 times the number of samples to achieve the same probabilities of detection and false alarm given by the MTM detector operating at the same conditions.

APPENDIX

For the case when the PR's signal $s_t(l)$, is modeled as a random Gaussian variable with zero mean and variance σ_s^2 (i.e., $s_t(l) \sim \mathcal{CN}(0, \sigma_s^2)$) [15]. Following the same derivation steps of the modulated signal case, it can be proved that the probabilities formulae are defined as follow:

$$P_d^{MTM}(f_i) = Q\left(\frac{\gamma - LK(\sigma_w^2+\sigma_s^2)}{\sqrt{2LC^2\lambda_\Sigma(\sigma_w^2+\sigma_s^2)^2}}\right) \quad (38)$$

$$P_f^{MTM}(f_i) = Q\left(\frac{\gamma - LK\sigma_w^2}{\sqrt{2LC^2\lambda_\Sigma\sigma_w^4}}\right) \quad (39)$$

$$P_m^{MTM}(f_i) = 1 - Q\left(\frac{\gamma - LK(\sigma_w^2+\sigma_s^2)}{\sqrt{2LC^2\lambda_\Sigma(\sigma_w^2+\sigma_s^2)^2}}\right) \quad (40)$$

and the number of samples L (i.e., OFDM blocks) is defined as follows:

$$L = \left(\frac{\sqrt{2C^2\lambda_\Sigma\sigma_w^4}Q^{-1}\left(P_f^{MTM}(f_i)\right) - \sqrt{2C^2\lambda_\Sigma(\sigma_w^2+\sigma_s^2)^2}Q^{-1}\left(P_d^{MTM}(f_i)\right)}{K\sigma_s^2}\right)^2 \quad (41)$$

REFERENCES

[1] J. Mitola and G. Q. Maguire, "Cognitive radio: making software radios more personal," *IEEE personal communications*, vol. 6, pp. 13-18, 1999.

[2] S. Haykin, "Cognitive radio: brain-empowered wireless communications," *IEEE journal on selected areas in communications*, vol. 23, pp. 201-220, 2005.

[3] W.-Y. L. I. F. Akyildiz, M. C. Vuran, and S. Mohantly, "Next Generation/dynamic spectrum access /cognitive radiowireless network: A survey," *Elsevier Computer Networks*, vol. 50, pp. 2127-2159, Septemper 2006.

[4] M. Jun, G. Y. Li, and J. Biing Hwang, "Signal Processing in Cognitive Radio," *Proceedings of the IEEE*, vol. 97, pp. 805-823, 2009.

[5] T. Yucek and H. Arslan, "A survey of spectrum sensing algorithms for cognitive radio applications," *Communications Surveys & Tutorials, IEEE*, vol. 11, pp. 116-130, 2009.

[6] D. B. Percival and A. T. Walden, *Spectral analysis for physical applications: multitaper and conventional univariate techniques*: Cambridge Univ Pr, 1993.

[7] D. J. Thomson, "Spectrum estimation and harmonic analysis," *Proceedings of the IEEE*, vol. 70, pp. 1055-1096, 1982.

[8] D. Slepian, "Prolate spheroidal wave functions, Fourier analysis, and uncertainty. V- The discrete case," *Bell System Technical Journal*, vol. 57, pp. 1371–1430, 1978.

[9] P. Stoica and T. Sundin, "On nonparametric spectral estimation," *Circuits, Systems, and Signal Processing*, vol. 18, pp. 169-181, 1999.

[10] D. J. Thomson, "Jackknifing multitaper spectrum estimates," *IEEE Signal Processing Magazine*, vol. 24, pp. 20-30, 2007.

[11] S. M. Kay, *Fundamentals of statistical signal processing: detection theory*: Prentice-Hall, 1998.

[12] W. Jun and Q. T. Zhang, "A Multitaper Spectrum Based Detector for Cognitive Radio," in *Wireless Communications and Networking Conference, 2009. WCNC 2009. IEEE*, 2009, pp. 1-5.

[13] Q. Zhi, C. Shuguang, and A. H. Sayed, "Optimal Linear Cooperation for Spectrum Sensing in Cognitive Radio Networks," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 2, pp. 28-40, 2008.

[14] Q. Zhi, C. Shuguang, A. H. Sayed, and H. V. Poor, "Optimal Multiband Joint Detection for Spectrum Sensing in Cognitive Radio Networks," *Signal Processing, IEEE Transactions on*, vol. 57, pp. 1128-1140, 2009.

[15] J. Hillenbrand, T. A. Weiss, and F. K. Jondral, "Calculation of detection and false alarm probabilities in spectrum pooling systems," *IEEE Communications letters*, vol. 9, pp. 349-351, 2005.

 ISBN: 978-1-61208-103-8

ASE-BAN, a Wireless Body Area Network Testbed

Jens Kargaard Madsen
Aarhus School of Engineering
Aarhus University
Aarhus, Denmark
jkm@iha.dk

Henrik Karstoft
Aarhus School of Engineering
Aarhus University
Aarhus, Denmark
hka@iha.dk

Finn Overgaard Hansen
Aarhus School of Engineering
Aarhus University
Aarhus, Denmark
foh@iha.dk

Thomas Skjødeberg Toftegaard
Aarhus School of Engineering
Aarhus University
Aarhus, Denmark
tst@cs.au.dk

***Abstract*— Miniature Body Area Networks used in health care support greater mobility to patients and reduces actual hospitalization. This paper presents the preliminary implementation of a wireless body area network gateway. It is designed to implement the gateway functionality between sensors/actuators attached to the body and a host server application. The gateway uses the BlackFin BF533 processor from Analog Devices, and uses Bluetooth for wireless communication. Two types of sensors are attached to the network: an electro-cardio-gram sensor and an oximeter sensor. The testbed has been successfully tested for electro-cardio-gram data collection, and using wireless communication in a battery powered configuration.**

Keywords-component; low power wireless sensor network; healthcare; ECG sensor; body area network; testbed; ASE-BAN

I. INTRODUCTION

The demand for health-related services in Europe is expected to grow in the near future, partly because of the relative increase in number of elders in the European region. Some demands will be on highly patient-centric and prevention-based health-related services. Technologies to cope with these demands are cheap real-time systems to monitor body functions of patients [1]. A ubiquitous computing network can be set up, where wireless technologies are applied to communicate accurate patient medical data to medical practitioners around the clock from the comfort of home; hence letting patients experience greater mobility and reduce hospitalization. This brings electronic health care support, known as m-Health in the literature [2], one step further.

A wireless real-time monitoring system can be organized in a wireless body area network (BAN) first coined by Van Dam et al. in 2001 [3]. The BAN in Fig. 1 consists of a number of different sensors and actuators connected using wireless communications to the intelligent personal node (body gateway). The sensors could e.g. be an electro-cardio-gram (ECG) sensor monitoring cardiovascular activity, a beat-to-beat sensor monitoring continuous blood pressure or an oximeter sensor observing the pulse and blood oxygen levels. An actuator could be a device stimulation muscle activator. The gateway communicates via a wireless link with a local or remote host application at a remote server. Such as system can provide ease in information-flow from the patient to the medical practitioners, in a convenient and secure way for the patient.

BAN's can acquire large quantities of patient medical information in real-time from the sensors. Such data should be communicated to the medical practitioners, in a suitable manner and data must be offloaded to the host for storage or post-processing from time to time.

Since wireless transmission is relatively energy costly, the gateway should only transmit context relevant data when needed, to minimize energy consumption. This give rise to several technical challenges such as, how should the sensors communicate wirelessly with the gateway? When should the gateway communicate data to the host? What data should be communicated to the host? How should the data be communicated to the host? This paper addresses some of these challenges.

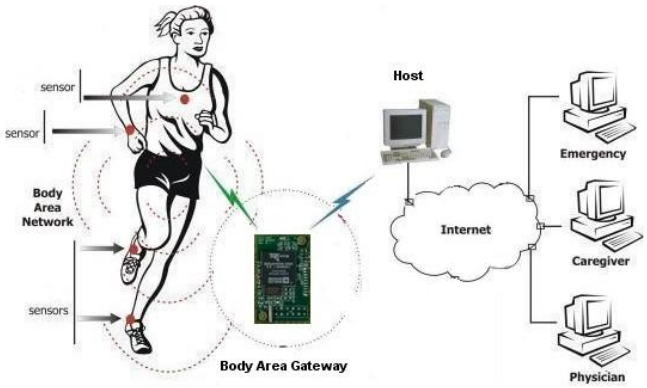


Figure 1. The wireless body area network.

A number of research groups have worked on implementing a BAN platform, typically for dedicated purposes. One example is the Human++ UniNode from the Netherlands [13] monitoring the autonomic nervous system. Here the network is build from a number of sensors communicating directly and the body area network is not connected to a separate gateway. In the same manner the MIT Media Lab [14] and the Fraunhofer Institute [15] have studied emotions using different portable monitoring systems. Additionally in [16] a prototype bio-potential sensor node is presented for monitoring a multitude of bio-potential signals. A system consisting of three of these nodes packaged in a headband enables wireless sleep stage

 ISBN: 978-1-61208-103-8

monitoring. What we try to do with ASE-BAN is to create a flexible platform that can be use in designing dedicated low power sensor nodes, more complex and resource demanding nodes e.g. including a signaling processor directly well as the gateway node itself where the protocol translation and connectivity of the BAN to the outside takes place.

Section II gives a brief discussion of BAN architectures with focus on design requirements/constraints.

Section III describes the low-cost Aarhus School of Engineering BAN (ASE-BAN) gateway prototype testbed, initially designed to monitor ECG signal and other patient medical signals over long time spans with no user intervention. This is work in progress.

The paper finalizes with a description of the future work of the ASE-BAN testbed.

II. THE BAN ARCHITECTURE

The network outlined in Fig. 2 illustrates a BAN consisting of a number of sensors/actuators nodes (motes), a body gateway and a host.

The motes and the body gateway are connected wirelessly within the body zone in a star or mesh network topology and relaying data (packets) to or from each other. Most (bi-directional) communication is between the gateway and the motes, but two motes could also communicate directly, e.g. in a sensor actuator setup where a measured parameter by sensor-mote-A (drop of glucose level) implies a consequent action to be performed real-time by actuator-mote-B (injection of insulin). Likewise, the gateway is connected wirelessly to the host in a bi-directional point-2-point connection.

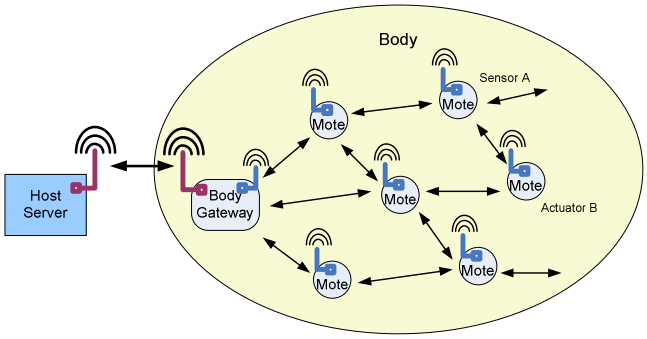


Figure 2. Wireless body area network topology and components such as sensors/actuators (motes), body gateway and host server.

A. The BAN Communication Protocol

Challenges with the design of a BAN communication protocol are issues like: noisy environment (RF noise and interference), variable traffic loads (dynamic bandwidth allocation), alarm situations (data priority/interruption), secure and accurate transmission (privacy and trustable), simple installation and service (plug and play), effortless adding/removing of motes and long time operation with minimal intervention (weeks/months/years). Therefore, key attributes for the BAN are: Reliability, scalability, security, power efficient, and easy of use and configure.

Since, single-hop communication (star topology) not always can be guaranteed (e.g. from front to back) and at the same time be power efficient, multi-hop communication (mesh topology) might be used to obtain optimized power efficient connectivity [4]. Interoperability meaning standard based protocols is also important to enable mote products from several vendors in the BAN.

An example of a protocol satisfying these requirements is the Time Synchronized Mesh Protocol (TSMP) proposed in [5], now being integrated into the emerging IEEE 802.15.4E standard. Key components of TSMP are:

- Time synchronized communication
- Frequency hopping
- Automatic mode joining and network formation
- Fully-redundant mesh routing
- Secure message transfer

In TSMP each transmission, transacted in a synchronized specific timeslot, contains a single packet and acknowledgements which are generated when a packet has been received unaltered and complete. Use of frequency hopping reduces the impact of interferences and increases the effective bandwidth. Aggressive use of duty-cycle and time-slot based principles makes the protocol very power efficient. A key attribute of TSMP is its self-organization mesh routing that makes it easy to add/remove motes. Finally, TSMP support encryption, authorization and integrity with regards to secure message transfer.

B. The BAN Gateway

A gateway (optionally two in case of redundancy) per BAN performs the following major tasks:

- Configuration and synchronization of the BAN
- Controlling and monitoring of the motes
- Collection and further processing of the sensors data
- Forwarding of processed and aggregated data wirelessly to the host for further processing and interpretation
- Reception of commands from the host

This requires hardware that includes: a transceiver supporting the communication within the BAN (motes), a transceiver supporting the communication with the Host, a powerful processor including memory and storage, and a power supply unit including a (rechargeable) battery. Optional, a display could be added for “on-site” monitoring.

The form factor and weight of the gateway should be tailored to be wearable with minimal impact on the body comfort, e.g. like a modern Smartphone or smaller.

 ISBN: 978-1-61208-103-8

Like most portable devices, the design should be optimized with low power consumption in mind. Weeks of operation without the need of recharging/replacing the battery would be acceptable.

C. The BAN Sensors

Sensors can be tiny patches worn on or implanted in the human body. The number and types of sensors in a BAN depends on the application. Examples of typical ones are:

- ECG sensor for monitoring heart activity
- Blood pressure sensor
- Oximeter sensor

The key functions of a BAN sensor are: physiological measurement and data collection, (optional) processing and forwarding wirelessly to the gateway. This requires specific physiological sensor hardware, a processor including memory, a transceiver (bi-directional communication) and a power supply source.

Small form factor, light in weight, and ultra low energy consumption are required with respect to physical comfort and minimal service. The latter one is with regards to extended battery lifetime (months or even years) without the need of intervention.

Therefore, sensor hardware should be designed and implemented with ultra low power consumption in mind, and with support for communication protocol supporting such operations. E.g. sensors kept in sleep mode when not performing any active tasks.

Integrated energy harvesting is another possibility to extend the rechargeable battery lifetime, potentially 'forever' in case the harvested energy is larger than the consumed energy over time. In body area networks body heat and body vibrations are obvious sources for energy harvesting. In [6] an example of using body heat is described that with proper energy management may eliminate the use of a battery.

III. IMPLEMENTATION & RESULTS

This section describes the low-cost BAN prototype testbed. This is work in progress and only preliminary results will be presented.

A. The ASE-BAN Testbed Model Overview

Fig. 3 illustrates the ASE-BAN functional diagram. Here one module describes the actual ASE-BAN gateway and two additional modules implement the ECG and the oximeter sensors.

In this initial version the wireless connection between the gateway and the host is implemented using Bluetooth whereas the connections (wired in this prototype) to the two sensor modules is only emulating the wireless channel. In the next version of the testbed the wireless ASE-BAN connections will be implemented proprietary low power radio communication modules.

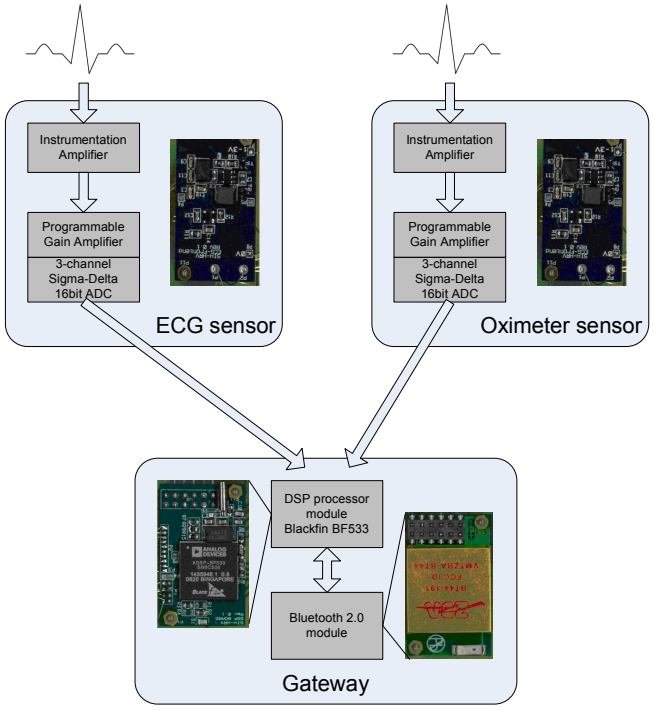


Figure 3. Block diagram of the ASE-BAN gateway with an ECG and an oximeter sensor module.

B. The ASE-BAN Gateway

The testbed prototype gateway consists of two modules, the processor module and the communication module, as illustrated in the lower part of Fig. 3. The modules are assembled to form a sandwich structure as show in Fig. 4.

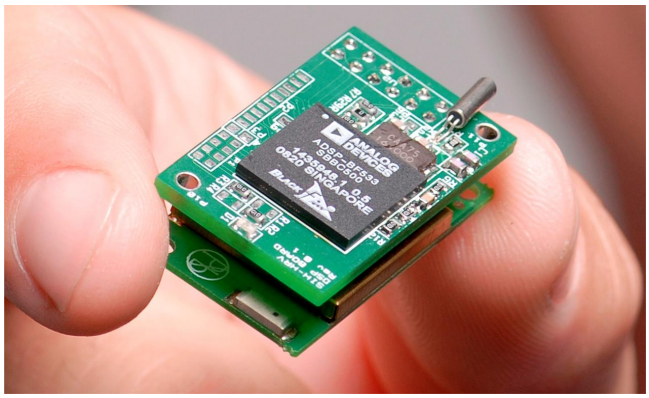


Figure 4. The physical ASE-BAN gateway module. The size is 13 mm x 18 mm x 30 mm. The weight is approximately 6 g.

The processor module is a small foot-print Digital Signal Processor platform equipped with a BlackFin BF533 signal processor from Analog Devices [7]. This signal processor is a high-performance fix-point processor with two 16 bits multiply-and-accumulate units, capable of parallel processing. The processor is capable of handle clock-speeds up to 600 MHz. The on-chip real-time-clock is connected to a 32 kHz crystal. The module also includes a M25P10-A

 ISBN: 978-1-61208-103-8

serial 1 MBit data flash for program storage and a secure digital memory card for on-board local data storage. On reset the processor boots the program form the flash. The processor may transfer data to the SD-card using the serial peripheral interface.

The wireless communication module, consist of a Bluetooth 2.0 module of class 2. The RF range is up to 100 meter. It supports data transfer rates up to 3 Mbits/s. The module is connected to the processor module trough a UART interface. The module facilitates connectivity to sensors in the ASE-BAN and to the host, in this setup a PC or cell phones. The module is easy to use, but costly in terms of energy consumption, especially in relation to the rather low bandwidth need for the current set of sensors (< 500 Hz @ 16 bit). The next generation will have a more energy efficient communication module.

The processor platform is used to process the signals from the sensors. Current software runs standard adaptive noise removing techniques to remove hum in the ECG. The R-peak in the ECG signal is calculated using the Pan Thomkins algorithm [8] and finally classic pNN50 Heart Rate Variability [9] is calculated for diagnostic purposes.

C. *The ASE-BAN Sensors*

The current prototype supports 2 types of sensors: an ECG sensor and an oximeter sensor.

The ECG sensor module measures 2 lead ECG signal on patients. The module includes an ECG amplifier, an analog-to-digital converter and a power supply unit.

Since ECG signal typically has peak to peak amplitude of approximately 2 mV amplification is needed prior to the analog to digital conversion. The amplification is done using the AD620 instrumentation amplifier from Analog Devices [10]. This amplifier has high bandwidth, low noise and providing high common-mode rejection, as such offers high quality amplification of the ECG signal.

The AD conversion is implemented, using the AD770 3-channel 16 bit ΣΔ-converter from Analog Devices [11]. The sampling rate for the ECG signal is set to 500 Hz.

The current power supply unit accepts input voltages in the range 0.8 – 3 V and enables warning on low battery.

The sensor is intended to be worn for a longer time span without intervention; hence appropriate electrodes must be selected, to provide for high signal quality and patient comfort. The prototype uses 2 lead insulated bio-electrodes which provide good signal quality and reduced risk for skin irritation [12]. Fig. 5 shows a recorded ECG on the host server. The power consumption can be as small as 500 µW, this means a battery on 0.5 Wh will operate a couple of weeks given continuous operation.

The oximeter measures the oxygen saturation in the patient blood. The module is being implemented in a similar way as the ECG module (same size etc.), but since the module is in the design phase no measurement results are at this point in time available.

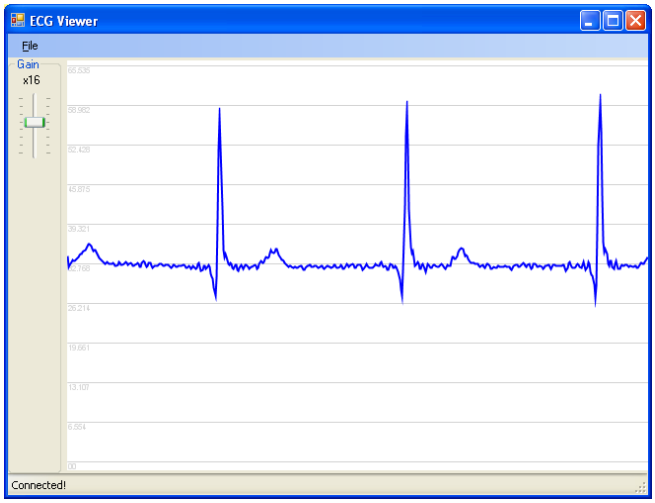


Figure 5. A real-life ASE-BAN ECG mesurement.

IV. CONCLUSION & FUTURE WORK

In this paper an initial prototype body area network testbed for measuring ECG and oxygen level in blood has been presented. The testbed has been implemented and successfully tested for ECG data collection. The oximeter module is in the design phase so no measurement results are available at this point in time.

Since this paper presents the initial implementation of the ASE-BAN testbed a larger number of activities have been postponed to further work in the near future. This work can be categorized in two levels, short term and long term. The immediate short of to implement and test the oximeter sensor. Additionally the short term issues are to work with power efficient signal processing techniques for robust detection of the R-peaks and accurate estimation of the heart rate variability for diagnostic purposes including compression of sensor data. In the longer term new sensor hardware for the body area network will be implemented to create low power solutions with little or no battery power. A number of energy harvesting techniques are being investigated. Additionally the gateway hardware itself will be optimized in terms of energy consumption. This includes the radio module as well as the software implementation which in the future will offer mechanisms to put the device to sleep when not necessary. Additionally the network communication protocol itself will be designed to reduce the power consumption of the BAN for continuous real-time monitoring. Finally the introduction of an ultra-small-scale IP stack in the gateway is being considered for connectivity to the host and further on a network infrastructure.

REFERENCES

[1] Frost and Sullivan. European Remote Patient Monitoring Markets, June 2008

[2] R. S. H. Istepanian, E. Jovanov, and Y. T. Zhang, “Guest editorial introduction to the special section on m-health: Beyond seamless mobility and global wireless health-care connectivity," Information

 ISBN: 978-1-61208-103-8

Technology in Biomedicine, IEEE Transactions on, vol. 8, no. 4, pp. 405-414, Dec. 2004.

[3] K. Van Dam, S. Pitchers, and M. Barnard, “Body area networks: Towards a wearable future," in Proceedings of WWRF kick off meeting, Munich, Germany, 6-7 March 2001.

[4] B. Latre, B. Braem, I. Moerman, C. Blondia, P. Demesester ”A Survey on Wireless Body Area Network”

[5] Dust Network “Technical Overview of Time Synchronized Mesh Protocol”,Doc.nr. 025-0003-01, June 20, 2006, http://www.dust-networks.com/technology (June 2010).

[6] Qadeer A. Khan, Sarvesh J. Bang “Energy Harvesting for Slef Powered Wearable Healt Monitoring System”, Oregon State University, http://en.wikipedia.org/wiki/Body_Area_Network (June, 2010).

[7] Analog Devices, “High Performance General Purpose Blackfin Processor”, 2008, http://www.analog.com/en/embedded-processing-dsp/blackfin/adsp-bf533/processors/product.html

[8] J. Pan and W. J. Tompkins, A real-time QRS detection algorithm. *IEEE Trans. Biomed. Eng.*, **BME-32** (3):230-236, 1985.

[9] “Guidelines. Heart rate variability”. European Heart Journal (1996) 17, 354-381.

[10] Analog Devices, “Low Drift, Low Power Instrumentation Amp with Set Gains of 1 to 10000”, 2008, http://www.analog.com/en/other/militaryaerospace/ad620/products/product.html (june 2010).

[11] Analog Devices, “3V/5V, 1 mW 3-Channel Pseudo Differential, 16-Bit Sigma-Delta ADC”, 2008, http://www.analog.com/en/analog-to-digital-converters/ad-converters/ad7706/products/product.html (May 2010).

[12] C. Park, P.H. Chou, Y. Bai, R. Matthews, A. Hibbs, "An ultra-wearable, wireless, low power ECG monitoring system", Biomedical Circuits and Systems Conference, Nov. 2006, pp. 241-244.

[13] L. Brown, B. Grundlehner, J. van de Molengraft, J. Penders and B. Gyselinckx, “Body Area Network for Monitoring Autonomic Nervous System Responses”, 3rd International Conference on Pervasive Computing Technologies for Healthcare, April 2009, London, UK.

[14] M. Strauss, “Handwave: Design and manufacture of a wearable wireless skin conductance sensor and housing,” Master Thesis, MIT June 2005.

[15] C. Peter, E. Ebert and H. Beikirch, “A wearable multi-sensor system for mobile acquisition of emotion-related physiological data”, in Proceedings of the 1st International Conference on Affective Computing and Intelligent Interaction, 2005.

[16] N. de Vicq, F. Robert, J. Penders, B. Gyselickx and T.Torfs, ”Wireless body area network for sleep staging”, in Proc. Int. Conf. on Biological Circuits and Systems, 2007.

 ISBN: 978-1-61208-103-8

Body Aura — A New Approach Towards Ambient Intelligence

Vision Paper

Peter H. Deussen, Edzard Höfig
Fraunhofer Institute for Open Communication Systems
Berlin, Germany
{peter.deussen|edzard.hoefig}@fokus.fraunhofer.de

Borbala Katalin Benko
Budapest University of Technology and Economics
Budapest, Hungary
bbenko@hit.bme.hu

***Abstract*—The 21st century will be characterized by a number of technical revolutions, which will not leave the ways in which humans interact unchanged. New forms of interactions will change our society in unforeseeable ways. On the other hand, pressing problems such as environmental protection, energy shortages, the spread of pandemic diseases, global crises, etc., demand novel approaches to collect and to correlate more precisely demographic data. To analyze this challenge and to investigate possible technological scenarios, this paper introduces the concept of a *Body Aura*, a digital extension of the body functions of a person or a group of persons to the environment.**

***Keywords*-Body Aura, ubiquitous systems, autonomic systems, self-representation, emergence**

I. Introduction

The 21st century will be characterized by a number of technical revolutions, which will not leave the ways in which humans interact unchanged. Current scenarios on pervasive devices and smart rooms, augmented reality, and even the (already outdated) idea of the "cyberspace" as a new type of human/computer interface barely scratch the surface. New forms of interactions will change our society in unforeseeable ways. On the other hand, pressing problems such as environmental protection, energy shortages, spreading pandemic diseases, global crises, etc., demand novel approaches to collect and to correlate more precisely demographic data; and this has to be done in a secure, reliable way, which respects issues such as privacy and trust.

Considering multiple achievements in areas such as sensor networks, augmented reality, distributed computing, and autonomic systems, the major question is not how to interact technologically with the digital computing, data, and communication resources, but to embed, to maintain, and, to relocate sufficient intelligence and functionality into pervasive devices and networks to perceive them as a useful and trusted extension of personal, professional, and societal spheres. Human activities do hardly occur in isolation but in almost all cases in relation to activities of other humans, organized by ad-hoc communications on various layers, work-flows, and with common as well as conflicting interests. The capability of pervasive and ubiquitous systems to perceive and to support not only individuals but also relations, objectives, interests, and patterns for (probably large) groups of humans becomes crucial for their practical applicability.

To analyze this challenge and to investigate possible technological scenarios, this paper introduces the concept of a *Body Aura*, which can be defined as a *digital extension of the body functions of a person or a group of persons to the environment, both to monitor, correlate and to effectively understand them in relation to the physical parameters, and to actually control those parameters, producing notifications, performance data, and alarms if necessary*.

The paper is organized as follows: Section II discusses the general concept of a Body Aura. This notion will be elaborated in more detail in Section III by means of a number of examples, which are used to identify challenges concerning the development of a Body Aura. Some technical aspects are discussed in Section IV. Section VI draws conclusions and gives an outlook on further work.

II. Concepts

The Body Aura vision assumes that in the near future a pervasive digitalization of the human environment (and bodies) will take place, which involves not only various sensors, control devices, communication facilities (wireless technologies of various kinds), data sources (such as RFID tags), but also storage capacities and processing power (a trend which is already visible today by the omnipresence of mobile devices). The notion of "processors per cubic meter" will become a meaningful unit. The deployment of pervasive applications and systems will become feasible, which today cannot even imagined, going beyond visions of augmented reality, pervasive and ubiquitous communication services. It will be the age of intelligent matter supporting the seamless extension of human bodies, minds, and interactions by information and communication technologies.

A Body Aura forms a vital personal field for both individuals and groups of people to provide various direct interactions between users and their environments, to support human activities both on a conscious and a sub-conscious level. A Body Aura establishes itself as a radically distributed system using resources (sensors, actuators, computing, storage, and communication) found in the proximity of its users without centralized elements. There is no "Body

 ISBN: 978-1-61208-103-8

Aura gadget" to be carried around all the time, a Body Aura results from the ongoing interaction of those resources. User movement results in a re-location of parts of the Aura while maintaining both function and integrity.

Human interactions result in activities which are meaningful only in the context of a group, hence the person-centric conception of a Body Aura needs to be extended to an activity-centric conception: Instead of monitoring (and predicting) the current "state" of a solitary user, common activities of a set of users need to be recognized and represented. We refer to this transformational process to as the *emergence* of new Auras. Emergent Auras will not replace personal ones, but complement and extend them.

Body Auras can be considered as exemplificative instances of systems which strongly depend on the internal quality of self-awareness. A Body Aura needs, at any instance in time, be aware of its own configuration and internal operation as well as its relationship to its users and their environment. Autonomic features such as self-configuration, self-optimization, self-healing, and self-protection, but also self-organization capabilities, are vital for the functionality and survivability of Body Auras [1], [2].

III. Examples and Challenges

To explain the idea of a Body Aura in more detail and to identify the accompanying challenges, let us analyze a number of scenarios from various application areas:

A. Sports Health Monitoring

As a first basic example, consider a cyclist in training (Fig. 1). To obtain optimal training results, his Body Aura takes measurements of his body functions such as heartbeat, body temperature, breathing frequency, blood pressure, etc. To obtain a better estimate of the cyclist's performance, it takes also environment data such as temperature, air moisture, etc., into account. The current position of the cyclist is determined using GPS. A (hypothetical) actuator controlling the resistance of the bottom bracket is used to emulate accelerations, road conditions, etc.; data to emulate a real racetrack are downloaded and continuously updated via WLAN. The necessary computations are performed on various computing nodes in the current proximity of the cyclists. The Body Aura resides only to a certain part at the devices carried by the cyclists or mounted on his bike; other functions are performed by devices in this proximity.

This example illustrates the first main challenge that has to be addressed in the Body Aura approach, namely:

Homeostasis [3] (also called *meta-stability*) refers to the capability of a system to maintain its identity, function, and structure in the presence of continuously changing conditions. For instance, the cyclist's Body Aura has to continuously re-locate parts of it when its user leaves the range of fixed devices (computing nodes, WLAN access points) and enters the range of others. It has in particular the

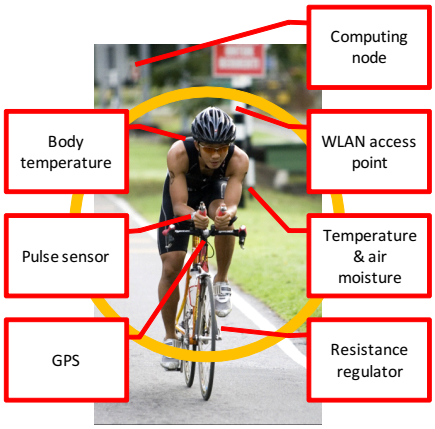


Figure 1. Health monitoring example

ability to predict the availability of resources for a certain time frame, and to degrade its functions in a graceful way if required resources are temporarily unavailable.

A Body Aura of a solitary person may comprise of only a handful of devices. Scalability becomes an issue if an Aura relates to more than one person. For instance, each member of a diving team needs to be aware of the body conditions of her partners, as well as of environmental conditions such as currencies, water temperature, etc. The selection of gas mixes and gas pressure, deceleration and acceleration speed, diving formation, and so on, are critical parameters for a successful and safe dive.

Context and situational perception. The above example illustrate that a precise detection of the current situation and context of a person in relation to other persons is another crucial capabilities of Body Auras. It also makes clear that contextual information can be obtained not only from environmental sensors but also from the interaction of several Body Auras.

B. Recreation, Entertainment, and Travel

Everybody who used to travel a lot immediately understands that localized information is of high importance, e.g., on public transportation, locations of recreational facilities and restaurants, current traffic conditions, and so on. Of course, the idea of localized information has been presented before. We use this example to identify another challenge:

Identification. How does a Body Aura identify the person it relates to? As already pointed out, there is no Body Aura device which allows for an authentication. If a Body Aura is not strongly tied to the body function of a specific person (where the actual configuration of worn sensors can be used as authentication criterion), the maintenance of the relation between Aura and user becomes a considerable challenge.

Consider a new class of role gaming applications where players assume roles in real life situations, based for instance on "spook" scenarios: Players try to investigate the pretended

 ISBN: 978-1-61208-103-8

activities of other players (e.g., spy out the delivery of—in the game—crucial information to another player). Body Auras can be used to mediate between players, to pass information related to the game situation. Moreover, game opportunities and limitations depend on the particular role a player assumes. Not all players are willing or capable to show James Bond like mannerisms, but prefer a more realistic style. A Body Aura hence needs not only be capable to distinguish between game and "first life" situations, but also needs to be able to adapt to the specifics of the role a player tries to assume.

Adaptability and Learning. Games can be scripted, but the specific reactions of a player to a certain game are not predictable by the game provider, and will change gradually over time when the player becomes more experienced. Therefore, Body Auras need to be capable to learn about the specifics of the person it relates to, and adapt itself accordingly with respect to its assessment of the user's state, and its reactions to it, as well as its internal structure (prioritizing of data and processes for prediction and reaction establishment).

C. Catastrophe and Emergency Management

Consider a major emergency (a large fire, an earth-quake). The communication infrastructure might be impaired or not available at all, making it hard for rescue teams to locate injured persons, coordinate and prioritize their activities. Devices powered by batteries or local energy sources (sun-light collectors), etc. might be still working, capable to establish an ad-hoc communication network. Body Auras of the rescue team, members and people to be rescued will not only interact, collecting, and delivering crucial information, but in effect "merge" and perform collective activities in an expanded context.

Another example of Body Aura emergence is the detection of the propagation of epidemic diseases. Body Auras can provide a city- or even country-wide instrument to monitor people's body functions by the responsible authorities.

Emergence. We refer to the shift from the focus of a singular person to a group of persons, together with a reinterpretation and reassignment of Body Aura functions (that is, data are still related to body functions, but interpreted in the context of different tasks) to as Body Aura "emergence". The emergency scenario used to introduce this capability is of course an extreme example, in fact every group related activity supported by the interaction of Body Auras to provide services which make sense only in a group context relates to Body Aura emergence.

Self-assessment. Some applications of the Body Aura concept just require "best effort". For instance, for the cyclist in the first example it is not of importance if his Body Aura stops working for a couple of seconds because the momentarily available resources are insufficient. The emergency scenario however illustrates that in some contexts best effort is not desirable. In such situation, a Body Aura has been able to assess their own ability to work on an appropriate level given the current restrictions and limitations. If necessary it has to indicate problems to a human operator.

Two additional problems are immanent:

Security. How can it be ensured that personal information is not distributed to parties not allowed to have this information? How can Body Auras be secured against misuse?

Scalability. Progressing from the example of a cyclist trying to optimize his training efforts to the emergence of a population-wide Body Aura is a considerable step which can only be achieved if the envisioned highly distributed architecture of Body Auras is utilized to avoid single points of processing (and failure) and to keep state information as local as possible.

D. Work Convenience

Consider a group of workers. Their Body Auras do not only monitor their body functions, but also feedback information to the workers. For instance, a worker in a dangerous environment may get a notification about the level of her exhaustion. But also, the work flow itself can be assisted by providing sub-conscious interfacing functions to "smart" environments: The room temperature and illumination may be automatically adjusted by the Body Auras of office workers, equipment may adjust to the worker's preferences and physical state, etc.

User/Environment Interaction. Body Auras are envisioned to be more than complex measurement instruments; they actually will perform interactions with the environment on behalf of their users without an explicit user action. In this way, Body Auras may be viewed not so much as tools, but as extensions of the sensoric and haptic body functions of their users to a computerized environment.

IV. Technical Concepts

To describe the main concepts of the Body Aura idea, consider the "high level architecture" displayed in Fig. 2. The general idea departs from the assumption that a Body Aura can be suitable defined by a "Representation of its Self", which describes all aspects, which need to be considered to perform its tasks.

Sensors both related to the Body Aura users and their physical environment acquire raw data which are interpreted using the self-representation as a contextual frame. User augmentation and environment related actuators are used to feed back information to the users, and to interact with their environment. Additional resources for computing, communication, and storage, are integrated as additional supportive elements.

Representation of the Self. Self-awareness of Body Auras is a basic precondition for the establishment of functions such as autonomic distribution, homeostasis, prediction, etc. For that, an internal self-representation of Body Auras is

 ISBN: 978-1-61208-103-8

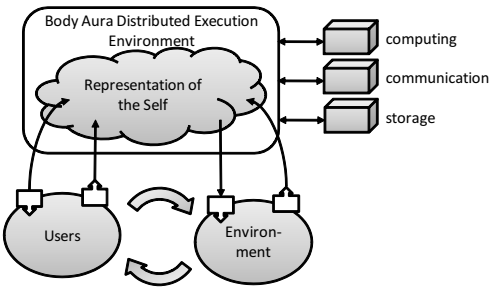


Figure 2. General Body Aura architecture

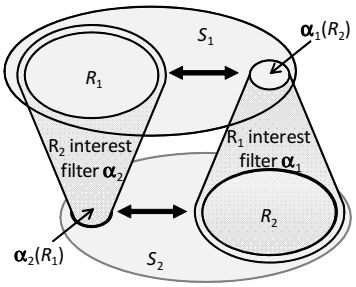


Figure 4. Abstracted composition of self-representations

required which has—with regard to the challenges identified in the previous Section III—to fulfil a number of properties (Fig. 3): It has to "store" data about the state of the users, their environment, and its own actual configuration (sensors, actuators, resources). We call this the representation of *situational knowledge*. Moreover, to provide for planning and prediction, it has to comprise a causal representation of the relevant (virtual and real-world) processes, accompanied by known facts about users, resources, etc. We refer to this type of information as *background knowledge*. Finally, a Body Aura performs a number of tasks to support or serve its users. These activities—defining the *purpose* of a Body Aura—need to be represented within the self-representation too (one might view them as "application layer" of a Body Aura).

Distribution and Scalability. Due to the distributed character of Body Aura there cannot be a centralized self-representation. Therefore, mechanisms are needed to manage a distributed version of the Body Aura model. We propose to use a holistic approach to represent the whole structure of the Body Aura on each contributing entity, but on different level of abstractions. An entity maintains model structures concerning its own functions and capabilities to a higher

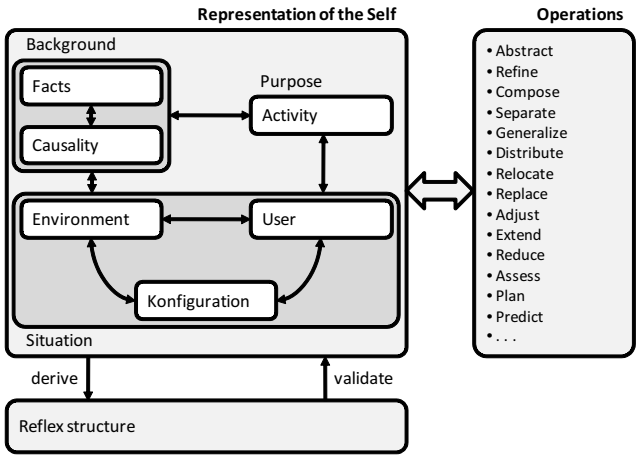


Figure 3. Self-representation, reflexes, and operations

degree of detail then the model structures describing adjacent entities. Body Auras are thus composed of various "knowledge fragments" and pre-defined or learned abstraction operations. The construction of composed representation and the determination of an appropriate level of abstraction with regard to a certain entity becomes an automatic process.

A comprehensive discussion of this idea of abstracted composition of self-representation is far beyond the scope of this vision paper; Fig. 4 provides an intuition by considering two sub-systems (devices, sensors, etc.) with self-representations R_1 and R_2, respectively, which are contributing to a Body Aura. Regarding of what is important for each of these systems to know from each other, a "filter" α_i is constructed which is used to scale down the self-representation of the other system. Thus the composite view of S_1 to the system comprising both system consists of R_1 composed with the abstracted view to R_2 according to α_1 (and similarly for S_2). A formal consideration of these mechanisms can be found in [4], [5].

Emergence of Auras is—on the level of the self-representation—performed by the application of composition/abstraction operations to the involved Auras. Mappings between group related activities and personal ones have to be understood as a reinterpretations of representation structures. This process has to be iterative: Joining of Body Auras will result in the emergence of another Body Aura, which can be merged with other Body Auras in a seamless process.

Reflex Establishment. Some activities of a Body Aura occur frequently and thus may be understood as general behaviour patterns which are applicable is several contexts. Thus instead of computing those patterns over and over again (including the validation of their effectiveness), "short cuts" can be established which apply those pattern without any further planning or analysis activities—they become "reflexes". We assume that the establishment of reflex patterns is likely to increase the efficiency of a Body Aura, and thus increase also the range of tasks which can be performed.

Context-awareness. The creation of self-representation requires adaptive, efficient and scalable methods for the exploitation and utilization of the information available about the user and the actual contextual conditions. An open system must be ready to detect relationships in a lightweight

 ISBN: 978-1-61208-103-8

manner, often from raw data, about trends, frequent episodes, timely relationships and anomalies or abnormal situations. Another important aspect is the active extraction of predictive relationships, which enables the proactive handling of predicted near-term situations or presumptive future actions.

Learning and Adaptation. A new-born Body Aura with its limited, default knowledge is of much less use than an evolved, optimized one that is highly adapted to its environment and its user. Hence, mechanisms are needed to adjust the Body Aura's self-representation: New self-representation fragments get "learnt", existing ones get refined, replaced or removed. The progress of self-representations is envisioned as an automatic process, without human interaction or intervention—based on the trends, situations, or predictive relationships detected in the local Aura. There are several ways to facilitate this kind of evidence based refinement the models, for example techniques may be borrowed from reinforcement learning [6], data mining [7] or immunology [8]. However, the sole detection of these relationships within the data is not enough to create self-representation fragments. Cognitive aspects, higher-level decisions and goals are also vital for a Body Aura in the process of transforming a detected relationship into an actively utilized, executable self-representational model. High-level aspects may simply seek optimality (I already know a better strategy than what this relationship suggests) or may include meta-reasons (I don't want to what the observations suggest because it would hurt my environment and that's not in line with my meta-goals). In other words, a Body Aura needs to be able to learn from its experience, but is not forced to use the knowledge blindly, without control. Knowledge both emerges on the lower level from the observations, and on the higher level, including cognitive aspects.

Networking. Due to the non-centralized, and self-organized character of the Body Aura system, where the amount of conscious control is reduced radically, the classical telecommunication approaches cannot be used. The lack of centralized entities means that the Body Aura system is similar to an opportunistic or a delay-tolerant network in terms of communication. However the well-known ad hoc networking solutions cannot be adapted to the Body Aura system, as it is not taking into account the emergence of Body Auras, together with the regrouping process.

This means that the Body Auras (which could be represented as mobile nodes in the communication perspective) form mobility groups, and the structure of these groupings could heavily effect the propagation and relaying of the information in the system. The information propagation and the dissemination process is not only significantly influenced by the group mobility patterns of the Body Auras but also by the common decisions of the groups, for example which information are important for the given group, or which information should be disseminated to the neighbouring groups, namely merged Auras.

Security. The Body Aura dia removes static architectural features and in favour of heterogeneous structures, and moreover increases the independence from the "physical layer". Thus, its architecture and as a consequence its security architecture is completely detached from single hardware instances. Due to this feature, Body Aura also allows its distribution on multiple devices, representing a distributed client application which acts on behalf of the body the Aura is associated with.

In Body Aura, complete sets of devices, and several services which are spread among potentially numerous devices will act on behalf of a "body". Thus, the question arise: Who will be the principal, i.e., who authenticates what and what does this imply for the system? Novel mechanisms (e. g., based on biometric pattern recognition) need to be developed to establish a trust relation between Body Auras, users, and other involved parties.

Augmenting the Person. A Body Aura follows the activities of the person (the atomic level) or a group of persons through passive observation. If some tasks are recognized the Body Aura pursues all activities of the person and compares them to corresponding activity patterns. If more than one activity representation applies, these are treated as competing. Some of the tasks may reduce the amount of activity models which correspond to the current actions of the person. In some cases the Body Aura is not able to recognize what the user is doing and what his intention is. In this case the Body Aura tries to interact with the user. Hence, models for the interaction of persons with their own Body Auras are needed. But interaction between the person and the Body Aura is an additional load for the person and should happen as rarely as possible. A model is required, which describes several levels of interaction complexity (yes/no, item selection, commands) and its form (gestures, voice, mouse, keyboard).

V. Related Work

Over the last two decades, a new understanding of the relationship between human users and computer systems has emerged, grounded in visions developed in the early 90's, of ubiquitous computing [9] (i.e. computation moving out of the box to be pervasive in the user's ambient), augmentation of the real world [10] (i.e. enhancing physical reality with digital interaction), and wearable computing [11] (i.e. augmentation of humans with always-there computational services). These original visions have largely converged in today's digital world, with personal mobile devices that continually support their users, pervasive devices and services available in the ambient, and widespread technologies for linking digital interaction to the entities and activity in the real world (passive and active tagging [12], location systems [13], mobile augmented reality [14]). However, in spite of the paradigm shift that has taken place, there is a total lack of concepts for how the intelligence that is deeply and densely

 ISBN: 978-1-61208-103-8

embedded in people's environments (in the form of sensors, actuators, digital media, smart objects, smart materials, etc.) can be harnessed to effectively support user activities and needs. Key barriers are the very limited representation of user context to the environment, and the centralized approach to interface complex computerized environments.

The distinct approach, in departure from the state of the art, is to develop the interface between people and complex computerized environments as a self-aware autonomic system conceived as a Body Aura around the user. In terms of high-level vision there have been precedents of conceptual work on auras or spheres around the user, e.g. Ferscha et al.'s Digital Aura as a thought model for interactions in a pervasive computing world [15] and Mynatt et al.'s Audio Aura providing serendipitous information via background auditory cues [16], but in practically all work on human-environment interaction, the sphere of interaction is implicitly defined by the nature of the infrastructure (network topology, communication range of user devices, sensor coverage).

VI. Conclusion and Further Work

In this paper, we have introduced the vision of a Body Aura, which acts as an information technological extension of the user's body functions to his or her physical environment, and is capable to provide a functions for (probably large) groups of persons as well, hence is applicable in a large variety of scenarios.

The Body Aura vision involves a number of research areas which cannot be addressed simultaneously. In our future research, we will concentrate of the following key issues:

1) One of the main concepts of the Body Aura approach is the definition and maintenance of a distributed "representation of the self" describing all operational aspects of a Body Aura at a "suitable degree of abstraction", defined by the knowledge requirements of devices contributing to the Aura. Although some work has been already conducted into this direction [4], [5], further research is needed.
2) Investigations on the interpretation of behavioral representations (UML Statecharts) of functions implemented on resource limited devices [17] indicate that it is possible to maintain computational models even on very small devices with suitable efficiency. Future research will address different modelling formalisms.
3) An open learning an adaptation model to systematically detect, refine and utilize relationships within the environment, with help of data mining techniques.

References

[1] J. Kephart and D. Chess, "The vision of autonomic computing," *IEEE Computer*, vol. 36, no. 1, pp. 41–52, 2003.

[2] S. Dobson, S. Denazis, A. Fernandez, D. Gaiti, E. Gelenbe, F. Massacci, P. Nixon, F. Saffre, N. Schmidt, and F. Zambonelli, "A survey of autonomic communications," *ACM Trans. on Autonomous and Adaptive Systems*, vol. 1, no. 2, pp. 223–259, 2006.

[3] W. R. Ashby, "Homeostasis," in *Encyclopedia of Biochemistry*, R. J. Williams and J. E. M. Lansford, Eds. New York, N.Y.: Reinhold Publishing Corp., 1967, pp. 411–412.

[4] P. H. Deussen, "Model based reactive planning and prediction for autonomic systems," in *Proc. Workshop on INnovative SERvice Technologies*, 2007, pp. 1–10.

[5] ——, "Supervision of autonomic systems — tutorial," in *Proc. Budapest Tutorial and Workshop on Autonomic Communications and Component-ware*. Budapest, Hungary: HTE (Scientific Association for Infocommunictaion Hungary), July 7–9 2008, published on CD, avail. at www.peterdeussen.net.

[6] D. Lanyi and B. K. Benko, "A nontraditional approach for a highly interactive collective-adaptive system," in *Proc. EMERGING 2010*, 2010, p. (in press).

[7] L. Huan and H. Motoda, *Feature Extraction, Construction and Selection, A Data Mining Perspective*. Springer, 1998.

[8] D. Flower, P. Andrews, J. Timmis, P. Guan, and I. Doytchinova, *In Silico Immunology*. Springer US, 2007.

[9] M. Weiser, "The computer for the twenty-first century," *Scientific American*, vol. 265, no. 3, pp. 94–110, Sept. 1991.

[10] P. Wellner, W. Mackay, and R. Gold, "Back to the real world," *Comm. of the ACM*, vol. 36, no. 77, pp. 24–27, 1993.

[11] B. Rhodes, "The wearable remembrance agent: A system for augmented memory," *Personal and Ubiquitous Computing*, vol. 1, no. 4, pp. 218–224, 1997.

[12] R. Want, K. P. Fishkin, A. Gujar, and B. L. Harrison, "Bridging physical and virtual worlds with electronic tags," in *Proc. SIGCHI Conf. on Human Factors in Computing Systems (CHI '99)*, 1999, pp. 370–377.

[13] J. Hightower and G. Borriello, "Location systems for ubiquitous computing," *Computer*, vol. 34, no. 8, pp. 57–66, Aug. 2001.

[14] D. Wagner, T. Pintaric, F. Ledermann, and D. Schmalstieg, "Towards massively multi-user augmented reality on handheld devices," in *Proc. Int. Conf. Pervasive Computing (Pervasive '05)*, 2005, pp. 208–219.

[15] A. Ferscha, M. Hechinger, R. Mayrhofer, M. dos Santos Rocha, M. Franz, and R. Oberhauser, "Digital aura," in *Proc. Advances in Pervasive Computing*, 2004, video paper at Pervasive 2004 conference.

[16] E. D. Mynatt, M. Back, R. Want, M. Baer, and J. B. Ellis, "Designing audio aura," in *In Proc. SIGCHI Conf. on Human Factors in Computing Systems (CHI '98)*, 1998, pp. 566–573.

[17] E. Höfig, P. H. Deussen, and H. Coskun, "Statechart interpretation on resource constrained platforms: a performance analysis," in *Proc. ACM/IEEE 12th Int. Conf. on Model Driven Engineering Languages and Systems*, Denver, Colorado, USA, Oct. 2009, pp. 99–108.

 ISBN: 978-1-61208-103-8