Infocommunications Journal

A PUBLICATION OF THE SCIENTIFIC ASSOCIATION FOR INFOCOMMUNICATIONS (HTE)

September 2014 | Volume VI | Number 3 | ISSN 2061-2079

Technically Co-Sponsored by

IEEE COMMUNICATIONS SOCIETY

hte

IEEE HUNGARY SECTION

Editorial Board

Indexing information

Infocommunications Journal is covered by Inspec, Compendex and Scopus.

Infocommunications Journal

Technically co-sponsored by IEEE Communications Society and IEEE Hungary Section

Supporters

GÁBOR BÓDI – president, National Council for Telecommunications and Informatics (NHIT)
GÁBOR MAGYAR – president, Scientific Association for Infocommunications (HTE)

Editorial Office (Subscription and Advertisements)**:**
Scientific Association for Infocommunications
H-1051 Budapest, Bajcsy-Zsilinszky str. 12, Room: 502
Phone: +36 1 353 1027, Fax: +36 1 353 0451
E-mail: info@hte.hu • Web: www.hte.hu

Articles can be sent also to the following address:
Budapest University of Technology and Economics
Department of Networked Systems and Services
Tel.: +36 1 463 3261, Fax: +36 1 463 3263
E-mail: szabo@hit.bme.hu

Subscription rates for foreign subscribers: 4 issues 50 USD, single copies 15 USD + postage

Publisher: PÉTER NAGY • Manager: ANDRÁS DANKÓ

HU ISSN 2061-2079 • Layout: MATT DTP Bt. • Printed by: FOM Media

www.infocommunications.hu

Special Issue on the Future Internet – *Guest Editorial*

Gyula Sallai, Wolfgang Schreiner, and János Sztrik

Recent dramatic changes such as the rising number of Internet users, the penetration of portable and mobile devices, or the Internet of Things, has motivated a number of research initiatives, labeled "Future Internet" worldwide, supported by NSF in the USA and EU research framework programs in Europe. In Hungary, the ongoing "Future Internet Research, Services and Technology – FIRST" project, supported by the European Social Funds focuses on key theoretical, modeling, planning, application and experimental aspects of Future Internet. The six papers published in this special issue demonstrate the research results achieved by the FIRST research community in various fields related to Future Internet.

Since the standardization of the TCP/IP 40 years ago, TCP is, after several modifications, still the protocol providing reliable end-to-end transport on the Internet. The first paper, "Towards the Transport Protocols of Future Internet", by *Z. Móczár and S. Molnár,* presents the evolution of transport protocols since the early days of the Internet, gives an overview of the main pitfalls the researchers faced with during the years, and suggests a promising approach which may be able to satisfy the diverse requirements of future networks.

The Internet Protocol by its nature does not guarantee the delivery of packets in the right order. Therefore, it is important to investigate the effects of packet reordering. Authors *A. Kuki, B. Almási, T. Bérczes and J. Sztrik,* in their paper "Modeling a QoS Classified Communication in a multiuser Wi-Fi Environment", propose a finite source queueing model that includes the packet reordering feature. The authors show how the packet reordering phenomenon influences the main waiting times compared to the FIFO discipline.

In next generation wireless telecommunications networks not only the movements of single mobile endpoints but also entire mobile network movements need to be managed (network mobility or NEMO). The paper "A study on the Performance of an Advanced Framework for Prediction-based NEMO Handovers in Multihomed Scenarios", by *L. Bokor, G. Jeney, J. Kovács,* provides an extensive performance evaluation of an advanced handover management solution that aims at providing ubiquitous IPv6 connection and seamless Internet access for NEMO scenarios.

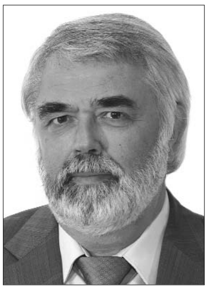
Despite the reliability feature of TCP, its relatively high complexity does not always enable to implement it in a hardware environment with constrained resources. The paper by *P. Orosz, T. Skopkó, and M. Varga,* titled "RCTP: A Low-complexity Transport Protocol for Collecting Measurement Data", introduces a low-complexity transport protocol dedicated to a real-time network monitoring system operating above 10 Gbps. The protocol may be suitable for measurement networks such as sensor networks.

The fifth paper ot this special issue, "Internet of Things: application areas and research results of the FIRST project", by *Z. Gál, B. Almási, S. Oniga, S. Baran, T. Dabóczi, R. Vida, and I. Farkas,* gives an overview of the research results achieved within the FIRST/IoT Project. The paper deals with the following six topics:

i) Integration of the IoT into the IPv4/IPv6 systems
ii) Cyber physical systems
iii) Self-optimizing and self-managing communication mechanisms of the IoT systems
iv) E-health powered by IoT
v) Weather prediction network tool development and analysis
vi) Development of testbeds and virtual service platforms

The last paper deals with the interesting phenomena on social network that has been in the focus of research during the past two decades or so. The authors, *G. Kocsis and I. Varga,* in their paper "Investigation of spreading phenomena on social networks", studied information spreading on different network topologies. Based on a novel complex network generating method several test cases were created for social simulations, focusing mainly on the case of declining social networks.

Guest Editors:



GYULA SALLAI received MSc degree from the Budapest University of Technology and Economics (BME), PhD and DSc degrees from the Hungarian Academy of Sciences (MTA), all in telecommunications. He was senior researcher in telecommunication network planning, then research director, strategic director, later deputy CEO with the Hungarian Telecom Company; then international vice president, after that executive vice president for the ICT regulation with the Communication Authority of Hungary. From 2002 to 2010 he was the head of the Department of Telecommunications and Media Informatics of the BME, and from 2004 to 2008 the vice-rector of the BME as well. From 2005 to 2011 he was also the chairman of the Telecommunication Committee of the MTA and the president of the Hungarian Scientific Association for Infocommunications (HTE).
Recently he is full-professor at the BME, Scientific Director of Future Internet Research Coordination Centre, member of the FIRST Project Council and honorary president of the HTE. His main fields of interest are the ICT trends, strategic, management and regulatory issues, Future Internet engineering.

WOLFGANG SCHREINER is since 2004 associate professor of the Research Institute for Symbolic Computation (RISC) at the Johannes Kepler University Linz, Austria. His research areas are formal methods and parallel and distributed computing; he has directed in these areas various national and international research and development projects, participated in the program committees of 90 conferences, served as evaluator for various European projects, and is member of the editorial board of the Journal of Universal Computer Science.
Prof. Schreiner has (co-)authored 13 book chapters, 9 journal publications, 46 other refereed publications, 12 non-refereed publications and 70 technical reports.

JÁNOS SZTRIK is a Full Professor at the Faculty of Informatics and Head of Department of Informatics Systems and Networks, University of Debrecen, Debrecen, Hungary. He received the M.Sc. degree in 1978, the Ph.D in 1980 both in probability theory and mathematical statistics from the University of Debrecen. He obtained the Candidate of Mathematical Sciences degree in probability theory and mathematical statistics in 1989 from the Kiev State University, Kiev, USSR, habilitation from University of Debrecen in 1999, Doctor of the Hungarian Academy of Sciences, Budapest, 2002.
His research interests are in the field of production systems modeling and analysis, queueing theory, reliability theory, and computer science.

Towards the Transport Protocols of Future Internet

Zoltán Móczár, Sándor Molnár

***Abstract*—End-to-end congestion control performed by the Transmission Control Protocol (TCP) is the main data transfer mechanism of today's Internet providing reliable communication between hosts. Since the deployment of TCP the Internet has gone through a significant change due to the evolving network technologies and the diversity of applications. This process has led to a heterogeneous environment with complex traffic characteristics raising the demand for working out different TCP versions to achieve better performance in various network conditions. In addition to the traditional congestion control scheme several alternative solutions have also been proposed for reliable transport. However, the current practice of continuous modification and refinement of TCP for specific network environments does not seem to be a viable option, hence there is an increasing need for a more efficient and flexible transport protocol. In this paper we present a survey of the major data transfer mechanisms developed in the last decades, and advocate a possible direction for future research.**

***Index Terms*—transport protocols, congestion control, fountain coding, future Internet.**

I. Introduction

THE success of the Internet partially stems from the algorithms implemented in the *Transmission Control Protocol (TCP)*. This transport protocol has guaranteed the reliable transfer between end hosts and the stable operation of networks for several decades. However, network environments, applications and user behavior have changed considerably during this long period making TCP suboptimal under different conditions. As a result, a huge number of versions and enhancements of TCP have been proposed for emerging environments mainly focusing on its congestion control scheme and the related mechanisms [1].

Transport layer protocols play a significant role in the efficient and fair utilization of available network resources, and also have a great impact on the quality of user experience. Due to the importance of this research topic, thousands of researchers and developers worldwide are working on more and more efficient transport solutions. The research, development and standardization processes are managed by two large open international communities, the Internet Research Task Force (IRTF) and the Internet Engineering Task Force (IETF), which are organized into different research and working groups, respectively. The main groups focusing on the area of data transport are the *Internet Congestion Control Research Group (ICCRG)* at IRTF and the working groups of the *Transport and Services Area (TSV)* at IETF. The key goal of ICCRG is to move towards consensus on which technologies can be considered as viable long-term solutions for the Internet congestion control architecture, and to identify the trade-off between potential benefits and costs. As opposed to ICCRG, the members of TSV work on mechanisms related to end-to-end data transfer to support various Internet applications and services that exchange potentially large volumes of traffic at high bandwidths.

This paper presents the evolution of transport protocols since the early days of Internet introducing the main pitfalls the researchers faced with during the years. Furthermore, we shed light on a promising approach, which may be able to satisfy the diverse requirements of future networks.

II. Overview of Transport Protocols

In current Internet the Transmission Control Protocol carries the vast majority of network traffic. The history of TCP dates back to 1981 when the official protocol specification was published by the IETF in RFC 793 [2]. Over the past three decades a significant research effort has been devoted to TCP in order to meet the requirements of the continuously evolving communication networks. This process has resulted in countless TCP versions aimed to provide high performance in various environments [1]. Although, TCP determined the mainstream of the research on transport protocols, in the last years many alternative proposals have also been published to serve as the basis of reliable data communication. In this section we give an overview of the most widely known protocols including the different types of TCP and other proposals, as well.

A. Transmission Control Protocol

TCP is a connection-oriented transport protocol that provides reliable data transfer in end-to-end communication. It means that lost packets are retransmitted, and therefore, each sent packet will eventually be delivered to the destination. One of the most important features of TCP is its *congestion control* mechanism, which is used to avoid congestion collapse [20] by determining the proper sending rate and to achieve high performance. To this end, TCP maintains a congestion window (*cwnd*) that controls the number of outstanding unacknowledged packets in the network. An important aspect in the context of congestion control protocols is how they can share the available bandwidth among competing flows, also known as *fairness* property. Fairness can be interpreted between the same and different TCP versions (intra- and inter-protocol), as well as on various time scales (transient and steady-state) [21].

TCP variants can be classified based on the type of congestion indication and the target environment as shown in Table I. Most congestion control methods use packet loss information to detect congestion also called as *loss-based* TCPs. In case

Manuscript received June 16, 2014, revised August 31, 2014.

Z. Móczár and S. Molnár are with the Department of Telecommunications and Media Informatics, Budapest University of Technology and Economics, Hungary. Authors are also with the Inter-University Centre for Telecommunications and Informatics, Kassai út 26., 4028 Debrecen, Hungary.

TABLE I
THE EVOLUTION OF TCP VARIANTS

Version	Congestion indicator		Target environment			Implementation					New features	Published
	Loss	*Delay*	*Wired*	*Wireless*	*High-speed*	*BSD*	*Linux*	*Win*	*Mac*	*ns-2*		
TCP Tahoe [2]	×		×			×	×				slow-start, congestion avoidance and fast retransmit	1988
TCP Reno [3]	×		×			×	×	×			fast recovery to mitigate the impact of packet losses	1990
TCP Vegas [4]		×	×				×				bottleneck buffer utilization as a congestion feedback	1995
TCP NewReno [5]	×		×			×	×		×		fast recovery, resistance to multiple losses	1999
Freeze-TCP [6]	×			×			×			×	considering radio signal quality in mobile networks	2000
TCP-Peach [7]	×			×			×				sudden start and rapid recovery for satellite networks	2001
TCP Westwood [8]	×			×						×	estimation of the available bandwidth	2001
ATCP [9]	×			×		×				×	detection of route changes in ad-hoc networks	2001
TCP Nice [10]		×	×				×				delay threshold as a secondary congestion indicator	2002
Scalable TCP [11]	×		×		×		×				MIMD congestion avoidance algorithm	2003
TCP-LP [12]	×		×				×			×	early congestion detection to react sooner than TCP	2003
HighSpeed TCP [13]	×		×		×		×			×	AIMD mechanism as the function of *cwnd*	2003
FAST TCP [14]		×	×		×		×			×	updating *cwnd* based on different equations	2003
BIC TCP [15]	×		×		×		×			×	binary search to find the proper *cwnd*	2004
Compound TCP [16]	×	×	×		×		×	×			calculation of *cwnd* using loss and delay components	2005
TCP-Illinois [17]	×	×	×		×		×				AIMD as the function of the queueing delay	2006
TCP Cubic [18]	×		×		×		×			×	control of *cwnd* by applying a cubic function	2008
LEDBAT [19]		×	×				×		×	×	congestion control for low-priority traffic	2012

of these algorithms packet loss is interpreted as the sign of a full network buffer, from which the last incoming packet was dropped, hence transmission rate should be reduced. Another group of congestion control mechanisms react to the increase observed in the round-trip time (RTT) of packets due to the building up of queues. This approach, often referred to as *delay-based* TCP, has the ability to detect congestion early rather than merely waiting until the network gets overutilized and packets are lost. In addition, *hybrid* solutions have also been proposed, which combine the beneficial properties of loss-based and delay-based algorithms.

During the years, the fast development of networks motivated researchers to optimize TCP for certain environments and purposes by modifying the traditional congestion control mechanism. Since standard TCP versions (like TCP Tahoe and Reno) failed to obtain full utilization in networks with high-bandwidth links, new algorithms have been introduced to improve the performance in such conditions. The most relevant *high-speed* TCP versions include Scalable TCP [11], HighSpeed TCP [13], FAST TCP [14] and TCP Cubic [18]. On the other hand, as TCP was primarily designed for wired networks, emerging wireless communication induced a considerable research work to develop TCP versions, which can provide better performance in different kinds of *wireless* networks [22]. The performance issues experienced in such environments stem from the unique characteristics of wireless links and the packet loss model used by TCP. The problems manifest in many applications as degradation of throughput, inefficiency in network resource utilization and excessive interruption of data transmissions. Modification of standard TCP for wireless communication has been an active research area in recent years, and many schemes have been proposed for various environments such as cellular (e.g. Freeze-TCP [6]), satellite (e.g. TCP-Peach [7]) and ad-hoc networks (e.g. ATCP [9]). In real networks a traffic mix consists of hundreds or thousands of flows generated by diverse applications and services. In order to treat low-priority traffic (e.g. background transfers like automatic software updates and data backups) differently from high-priority traffic, *low-priority* congestion control methods have been introduced. These protocols, such as TCP Nice [10], TCP-LP [12] and LEDBAT [19], respond to congestion earlier than standard TCP yielding bandwidth to competing TCP flows with higher priority.

1) Loss-based TCP Versions: One of the earliest approaches to handle congestion was introduced in TCP Tahoe [20], which was also served as the first practical implementation of these control schemes in the BSD operating system. The proposal is based on the original TCP specification [2] and introduces new algorithms called slow-start, congestion avoidance (AIMD: additive increase multiplicative decrease [23]) and fast retransmit, as well as an improved method for round-trip time estimation. These mechanisms allow the sender to detect available network resources and adjust the transmission rate accordingly. However, reducing the congestion window to one packet when a packet loss occurs, as done by Tahoe, is a very aggressive solution.

TCP Reno [3] tackles this problem by applying a novel method referred to as fast recovery algorithm. In case of Reno a lost packet is detected and retransmitted if triple duplicate acknowledgements are received or a timeout event occurs at the sender. This mechanism makes TCP Reno effective to recover from a single packet loss, but it still suffers from performance degradation when multiple packets are dropped from a window of data. To overcome this limitation a selective acknowledgement (SACK) option has been proposed in [24].

TCP NewReno [5] is a variant of TCP Reno intended to improve its performance when a burst of packets is lost. To this end, NewReno modifies Reno's fast recovery algorithm making it possible to recover without a retransmission timeout by resending one packet per each round-trip time until all of the lost packets from the window have been retransmitted.

TCP Cubic [18], being an enhanced version of its predecessor, BIC TCP [15], is one of the most widely used TCP versions today since it serves as the default congestion control algorithm of Linux operating systems. BIC TCP was originally designed to solve the well-known RTT unfairness problem by combining two schemes called additive increase and binary search. TCP Cubic simplifies the window control of BIC and it applies a cubic function in terms of the elapsed time from the last loss event, which provides good stability and scalability. Furthermore, it keeps the window growth rate independent of RTT making the protocol TCP-friendly under both short and long RTT paths.

Beside the congestion control algorithms described above, many other solutions have been worked out to improve the performance of standard TCP. One of the main issues is that it takes a long time to make a full recovery from packet loss for high-bandwidth, long-distance connections, because the congestion window builds up very slowly. In order to cope with this limitation HighSpeed TCP (HSTCP) [13] was proposed, which can achieve better performance on high-capacity links by modifying the congestion control algorithm for use with large congestion windows. Scalable TCP (STCP) [11] applies a multiplicative increase and multiplicative decrease (MIMD) algorithm to obtain performance improvement in high-speed networks and it can also guarantee the scalability of the protocol. TCP Westwood [8] is a sender-side modification of the congestion control mechanism that improves the performance of TCP Reno both in wired and wireless networks. The main problem is that TCP Reno equally reacts to random and congestion losses, thus cannot distinguish between them. In fact, TCP Westwood shows moderate sensitivity to random errors, therefore the improvement is most significant in wireless networks with lossy links. MultiPath TCP (MPTCP) [25] is a recent approach for enabling the simultaneous use of multiple IP addresses or interfaces by a modification of TCP that presents a regular TCP interface to applications, while in fact spreading data across several subflows.

2) Delay-based TCP Versions: TCP Vegas [4], as a pioneer of delay-based TCPs, measures the difference (δ) between the expected and actual throughput based on round-trip delays. If δ is less than a lower threshold denoted by α, Vegas assumes that the path is not congested and increases the sending rate. If δ is larger than an upper threshold denoted by β, it is regarded as the strong indication of congestion, hence Vegas reduces the transmission rate. The expected throughput is calculated by dividing the current congestion window by the minimum RTT.

FAST TCP [14] is a congestion avoidance algorithm especially targeted for long-distance, high-latency links. FAST determines the current congestion window size based on both round-trip delays and packet losses over a path. The algorithm estimates the queueing delay of the path using RTTs and if the delay falls below a threshold, it increases the window aggressively. If it gets closer to the threshold, the algorithm slowly reduces the increasing rate.

3) Hybrid Solutions: Compound TCP (CTCP) [16], implemented in several Microsoft Windows operating systems, is a synergy of delay-based and loss-based approaches extending the standard TCP Reno congestion avoidance algorithm by a scalable, delay-based component. CTCP exploits the information about both packet loss and delay to control the transmission rate. The delay-based component can rapidly increase the sending rate when the network path is underutilized, but ease if the bottleneck queue becomes full. This mechanism provides good scalability in terms of bandwidth, and a reasonably fair behavior.

TCP-Illinois [17] uses packet loss information to determine whether the congestion window size should be increased or decreased, and measures the queueing delay to determine the amount of increment or decrement. This hybrid solution makes it possible to obtain high throughput and fair resource allocation, while being compatible with standard TCP.

B. Other Proposals

Beyond the Transmission Control Protocol several approaches have also been suggested for reliable data transport in communication networks. Some of these protocols are partially based on the concept of TCP, or use similar mechanisms.

Internet traffic has a complex characteristics investigated in many papers in the last decade. Recent studies showed that most flows are small carrying *only several kilobytes* of data and short lasting *less than a few seconds* [26]. Rate Control Protocol (RCP) [27] is a congestion control algorithm designed to significantly speed up the download of short-lived flows generated by typical applications. For example, a mid-size flow contains 1000 packets and TCP makes them last nearly 10 times longer than it would be necessary. RCP enables flows to finish close to the minimum possible, leading to a notable improvement for web users and distributed file systems.

eXplicit Control Protocol (XCP) [28] uses direct congestion notification instead of the indirect congestion indicators such as packet loss or delay. XCP delivers the highest possible application performance over a broad range of network infrastructures including high-speed and high-delay links where

TCP performs poorly. It also introduces a novel way for separating the efficiency and fairness policies of congestion control, enabling routers to quickly make use of available bandwidth while conservatively managing the allocation of the available bandwidth to competing flows. XCP carries the per-flow congestion state in the packet header allowing the sender to request a desired throughput for its transmission, and XCP-capable routers inform the senders about the degree of the congestion at the bottleneck.

Stream Control Transmission Protocol (SCTP) [29] is a reliable transport protocol that provides stable, ordered delivery of data between two endpoints by using congestion control like TCP and also preserves data message boundaries like UDP. However, unlike TCP and UDP, SCTP offers additional services such as multi-homing, multi-streaming, security and authentication.

Quick UDP Internet Connections (QUIC) [30] is a recent approach for data transfer announced by Google in 2013. QUIC is currently under development and has been integrated in Google Chrome for evaluation purposes. The new protocol supports a set multiplexed connections over UDP, and was designed to provide security protection equivalent to TLS/SSL, along with reduced connection and transport latency. It also implements and applies a bandwidth estimation algorithm in each direction in order to avoid network congestion. QUIC's main goal is to optimize the performance of connection-oriented web applications and services by reducing the connectivity overhead to zero RTT.

III. Digital Fountain Based Communication

Over the years, the issues of TCP motivated researchers to find alternative ways for data transfer beside the traditional congestion control based approach. In 2007, GENI (Global Environment for Network Innovations) [31] published a research plan, in which they recommend the omission of the congestion control mechanism and suggest to use efficient erasure coding to cope with network congestion. Since then, the questions related to this idea have been investigated only in a few papers. Raghavan and Snoeren argue in [32] that it may not be necessary to keep the network uncongested to achieve good performance and fairness. They introduce a concept called decongestion control and presume that a protocol relying upon greedy, high-speed transmission has the potential to perform better than TCP. Bonald et al. studied the consequences of operating a network without congestion control [33], and concluded that it does not inevitably lead to congestion collapse as believed earlier.

In this section we review the related work carried out in the field of erasure code driven data transport, then introduce and describe a possible data transfer paradigm for Future Internet, which applies a fundamentally different principle compared to that of TCP. According to our concept presented in [34], congestion control can be completely omitted from the transport layer if efficient fountain coding is used as a replacement. We propose a novel network architecture where each host communicates by a digital fountain based transport protocol, while fair schedulers deployed in routers are responsible for fair bandwidth sharing among competing traffic flows. We show that this new paradigm has many benefits and also introduce some possible future application areas.

A. Error-Correcting Codes in Data Transport

In recent times, many research works have focused on the application of erasure codes in data transport. A theoretical fountain based protocol (FBP) was investigated in [35]. The authors showed that a Nash equilibrium can be reached in a network with FBP-based hosts resulting in a performance similar to the case when each host uses TCP. Kumar et al. proposed a transport protocol for wireless networks using fountain codes [36] and analyzed its performance by a Markovian stochastic model. They demonstrated through packet-level simulations that their protocol may perform better or worse than TCP depending on the redundancy parameter, the number of nodes in a WLAN cell and the wireless channel conditions. The authors of [37] designed a new TCP version on the basis of rateless erasure codes to enhance its operation in lossy environments. According to their results, such modification of TCP has proven to be effective in case of high packet loss rate. Y. Cui and his colleagues proposed FMTCP (Fountain code-based Multipath TCP) in [38], which exploits the advantage of the fountain coding scheme to avoid the performance degradation caused by frequent retransmissions applied in MPTCP. The authors introduced an algorithm to flexibly allocate encoded symbols to different subflows based on the expected packet arrival time over different paths.

B. A Novel Data Transfer Paradigm

1) Architecture and Protocol: The key component of our recent proposal is a new transport mechanism called *Digital Fountain based Communication Protocol (DFCP)* [34], which uses digital fountain codes to recover lost packets instead of traditional retransmissions. Fountain codes [39] are rateless erasure codes with the property that a potentially limitless sequence of encoded symbols can be generated from a given set of source symbols, such that the original source symbols can ideally be recovered from any subset of the encoded symbols of size equal to or only slightly larger than the number of source symbols. Raptor codes [40] are the most efficient ones in the family of fountain codes as they offer linear time encoding and decoding complexity, hence the latest version of DFCP implements this scheme.

Our vision of the future network architecture relying on DFCP is shown in Figure 1. There are multiple senders exchanging information with the corresponding receivers, and each host is allowed to send at its maximum transmission rate. Senders generate a potentially infinite stream of encoded symbols from the original message of size k by adding a redundancy of $\epsilon > 0$. When any subset of size $\lceil(1+\varepsilon)k\rceil$ encoded symbols arrive to the receiver, high probability decoding becomes possible, and fountain coding ensures that each received packet at the destination increases the probability of successful decoding. This approach makes it possible to leave the network congested resulting in fully utilized links. To ensure equal bandwidth sharing among competing flows we

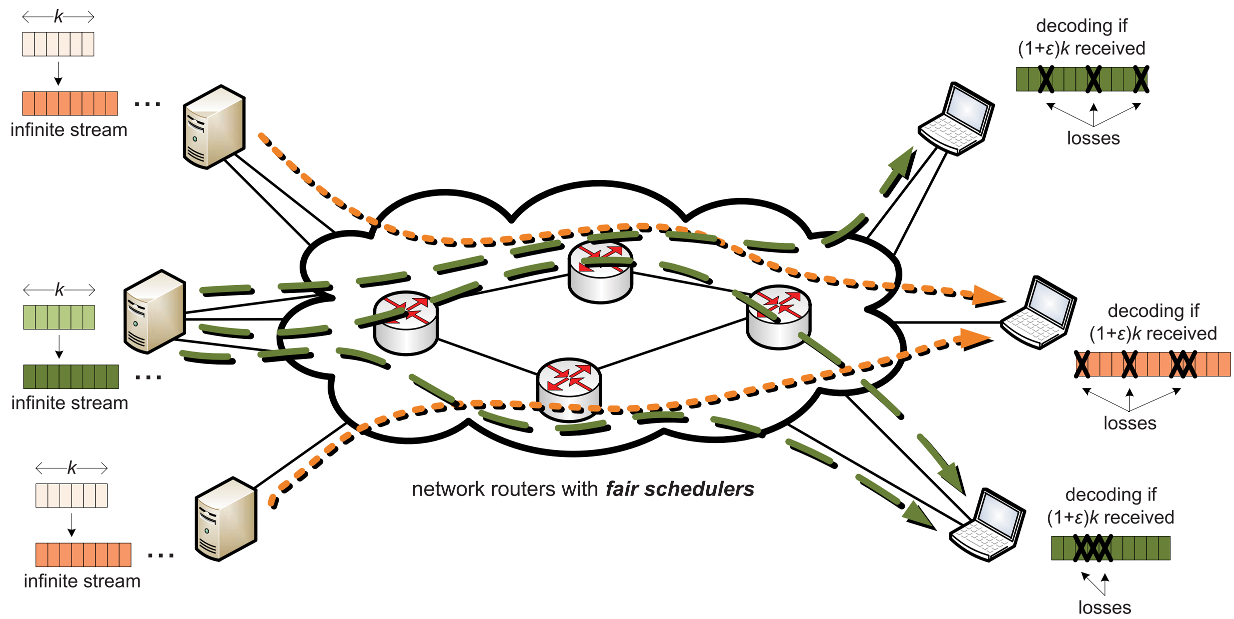


Fig. 1. The network architecture built upon DFCP

suggest the use of *fair schedulers* in the network nodes. Several implementations approximating the ideal fair scheduling, such as Deficit Round Robin (DRR) [41], are available and can be configured easily in routers. The feasibility of this solution is supported by the scalability of per-flow fair queueing [42]. We note that maximal rate sending does not mean the full utilization of the transmission capacity available at the sender side in all cases since it would lead to the so-called dead packet problem. This phenomenon happens when a source transmits at a higher speed than its fair share of the bottleneck link needlessly wasting the bandwidth on the whole path from concurrent flows. However, there are many possible ways to avoid this undesirable behavior, for example, by carrying a feedback signal in the acknowledgements about the degree of congestion at the bottleneck (see, e.g. [43]), or by estimating the available bandwidth on the path [44].

2) Potential Benefits: The operation of DFCP has been validated on three independent testing platforms [45] including a laboratory testbed, the Emulab network emulation environment [46] and the ns-2 network simulator [47]. In addition, a comprehensive performance evaluation study comparing DFCP to the most relevant TCP versions was also presented in [45]. The main purpose of our investigation was to understand the nature of digital fountain based communication and to reveal its features. Measurements were performed both on simple topologies and in multi-bottleneck networks. The results pointed out that the new paradigm has many potential benefits. One of the main fundamental properties of DFCP is its high resistance to packet losses while it also shows a moderate sensitivity to delay. The latter feature makes it possible to eliminate the well-known RTT unfairness problem of TCP as DFCP can provide fair bandwidth sharing among competing flows independently of their RTTs. Another excellent improvement compared to congestion control based data transmission is that the concept of DFCP avoids the issues introduced by TCP's slow-start algorithm, and hence enables both short-lived and long-lived flows to complete faster [48]. Furthermore, our protocol is able to obtain maximum performance even with small buffers, which could make it attractive for all-optical networks. Finally, digital fountain based transport guarantees good scalability and stability as well, both in terms of performance and fairness for increasing number of flows and link capacity.

3) Possible Applications: The proposed network architecture can provide an efficient framework for numerous applications. For example, our scheme supports *multipath communication*, which makes it possible to achieve better network resiliency and load balancing. Since DFCP is insensitive to packet loss and delay, it is a good candidate for *wireless networks*, as well. Moreover, the new data transfer paradigm is closely aligned to the high utilization requirement of *data centers* and the concept of *all-optical networking* where only small buffers can be realized.

IV. SUMMARY

In this paper we have reviewed the evolution of transport protocols since the introduction of TCP till today, and discussed the main principles of different congestion control schemes optimized for various target network environments. We have also presented several alternative data transport mechanisms developed in the past decades including recent approaches. We claim that the main lesson learned from this long research history of transport protocols is a need for a paradigm shift. We advocate a promising data transfer method for Future Internet based on digital fountain codes, which can provide more efficient and flexible operation than TCP and may open the way to a broad range of application areas.

Acknowledgement

This work was partially supported by the European Union and the European Social Fund through project FIRST (grant no. TÁMOP-4.2.2.C-11/1/KONV-2012-0001) organized by ETIK Debrecen. The research also received support from the High Speed Networks Laboratory (HSNLab) at Budapest University of Technology and Economics.

References

[1] A. Afanasyev, N. Tilley, P. Reiher, L. Kleinrock, "Host-to-Host Congestion Control for TCP", *IEEE Communications Surveys and Tutorials*, vol. 12, no. 3, pp. 304–342, 2010.
[2] J. Postel, "Transmission Control Protocol", *RFC 793, IETF*, 1981.
[3] K. Fall, S. Floyd, "Simulation-based Comparisons of Tahoe, Reno, and SACK TCP", *ACM SIGCOMM Computer Communication Review*, vol. 26, no. 3, pp. 5–21, 1996.
[4] L. S. Brakmo, L. L. Peterson, "TCP Vegas: End-to-End Congestion Avoidance on a Global Internet", *IEEE Journal on Selected Areas in Communications*, vol. 13, no. 8, pp. 1465–1480, 1995.
[5] S. Floyd, T. Henderson, A. Gurtov, "The NewReno Modification to TCP's Fast Recovery Algorithm", *RFC 3782, IETF*, 2004.
[6] T. Goff, J. Moronski, D. Phatak, "Freeze-TCP: A True End-to-End Enhancement Mechanism for Mobile Environments", *Proceedings of the 19th IEEE International Conference on Computer Communications*, pp. 1537–1545, Tel-Aviv, Israel, 2000.
[7] I. F. Akyildiz, G. Morabito, S. Palazzo, "TCP-Peach: A New Congestion Control Scheme for Satellite IP Networks", *IEEE/ACM Transactions on Networking*, vol. 9, no. 3, pp. 307–321, 2001.
[8] S. Mascolo, C. Casetti, M. Gerla, M. Y. Sanadidi, R. Wang, "TCP Westwood: Bandwidth Estimation for Enhanced Transport over Wireless Links", *Proceedings of the 7th ACM MOBICOM Conference*, pp. 287–297, 2001.
[9] J. Liu, S. Singh, "ATCP: TCP for Mobile Ad Hoc Networks", *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 7, pp. 1300–1315, 2001.
[10] A. Venkataramani, R. Kokku, M. Dahlin, "TCP Nice: A Mechanism for Background Transfers", *ACM SIGOPS Operating Systems Review*, vol. 36, pp. 329–344, 2002.
[11] T. Kelly, "Scalable TCP: Improving Performance in High-Speed Wide Area Networks", *ACM Computer Communication Review*, vol. 33, no. 2, pp. 83–91, 2003.
[12] A. Kuzmanovic, E. W. Knightly, "TCP-LP: A Distributed Algorithm for Low Priority Data Transfer", *Proceedings of the 22th IEEE International Conference on Computer Communications*, pp. 1691–1701, San Francisco, CA, USA, 2003.
[13] S. Floyd, "HighSpeed TCP for Large Congestion Windows", *RFC 3649, IETF*, 2003.
[14] C. Jin, D. X. Wei, S. H. Low, "FAST TCP: Motivation, Architecture, Algorithms, Performance", *IEEE/ACM Transactions on Networking*, vol. 14, no. 6, pp. 1246–1259, 2006.
[15] L. Xu, K. Harfoush, I. Rhee, "Binary Increase Congestion Control (BIC) for Fast Long-Distance Networks", *Proceedings of the 23rd IEEE International Conference on Computer Communications*, vol. 4, pp. 2514–2524, Hong Kong, China, 2004.
[16] K. Tan, J. Song, Q. Zhang, M. Sridharan, "A Compound TCP Approach for High-Speed and Long Distance Networks", *Technical Report, Microsoft Research*, pp. 1–12, 2005.
[17] S. Liu, T. Basar, R. Srikant, "TCP-Illinois: A Loss and Delay-Based Congestion Control Algorithm for High-Speed Networks", *Proceedings of the 1st International Conference on Performance Evaluation Methodologies and Tools*, pp. 1–13, Pisa, Italy, 2006.
[18] I. Rhee, L. Xu, "CUBIC: A New TCP-Friendly High-Speed TCP Variant", *Proceedings of the 3rd International Workshop on Protocols for Fast Long-Distance Networks*, pp. 1–6, Lyon, France, 2005.
[19] S. Shalunov, G. Hazel, J. Iyengar, M. Kuehlewind, "Low Extra Delay Background Transport (LEDBAT)", *RFC 6817, IETF*, 2012.
[20] V. Jacobson, "Congestion Avoidance and Control", *Proceedings of the 1988 ACM SIGCOMM Symposium*, pp. 314–329, Stanford, CA, USA, 1988.
[21] S. Molnár, B. Sonkoly, T. A. Trinh, "A Comprehensive TCP Fairness Analysis in High Speed Networks", *Computer Communications, Elsevier*, vol. 32, no. 13–14, pp. 1460–1484, 2009.
[22] Y. Tian, K. Xu, N. Ansari, "TCP in Wireless Environments: Problems and Solutions", *IEEE Communications Magazine*, vol. 43, no. 3, pp. 27–32, 2005.
[23] D.-M. Chiu, R. Jain, "Analysis of the Increase and Decrease Algorithms for Congestion Avoidance in Computer Networks", *Journal of Computer Networks and ISDN Systems*, vol. 17, no. 1, pp. 1–14, 1989.
[24] M. Mathis, J. Mahdavi, S. Floyd, A. Romanow, "TCP Selective Acknowledgment Options", *RFC 2018, IETF*, 1996.
[25] A. Ford, C. Raiciu, M. Handley, S. Barre, J. Iyengar, "Architectural Guidelines for Multipath TCP Development", *RFC 6182, IETF*, 2011.
[26] S. Molnár, Z. Móczár, "Three-dimensional Characterization of Internet Flows", *Proceedings of the 2011 IEEE International Conference on Communications*, pp. 1–6, Kyoto, Japan, 2011.
[27] N. Dukkipati, N. McKeown, "Why Flow-Completion Time is the Right Metric for Congestion Control", *ACM SIGCOMM Computer Communication Review*, vol. 36, no. 1, pp. 59–62, 2006.
[28] D. Katabi, M. Handley, C. Rohrs, "Congestion Control for High Bandwidth-Delay Product Networks", *ACM SIGCOMM Computer Communication Review*, vol. 32, no. 4, pp. 89–102, 2002.
[29] R. Stewart, Q. Xie, K. Morneault, C. Sharp, H. Schwarzbauer, T. Taylor, I. Rytina, M. Kalla, L. Zhang, V. Paxson, "Stream Control Transmission Protocol", *RFC 2960, IETF*, 2000.
[30] J. Brodkin, "Google Making the Web Faster with a Protocol that Reduces Round-Trips", http://arstechnica.com/information-technology/2013/06/google-making-the-web-faster-with-protocol-that-reduces-round-trips/, 2013.
[31] D. Clark, S. Shenker, A. Falk, "GENI Research Plan (Version 4.5)", April 23, 2007.
[32] B. Raghavan, A. C. Snoeren, "Decongestion Control", *Proceedings of the 5th ACM Workshop on Hot Topics in Networks*, pp. 61–66, Irvine, CA, USA, 2006.
[33] T. Bonald, M. Feuillet, A. Proutiere, "Is the 'Law of the Jungle' Sustainable for the Internet?", *Proceedings of the 28th IEEE Conference on Computer Communications*, pp. 28–36, Rio de Janeiro, Brazil, 2009.
[34] S. Molnár, Z. Móczár, A. Temesváry, B. Sonkoly, Sz. Solymos, T. Csicsics, "Data Transfer Paradigms for Future Networks: Fountain Coding or Congestion Control?", *Proceedings of the IFIP Networking 2013 Conference*, pp. 1–9, New York, NY, USA, 2013.
[35] L. López, A. Fernández, V. Cholvi, "A Game Theoretic Comparison of TCP and Digital Fountain Based Protocols", *Computer Networks, Elsevier*, vol. 51, no. 12, pp. 3413–3426, 2007.
[36] D. Kumar, T. Chahed, E. Altman, "Analysis of a Fountain Codes Based Transport in an 802.11 WLAN Cell", *Proceedings of the 21st International Teletraffic Congress*, pp. 1–8, Paris, France, 2009.
[37] A. Botos, Z. A. Polgar, V. Bota, "Analysis of a Transport Protocol Based on Rateless Erasure Correcting Codes", *Proceedings of the 2010 IEEE International Conference on Intelligent Computer Communication and Processing*, vol. 1, pp. 465–471, Cluj-Napoca, Romania, 2010.
[38] Y. Cui, X. Wang, H. Wang, G. Pan, Y. Wang, "FMTCP: A Fountain Code-Based Multipath Transmission Control Protocol", *Proceedings of the 32nd IEEE International Conference on Distributed Computing Systems*, pp. 366–375, Macau, China, 2012.
[39] D. J. C. MacKay, "Fountain Codes", *IEE Proceedings – Communications*, vol. 152, no. 6, pp. 1062–1068, 2005.
[40] A. Shokrollahi, "Raptor Codes", *IEEE Transactions on Information Theory*, vol. 52, no. 6, pp. 2551–2567, 2006.
[41] M. Shreedhar, G. Varghese, "Efficient Fair Queuing Using Deficit Round-Robin", *IEEE/ACM Transactions on Networking*, vol. 4, no. 3, pp. 375–385, 1996.
[42] A. Kortebi, L. Muscariello, S. Oueslati, J. Roberts, "On the Scalability of Fair Queuing", *Proceedings of the 3rd ACM Workshop on Hot Topics in Networks*, pp. 1–6, San Diego, CA, USA, 2004.
[43] B. Briscoe, R. Woundy, A. Cooper, "Congestion Exposure (ConEx) Concepts and Use Cases", *RFC 6789, IETF*, 2012.
[44] C. D. Guerrero, M. A. Labrador, "On the Applicability of Available Bandwidth Estimation Techniques and Tools", *Computer Communications, Elsevier*, vol. 33, no. 1, pp. 11–22, 2010.
[45] Z. Móczár, S. Molnár, B. Sonkoly, "Multi-Platform Performance Evaluation of Digital Fountain Based Transport", *IEEE Science and Information Conference 2014*, London, UK, 2014.
[46] Emulab Network Emulation Testbed, http://www.emulab.net/
[47] ns-2 Network Simulator, http://www.isi.edu/nsnam/ns/
[48] S. Molnár, Z. Móczár, B. Sonkoly, "How to Transfer Flows Efficiently via the Internet?", *Proceedings of the 3rd IEEE International Conference on Computing, Networking and Communications*, pp. 462–466, Honolulu, HI, USA, 2014.

Zoltán Móczár received his M.Sc. degree in computer science from the Budapest University of Technology and Economics (BME), Hungary in 2011. Currently he is a Ph.D. student at the High Speed Networks Laboratory, Department of Telecommunications and Media Informatics, BME. His research interests include traffic engineering, performance evaluation of high-speed transport protocols and Future Internet design. He has participated in several academic and applied research projects.

Sándor Molnár is an Associate Professor at the Department of Telecommunications and Media Informatics, Budapest University of Technology and Economics (BME). In electrical engineering he received his M.Sc. in 1991 and Ph.D. in 1996 from BME. Dr. Molnár is serving in the editorial board of international journals and is working in the TPC of IEEE conferences. He has been participating in various EU projects and was the BME project leader of the Gold Award winner 2009 CELTIC project titled "Traffic Measurements and Models in Multi-Service Networks (TRAMMS)". His main interests include teletraffic analysis and performance evaluation of modern communication networks.

CALL FOR PAPERS

Special Issue on Applied Cryptography and Security

This special issue will focus on the area of applied cryptography, bringing up selected papers from *Santa Crypt 2014*. Santa's Crypto Get-Together (SantaCrypt) started in 2001 as the first annual Czech and Slovak workshop aiming to facilitate closer cooperation of professionals working in the field of applied cryptography and related areas of security. This get-together of experts is organised in order to foster exchange of information and ideas on past, ongoing, and also future projects.

The special issue of the Infocommunications Journal will focus on the following areas:

- Applied Cryptography
- Practical Cryptanalysis
- Cryptographic Protocols
- Security Mechanisms Deploying Cryptography

Submissions to Santacrypt 2014 should be mailed to matyas@fi.muni.cz, and clearly marked with Subject "SantaCrypt 2014". The manuscripts should follow the MKB format (http://mkb.tns.cz/cfp.htm.en) with the maximum length of 10 pages (corresponds to 8 pages in the IEEE format). The final deadline for the submissions is 30th September 2014. Submissions will be evaluated by the program committee and authors will be informed about the evaluation results by 30th October. Camera-ready versions for the workshop proceedings have to be delivered by 11th November. The workshop takes place in Prague, Czechia, on November 27-28, 2014.

No more than 4 papers from the workshop shall be then selected for the special issue of the Infocommunications Journal, and authors of these papers will have the opportunity to revise their papers (including typesetting in the IEEE format) after the workshop – final versions for the special issue will be due January 11, 2015.

Guest Editors:

VÁCLAV (VASHEK) MATYÁS is a Professor at the Masaryk University, Brno, CZ, and Vice-Dean for Foreign Affairs and External Relations, Faculty of Informatics. His research interests relate to applied cryptography and security, where he published over 150 peer-reviewed papers and articles, and co-authored several books. He was a Fulbright-Masaryk Visiting Scholar with Harvard University, Center for Research on Computation and Society in 2011-12, and previously he worked also with Microsoft Research Cambridge, University College Dublin, Ubilab at UBS AG, and was a Royal Society Postdoctoral Fellow with the Cambridge University Computer Lab. Vashek edited the Computer and Communications Security Reviews, and worked on the development of Common Criteria and with ISO/IEC JTC1 SC27. He received his PhD degree from Masaryk University, Brno and can be contacted at *matyas@fi.muni.cz*.

ZDENEK ŘÍHA is an Assistant Professor at the Masaryk University, Faculty of Informatics, in Brno, Czech Republic. He received his PhD degree from the Faculty of Informatics, Masaryk University. In 1999 he spent 6 months on an internship at Ubilab, the research lab of the bank UBS, focusing on security and usability aspects of biometric authentication systems. Between 2005 and 2008 he was seconded as a Detached National Expert to the European Commission's Joint Research Centre in Italy, where he worked on various projects related to privacy protection and electronic passports. He was involved in the ePassport interoperability group known as the Brussels Interoperability Group. Zdeněk has been working with the WG 5 (Identity management and privacy technologies) of ISO/IEC JTC 1/SC 27. Zdeněk's research interests include Android security, smartcard security, PKI, security of biometric systems and machine readable travel documents. Zdeněk can be contacted at *zriha@fi.muni.cz*.

PAVOL ZAJAC is an Assistant Professor at the Slovak University of Technology in Bratislava, Faculty of Electrical Engineering and Information Technology, Slovakia (FEI STU). He received his PhD degree in Applied Mathematics from FEI STU. His main research focus is in the area of mathematical cryptology, including mathematical principles of cipher design, and algebraic cryptanalysis. He also works on various application projects, including efficient implementation of post-quantum cryptosystems, and improving privacy of user data on mobile devices. Pavol can be contacted at *pavol.zajac@stuba.sk*.

Modeling a QoS classified communication in a multiuser Wi-Fi environment

Attila Kuki, Béla Almási, Tamás Bérczes, János Sztrik

Abstract—**The aim of the present paper is to propose a finite source queueing model in order to include the random backoff feature of the wireless communication. Two classes of sources (high priority and low priority traffic) are included. The random backoff feature is implemented by using retrial queues for each traffic class. Supposing exponentially distributed inter-event times the MOSEL software tool is used to develop the special software to calculate the most important steady-state performance characteristics of the system, such as utilizations, mean orbit sizes, mean waiting times, that is mean time spent in the orbit. It is showed how the retrial discipline effects the mean waiting times (compared to the FIFO discipline): not only the values, but also the form of the curve is quite different in the case of packet reordering.**

I. Introduction

The investigation of the users' connection to a Wi-Fi access point is a focused research area today.

For the mathematical analysis, queueing models are widely used to create stochastic models which can be used to calculate the most important performance characteristics of the communication systems (e.g. utilizations, response times) (see e.g. [3], [4], [5], [6], [7], [8], [9]). Using the classic FIFO service discipline in these models are not suitable for modeling the wireless network environment, because the random backoff feature of the wireless access is completely omitted. Although the queueing model technology can be used efficiently to model and study the influence of the priorities on the QoS performance between the traffic classes (see [10]), but it can be hardly used to model the wireless access feature.

In [1] the authors created a queueing model to evaluate a sensor network environment with two quality classes of sources. The Emergency class represents the sources of very important communication (e.g. fire alarm), and the Normal class represents the standard communication (e.g. temperature data). In the model of [1] the Emergency class is served by a FIFO queue. The FIFO discipline absolutely excludes the ability of request reordering. The random backoff feature of the Wi-Fi access can not be precisely described by the FIFO discipline.

In this paper we would like to introduce a new element in the queueing modeling of the QoS performance. The basic idea of our solution appeared in [2], where the authors used two orbits (retrial queues) in an infinite queueing model. The requests staying in the orbit are randomly retrying the transmission, so opening the possibility of describing the random backoff feature of the wireless network access. We establish a finite source queueing model containing two QoS traffic classes. For the service of the requests in each class we use retrial systems, random backoff may occur inside the classes too. In this paper we use finite source queueing system to model small sized network environments, where the number of users may not be considered as infinite.

The main contribution of the present paper is to introduce a finite source queueing system with two orbits, which is more suitable for modeling small environments (e.g. a picocell or femtocell sized radio networks, where the number of sources may not be considered as infinite). We would like to study the system parameters of the multiuser Wi-Fi access environment (queue length, service time etc).

The rest of the paper is organized as follows. Section 2 describes the precise mathematical model when a multi-dimensional continuous time Markov-chain is defined for describing the system's dynamics. Formulas of the most important performance measures are also discussed here. For presentation of numerical results the MOdeling Specification and Evaluation Language (MOSEL see [11]) tool is used. In Section 3 we investigate a concrete environment, and we study the effect of different parameters on the systems performance. Finally, the conclusion closes the paper.

II. System Model

For modeling the effect of the wireless access problem, let's see a queueing model with a single server unit, where the jobs come from two classes of finite sources. These sources represent the incoming packets of the Wi-Fi connected users. The first class of sources corresponds to the high priority sessions (eg. interactive voice or video stream), and the second class of sources refers to the low priority sessions (eg. file transfer or database transaction). The number of sources of the high priority class is denoted by N, and the number of sources of the low priority class is denoted by K. The sources of both classes may send a new service request (ie. a new packet is sent through the communication channel). The distribution of the inter-request times is exponential with parameter λ_1 for the high priority packets and with parameter λ_2 for the low priority class.

Because there are no queues (buffers) for either high priority class or low priority class, there can be at most one request in the service area. So, the server can be engaged with a request from the high priority class, or with a request from the low priority class (or it can be idle).

Manuscript received June 13, 2014, revised August 31, 2014.

The authors are with the Department of Informatics Systems and Networks, Faculty of Informatics, University of Debrecen, Hungary, e-mail: {berczes.tamas,almasi.bela,sztrik.janos,kuki}@inf.unideb.hu

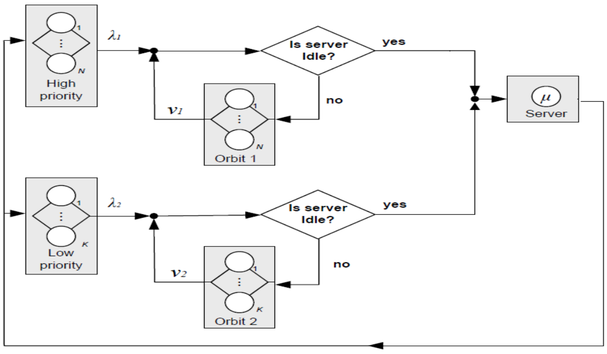


Fig. 1. A retrial queue with components

For a request generated from the high priority sources there are two possibilities. In case of the idle server the request is transmitted directly to the server, where the request will be served with an exponentially distributed service time. The service rate is denoted by μ. When the request is served, it goes back to the high priority source. In case of the busy server, the request is sent to the orbit from the high priority sources. From the orbit the request will retry to be served after an exponentially distributed retrial time, the retrial rate is denoted by ν_1. Based on these random times, the order of high priority packets arriving to the orbit may differ from the order of packets leaving the orbit.

For a request generated from the low priority sources there are these two possibilities, as well. In case of the idle server the request is transmitted directly to the server, where the request will be served with an exponentially distributed service time, with the same service rate μ. When the request is served, it goes back to the low priority source. In case of the busy server, the request is sent to the orbit from the low priority sources. From the orbit the request will retry to be served after an exponentially distributed retrial time, the retrial rate is denoted by ν_2. Again, based on these random times, the order of low priority packets arriving to the orbit may differ from the order of packets leaving the orbit.

The functionality of this communication network is presented on Fig. 1.

To create a stochastic process describing the behavior of the system the following notations are introduced (Table I contains the overview of parameters of the network):

- $k_1(t)$ is the number of active sources in the high priority class at time t,
- $k_2(t)$ is the number of active sources in the low priority class at time t,
- $o_1(t)$ is the number of requests in the orbit for high priority requests at time t,
- $o_2(t)$ is the number of requests in the orbit for low priority requests at time t,
- $y(t) = 0$ if there is no request in the server at time t. The server is available and ready to receive a job. $y(t) = 1$ if the server is engaged with a request coming from the high priority class, and $y(t) = 2$ if the server is engaged with a job coming from the low priority class at time t.

It is easy to see that:

$$k_1(t) = \begin{cases} N - o_1(t), & y(t) = 0, 2 \\ N - o_1(t) - 1, & y(t) = 1 \end{cases}$$

and

$$k_2(t) = \begin{cases} K - o_2(t), & y(t) = 0, 1 \\ K - o_2(t) - 1, & y(t) = 2 \end{cases}$$

TABLE I
LIST OF NETWORK PARAMETERS

Parameter	Maximum	Value at t	Unit
Active high priority s.	N	$k_1(t)$	-
Active low priority s.	K	$k_2(t)$	-
High priority gen. rate		λ_1	1/s
Low priority gen. rate		λ_2	1/s
Total gen. rate	$\lambda_1 N + \lambda_2 K$	$\lambda_1 k_1(t) + \lambda_2 k_2(t)$	1/s
Requests in high pr. orbit	N	$o_1(t)$	-
Requests in low pr. orbit	K	$o_2(t)$	-
Ret. rate in high pr. orbit		ν_1	1/s
Ret. rate in low pr. orbit		ν_2	1/s
Service rate		μ	1/s

In order to obtain the steady-state probabilities and performance measures within the Markovian framework, the mathematical tractability of the proposed model should be preserved. Therefore, we follow the classical approach frequently applied in the theory of retrial queues for the performance evaluation of infocommunication systems, namely, the distributions of inter-event times (i.e., request generation times for low and high priority packets, service time, retrial times) presented in the system are assumed to be exponentially distributed and totally independent of each other.

Consequently, the state of the network at a time t can be described by a Continuous Time Markov Chain (CTMC) with 3 dimensions:

$$X(t) = (y(t); o_1(t); o_2(t))$$

The steady-state distributions are denoted by

$$P(y, o_1, o_2) = \lim_{t\to\infty} P(y(t) = y, o_1(t) = o_1, o_2(t) = o_2)$$

Note that in the present case, the unique stationary distribution always exists, because the underlying CTMC is irreducible and the state space of the CTMC is finite. For computing the steady-state probabilities and the system characteristics, the MOSEL-2 software tool is used. These computations are similar to the ones described in, for example [12], [13].

As soon as we have calculated the distributions defined above, the most important steady-state system performance measures can be obtained in the following way:

- *Utilization of the Server with respect to high priority packets*

$$U_{S_1} = \sum_{o_1=0}^{N-1} \sum_{o_2=0}^{K} P(1, o_1, o_2)$$

- *Utilization of the Server with respect to low priority packets*

$$U_{S_2} = \sum_{o_1=0}^{N} \sum_{o_2=0}^{K-1} P(2, o_1, o_2)$$

- *Overall utilization of the Server*

$$U_S = U_{S_1} + U_{S_2}$$

- *Mean number of jobs in the orbit for high priority requests*

$$\overline{O_1} = E(o_1(t)) = \\ = \sum_{y=0}^{2} \sum_{o_1=0}^{N} \sum_{o_2=0}^{K} o_1 P(y, o_1, o_2)$$

- *Mean number of jobs in the orbit for low priority requests*

$$\overline{O_2} = E(o_2(t)) = \\ = \sum_{y=0}^{2} \sum_{o_1=0}^{N} \sum_{o_2=0}^{K} o_2 P(y, o_1, o_2)$$

- *Mean number of high priority jobs in the network*

$$\overline{M_1} = \overline{O_1} + U_{S_1}$$

- *Mean number of low priority jobs in the network*

$$\overline{M_2} = \overline{O_2} + U_{S_2}$$

- *Mean number of jobs in the network*

$$\overline{M} = \overline{M_1} + \overline{M_2}$$

- *Mean number of active high priority sources*

$$\overline{\Lambda_1} = N - \overline{M_1}$$

- *Mean number of active low priority sources*

$$\overline{\Lambda_2} = K - \overline{M_2}$$

- *Mean generation rate of high priority sources*

$$\overline{\lambda_1} = \lambda_1 \overline{\Lambda_1}$$

- *Mean generation rate of low priority sources*

$$\overline{\lambda_2} = \lambda_2 \overline{\Lambda_2}$$

- *Mean waiting time in orbit for high priority requests*

$$\overline{W_1} = \frac{\overline{O_1}}{\overline{\lambda_1}}$$

- *Mean waiting time in orbit for low priority requests*

$$\overline{W_2} = \frac{\overline{O_2}}{\overline{\lambda_2}}$$

- *Mean response time of high priority requests*

$$\overline{T_1} = \frac{\overline{M_1}}{\overline{\lambda_1}}$$

- *Mean response time of low priority requests*

$$\overline{T_2} = \frac{\overline{M_2}}{\overline{\lambda_2}}$$

III. NUMERICAL RESULTS

Investigating the functionality and the behavior of the system several numerical calculations were performed. From the steady-state probabilities computed by MOSEL-2 tool the most interesting performance characteristics were obtained, which are graphically presented in this section. The numerical values of the model parameters are described in Table I and in Table II. In the calculations a single variable λ is used for the generation rates. The high priority generation rate is $\lambda_1 = \lambda$, and the low priority generation rate is $\lambda_2 = 2\lambda$.

On the Figures 2 - 8 the three different lines represent the effects of different retrial rates in the orbit for the high priority requests ($\nu_1 = 2$: blue lines, dotted with rhombus, $\nu_1 = 4$: red lines, dotted with squares, $\nu_1 = 8$: green lines, dotted with triangles). On Figure 9 the two lines correspond to the different functionality of the server (described in [1]).

TABLE II
NUMERICAL VALUES OF MODEL PARAMETERS

Case studies							
No.	N	K	λ	ν_1	ν_2	μ	y-axis
Fig. 2	50	50	$x - axis$, [0.1..1]	$2, 4, 8$	2	20	$\overline{O_1}$
Fig. 3	50	50	$x - axis$, [0.1..1]	$2, 4, 8$	2	20	$\overline{O_2}$
Fig. 4	50	50	$x - axis$, [0.1..4.6]	$2, 4, 8$	2	20	$\overline{W_1}$
Fig. 5	50	50	$x - axis$, [0.1..4.6]	$2, 4, 8$	2	20	$\overline{W_2}$
Fig. 6	50	50	$x - axis$, [0.1..1]	$2, 4, 8$	2	20	U_s
Fig. 7	50	50	$x - axis$, [0.1..1]	$2, 4, 8$	2	20	U_{s_1}
Fig. 8	50	50	$x - axis$, [0.1..1]	$2, 4, 8$	2	20	U_{s_2}
Fig. 8	50	50	$x - axis$, [0.1..4.6]	2	2	20	$\overline{O_2}$

On Figure 2 the mean orbit size is displayed for high priority requests. When the generation rate is increased, the size of the orbit will be larger. The size of the orbit depends on the retrial rate, as well. In case of lower retrial rate, the size of orbit will be significantly larger.

Figure 3 shows the size of the orbit for low priority requests. Compared to Figure 2, a reverse tendency can be observed here. As we increase the retrial rate of the orbit for high priority requests, the low priority requests will have difficulties reaching the server (higher value of ν_1 implies larger $\overline{O_2}$). In addition, the size of orbit for low priority request fills up faster than the size of the other orbit. This is a straight consequence of the larger generation rate for low priority requests.

Figure 4 displays the Overall Utilization of the Server as function of λ. The utilization of the server (U_s) rises dramatically at the beginning: at value of $\lambda = 0.3$ the utilization equals to 85 percent.

On Figures 5 and 6 the utilization of the server with respect to high and low priority requests are shown. In cases of higher values of ν_1 the U_{s_1} values will be greater. Reverse effect stands for low priority requests, because for higher values of ν_1 (high priority retrial rate), larger number of high priority service demand will be present in the system. If $\nu_2 = \nu_1 = 2$, the server utilization curves for the two priority classes are the same (the blue lines. When the retrial rate of ν_1 is greater than ν_2 (red and green lines), the lines have maximum points.

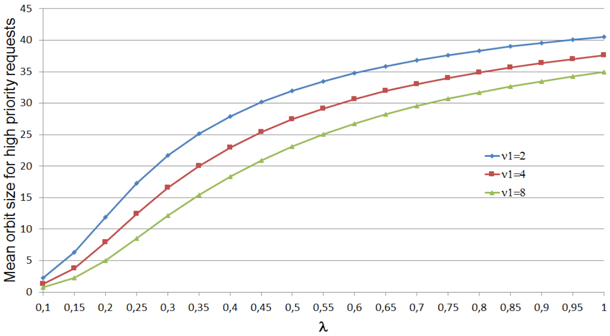


Fig. 2. Mean orbit size for high priority requests vs λ

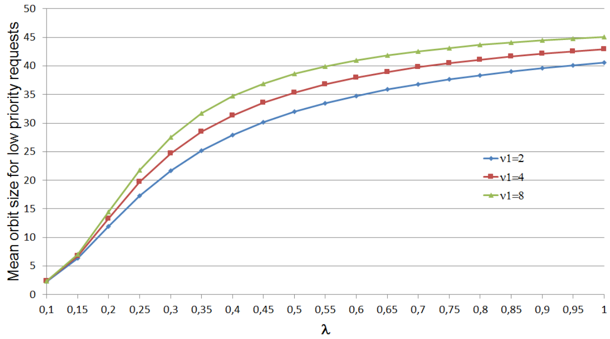


Fig. 3. Mean orbit size for low priority requests vs λ

For small values of generation rates the probabilities of CPU serves low priority requests are increasing, and higher values of generation rates the large number of high priority retrials will lower the considered probabilities. On Figure 4 it can be seen, that for higher values of generation rate (greater than 0.4) the utilization is almost constant (0.9). As U_{s_1} increases with the higher generation rates, too, consequently U_{s_2} has to be decreased.

In the Figures 7 and 8 we would like to investigate the effect of the random backoff feature on the mean waiting times (i.e. the mean times spent in the orbits). The parameters of these performance measures were the same like it was in [1]. Obviously, the larger generation rates will cause more time spent in the orbits. Because of the larger generation rate, for low priority requests $\overline{W_2}$ increases faster than $\overline{W_1}$ for high priority requests. Comparing the Figure 8 to the one presented in [1] (see Figure 9) we can see big differences caused by replacing the FIFO queue with an orbit: the values of the mean waiting times in the case of using FIFO discipline for the high priority requests are quite larger than in the case of using two orbits (the case for modeling the random backoff feature of the wireless access). On the other hand, it can be stated, that the functionalities of orbits connected with an other orbit and connected with a FIFO queue are different. In model of [1] the curves are first concave, then turn into convex, in recent model they are first convex, then change into concave. Thus the presence of the wireless access issue in the high priority class (which are not present in model of [1]) changes significantly the working conditions for the low priority class.

IV. Conclusion

In this paper a finite source queueing model was created in order to include the wireless network access feature of the communication. Two classes of sources (high priority and low priority traffic) were investigated. The random backoff feature of the Wi-fi access was implemented by using retrial queues for each traffic class. The conceptual working scheme of the model was described by a multidimensional Markov Chain, and the MOSEL software tool was used to develop the special software in order to calculate the most important steady-state performance characteristics of the system. At the end of the

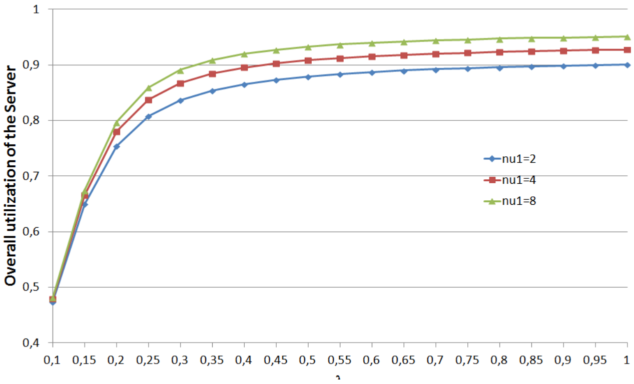


Fig. 4. Overall utilization of Server vs λ

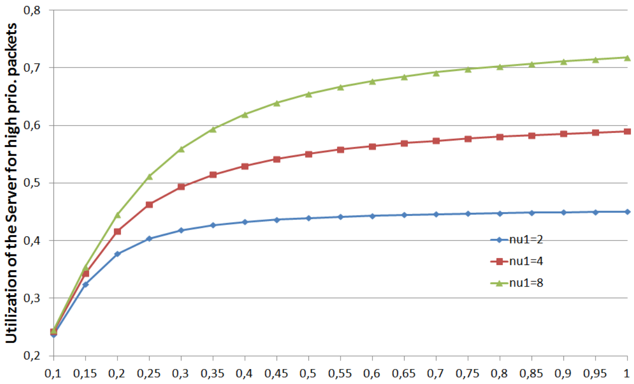


Fig. 5. Utilization of server for high priority packets vs λ

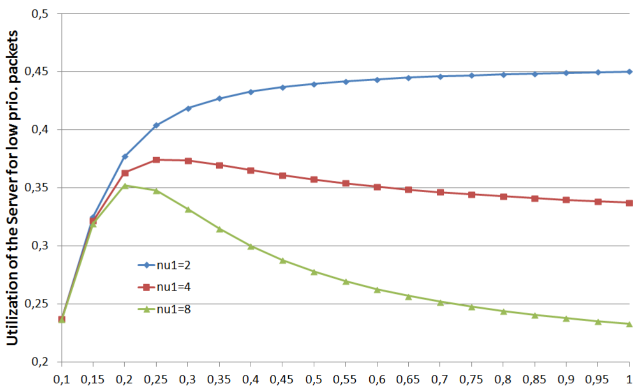


Fig. 6. Utilization of server for low priority packets vs λ

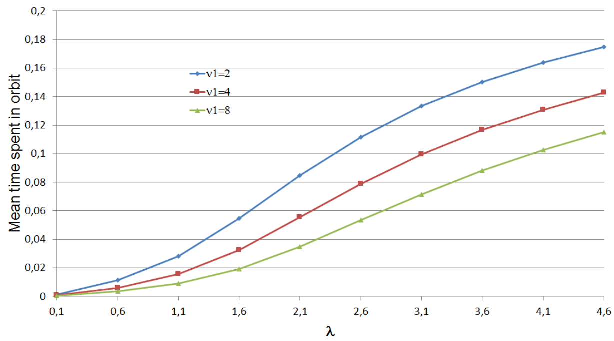


Fig. 7. Mean time spent in orbit for high priority requests vs λ

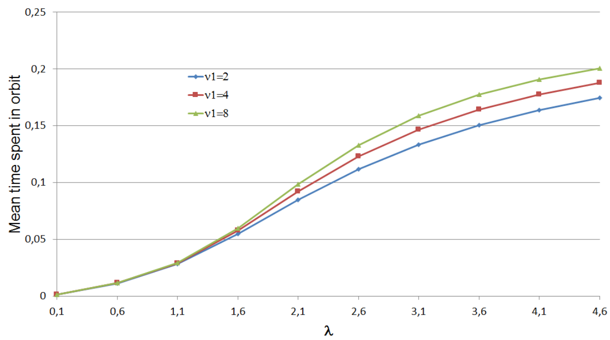


Fig. 8. Mean time spent in orbit for low priority requests vs λ

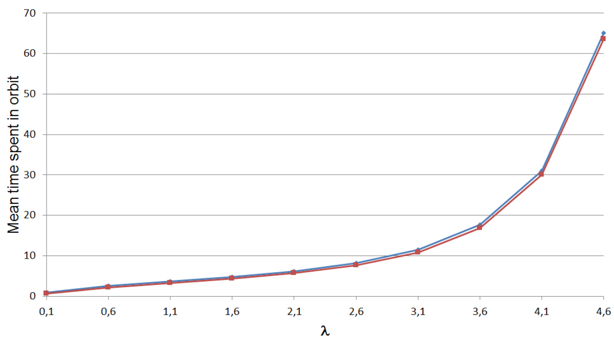


Fig. 9. Mean time spent in orbit for low priority requests vs λ (FIFO for high priority requests)

paper we showed how the feature of the random backoff effects the mean waiting times (compared to the FIFO discipline): not only the values, but also the form of the curve is quite different in the case of the considered wireless communication environment.

ACKNOWLEDGMENT

The research of A. Kuki and B. Almási was supported by the TÁMOP 4.2.2. C-11/1/KONV-2012-0001 project. The project has been supported by the European Union, co-financed by the European Social Fund.

The work of János Sztrik was realized in the frames of TÁMOP 4.2.4. A/2-11-1-2012-0001 National Excellence Program - Elaborating and operating an inland student and researcher personal support system. The project was subsidized by the European Union and co-financed by the European Social Fund.

The work of Tamás Bérczes was realized in the frames of TÁMOP 4.2.4. A/2-11-1-2012-0001 National Excellence Program - Elaborating and operating an inland student and researcher personal support system. The project was subsidized by the European Union and co-financed by the European Social Fund.

REFERENCES

[1] Bérczes, T., Almási, B., Sztrik, J., Kuki, A. "A contribution to modeling sensor communication networks by using finite-source queueing systems," Proc. 8th IEEE International Symposium on Applied Computational Intelligence and Informatics, Timisoara, Romania pp. 89-93, 2013.

[2] Avrachenkov, K., Nain, P., Yechiali, U. "A retrial system with two input streams and two orbit queues," Research Report no. 7999, Research Centre SOPHIA ANTIPOLIS MDITERRANE France, June, 2012.

[3] H. T. Tran, T. V. Do, László Pap "Analysis of a queue with two priority classes and feedback controls," Vietnam J. Computer Science 1(2): 71-78 (2014).

[4] Artalejo, J., and Gomez-Corral A. "Retrial Queueing Systems," Springer, Berlin, 2008

[5] Basharin, G. P., Samouylov, K. E., Yarkina, N. V. and Gudkova, I. A. "A new stage in mathematical teletraffic theory," Automation Remote Control, vol. 70, no. 12 pp. 1954-1964, 2009.

[6] Feng, W., and Kowada, M. "Performance analysis of wireless mobile networks with queueing priority and guard channels," International Transactions on Operational Research, no. 15, pp. 481- 508, 2008.

[7] Lakatos, L., Szeidl, L., and Telek, M. "Introduction to queuing systems with telecommunication applications," Springer, Heidelgerg, 2013.

[8] Ponomarenko, L., Kim, C. S., and Melikov, A. "Performance analysis and optimization of multi-traffic on communication networks," Springer, Heidelgerg, 2010.

[9] Stasiak, M., Glabowski, M., Wishniewski, A., and Zwierzykowski, P. "Modeling and dimensioning of mobile networks," John Wiley, Chichester, 2011.

[10] Szilágyi, Sz. Almási, B. "A Review of Congestion Management Algorithms on Cisco Routers," Journal of Computer Science and Control Systems vol. 5, no. 1, pp. 103-107, 2005.

[11] Begain, K., Bolch, G., and Herold, H. "Practical Performance Modeling, Application of the MOSEL Language," Kluwer Academic Publisher, Boston, 2001.

[12] Wüchner, P., Sztrik, J., and de Meer, H. "Modeling wireless sensor networks using finite-source retrial queues with unreliable orbit," Springer Lecture Notes in Computer Science, no. 6821, pp. 275-285, 2011.

[13] Bolch, G., Greiner, S., de Meer, H., and Trivedi, K. "Queueing Networks and Markov Chains," John Wiley & Sons, New York, 2nd edition, 2006.

Tamás Bérczes received his MSc Degree in Mathematics in 2000. at the University of Debrecen, Hungary. He is currently assistant professor at the Department of Informatics Systems and Networks of the same university. He received his Ph.D. in 2011. His primary research interests are performance analysis of retrial queues and their application.

Béla Almási is a associate professor at the Department of Informatics Systems and Networks at the Faculty of Informatics, University of Debrecen, Hungary. He received his Ph.D. in 1998. at the University of Debrecen, Hungary and Habilitation from University of Debrecen in 2006. His primary research interests are network systems, performance analysis of retrial queues and their application.

Attila Kuki He is currently an assistant professor at the Department of Informatics Systems and Networks at the Faculty of Informatics, University of Debrecen. H e received his Ph.D. in 1997 in Mathematics. His primary research interests are performance analysis of retrial queues and their application.

János Sztrik is a Full Professor at the Faculty of Informatics, University of Debrecen, Hungary. He obtained the Candidate of Mathematical Sciences Degree in Probability Theory and Mathematical Statistics in 1989 from the Kiev State University, Kiev, USSR, Habilitation from University of Debrecen in 1999, Doctor of the Hungarian Academy of Sciences, Budapest, 2002. His research interests are in the field of production systems modeling and analysis, queueing theory, reliability theory, and computer science.

A Study on the Performance of an Advanced Framework for Prediction-based NEMO Handovers in Multihomed Scenarios

László Bokor†, Gábor Jeney†, József Kovács‡

***Abstract* — In this paper we provide an extensive performance evaluation of an already introduced advanced handover management solution which aimed at providing ubiquitous IPv6 connection and seamless Internet access for network mobility (NEMO) scenarios. The novel solution makes use of geographic location information and previous records of access network parameters. The method exploits the benefits of multihomed mobility configurations by introducing a special handover execution protocol entirely based on flow bindings. Using actual location information and previously recorded context data, the system is able to predict handovers and proactively prepare itself for the appearance of access networks. We studied the performance of our proposal by implementing the framework and handover execution scheme in a real-life 3G/Wi-Fi multi-access testbed environment, and showed that handover latency is almost totally eliminated. As our solution strongly relies on the prediction accuracy, we have also developed a probabilistic system model and evaluated of the probability of wrong positioning on the prediction raster.**

***Keywords:* IPv6, cross-layer optimization, mobility management, NEMO, MCoA, geographic position information, flow bindings, policy exchange, multihoming, predictive handover, proactive handover, real-life implementation, probabilistic system model**

I. INTRODUCTION

Trends in information technology show that heterogeneous, IP-based wireless networks will support mobility for the widest range of single end terminals (e.g., mobile phones, SmartPhones, PDAs, tablets and other handhelds), and even Personal Area Networks (PANs), Vehicle Area Networks (VANs) [1], Intelligent Transportation Systems (ITSs) and Cooperative ITS (C-ITS) architectures [2]–[4], networks of RFID (Radio Frequency Identification) devices and sensors, and various mobile ad hoc networks [5] will have permanent Internet connectivity during movement. Hence, in next generation wireless telecommunication not only single mobile entities have to be taken into account (host or terminal mobility), but also entire mobile networks moving between different subnets need to be maintained as a whole (i.e., network mobility or NEMO). IPv6 has introduced support for both mobility cases by defining Mobile IPv6 (MIPv6) [6] and Network Mobility Basic Support (NEMO BS) [7]. With these mobility supporting mechanisms all sessions remain active, even when the mobile node/network changes its subnetwork. When a host or a moving network has multiple interfaces and/or several IPv6 addresses, it is regarded multihomed. Multihomed mobile hosts/networks need special protocols to support their mobility management (e.g., MCoA [8], Flow Bindings [9], [10]). Handover at network layer usually takes several seconds due to the large number of L1/L2/L3 processes, the lack of interaction between them, and their complexity. The overall time needed to complete these procedures could go up to several seconds. In order to ensure seamless, continuous communication, this huge outage should be avoided by applying optimized handover solutions in the architecture.

Several improvement proposals exist to overcome the huge delay. All of them aim to speed-up the handover process. Mobile IPv6 Fast Handovers [11] is one example, and there are plenty of other proposals as well [12]–[16]. However, according to our best knowledge, only our previous works [17]–[19] exploit the benefits of overlapping radio access coverages by proactively managing multiple tunnels and executing predictive tunnel switching based on generalized flow binding policy exchange. Our solution extends standard IPv6-based network mobility by forming an advanced and complete framework based on a special, multi-tunnel based, predictive, seamless handover solution. In this paper we further elaborate our previous work and provide an extensive performance analysis of the scheme. In order to do this, we provide a more detailed, broad background on the different handover schemes, show more implementation details, introduce new measurement results and also provide a novel, probabilistic system model for analytical evaluation of the prediction system.

The paper is structured as follows. Section II introduces the scientific background and related work, also summarizing the existing predictive mobility management schemes. Section III recaps our existing solution, and further details the protocol operation. Testbed and measurements results are explained in Section IV. Section V presents the probabilistic system model and our evaluation results on the probability of wrong positioning. Finally, we conclude the paper and show some possible future work in Section VI.

† Inter-University Centre for Telecommunications and Informatics (ETIK), Kassai u. 26. H-4028, Debrecen, Hungary, and Department of Networked Systems and Services, Budapest University of Technology and Economics, {bokorl, jeneyg}@hit.bme.hu

‡ Computer and Automation Research Institute, Hungarian Academy of Sciences (MTA), Kende u. 13-17, H-1111, Budapest, Hungary, jk@sztaki.hu

Manuscript received June 30, revised August 31.

II. BACKGROUND AND RELATED WORK

Mobile clients continuously change their position, which could yield access network failures or connection drops. Mobility management in heterogeneous access architectures is aware of handling the mobility related procedures. IPv6 has built-in support for terminal and network mobility, but these basic solutions do not tackle the problem when handovers provide serious communication outages due to the large number of L1/L2/L3 duties. First, the mobile terminal/router has to find and connect to the new network at L1 and L2 (PHY and MAC), and only after the successful L1/L2 connection it could launch the necessary L3 procedures to obtain the new IPv6 address(es) (with stateless [20] or stateful autoconfiguration [21]). After the new IPv6 address is set, the binding procedure starts: it binds (registers) its address(es) in the Home Agent (HA), which provides global accessibility. These procedures could easily result in several seconds of handover delay. Figure 1 introduces how it happens in case of NEMO BS [7].

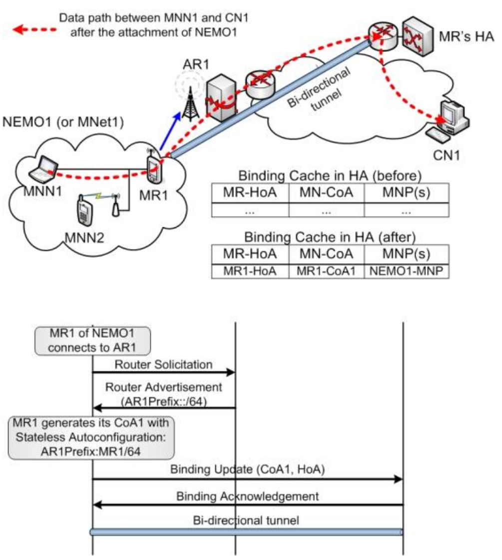


Fig. 1. NEMO BS handover management

The solution used by NEMO BS is similar to Mobile IPv6 but without routing optimization: when a Mobile Router (MR) leaves its home link, it configures a Care of Address (CoA) in the visited network and registers this CoA with its HA using the binding procedure. However, the Binding Update (BU) message in NEMO BS is quite different from that in MIPv6. While a BU message in MIPv6 contains the Care-of and the Home Address (HoA) of a mobile node, till a BU of an MR contains additional information: the IP subnet prefix or prefixes of the moving network. These so called Mobile Network Prefixes (MNPs) in the Binding Updates instruct the Home Agent to create a binding cache entry linking the MNPs to the MR's Care-of Address. After a successful registration, the HA intercepts and forwards packets destined not only to the MR, but also to any MNNs that have acquired an address from one of the Mobile Network prefixes of the MR. When the moving network changes its actual network point of attachment, only the MR configures new CoA and sends BU (containing the MNPs) to the HA. Since the Home Addresses of the MNNs inside a moving network are associated with the MNPs registered in the HAs, the HA of the network's MR intercepts all the packets addressed to MNNs and forwards them towards the MR's CoA. The MR decapsulates the packets destined to MNNs and forwards them on its appropriate ingress interfaces. Packets originated from inside the moving network will follow the same routes but in the reverse direction. It is obvious that the big number of encapsulations cause header overhead, and the fact that all the HAs should be involved in the communication path results using traffic routes far from the optimal ones. NEMO BS is not applicable for multi-access or multihoming scenarios as it supports only one interface that has to be configured before starting the NEMO stack. In a heterogeneous environment such as a 3G/Wi-Fi architecture, reconfiguration and restarting of the NEMO BS implementation is required to use a different interface than the configured one. Moreover, the handovers are handled "blind" as no network context information is available during the operation: a strictly reactive behaviour is used.

Multihoming is an advantageous method to provide always-on connectivity in a wireless environment and support multi-access scenarios. In order to exploit such possibilities, Multiple Care-of Addresses Registration (MCoA) [8] was introduced to the Mobile IPv6 protocol family. By utilizing that mobile nodes or routers can connect to multiple access networks simultaneously, it is now possible to enhance handover latency, network redundancy and perform policy based routing even in multi-access environments. Figure 2 depicts a typical scenario, where the MR has two external interfaces, where each interface is connected to an access network with a CoA, and through each CoA a Mobile IPv6 tunnel is created to the Home Agent. While with NEMO BS, identifying a binding was enough using the CoA and the HoA, it is no longer the case with NEMO MCoA as each mobility tunnel endpoint uses the same Home Address on the MR. Using network layer information, the MR can no longer perform an exact routing decision to select an individual tunnel. To solve this issue another identifier, known as Binding Identifier (BID) was introduced to identify the network interface over which the tunnel is established. As the BID is sent to the HA in the BU signalling message, the HA can differentiate between tunnels originating from the same MR. To identify and route packets toward the desired tunnel, policy routing must be used, which allows fine grained diversification among data packets and streams based on network layer and upper layer information. To avoid asymmetric routing where packets belonging to the same packet flow are routed on different tunnels, a flow binding mechanism has to be implemented. Using flow binding control messages, the MR registers flow descriptor and BID pairs at the Home Agent, so the HA would properly know which tunnel to use when it forwards packets of the data flow back to the mobile node [9], [10].

Using the above introduced multihoming solution, routing of individual media streams can be easily solved, enhancing the experience for not only moving, but stationary mobile nodes as the presence of multiple egress interfaces makes content delivery more reliable and robust.

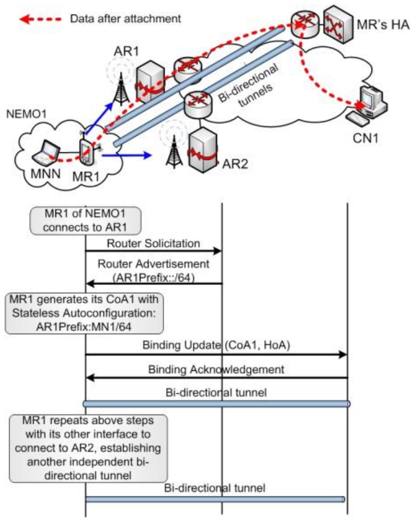


Fig. 2. NEMO multihoming solution with MCoA

Based on MCoA, a special type of handover scheme can be defined (we call it MCoA handover [18]), which relies on overlapping radio access networks (RANs), and in case of an appearing new access network on an unused MR interface, moves every traffic to this new RAN by activating symmetric policy rules for all the MR transmissions (Figure 3).

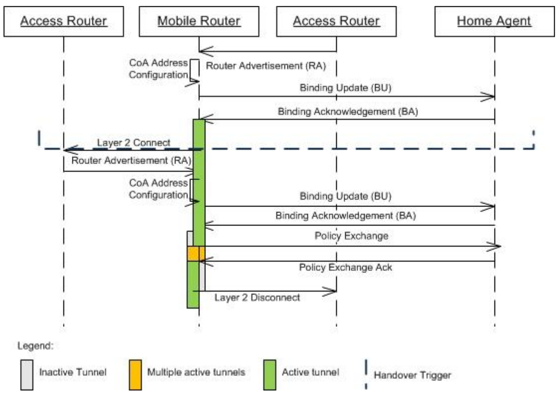


Fig. 3. A possible NEMO MCoA handover solution

This MCoA handover reduces Layer 2 and Layer 3 latencies by connecting to the new network with an MR interface, which does not used for any of the actual communication, and after the initiation of this new connection (i.e., L2 and L3 tasks like physical connection and NEMO tunnel creation), a simple policy exchange executes the NEMO handover. Of course handovers are still handled "blind" (i.e., no context information is available about the new network before the handover). Now it is clear why this idea requires overlapping RANs and creates the possibility for further optimization aiming at shrinking the handover latency.

Various proposals have been published to shrink the delay caused by handovers. In the next section, we recap how to use location information coming from e.g., Global Navigation Satellite System (GNSS), or Geographical Positioning System (GPS) data to speed-up the handover process of multihomed NEMO architectures in heterogeneous access environments. Our proposed method is only usable when the mobile terminal or mobile network moves on nearly the same path every time. Public transportation vehicles (trains, trams, buses, trolleys, etc.), cars and trucks travelling on highways/main roads are examples, when this assumption is valid. Random walks in city centres are beyond applicability, and thus our method cannot be applied there without modification.

Using location information for preparing handovers dates back to 2001. In [15] Wang et al. propose to use location information to improve the performance of inter-cell handovers. Their method is limited to L1/L2 handovers; they did not consider IP connectivity. Dutta et al. [16] extended Wang's work recently, also concentrating on the L1/L2 handovers only. Hee-Dong Park et al. [14] proposed first a NEMO scheme which can be used in vehicles travelling on a predetermined route. They store access network information in a database which is used to predict handovers. They have not considered MCoA scenario, though. Our method makes use of multihoming as we propose to use MCoA with advanced policy exchange mechanisms. The policy exchange mechanism is based on the recommendations of the IETF's Flow Bindings RFC in Mobile IPv6 and NEMO BS [8], [9]. Our solution supports IPv6 only, due to the fact that all related protocols are better implemented in IPv6.

In [22], the authors propose a similar scheme to ours, however this paper lacks the technical description of the system and the handover execution scheme.

Finally, we have published our solution in [17]–[19], where we provided a complete description of our methodology with some preliminary results. This article extends our previous papers with a more complete description also including the details of the applied tunnel management/handover execution scheme, and with a more complete real-life evaluation based on a comparison of our framework with other two different Mobile IPv6 based handover management techniques, i.e., with NEMO BS (Fig. 1) and NEMO MCoA (Fig. 3) handover solutions. We also provide an analysis of prediction accuracy in the proposed solution by studying the limitations of the overall architecture inherited by possible wrong positioning on the prediction raster network. We show that our proposed solution outperforms all existing implementations, and prove that an appropriate prediction raster can keep the probability of wrong positioning below 1%.

III. PREDICTIVE HANDOVER MANAGEMENT FOR NEMO MOBILE ROUTERS IN MULTI-ACCESS ENVIRONMENTS

In this section we recap the main considerations of our already introduced framework including the predictive handover management scheme designed for multihomed NEMO configurations, and also present the details of our multi-tunnel based efficient handover execution scheme which combines the benefits of MCoA with a new prediction-driven cross-layer management entity.

A. General considerations of the proposed solution

There are two levels of handovers which should be considered independently: L1/L2 handovers, and L3 handovers. L1/L2 level handovers are determined by the access technology currently in use. 3G/HSPA, Wi-Fi, etc. handover delays are due to their respective standard and implementation. However, if the mobile terminal or router contain more than one egress interface, it is possible to use one interface for communication and another one for preparation and execution of L1/L2 handovers. In such a handover scenario sessions should be re-directed between interfaces. Since our solution is based on IPv6, native IPv6 support is a must in all access networks.

L3 level handovers are handled by IP mobility solutions (e.g., MIPv6 or NEMO BS). L3 handovers can be speed-up by launching L3 procedures before L1/L2 handover happens. For this reason we use location information. It could be possible to launch the L3 procedure such that it just finishes as the new network appears. If so, the handover latency becomes lower (down to L1/L2 handover delay) and the service becomes almost ubiquitous. Under perfect circumstances (exploiting overlapping coverage areas and benefits of multi-access devices) the latency can be totally eliminated if the L1/L2/L3 preparations are executed in a predictive and timed manner.

In our IPv6-based NEMO BS extension we propose to use location information coming from e.g., Global Navigation Satellite System (GNSS), or Geographical Positioning System (GPS) data to speed-up the handover process of multihomed NEMO architectures in heterogeneous access environments.

The idea behind predictive handover management is very simple: as the node/network moves along a path, it records all access network related data in a database together with the geographical location information. The next time the node/network moves along the same path, based on the geographical information and speed vector, the stored information can be used to predict and prepare handovers before the actual availability of the networks. In the appropriate time, ongoing communication sessions can be seamlessly redirected to some other interface(s) – thus successfully finishing handovers.

The following information should be stored. Network type (WLAN, 3G, WiMAX, etc.), network identifier (e.g., BSSID of WLAN AP, 3G cell identifier, etc.) and IP level information (e.g., network prefix, which can be used to gather the IPv6 address of the node). The first three fields are required: without them it is not possible to prepare handovers in L1/L2/L3 relations. Some additional information, e.g., Signal-to-Noise Ratio (SNR), BandWidth (BW), reliability (how often the network appears at a given geographical location) and Round Trip Time (RTT) are useful for further intelligence and more sophisticated decisions. For instance, the more reliable network should be chosen if several networks are available.

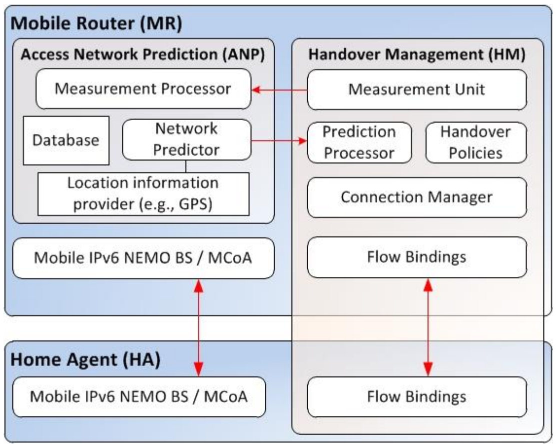


Fig. 4. The proposed predictive handover management framework

When multiple interfaces are available, MCoA [8] and Flow Bindings [9], [10] solutions can be of use. L3 handover preparation consists of the following components. First of all, a BID is created for each egress interface the MR possesses. These BIDs are used as unique identifiers of the interfaces. BIDs are sent in BU messages to the Home Agent in order to identify individual bindings of the MR. The HA that receives the BU messages creates a separate binding for each BID (i.e., for each egress interface of the MR). Therefore the MR owns only one Home Address but the bidirectional NEMO tunnels will be distinguishable based on the BIDs. The sole Home Address of the MR requires the introduction of Flow Bindings which directs packet flows to specific egress interface. In the proposed scheme we use Flow Bindings to direct the whole traffic of the MR through one active egress interface. In this way the solution loses the benefits of redundant interfaces, but gains the possibility to use inactive interfaces for handover preparation, i.e., selecting appropriate access network, performing lower layer connections and acquiring new IPv6 addresses.

Therefore the scheme requires several interfaces for operation. Some of the interfaces are used for normal communication (they will be referred as “active”); the others are used for handover preparation (they are termed as “inactive”). The activation of a new interface must be accurately synchronized with the deactivation of the old one. The activation/ deactivation procedure means simultaneous reallocation of NEMO BS tunnels. It is performed by properly scheduled flow binding policy control messages on the HA and the MR. The control messages are called Predictive Policy Exchange Messages.

B. The proposed framework and handover execution/tunnel management protocol

The proposed framework (Fig. 4) has three main components: Access Network Prediction (ANP), Handover Manager on the MR (HM-MR) and on the Home Agent (HM-HA). I do not claim all the functional entities; however the overall framework and the design of the predictive handover execution scheme are my results.

The left most module on Figure 4 running inside the Mobile Router is the ANP. The ANP is responsible for 1) maintaining the access network database; 2) sending prediction messages to HM-MR module; 3) reading GNSS information from the GNSS receiver; and 4) processing the network measurement messages received from the HM-MR module. Tasks 3) and 4) are for the maintenance of the access network database which should be continuously extended/updated during the movements of the MR. Based on up-to-date database records and current, precise position/speed information the ANP is able to provide candidate network parameter prediction. The predicttion vector is sent to the HM-MR module in an XML message, and the HM-MR measurement messages are also transmitted in an XML format. In order to avoid the explosion of the access network database size, the received GNSS coordinates are rounded in the following way: the longitude and latitude values are multiplied by 10,000 and rounded to the closest integer. Therefore instead of a continuous space they form a limited set with members called raster points inside a raster net, which plays an important role in the prediction precision (see later).

The Handover Management (HM) module can be divided into two parts depending on which node hosts it. The HM-MR runs on the Mobile Router (Fig. 4) and is responsible for two main tasks. On one hand HM-MR measures the channel state information and other network parameters of the actually available access networks during the movement of the MR. The scale of the measurable parameters is wide and depends on the decision algorithm to be applied. In our proposal the following parameters are measured, collected and sent periodically to the ANP module for further processing and storing in the database:

- Receive Signal Strength Indicator (RSSI) of UMTS
- Signal/Noise Ratio (SNR) and Basic Service Set Identifier (BSSID) of WLAN
- IPv6 prefix information

On the other hand, HM-MR also prepares predictive handovers by handling MCoA tunnels in a timed manner based on the prediction XML messages received from ANP and the indirect interaction with the NEMO MCoA implementation. In order to achieve this, we proposed a special predictive policy exchange scheme which can inform the Home Agent (i.e., the HM-HA module) about the Mobile Router's intents of future handovers. The periodically received candidate access network predictions supply all the necessary information required for handovers to be initiated by the HM-MR. If a handover event is predicted for the near future (e.g., prediction data reveal that the currently used access coverage will disappear soon), the decision algorithm will choose the destination network and initiate the handover mechanisms. In the proposed framework HM-MR follows a simple rule set to select the designated network from multiple candidates:

- an available WLAN network always has higher priority than 3G/UMTS
- the WLAN with the best SNR value has the highest priority among simultaneously available WLAN networks

The HM-HA module is located on the Home Agent (Fig. 4). The HA itself represents the same functional entity as in the case of standard MIPv6/NEMO/MCoA protocols, but in our scheme it also interacts with the HM-MR module through the HM-HA instance for predictive, timed and flow binding based NEMO MCoA handovers using the Predictive Policy Exchange Messages. The HA is informed about the predicted network prefixes and timing information, and thus changes in flow binding policies can be executed and scheduled before the handover event actually happens.

After the decision is made based on the rules defined at the HM-MR module, the designated network will be chosen and passed over to the Flow Bindings submodule at the MR side. This submodule handles the signalling between the MR and the HA for defining which MR-HA tunnel shall the system switch to and when. It is important to note, that before executing these timed and synchronized policy exchange commands for tunnel/routing adjustments on the MR and the HA entities, the designated network (i.e., a new interface) must be chosen and the L1/L2/L3 preparations must be finished for the selected network. Thanks to the prediction based and multihomed NEMO operation, the MR will be able to finish these preparations (including also the L3 NEMO MCoA tunnel build-up using the binding procedure) before the actual handover event occurs. However, it requires that the candidate networks are overlapping in their coverage.

Based on the GNSS aided predictions the policy exchange commands can be executed at exactly the same time both in the HM-MR and HM-HA modules. It means that all the NEMO traffic will be redirected to the new network defined by the new MCoA tunnel without noticeable packet loss or other QoS disruption. This is only possible because we already have a working Mobile IPv6 connectivity through the new network and all L2/L3 configurations are already performed. The Predictive Policy Exchange message would only carry timed commands to switch the packet flow to a functional, but inactive tunnel. Upon disconnect or failure of the active access network, routes and tunnel interfaces are deleted and the next default route with the highest priority is taken to ensure seamless connectivity. Recovering from such failure based on the enforcement of handover policies is out of scope of this document.

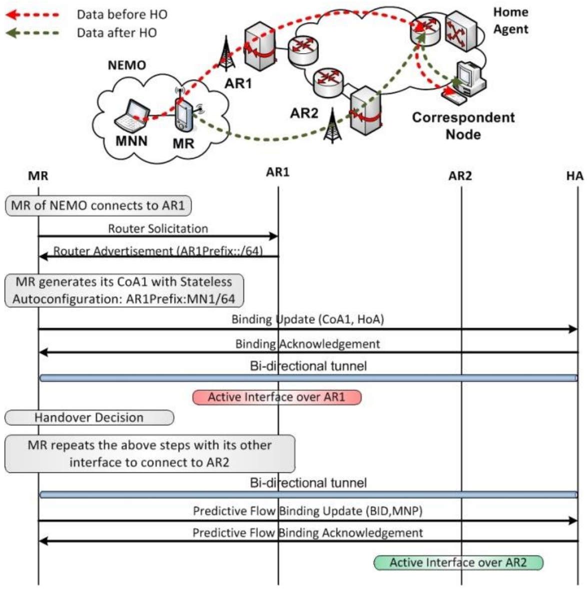


Fig. 5. Details of the proposed handover execution/tunnel management protocol

The proposed handover execution protocol is detailed in Figure 5. When the HM decides to perform a handover, in order to use the benefits of MCoA, the following steps are executed. Using one of the inactive interfaces the HM connects to the new access network and establishes a new Mobile IPv6 binding. At this stage, the current and new access networks are both connected and Mobility Tunnels are established between the MR and the HA. Handing over to the new access network is entirely based on Flow Bindings, which in this case means that all flows are moved from one interface to another. To avoid asymmetric routing, the MA and HA has to modify their bindings simultaneously, in a timely manner. The schedule is communicated by the Flow Binding modules in predictive Flow Binding Update/Acknowledgement messages (Figure 5). When the changes of flow bindings are executed, the new interface is marked as active, while the rest of the communication interfaces are set to inactive mode. The mobile network nodes (MNNs) inside the NEMO will always and transparently use the communication path spanned by the active interface. Different Handover Policies may have different effects on handover strategies.

IV. Implementation details

In this section we introduce the most important implementation details, our real-life, loosely-coupled UMTS/Wi-Fi heterogeneous testbed, the scenario created to evaluate our GNSS aided predictive NEMO handover management framework, and the measurement results. We compared our method against two other handover solutions (standard NEMO BS and NEMO MCoA handovers as detailed in Section II) in one networking scenario using four main handover performance metrics (Handover latency, UDP packet loss, TCP throughput and RTT).

A. Implementation details

This section is devoted to present some additional implementation details of some crucial modules of the overall framework.

1) Access Network Prediction (ANP)

The ANP has two main roles. On one hand it processes the measurements and records the data into the database. On the other hand it sends the periodic prediction messages towards the HM module. Measurement data

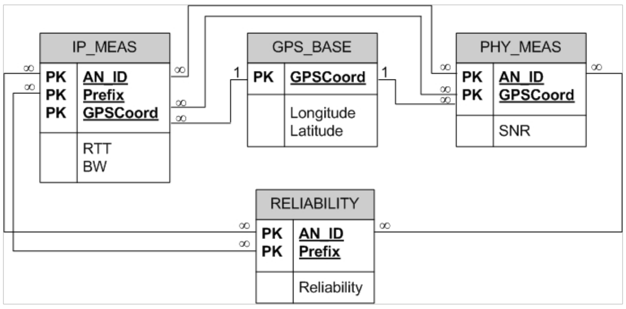


Fig. 6. Database scheme designed for the ANP module

Data from the measurement unit will be processed based on their time stamp `<MeasTimeStamp>`: GPS coordinates must be paired based on the measurement time, and then an entry should be inserted into the PHY_MEAS table about the followings:

- AN_ID: access network identification
- GPSCoord: the closest GPS coordinate in the raster to the measurement time
- SNR: Signal-to-Noise Ratio [dBm]
- IP_MEAS table is filled only if IP level measurement was also received (bandwidth and round-time-trip values)
- The Prefix entry could be empty or also could contain multiple prefix values
- Reliability table stores the level of reliability of a particular access network: as our vehicle could pass on the same route multiple times, we can summarize the measurement snapshots and using a special rating technique we can classify the stored data according to its appropriateness in a longer session of measurements. E.g., if the BW is low or the Home Agent is not available on a particular link during one measurement snapshot, the system will not immediately remove that access network but will start to degrade the level of reliability for that entry.

In order to create the prediction, first we filter the PHY_MEAS table based on GPS coordinates and SNR values: e.g., if we would like to implement a 10 second prediction window, then we will query AN_IDs with GPS coordinates in the next 10 seconds, and then unusable networks with terrible SNR values should be left out from the answer. Then we search the IP_MEAS table for prefixes for the given GPS coordinates and AN_ID values. In that way only networks with appropriate physical and IP level parameters will be sent towards the HM module. Based on the prefixes and the AN_ID all the stored values (Reliability, RTT, BW, etc.) can be gathered, which can be used by the HM module to make the handover decisions.

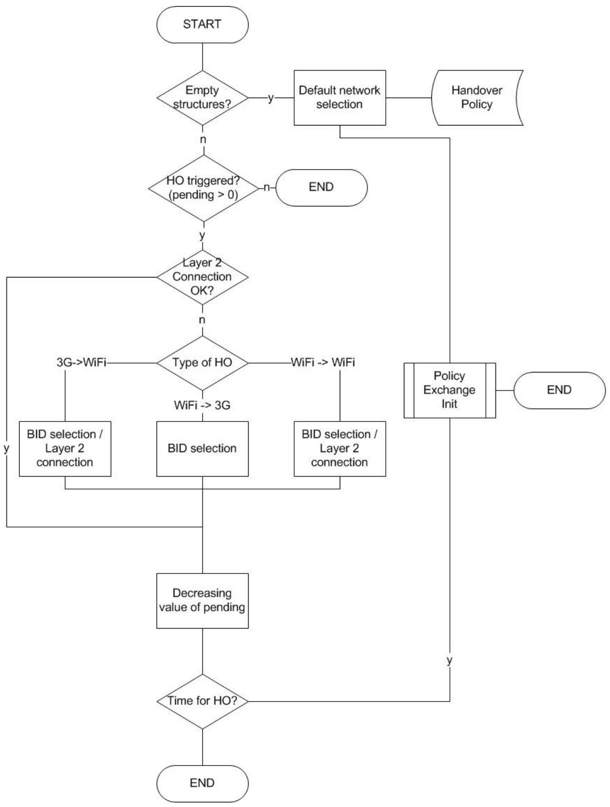


Fig. 7. Detailed operation of the Connection Manager module

2) Connection manager module

This module handles the execution of handovers, deals with the medium-dependent preparation tasks, maintains available interfaces and actual connections, and initiates Policy Exchange operations. As Figure 7 depicts, the Connection Manager loads up a default setting during the startup process. After that it awakes periodically and checks whether there is a need for handover or not. Depending on the decision, it will react according to the type of HO.

- 3G – WiFi: select one Wi-Fi from the list, puts the BID of that network into the Next network structure, then calls the appropriate OS function to perform the L2 connection and finally initiates the MIPv6 MCoA binding procedure.
- WiFi – WiFi: the unused interface must be selected and configured for usage. Then comes the connection initiation similarly to the 3G–WiFi case.
- WiFi – 3G: there is no need to explicitly handle L2 connections as the framework handles it transparently. Other duties are the same as before.

Important task of this module to continuously check the value of the *pending* variable, which is used to manage timely handovers: working with prediction window we can see the future and want to use the appropriate network as long as possible.

B. Testbed architecture

The basis of our evaluation efforts was a heterogeneous, native IPv6 UMTS/Wi-Fi testbed (Fig. 8) built on the existing hardware elements of Mobile Innovation Centre (MIK) [23]. As the figure shows, the 3G part of the access infrastructure is a standard, packet switched UMTS core running an IPv6-compatible GGSN implementation [24] in order to provide native IPv6 UMTS experience to the NEMO. This GGSN is connected to the outside IPv6 network through its Gi interface using native IPv6 transport. The WLAN part of the testbed comprises IEEE 802.11b/g compatible Linksys WRT54GL Access Points (APs) also with native IPv6 backhaul.

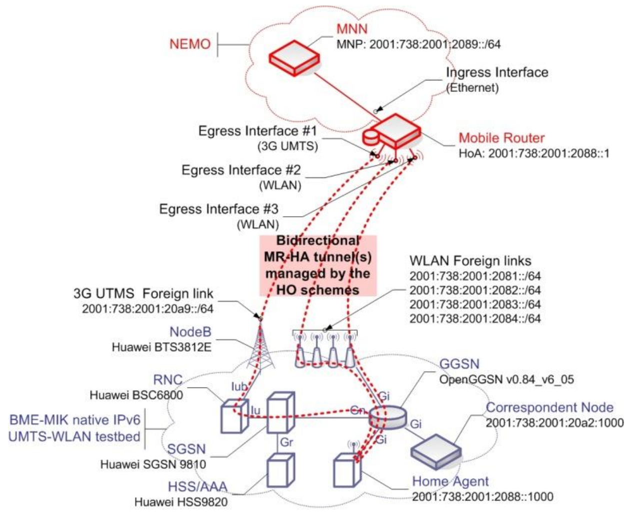


Fig. 8. Real-life testbed architecture

For accessing the above multi-access infrastructure and to provide advanced multihoming features with support for our handover solution, the Mobile Router has been equipped with three egress interfaces; one for UMTS and two for WLAN access, respectively. The MR controlled NEMO in our testbed comprises only one Mobile Network Node (MNN) which connects to the ingress interface of the MR over Ethernet. The Correspondent Node (CN) communicating with the NEMO from the outside network for testing purposes also uses Ethernet connection for IPv6 communication. These two latter nodes (i.e., the MNN and the CN) were running our measurement softwares: synthetic traffic generation was achieved by Netperf [25] while the packet capture and analysis was based on Tshark [26] and some additional shell scripts.

The HA and the MR – besides the NEMO BS and MCoA HA/MR functions implemented by appropriately patched UMIP 0.4 [27] instances – also can deal with the introduced tasks of our predictive policy exchange scheme (i.e., run the HM/ANP modules if needed).

Since the coverage limitations of this laboratory environment did not allow real, open-air motion of the NEMO, a special solution for movement emulation was introduced in our testbed using pre-recorded GPS traces. The traces were played back by the gpsfake component of gpsd [28] during every measurement run. This component is located on the MR together with the Access Network Prediction and Handover Manager modules of the predictive handover management system, and contains a virtual test track with pre-defined coverage structure along the path (Figure 9). The different coverage areas of this test track were emulated by a prepared database at the ANP, but the access networks were real.

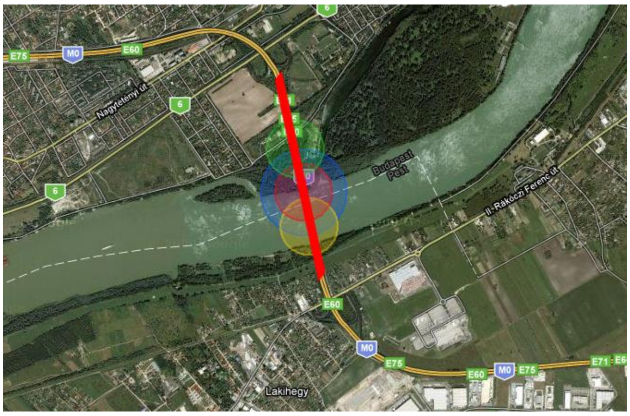


Fig. 9. Virtual test-track defined for the evaluation

The evaluation scenario we used was implemented based on the above testbed details strongly relying on this virtual motion/coverage information scheme. The yellow, two-lane road represents the left-to-right virtual path of network mobility executed during our measurement runs using the pre-recorded GPS trace. We assume that 3G UMTS coverage is available during the whole route, while WLAN access networks – represented with colored circles – are to appear and disappear according to Figure 9 when the mobile network moves. The green, red and yellow circles represent overlapping WLAN networks with similar range, quality and transmission power, while the blue circle is for an umbrella WLAN coverage with bigger range but worst quality (in means of SNR). On this path consisting of several heterogeneous overlapping access networks, every handover type (3G=>Wi-Fi, Wi-Fi=>Wi-Fi and Wi-Fi=>3G) appears.

V. MEASUREMENT RESULTS

In order to evaluate the performance of our proposed handover mechanism in the presented testbed, an extensive comparison with existing implementations of different Mobile IPv6 based handover management schemes is necessary.

NEMO BS was chosen as basis to emphasize the drawbacks of using only a single media for horizontal handovers without prediction. Due to the limited functionalities of this mobility protocol the presence of network outage during handovers is expected, causing error-prone transport protocols such as UDP to perform suboptimal.

To fully take advantage of our heterogeneous test environment NEMO MCoA handovers were used to demonstrate the benefits of inter-media handovers in multihomed networks by manually changing data flows among multiple network interfaces and applying Flow Bindings and Policy Exchange for handover execution. The chance of packet loss and the handover delay is expected to be less compared to the NEMO BS case, as the MCoA protocol extension allows switching

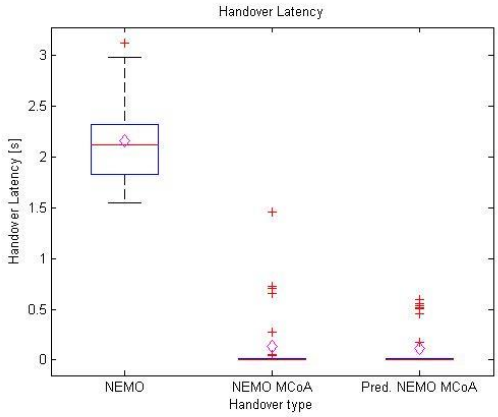


Fig. 10. Handover latency measurement results of multiple runs

data flows on already configured network interfaces. As this method lacks the prediction and automation features our approach has, we simulate handovers by changing to the first new available network blindly, emphasizing the risks what the absence of pre-recorded information can bring.

This second approach however is still expected to be outperformed by our GNSS aided predictive NEMO MCoA handover solution which uses automatic handover decisions based on various likelihood criteria, such as SNR/RSSI and reliability. The overall data throughput is expected to be the highest with our approach, as the amount of time spent on networks with good QoS parameters is maximized and the handover latency is minimized during any mobility cases.

The objective of our evaluation was to compare the above three handover methods and show the power of our framework. Therefore four main parameters were analyzed. In each scenario fifty measurements were executed. Netperf [25] was used for TCP and UDP packet flow generation while Tshark [26] was responsible for packet capture and analysis. The results are presented in box-and-whisker diagrams to display the collected numerical data groups in a compact way. The depicted statistical information are as follows: the lowest sample value (lower line), the lower quartile called Q1 (the lower edge of the box), the median called Q2 (the delimiter of the two distinctive colors of the box), the upper quartile called Q3 (the upper edge of the box), the largest sample value (the upper line), and the mean of the collected data (red lined rhombus). Red crosses are depicting outliers (i.e., measurement data if they are larger than Q3 + 1.5*(Q3 – Q1) or smaller than Q1 – 1.5*(Q3 – Q1)). The length of boxes (i.e., the interquartile range) represents the middle fifty percent of the measured data. Diamonds show the mean (average) value of the measurements, the solid line in the background depicts the range of measured data.

Figure 10 shows the results of handover latency measurements of multiple runs. Time stamped log messages and kernel events provide the measured latency in seconds passed between the handover decision and the availability of the new Mobile IPv6 tunnel interface on the Mobile Router. As NEMO BS only uses a single interface for handover operations, it

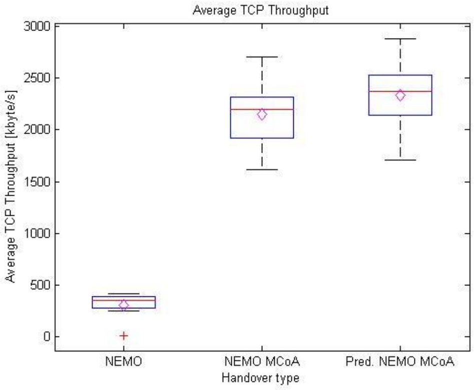


Fig. 11. TCP Throughput measurement results of multiple runs

showed significant delays while changing between different wireless networks. The gap in dataflow is partly caused by Layer 2 connection delays and Layer 3 operations such as IPv6 address acquiring from the access router and mobility signalling between the MR and the HA. In both MCoA cases the handovers took place on already configured network interfaces, where the mobility tunnels were already established. The only signalling on the channel was the above introduced policy exchange mechanism between the Mobile Router and the Home Agent. The handover latency in this scenario was measured based on the round-trip-time of the Predictive Policy Exchange Messages. As visible, there is not much improvement between NEMO MCoA and our method. The small improvement of Predictive NEMO MCoA comes from better network QoS parameters as it is capable of selecting the best networks along the path.

Figure 12 depicts our HO latency measurement results focusing on one single run on the virtual test-track. The figure shows that a simple NEMO BS system provides much higher HO delays compared to the advanced MCoA based solutions. It is also highlighted that our prediction based framework adapts to the actual network coverage more precisely while also maximising the time spent on a satisfactory RAN. The prediction based system chooses alternative networks, and follows the handover policy aiming to use 3G network only if Wi-Fi is not available. Our system does not connect to the Wi-Fi with bad SNR (i.e., Blue WLAN) while the MCoA handover is not able to differentiate between such parameters.

Our second test case in our evaluation was the measurement of TCP throughput between a Mobile Network Node and a Correspondent Node. The five-number-representation of TCP throughput is shown on Figure 11. The connection was not lost during the tests due to the error detection and flow control feature of the applied transport protocol.

The results justified our assumptions that the gap during the NEMO BS handovers significantly slows down the TCP stream, while during the MCoA handovers it remains stable on a much higher transfer rate. The Predictive NEMO MCoA handover outperformed all the others, thanks to the automated optimal network selection.

Figure 14 shows the boxplot of packet loss when transmitting a unidirectional UDP stream that originated from a MNN towards a CN (our third test scenario). The UDP stream was captured on both the MNN and the CN. After each run the packet loss ratio was calculated by dividing the sum of captured packages on both ends of the communication.

As the single communication medium is not ready during the NEMO BS handover for several seconds, and the transport layer protocol has no error recovery, there is a substantial amount of lost packets in that case. However, in both MCoA scenarios the packet loss remained below the acceptable 5% percent. The usage of Predictive NEMO MCoA handover clearly converges towards the ideal 0% in our evaluation scenario. Note that wireless transmission itself implies some packet loss, so 1% should be regarded really low.

The last test case was focusing on the RTT measurements. Figure 13 depicts our results gathered with one single run of test. The blue coloured Wi-Fi network was artificially degraded with Radio level settings and IP traffic shaping. In case of NEMO BS we can see that there are significant gaps in the connection around the handover points, meaning packet losses caused by managing handovers with only one active interface. In case of NEMO MCoA handovers our MR could also use the 3G access network. The figure clearly shows that without prediction the MR connects to every single Wi-Fi network sensed during the path. The continuity of the graph proves that no significant packet loss occurred, but the MR also used a bad quality Wi-Fi network with 400ms RTT (however, only for a limited amount of time, until the next Wi-Fi network was sensed by the MR on the road). Our predictive scheme avoids Wi-Fi networks with bad SNR if possible: the Blue Wi-Fi will not be used, the MR chooses the Red network right after the

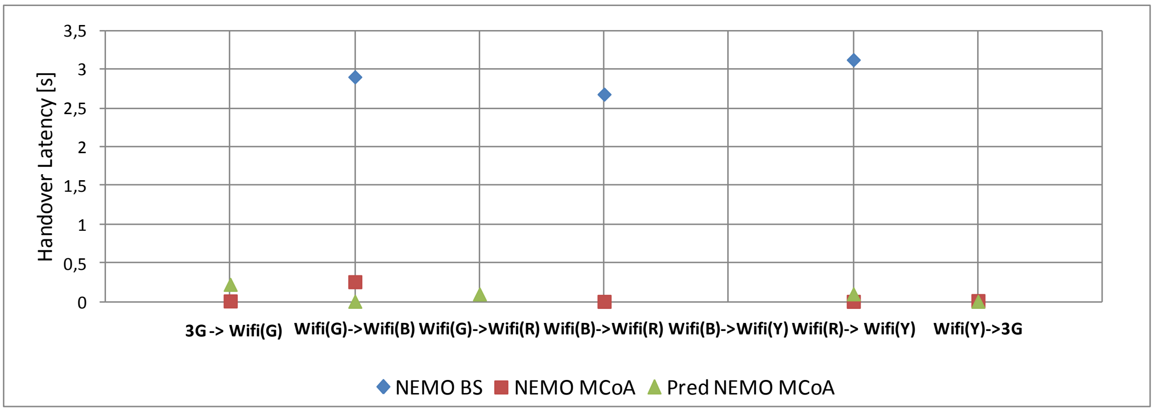


Fig. 12. Handover latency measurement results of one single run (i.e., one virtual path according to Fig. 9)

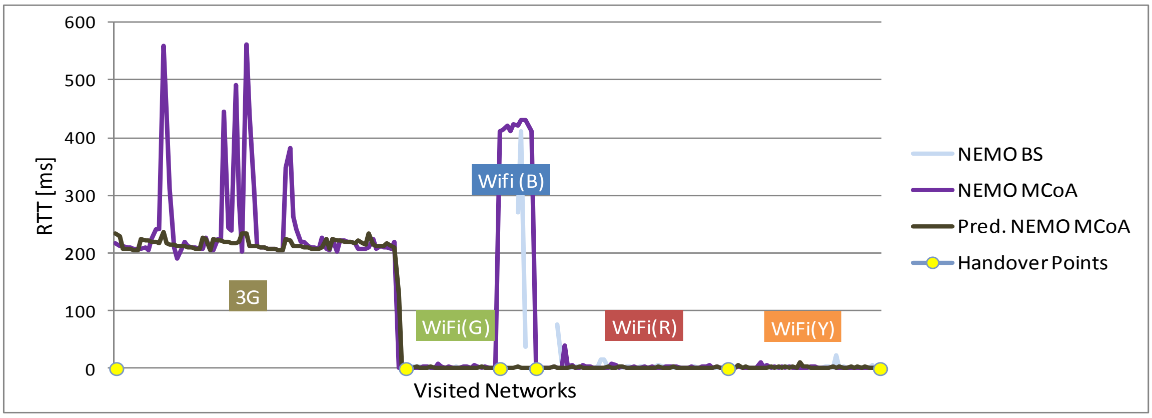


Fig. 13. RTT measurement results of one single run (i.e., one virtual path according to Fig. 9)

Green one, despite the fact that the Blue appears sooner during the movement. Networks with insufficient performance can be avoided.

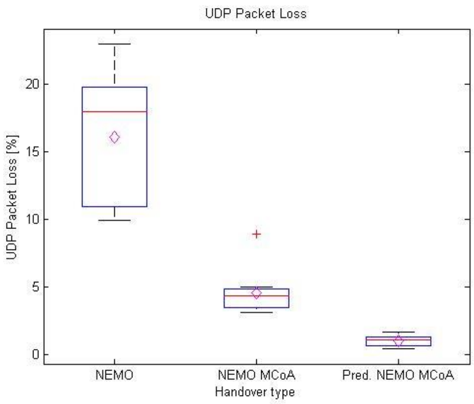


Fig. 14. UDP packet loss measurement results of multiple runs

VI. ANALYSIS OF PREDICTION ACCURACY

The proposed framework and handover execution protocol strongly relies on the prediction accuracy which depends on the rasterization scheme working inside the ANP module. That is why we have started to analyze the limitations of the overall architecture inherited by possible wrong positioning on the raster net inside the ANP. We have developed a probabilistic system model for the ANP module and proposed an appropriate rasterization scheme where the probability of wrong positioning on the raster remains below 1%.

Assume that we have a set of raster points given as $\mathcal{S} = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \dots \boldsymbol{x}_\infty\}$. $\boldsymbol{x}_i$ represents the ith point which is a geographical position with two coordinates: one on the west-east axis and one on the north-south axis. $\mathcal{S}$ is an infinite but countable set. The members of the set are constant: they are given by the actual raster size.

Assume that we are at a geographical position $\boldsymbol{x}_0$ ($\boldsymbol{x}_0$can be given by god – no possibility to measure it exactly). We have a GNSS measurement equipment and want to figure out, what $\boldsymbol{x}_0$ is. We make measurements and we get $\boldsymbol{\eta}$ as an estimate, which is not exact of course. $\boldsymbol{\eta}$ is a random number (Gaussian, due to the large number of independent effects), with expectation of $\boldsymbol{x}_0$ and covariance matrix $\boldsymbol{C}$:

$$\mathbb{E}\{\boldsymbol{\eta}\} = \boldsymbol{x}_0 \qquad (1)$$

$$Pr\{\boldsymbol{\eta} \le \boldsymbol{y} | \boldsymbol{x}_0, \boldsymbol{C}\} = \Phi(\boldsymbol{y}, \boldsymbol{x}_0, \boldsymbol{C}) = \int_{-\infty}^{y_1}\int_{-\infty}^{y_2} \frac{1}{\sqrt{2\pi}^2 (\det\boldsymbol{C})^{1/2}} e^{(\boldsymbol{z}-\boldsymbol{x}_0)^{\mathsf{T}}\boldsymbol{C}^{-1}(\boldsymbol{z}-\boldsymbol{x}_0)} dz_2 dz_1 \qquad (2)$$

Note that in the last equation we introduced a new notation, Φ. Also note that $\boldsymbol{\eta} \le \boldsymbol{y}$ means all $\boldsymbol{\eta}$ points where both coordinates are less than or equal to the ones of $\boldsymbol{y}$.

The database uses the raster points only. Thus, based on the measured value $\boldsymbol{\eta}$ we can choose the closest raster point as

$$\boldsymbol{\xi}(t) = \operatorname{argmin}_{\boldsymbol{x}\in\mathcal{S}} \parallel \boldsymbol{\eta}(t) - \boldsymbol{x} \parallel \qquad (3)$$

Here, the time dependence have been also added as (t), and $\parallel . \parallel$ measures the absolute distance. With the help of god (knowing $\boldsymbol{x}_0(t)$), we would get the perfect estimate $\boldsymbol{x}(t)$ as

$$\boldsymbol{x}(t) = \operatorname{argmin}_{\boldsymbol{x}\in\mathcal{S}} \parallel \boldsymbol{x}_0(t) - \boldsymbol{x} \parallel \qquad (4)$$

The first question is the following. What is the probability of making a wrong estimate?

$$Pr\{\boldsymbol{\xi}(t) \ne \boldsymbol{x}(t)\} = ? \qquad (5)$$

Note that both (4) and (5) are non-linear operations, making it difficult to analyse the problem. The following subsection is about evaluating this probability.

Fig. 15 shows the general geographical setup. As the raster net is self similar, we can put it into the centre of the coordinate system. The area of $\mathcal{T}$ is defined as

$$\mathcal{T} = \{(i,j),\ where\ -a \le i \le +a\ and\ -b \le j \le +b\} \qquad (6)$$

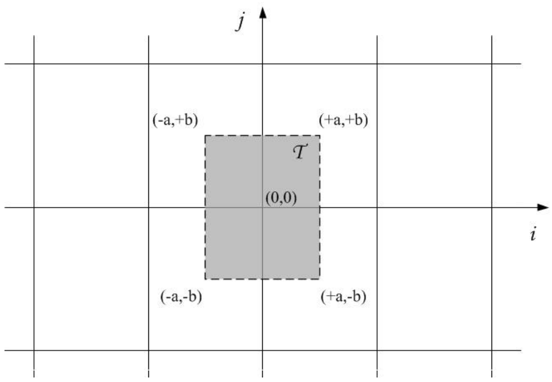


Fig. 15. Raster net setup of the probability model

Assuming that the real geographical position ($\boldsymbol{x}_0$) is equally probable at any position, the probability of making a wrong estimate ($Pr\{\boldsymbol{\xi}(t) \neq \boldsymbol{x}(t)\}$), equals the probability that the real position is inside the grey area $\mathcal{T}$ ($\boldsymbol{x}_0(t) \in \mathcal{T}$), and the measured point is outside of $\mathcal{T}$ ($\boldsymbol{\eta}(t) \notin \mathcal{T}$):

$$Pr\{\boldsymbol{\xi}(t) \neq \boldsymbol{x}(t)\} = Pr\{\boldsymbol{x}_0(t) \in \mathcal{T} \cap \boldsymbol{\eta}(t) \notin \mathcal{T}\} = 1 - Pr\{\boldsymbol{x}_0(t) \in \mathcal{T} \cap \boldsymbol{\eta}(t) \in \mathcal{T}\} \quad (7)$$

The probability of $\boldsymbol{\eta}$ falling into $\mathcal{T}$ can be computed as

$$Pr\{\boldsymbol{\eta} \in \mathcal{T} | \boldsymbol{x}_0, \boldsymbol{C}\} = \Phi\left(\begin{pmatrix}+a\\+b\end{pmatrix}, \boldsymbol{x}_0, \boldsymbol{C}\right) + \Phi\left(\begin{pmatrix}-a\\-b\end{pmatrix}, \boldsymbol{x}_0, \boldsymbol{C}\right) - \Phi\left(\begin{pmatrix}+a\\-b\end{pmatrix}, \boldsymbol{x}_0, \boldsymbol{C}\right) - \Phi\left(\begin{pmatrix}-a\\+b\end{pmatrix}, \boldsymbol{x}_0, \boldsymbol{C}\right) \quad (8)$$

Following equation (7), the Bayes' rule and our positioning error constraint, we get

$$0.01 \geq Pr\{\boldsymbol{\xi}(t) \neq \boldsymbol{x}(t)\} = 1 - \frac{1}{4ab}\int_{-a}^{+a}\int_{-b}^{+b} Pr\{\boldsymbol{\eta} \in \mathcal{T} | \boldsymbol{x}_0, \boldsymbol{C}\}_{\boldsymbol{x}_0 = \binom{i}{j}} dj\,di \quad (9)$$

Considering a GPS system for our GNSS measurements with a horizontal positioning error of $\sigma = 5m$ (standard deviation), and taking into consideration that the length in one minute of longitude depends on the latitude (which is ~47.5° N for Budapest) our final equation is

$$0.01 \geq 1 - \operatorname{erf}\left(\frac{\frac{x}{2} \cdot cos(47.5°)}{5}\right) \cdot \operatorname{erf}\left(\frac{\frac{x}{2}}{5}\right) \quad (10)$$

Solving (10) where erf is the error function ($\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}}\int_0^x e^{-t^2}\,dt$) we get that the appropriate raster net to be used in our predictive NEMO handover framework for ANP implementation is larger or equal to $18.2m$ x $27m$.

VII. CONCLUSION

This paper shows that GNSS aided prediction along with other optimization techniques provide the best results in multihomed NEMO configurations. The implemented system with the decision and prediction module outperforms all existing handover solutions in Mobile IPv6 scenarios. Handover latency has been shrunk to L1/L2 handover times or even below. Figure 16 compares the evaluated schemes. Horizontal handover –a basic function – is available for all solutions. However, vertical handover support requires multiple tunnel management, which requires MCoA capability. Automatic layer 2 connection is a function to manage interface connections in a cross-layer manner to provide adaptivity in mobility management. In case of simple NEMO MCoA handover it is only manually achievable. Access Network Prediction is a higher layer intelligence also requiring efficient cross-layer communication. With the help of our probability model based analysis, the details of the crucial ANP module can be wisely selected: we provided analytical method for optimal design of the ANP's raster network.

For the above advanced functions, a well-prepared policy manager is required to handle the rules of handover initiation and execution even in a dynamically changing environment. The system of such rules creates an adaptive order of priority like in our example, where in case of overlapping 3G/Wi-Fi networks Wi-Fi is preferred, but from multiple Wi-Fi RANs the one with the best SNR should be chosen. To execute the handovers in our proposed scheme, Flow Bindings and Policy Exchange is required, which is also mandatory for the simple MCoA handover solution.

In the future we plan to implement pluggable decision and prediction modules that further optimize network selection on various types of transportation scenarios.

We also plan to extend the solution with possible improvements in the sub modules of the system. For instance, gyroscope could be used to improve localization in places where GNSS systems are not able to work (e.g., in tunnels). Databases could be built quicker if P2P sharing of the database is supported. That is, Mobile Routers could improve their knowledge if they share their database with neighboring routers, or infrastructure based information arrives (e.g., from the road operator). Obviously, security considerations should be addressed first, before opening the databases.

Function/Handover	NEMO BS	NEMO MCoA	Predictive NEMO MCoA
Horizontal handover	☑	☑	☑
Vertical handover	☒	☑	☑
Automatic Layer 2 connection	☒	☒	☑
Access Network Prediction	☒	☒	☑
Handover Policy Manager	☒	☒	☑
Flow Binding / Policy Exchange	☒	☑	☑

☑	Available	☒	Out of scope		Implemented		Not implemented

Fig. 16. Functional comparison of the evaluated schemes

ACKNOWLEDGEMENT

The publication was supported by the TÁMOP-4.2.2.C-11/1/KONV-2012-0001 project. The project has been supported by the European Union, co-financed by the European Social Fund. Initial parts of this work also received funding from the NKTH Déri Miksa BOSS project Grant OMFB-01467/2006. The authors would like to acknowledge the contributions of their colleagues in both projects, especially to Zsigmond Mihály for his valuable comments and remarks.

REFERENCES

[1] T. Ernst, "The Information Technology Era of the Vehicular Industry," *SIGCOMM Comput Commun Rev*, vol. 36, no. 2, pp. 49–52, Apr. 2006.

[2] A. Stevens and J. Hopkin, "Benefits and deployment opportunities for vehicle/roadside cooperative ITS," 2012, pp. 1–6.

[3] L. Bokor, N. Montavont, P. Di Francesco, T. Ernst, T. Hof, and J. Korva, "ANEMONE: A Pan-European Testbed to Validate IPv6 Mobility Technologies," 2007, pp. 44–44.

[4] T. Ernst, L. Bokor, A. Boutet, and Y. Lopez, "An Open Network for Testing, Verification and Validation of IPv6-based ITS Components," in *Proc. of ITST 2007*, Sophia Antipolis, 2007, pp. 1–6.

[5] L. A. DaSilva, S. F. Midkiff, J. S. Park, G. C. Hadjichristofi, I. Davis, N.J., K. S. Phanse, and T. Lin, "Network mobility and protocol interoperability in ad hoc networks," *Commun. Mag. IEEE*, vol. 42, no. 11, pp. 88–96, Nov. 2004.

[6] C. Perkins, D. Johnson, and J. Arkko, *Mobility Support in IPv6*. IETF, 2011.

[7] V. Devarapalli, R. Wakikawa, A. Petrescu, P. Thubert, *Network Mobility (NEMO) Basic Support Protocol*. IETF, 2005.

[8] R. Wakikawa, V. Devarapalli, G. Tsirtsis, T. Ernst, and K. Nagami, *Multiple Care-of Addresses Registration*. IETF, 2009.

[9] G. Tsirtsis, H. Soliman, N. Montavont, G. Giaretta, and K. Kuladinithi, *Flow Bindings in Mobile IPv6 and Network Mobility (NEMO) Basic Support*. IETF, 2011.

[10] G. Tsirtsis, G. Giarreta, H. Soliman, and N. Montavont, *Traffic Selectors for Flow Bindings*. IETF, 2011.

[11] R. Koodli, *Fast Handovers for Mobile IPv6*. IETF, 2005.

[12] Y.-H. Wang, K.-F. Huang, and H.-Y. Ho, "A Seamless Handover Scheme with Pre-registration in NEMO," in *Advanced Information Networking and Applications Workshops, 2009. WAINA '09. International Conference on*, 2009, pp. 338–344.

[13] C.-W. Lee, Y. S. Sun, and M.-C. Chen, "HiMIP-NEMO: Combining Cross-Layer Network Mobility Management and Resource Allocation for Fast QoS-Handovers," in *Vehicular Technology Conference, 2008. VTC Spring 2008. IEEE*, 2008, pp. 2282–2286.

[14] H.-D. Park, Y.-H. Kwon, K.-W. Lee, Y.-S. Choi, S.-H. Lee, and Y.-Z. Cho, "Network Mobility Management Using Predictive Binding Update," in *Distributed Computing – IWDC 2005*, vol. 3741, A. Pal, A. Kshemkalyani, R. Kumar, and A. Gupta, Eds. Springer Berlin Heidelberg, 2005, pp. 560–565.

[15] S. S. Wang and C.-H. Wu, "Effective handoff method using mobile location information," in *Vehicular Technology Conference, 2001. VTC 2001 Spring. IEEE VTS 53rd*, 2001, vol. 4, pp. 2585–2589 vol.4.

[16] A. Dutta, S. Chakravarty, K. Taniuchi, V. Fajardo, Y. Ohba, D. Famolari, and H. Schulzrinne, "An Experimental Study of Location Assisted Proactive Handover," in *Global Telecommunications Conference, 2007. GLOBECOM '07. IEEE*, 2007, pp. 2037–2042.

[17] Jeney Gabor, Bokor Laszlo, and Mihaly Zsigmond, "GPS aided predictive handover management for multihomed NEMO configurations," 2009, pp. 69–73.

[18] Kovacs Jozsef, Bokor Laszlo, and Jeney Gabor, "Performance Evaluation of GNSS Aided Predictive Multihomed NEMO Configurations," in *ITST-2011: 2011 11th International Conference on ITS Telecommunications*, Institute of Electrical & Electronics Engineers (IEEE), 2011, pp. 293–298.

[19] Jozsef Kovacs, Laszlo Bokor, Zoltan Kanizsai, and Sandor Imre, "Review of Advanced Mobility Solutions for Multimedia Networking in IPv6," in *Intelligent Multimedia Technologies for Networking Applications: Techniques and Tools*, D. Kanellopoulos, Ed. Hershey: IGI Global, Information Science Reference, 2013, pp. 25–47.

[20] S. Thomson, T. Narten, and T. Jinmei, *IPv6 Stateless Address Autoconfiguration*. IETF, 2007.

[21] R. Droms, J. Bound, B. Volz, T. Lemon, C. Perkins, and M. Carney, *Dynamic Host Configuration Protocol for IPv6 (DHCPv6)*. IETF, 2003.

[22] M. Han, K.-S. Han, and D.-J. Lee, "Fast IP Handover Performance Improvements Using Performance Enhancing Proxys between Satellite Networks and Wireless LAN Networks for High-Speed Trains," in *Vehicular Technology Conference, 2008. VTC Spring 2008. IEEE*, 2008, pp. 2341–2344.

[23] "Mobile Innovation Centre Hungary," Mar-2014. [Online]. Available: http://mik.bme.hu/.

[24] L. Bokor, Z. Kanizsai, and G. Jeney, "Setting up Native IPv6 3G UMTS Access" ANEMONE Tutorial," Mar-2014. [Online]. Available: http://anemone.mik.bme.hu/index.php?title=Setting_up_native_IPv6_UMTS_access_with_open-source_GGSN_implementation.

[25] "NetPerf," Mar-2014. [Online]. Available: http://www.netperf.org.

[26] "Tshark," Mar-2014. [Online]. Available: http://www.wireshark.org.

[27] "UMIP - USAGI-patched Mobile IPv6 for Linux," Mar-2014. [Online]. Available: http://umip.org.

[28] "GPSD - a GPS service daemon," Mar-2014. [Online]. Available: http://www.catb.org/gpsd/.

László Bokor graduated in 2004 with M.Sc. degree in computer engineering from the Budapest University of Technology and Economics (BME) at the Department of Telecommunicatons. In 2006 he got an M.Sc.+ degree in bank informatics from the same university.s Faculty of Economic and Social Sciences. He is a Ph.D. candidate at BME, member of the IEEE, member of Multimedia Networks and Services Laboratory (MEDIANETS) and Mobile Innovation Centre Hungary (MIK) where he currently participates in researches of wireless protocols and works on advanced mobility management related projects such as FP7-ICT CONCERTO, EUREKA-Celtic SIGMONA and TÁMOP-4.2.2.C-11/1/KONV-2012-0001 FIRST. His research interests include IPv6 mobility, mobile computing, next generation networks, mobile broadband networking architectures, network performance analyzing, and heterogeneous networks. He is an author of several referred papers in international scientific journals and conferences.

Gábor Jeney was born in Miskolc, Hungary, in 13th January 1975. He received the M.Sc. degree in Electrical Engineering in 1998 from the Technical University of Budapest, Hungary in the field of mobile and optical communications. He received the Ph.D. degree in Electrical Engineering in 2005 from the same university, renamed meanwhile as Budapest University of Technology and Economics, Hungary in the field of multi-user detection. He also holds an M.Sc.+ Engineer-economist degree from the Budapest University of Economic Sciences and Public Administration, Hungary. He is currently a research fellow at the Mobile Innovation Centre, Budapest University of Technology and Economics, Budapest, Hungary. He is involved in several national and European Commission funded project. He was the author/coauthor of several dozens of articles published in conferences and journals. His research interest includes neutral networks, mobile and wide-band communication systems, networking and transport protocols. Dr. Jeney is a member of HTE (Scientific Association for Info-communications), Hungary since 1999, and he is a member of IEEE since 2000.

József Kovács (M.Sc.) is a research associate at Institute for Computer Science and Control, Hungarian Academy of Sciences (MTA SZTAKI), Budapest, Hungary. He is also a Ph.D. student in Computer Science and Information Technology at the Széchenyi István University, Győr. He received his M.Sc. degree in Information Technology at Budapest University of Technology and Economics (BME). His main interest is communication protocol research and development for vehicular (VANET) and ad-hoc sensor networks. In the past years he participated in several mobility related research projects with special focus on the use of Mobile IPv6 protocols, non-IP V2X protocols and performance analysis of VANETs. Currently he is working with cooperative systems as part of his research at MTA SZTAKI.

RCTP: A Low-complexity Transport Protocol for Collecting Measurement Data

Péter Orosz, Tamás Skopkó, Máté Varga

***Abstract*— Transmission Control Protocol (TCP) and User Datagram Protocol (UDP) are the essential transport protocols of the Internet Protocol (IP) networks. Both of them are dedicated for certain purposes and face their obvious limitations. Thus, there is a constant endeavor for developing alternatives. Despite the reliability feature of TCP, its relatively high complexity does not always enable to implement it in a hardware environment with constrained resources. Our paper introduces a low-complexity transport protocol dedicated to a real-time network monitoring system operating at 10+ Gbps. Its task is to transport the preprocessed IP packets from the monitoring device to the post-processing hosts without loss over a dedicated LAN. Resource requirement on sender side has to be reduced as much as possible while trying to maintain high throughput. Although RCTP is intended to serve in a network measurement system, it may be suitable for other measurement infrastructures such as sensor networks, where data provided by the sensors with limited resources have to be collected at a central node.**

Keywords- Transport protocol; datagram; UDP; monitoring system.

I. INTRODUCTION

Passive traffic monitoring is an essential task for network infrastructure providers. There is a growing demand on high processing capacity to monitor core networks. In order to support lossless packet-level monitoring, dedicated hardware-accelerated equipments should be deployed on certain links of the infrastructure. Besides network processors, an increasing number of solutions are based on Field Programmable Gate Arrays (FPGA) as central building block. This technology ensures flexibility for integrating services into one device and performing real-time processing, while its resources (e.g., number of logical units and available on-chip memory) are always limited. The transport protocol introduced in this paper is optimized to distributed measurement architectures with constrained hardware resources and is part of a high performance monitoring system using FPGA-acceleration for packet capturing, parsing and filtering.

Each probe device is tapping one specific direction of a 100 Gbps link with passive optical splitter as seen in Fig. 1. Received packets are preprocessed within the FPGA. These tasks include high precision timestamping, protocol parsing and classification of the packets.

Manuscript received June 19, 2014, revised August 31, 2014. P. Orosz, T. Skopkó and M. Varga are with the Faculty of Informatics, University of Debrecen, Debrecen, H-4028 Hungary, (emails: oroszp@unideb.hu, skopkot@unideb.hu, varga.mate.unideb@gmail.com).

Packet flows are also identified and forwarded to collector agents, where deeper inspection, heuristic recognition and statistical analysis are done. Up to 10 collectors are connected with 10 Gbps links to one probe device and the monitored traffic of the 100 Gbps link is distributed uniformly between the 10 Gbps links. One specific flow is forwarded to one specific collector only. The transport protocol introduced is intended to serve the transmission of captured traffic from the probe to the collectors.

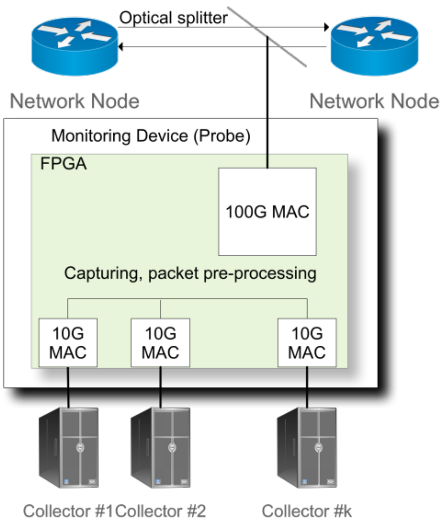


Figure 1. The monitoring system tapping a 100 Gbps link and distributing monitored data to collectors after preprocessing tasks are done

Nevertheless, the proposed protocol can be integrated to any measurement infrastructure where multiple monitoring devices are generating and transmitting measurement data towards processing agents over the network. In an FPGA-accelerated system, the high processing capacity is achievable by using pipelining technique, hence a number of modules (interface MACs, transport protocols, etc) are placed into the chip in multiple instances [1]. Therefore, a transport protocol that is implemented in hardware should have a low resource usage. Many TCP variants provide reliable transmission with high throughput [2][3][4][5][6][7]. However, implementing multiple instances of the protocol inside an FPGA chip results in a significant resource utilization. In our measurement system the captured and preprocessed traffic of an aggregated 10+ Gbps link is distributed to multiple 10 Gbps outgoing connections towards post-processing hosts. Considering the required memory, TCP retransmission buffer has to be implemented in an off-chip RAM. However, accessing an external memory by many concurrent TCP instances would build up a serious bottleneck at line rate. To overcome processing bottlenecks we commonly used offloading techniques in our former research projects; converting the complex and resource-intensive operations to delay-tolerant processes generally improve the overall performance [8]. In our system the monitoring device itself is connected to the agents directly or through a dedicated LAN. This private infrastructure can forward packets reliably. In our design, transmission control is offloaded from the sender to the receiver in the form of rate control. Therefore, sender side is simple enough to be implemented even in a resource- and energy-constrained hardware. We improved its performance further by packetizing the data into MTU-sized packets before transmission, thus reducing packet and interrupt rates at the receiver.

In Section II, TCP, the most feature-rich and widely applied transport protocol is set in contrast with a more lightweight UDP as well as a number of alternative protocols with reliability feature. Linux packet processing is reviewed from point of lossless packet reception in Section III. In Section IV our custom protocol is introduced in detail. Section V is about the performance evaluation of the protocol using a hybrid FPGA and Linux implementation at 10 Gbps.

II. RELATED WORKS

TCP is a set of algorithms ensuring the adaptation of the protocol to various traffic situations. Reliable transmission is its basic feature, which is implemented using a positive acknowledgement mechanism. Buffering unacknowledged data makes possible to retransmit the whole buffered byte stream or a part of it when packet loss occurs. Researches proposed formulas for optimal sizing of buffers which is commonly specified by the bandwidth-delay product, to increase the performance of the protocol [9]. The size of the retransmission buffer at 10+ Gbps bandwidth, adjusted with this formula, exceeds the megabyte limit. This renders an efficient hardware implementation difficult to design. The complex logic required by TCP's algorithms make the situation even challenging. T. Uchida implemented TCP in FPGA for Gigabit Ethernet [10]. His circuit took up about 3000 slices. It uses block RAM buffering, which is sufficient only for storing small amount of data. We ruled out this implementation since our application requires wire speed transmission at 10+ Gbps, which implies large memory buffers.

By its simplicity, UDP is an optimal basis for a transport protocol dedicated for hardware implementation. Reliability feature is normally implemented in an upper protocol layer or in the application itself, by using extra buffer. UDP Lite is a variant of UDP that enables the forwarding of corrupted packets as well. One of the first attempts of combining simple message-oriented communication with reliable transmission into a lightweight transport protocol was Reliable UDP (rUDP) developed at Bell Labs [11]. It resembles TCP's retransmission mechanism with a windowing technique very similar to TCP resulting in the demand on extra buffer. Although Cisco and Microsoft have their own implementation of this protocol, it remained in draft status at IETF.

Y. Gu and R. L. Grossman worked out an UDP-based transport protocol for WAN purposes [12]. It uses acknowledgements similar for TCP but it is timer-based and makes possible to acknowledge more frequently on slow or unreliable connections. Its main bottleneck for our application is its requirement for sender buffer.

E.Kohler *et al.* introduced Datagram Congestion Control Protocol (DCCP) [13]. It provides a TCP-like feature set with reliable packet transmission and focuses on congestion control using Explicit Notification Control (ECN). Though it does not support in-order delivery of packets, it is not an explicit requirement for our application since packets are timestamped by the monitor hardware. Its real bottleneck is the recovery process triggering by a series of lost packets that vindicates the demand for buffering high amount of data.

An effective combination of TCP and UDP called Stream Control Transmission Protocol (SCTP) was presented by R. Stewart and C. Metz [14]. It supports fragmentation of transmitted data into chunks and thus enabling to overcome some limitations of TCP by making some of its features disabled. It can reliably transmit data but since no system bottlenecks are taken into account, it retransmits lost data. This makes data buffering necessary and thus this protocol is also ruled out.

The same limitation is present if we consider Reliable Datagram Sockets (RDS) developed at Oracle [15]. This transport protocol is designed for inter-process communication (IPC) on InfiniBand and primarily benefits from bidirectional dataflow because acknowledgements are placed at the end of data payload. Although it avoids losses caused by a saturated socket buffer, it does not care about other factors such as processing capacity bottlenecks.

Scalable and Secure Transport Protocol (SSTP) is intended to serve as a scalable transport for metering systems [16]. The protocol presented by Y.J. Kim *et al.* ensures reliable transfer of data for short flows as well as data encryption. Although it is less complex compared to TCP it still contains functions (i.e., fairness and secure layer) unnecessary in our environment.

Y.J. Kim and M. Thottan also presented Smart Grid Transport Protocol (SGTP) as an alternative transport protocol for collecting periodic measured data but it focuses on short-term flows consisting of small packets [17]. Our monitoring system produces long flows and large packets.

Lossless transmission of metered or monitored data through various network conditions (e.g. 802.11x, etc.) is a general requirement. Thus transport protocols intended for these applications generally include buffering and retransmission of lost data. This is not a problem when only short packet flows are present. At the same time it is not likely to achieve line rate transmission performance. Probes collecting the measurement information are generally low performance hardware with a requirement of low resource consumption.

In contrast, data center transport protocols take aim at high throughput besides the demand on reliable transmission. Generic hardware is applied in this environment with the feasibility to use more complex protocols and larger retransmission buffers. Replacement of TCP is researched to achieve better transmission performance.

Our monitoring system differs from the aforementioned monitoring and data center applications. Although information flow is unidirectional, network probes can collect high amount of measurement data. Even though FPGAs have high processing performance, their resources are limited. Amongst monitoring and preprocessing functions, many instances of a chosen transport protocol have to be compiled in. This transport protocol has to occupy as low resources as possible. Throughput has to be maximized and no loss of measured data should occur.

Like in most data centers we also have a dedicated packet forwarding infrastructure that ensures lossless transmission. This implies that a packet loss can be caused by the collector's operating system or data processing application. Also there is only one packet flow per agent, hereby no flow will be interfering and no fairness has to be implemented.

Since we did not find any suitable protocol to fit our special requirement, we created a new datagram-based protocol. The design was driven by the sender-side simplicity and reception-side flow control.

III. Background

In a packet processing system packet loss occurs when any of the buffers in the data path gets saturated. There are three critical buffers for UDP reception in the Linux kernel. Since these buffers are critical factors from the view of our protocol, we inspected their behavior more closely. The first one is related to the network interface card (NIC) hardware, it is the RX ring buffer [18]. This buffer is relatively small and its size is fixed or limited to a certain range, if configurable at all. A high performance NIC and its driver can process this queue at wire speed therefore no extra attention needed to take when using enterprise grade NIC such as Intel's server cards. Modern NICs can handle multiple RX queues. Each RX queue can be assigned to a different flow. In our measurement system there is only one packet flow to a specific agent, consequently advances of multiple RX queues cannot be exploited. Moreover current packet processing architectures in the generic operating systems are not designed to parallelize the processing of one packet flow between multiple CPU cores. However utilizing multiple cores is an important factor for preparing our architecture.

The second RX buffer is the ingress buffer or backlog queue. Packets are copied to this queue from the RX ring by the interrupt handler. Modern NAPI based drivers omit this step, hardware interrupts (hard IRQs) are performing critical tasks instead. Our protocol implementation is designed for NAPI supported NICs, thus no consideration is made to the load of the backlog queue. In case of NAPI drivers executing non-critical packet processing tasks are up to the software interrupts (soft IRQs), hereby minimizing the CPU time spent in hardware IRQ context. Moreover soft IRQs can be run simultaneously (with the limitation of single flow per CPU). The most important constraint for this kernel task is the processing capacity of the CPU core executing the soft IRQ for the corresponding RX interface (soft RX). On contemporary generic PC architectures at 1 Gbps or higher data rate, processing high intensity traffic containing small sized packets saturates the CPU. To maximize the performance of these architectures, it is advisable to use MTU-sized packets, especially if latency is not a critical factor.

The last kernel-related buffer is the sockets' receive buffer. A user space application receives data through this queue. The socket buffer is not a ring buffer. If the application cannot keep up with processing, the kernel will drop packets not fitting into the socket buffer. Size and occupancy of this buffer have to be considered if reliable transmission is an aim.

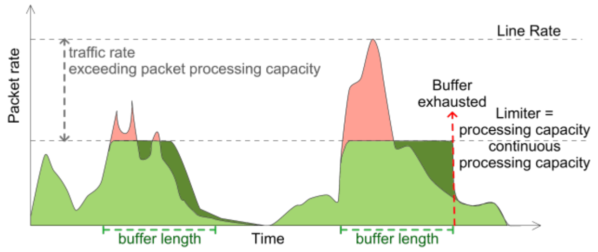


Figure 2. Packet processing performance behaving like a traffic shaper

Performance of software based packet processing is strictly limited by the processing capacity of the CPU core running the processing task. Intensive bursts overwhelm the processing CPU core. In this case, a smaller amount of packets are processed immediately, while the remaining packets can be processed when no new packets arrive. By using larger buffers the number of these temporarily unprocessed packets can be increased. When processing longer buffers we can also expect higher latency in packet forwarding. Also, larger buffers compensate processing bottleneck temporarily. In this aspect, packet processing subsystem of the operating system acts like a traffic shaper as seen in Fig. 2 where CPU's processing capacity corresponds to a limited bandwidth. If there is a higher rate of traffic at its input, it lengthens the traffic in time with a lower de-queuing rate.

We also examined the Ethernet flow control mechanism as a possible method for controlling the packet rate. We concluded that it is not efficient enough because it allows only a pulse-width modulation (PWM) type shaping of the traffic and control packets should be sent out at a very high rate.

IV. Introduction to RCTP

The control mechanism in RCTP is based on a closed-loop controller. Fig. 3 shows the feedback mechanism of the protocol. Sender is instructed by the receiver to send packets at a specified rate. The rate is adjusted by a control function based on resource usage. Sender side (inside a monitoring hardware) is low-complexity. Its task is the send UDP packets at the current limiting rate. Software-based reception side (processing agent) is not strictly limited in resources, thus it can implement and handle more complex tasks. We have chosen the generic Linux operating system as implementation ground.

Continuous monitoring of critical system resources enables us to avoid packet loss. As we unfolded in Section III, the most important metrics are the current usage of the RX socket buffer and the current load of CPU core running the soft RX. These parameters are monitored at a fixed sampling frequency specified at the initialization phase. These values are fed into the controller function. For output, the function calculates an enforced packet rate that the agent is able to keep up with the processing considering the current system load and buffer occupancy. Our experiments show that different control functions and parameters are not equally efficient for every metrics. They are detailed in subsection IV.D.

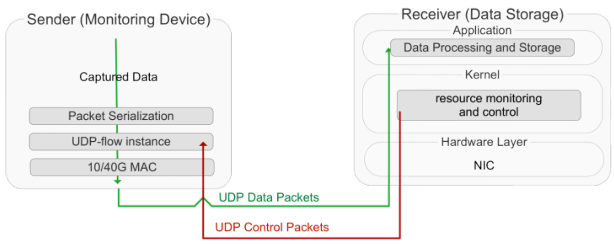


Figure 3. Rate and control mechanism of RCTP transported data

A. Sender-side operation

Data sender implements a simple logic packetizing the byte stream to MTU-sized UDP packets. These are transmitted to the agent. Transmission of a packet is scheduled by the enforced sending rate. Fig. 4 displays transmission loop based on the forced packet rate and the action to take upon reception of a control packet. When a rate control packet is received, a new packet rate is calculated to schedule later packet transmissions. Only a small-size FIFO is required for packing the payload and shaping the traffic to the enforced rate. This can be implemented in block RAMs inside the FPGA.

A timeout parameter built into the packetizer can trigger flushing data from the FIFO even if there is not enough data to construct an MTU-sized packet. The latency caused by packetizing and delayed flushing is not a bottleneck in the monitoring system.

B. Receiver-side operation

Reception side consists of two parts. The receiving application is a user space program capable of receiving UDP packets. Its socket buffer size has to be adjusted considering connection bandwidth and network delay.

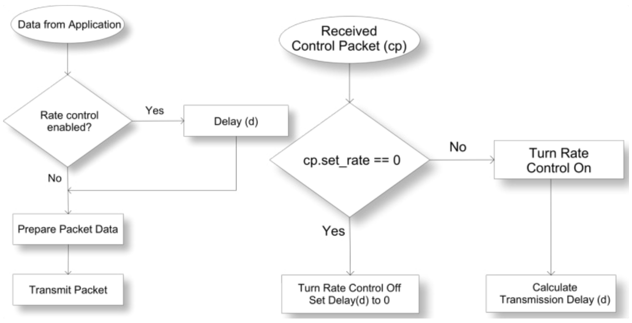


Figure 4. Sender-side packet transmission based on current delay (left) and control packet feedback mechanism (right)

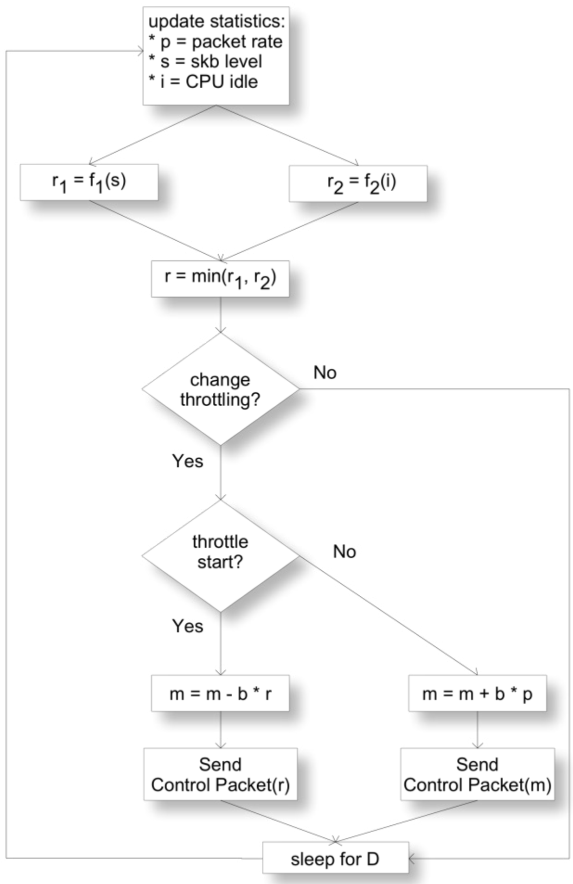


Figure 5. Rate control mechanism operating at the receiver side

Resource monitoring and packet loss avoidance is done by the controller logic. Since its task is time-critical, the controller should be implemented as a kernel process. This ensures a faster feedback. Although we implemented the protocol's control logic in a Linux kernel module, it can be ported to any operating systems capable of reporting the corresponding performance metrics for the controller's input and its exclusive execution can be assured.

Fig. 5 depicts the rate calculation and the signaling mechanism. Current packet rate, socket buffer occupancy and idle CPU cycles are monitored. Both for the socket buffer usage and the CPU load, the desired packet rates are calculated by the control function (denoted by f_1 and f_2 respectively). The minimum of these rate values is taken as the effective packet rate. If throttling state is changed, control packet will be transmitted. In the throttling start phase, the previously calculated new packet rate is sent out, while in throttling stop phase, sender is informed about the maximal packet rate. The initial value of the maximal packet rate is started at the nominal packet rate derived from the bandwidth. On throttle start, it is decreased based by the product of the new rate and the constant b_1. On throttle stop, it is increased by the product of the current packet rate and the constant b_2. This avoids flapping effect and helps to stabilize throughput. The values for b_1 *and* b_2 were determined empirically (see Section V).

C. Preparing the Operating System

Performance of RCTP depends on certain hardware and operating system factors. Bandwidth is an important factor. As we unfolded in Section III, buffer sizes have to be adjusted considering the link bandwidth and the processing capacity.

The minimum network buffer (skbuff) and socket buffer size can be determined by (1):

$$S = 2 \times B \times \frac{1}{F} \times D / 8 \tag{1}$$

where S is the size in bytes, B is the connection bandwidth in bps, F is the sampling frequency and D denotes the roundtrip time (seconds), since control information should also reach the sender. Multiplication by 2 ensures the elasticity to compensate jitter caused by scheduling the sampling task.

On a Linux-based reception side, the packet processing related system control (sysctl) parameters to adjust are: *net.core.rmem_default*, *net.core.rmem_max*, *net.ipv4.udp_rmem* and *net.ipv4.udp_rmem_min*.

Optimizing of *net.core.netdev_max_backlog* is not necessary when using NIC drivers with NAPI support. Recent 10+ GbE drivers are prepared to polling mode since high bandwidth cannot be processed efficiently using the traditional 1 interrupt per *k* packet technique.

Modern operating systems utilize an IRQ balancer to distribute soft IRQs among CPU cores efficiently. Its aim is a uniform distribution of soft IRQ load in a multiprocessor system. There are some critical tasks for RCTP that are intended to be explicitly executed on dedicated CPU cores and increases the probability of leaving the information for execution in a lower level CPU cache. This ensures a more predictable scheduling and processing of the corresponding tasks. On a Linux system, the affinity mask specifies the cores that are allowed to run a certain soft IRQ, and can be configured using the *smp_affinity* interface in the */proc/irq* interface tree. For a user space application the *taskset* utility can be used for the same purpose. For an explicit assignment, the *cgroup/cpuset* toolset should be invoked.

The optimal distribution strategy for a multicore system is the following: one core for the RCTP kernel process, one core for the interface's soft RX, one other core for the transmission software IRQ (soft TX), and at least one core is dedicated for the user space application. If soft TX is separated, there is more chance to transmit the control packet with shorter delay resulting in a shorter feedback time.

There is no point for using real-time and preemptive kernel. Since time critical performance monitoring is done in kernel space on a carefully configured system, no user space application will disturb the execution of the control process. Using large buffers eliminates the advantages of a real-time kernel.

Performance metrics feeding into the control function are monitored at a fixed sampling rate. Resolution of statistics is directly affected by the kernel ticks (jiffies), especially for the CPU related metrics. For optimal result a 1000 Hz kernel should be used. This allows a sampling period down to 1 ms. However, an optimal trade-off between sampling frequency and feedback accuracy should be determined. The higher the sampling frequency, the lower the resolution of CPU idle statistics report, resulting lower accuracy of optimal transmit rate prediction. Though reducing scheduling-clock ticks (NO_HZ kernel option) helps to save energy, it deforms CPU idle statistics therefore it should be turned off.

Modern network interface cards support offloading of certain packet processing operations to leave more CPU time for other tasks. These features affect the performance of RCTP also. RX checksumming offloads the verification of UDP checksum to hardware. Enabling this feature is recommended. Since no fragmentation feature is used at the construction of packets, UDP fragmentation offload (UFO) feature cannot be exploited, it can be left disabled. Generic receive offload (UFO) and large receive offload (LRO) allow combining smaller sized UDP packets into a larger, near MTU-sized one. These features cannot be exploited since we are already using MTU-sized packets. Therefore, they would deform the packet counter in the kernel rendering the protocol's control function unusable.

D. Control functions

Current packet rate and monitored performance metrics are input parameters of the control function to calculate the new advertised packet rate.

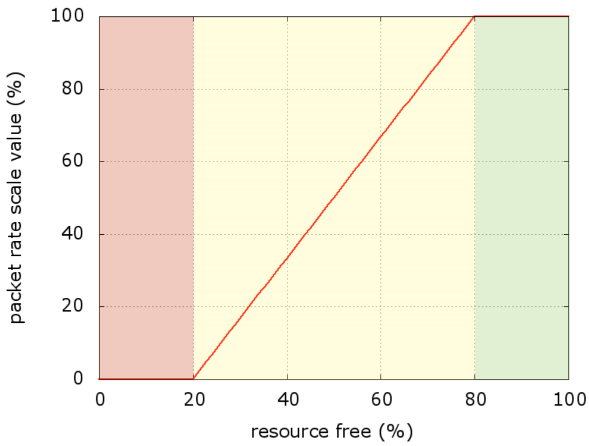


Figure 6. Linear control function with 20% and 80% thresholds

Seven different control functions were evaluated. Each of them has its green, yellow and red intervals as seen in Fig. 6.

In the green range there is no throttling necessary, the sender can operate at an arbitrary packet rate (100% scale value, denoted by m in subsection B). In the red range traffic should be immediately stopped (0% scale value). In the middle yellow interval, a new throttling rate is enforced. Fig. 7 shows the shape of the implemented control functions.

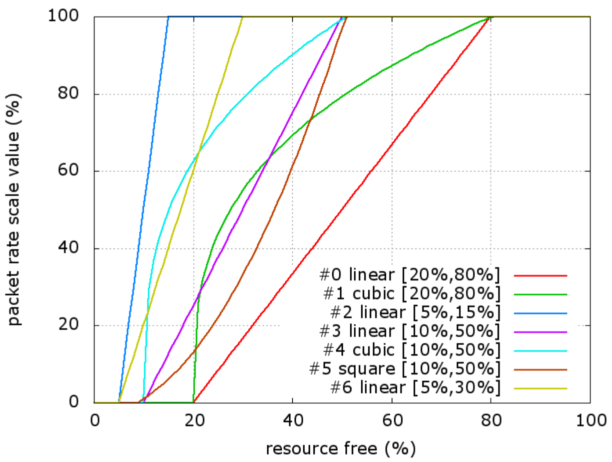


Figure 7. Throttling curves of different control functions

While cubic-based functions throttle less aggressively near the red interval, square-based one throttles more.

The effectiveness of the applied control functions depends on the resolution of the metric the function is applied on. Since socket buffer size is at least in the 10^5 bytes order, metric for socket buffer usage provides higher resolution. We should consider that there is no floating point arithmetic available in the Linux kernel by default, for performance reason we should operate on integer values. Since the most obvious method for monitoring CPU load is to query the number of idle cycles, the resolution (R) of this metric is defined by (2)

$$R = f\ /\ (1000\ /\ F) \tag{2}$$

where F is the kernel frequency (jiffies) and f denotes the sampling period (in ms).

V. Performance Evaluation

We evaluated the performance of the protocol in a 10 Gbps measurement setup in a configuration similar to the network monitoring system. Sender side was implemented in C programming language (as a user space application) as well as in VHSIC Hardware Description Language (VHDL) on NetFPGA-1G (for testing Gigabit Ethernet performance) prototyping on NetFPGA-10G. Reception side is a generic x64 Linux system running on an Intel Core i7 architecture equipped with an Intel 82599ES 10 GbE NIC. The system was configured as a non-preemptive 1000 Hz server host and critical parameters were set and resource allocation was done as described in Section IV. The controller was implemented as a Linux kernel module and the user space application was implemented in C.

Using the user space application the system was not able to generate a single flow of packets at wire speed even on a 3 GHz CPU. To eliminate this bottleneck, we continued the validation using an FPGA implementation of the sender side.

Our implementation includes a reporting interface for monitoring all the metrics that are inputs of the controller as well as other metrics such as throughput and the rate of soft RX execution.

A. Control functions

Performance metrics are fed into the controller as input parameters. For each metric a desired packet rate is calculated using a pre-defined control function. We evaluated linear as well as nonlinear functions to investigate the effectiveness of the protocol. In the following measurements CPU metric was the most frequent throttling factor. Fig. 8 represents the *'cubic 20-80%'* control function described by (3). In contrast, Fig. 9 shows the behavior of the most effective *'linear 5-30%'* control function described by (4) and activated in the range of t_0=5 and t_{max}=30.

$$F(x) = 25.5\ \times\ \sqrt[3]{x - t_0} \tag{3}$$

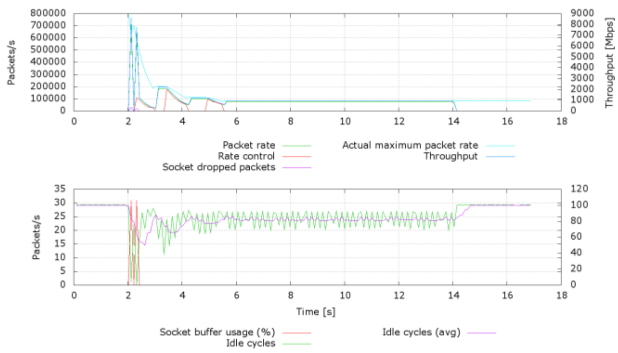


Figure 8. Behaviour of control mechanism using the *'cubic 20-80%'* function

Transmission starts with a close to the wire speed packet rate but after CPU resource drops below the throttling threshold the controller enforces a lower packet rate. After being able to get some more idle cycles it carefully increases the allowed maximum packet rate. It soon reaches the steady state when no rate change is necessary. The achieved throughput is much higher than in the case of the other involved control functions we evaluated, and is comparable to TCP's throughput as well.

$$F(x) = 100 - (t_{max} - t_0) \times (x - t_0) \tag{4}$$

This function is to be evaluated between the transfer stop threshold denoted by t_0 and unconstrained transfer threshold marked with t_{max} formerly.

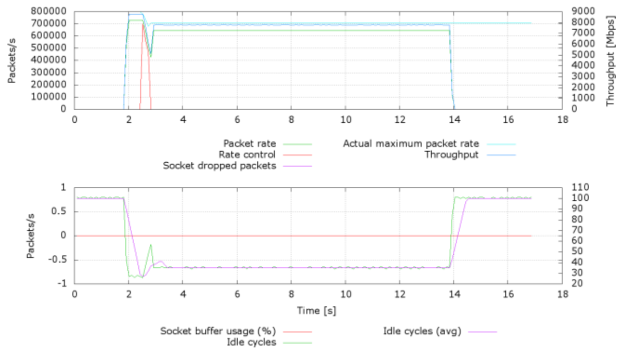


Figure 9. Behaviour of control mechanism using the *'linear 5-30%'* function

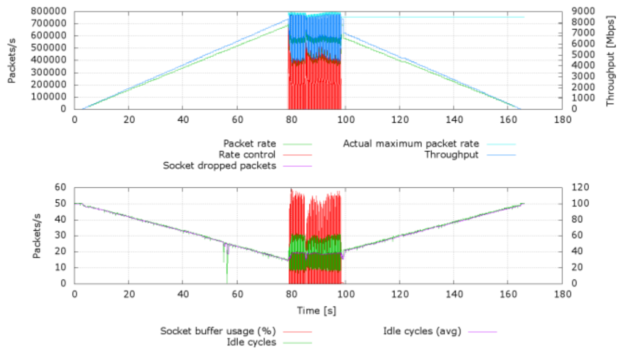


Figure 10. Flapping effect in controller caused by constants b_1=0.1 and b_2=0.1

Fig. 10 shows the flapping effect. During the measurement, MTU-sized packets were transmitted. Packet rate was continuously increased from 1 to $8*10^6$. A '*cubic 10-50%'* control function was used. Using both 0.1 for penalty and premium constant caused a flapping of control and destabilized the packet rate. Using a penalty constant value of 0.2 stabilized the transmission rate (Fig. 11).

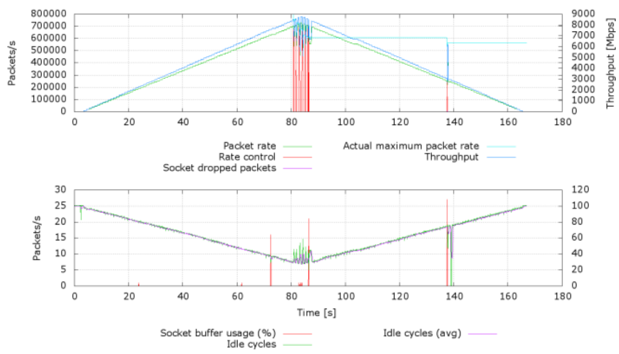


Figure 11. Stabilized throughput using constants b_1=0.2 and b_2=0.1

B. Sampling frequency

As we referred it in Section IV B, the resolution of CPU load metrics is affected by the sampling frequency. The lower sampling frequency the higher metric resolution due to the maximum number of idle CPU cycles between two sampling moments. At the same time, lower sampling frequency implies slower feedback in the control loop. This is analogous with using a longer connection. A lower frequency also implies the need for larger socket buffers. Advantage of higher sampling frequency is the shorter time to converge the maximum packet rate limit. We performed measurements using 10, 25, 50, 75 and 100 ms sampling periods.

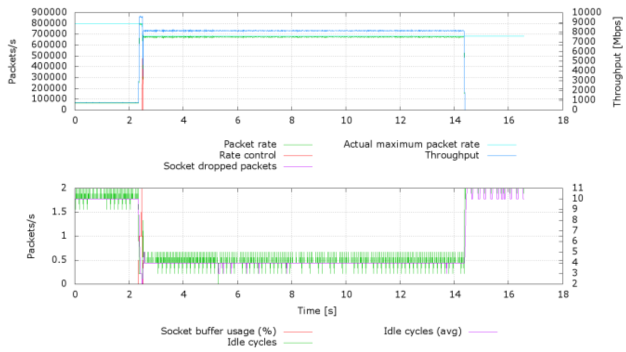


Figure 12. Faster convergence using a higher sampling frequency

Fig. 9 shows a measurement at a requested packet rate of $8*10^6$ and sampling period of 100 ms. The control function was the most effective with the '*linear 5-30%'* configuration. After starting the transmission, idle CPU time immediately dropped. The controller enforced a relatively low packet rate. This resulted in an increased number of idle CPU cycles. This allowed increasing the maximum packet limit. It took about 1 second to reach the steady state at a not significantly lower packet rate than the original one.

Fig. 12 depicts the results of a measurement with the same parameters but at a higher sampling frequency, corresponding to a 10 ms sampling period. This also resulted in as a high throughput as with 100 ms but the convergence time was much shorter.

Although the resolution of CPU metric for the 10 ms sampling period is not as fine as for 100 ms, the protocol's algorithm could efficiently control the packet rate. However, convergence period is shorter while it enables higher throughput.

C. NIC offloading functions

We did not implement UDP checksum generation, thus the performance improvement of RX offloading could not be tested. But in a production system this could be exploited, especially when measurement data has to be collected at higher packet rate.

D. RCTP versus TCP

Although TCP provides a reliable packet transmission at a respectably high throughput, our implementation of RCTP has also a comparable performance. Table I summarizes the most important features of TCP. Despite RCTP does not have such a wide range of abilities, packet losses of the receivers can be totally avoided in our system. TCP's feature set could be replaced by a sophisticated rate control mechanism.

TABLE I. FEATURE COMPARISON OF TCP AND RCTP

Resource	TCP	RCTP
Session initialization	Yes	No
Reliable transmission	Retransmission, congestion control, flow control	Rate control
Protection against packet loss	Yes	Yes *
Protection against reordering or duplication	Yes	No
Protection against packet corruption	Yes	No
Fairness	Yes	No **
Full duplex	Yes	No

* Packet losses caused by the forwarding infrastructure is not handled
** Not needed because of one flow per agent

In Table II, an enlightened implementation of TCP served in a very similar monitoring system is set in contrast with our RCTP implementation on the same FPGA chip. Even with this simplified design of TCP, the improvement of demand on resources RCTP is significant. Note that more instances of the protocol are present in the production system at the same time, thus multiple savings on resources can be expected.

TABLE II. SIMULATION COMPARISON OF RESOURCE OCCUPATION OF TCP AND RCTP ON THE SAME XILINX XC5VLX110T FPGA CHIP

Resource	TCP	RCTP	Total Available
Number of Slice Registers	1350	449	69120
Number of Slice LUTs	2322	647	69120
Number of Block RAM/FIFO	9	0	148

VI. CONCLUSION AND FUTURE WORK

We introduced a low-complexity transport protocol called RCTP optimized for transporting monitored data to multiple post-processing agents in a monitoring system using low hardware resources considering the resource-constrained design of the data sender. It can avoid packet loss introduced by the receiver, while its throughput can be close to the performance of TCP. Its demand for resources, comparing to TCP, is much lower, which is an important factor for implementing multiple functions in a reprogrammable architecture array such as an FPGA. We evaluated the protocol at 10 Gbps using its FPGA and Linux implementations.

Although it does not seem to be a problematic point, the relatively low resolution of the CPU metric could be improved by an adaptive sampling technique where the sampling frequency could be changed in runtime. By using an initialization phase or by collecting statistics data, an optimal sampling period could be determined. The aim of optimization is to achieve a shorter convergence time or to decrease the time while metrics show the resources being low.

ACKNOWLEDGMENT

The research of Péter Orosz was supported by the European Union and the State of Hungary, co-financed by the European Social Fund in the framework of TÁMOP-4.2.4.A/ 2-11/1-2012-0001 'National Excellence Program'.

REFERENCES

[1] W. Jiang, and V. K. Prasanna. "Large-scale wire-speed packet classification on FPGAs," Proceedings of the ACM/SIGDA international symposium on Field Programmable Gate Arrays, ACM, 2009.

[2] S. Floyd, "HighSpeed TCP for Large Congestion Windows," RFC Editor, 2003.

[3] L. Xu, K. Harfoush, and I. Rhee. "Binary increase congestion control (BIC) for fast long-distance networks," *INFOCOM 2004. Twenty-third Annual Joint Conference of the IEEE Computer and Communications Societies*. Vol. 4. IEEE, 2004.

[4] S. Ha, S, I. Rhee, and L. Xu. "CUBIC: a new TCP-friendly high-speed TCP variant." *ACM SIGOPS Operating Systems Review* 42.5 (2008): 64-74.

[5] K. T. J. Song, Q. Zhang, and M. Sridharan, "Compound TCP: A scalable and TCP-friendly congestion control for high-speed networks," *Proceedings of PFLDnet 2006* (2006).

[6] H. Shimonishi, T. Hama, and T. Murase, "TCP-adaptive reno for improving efficiency-friendliness tradeoffs of TCP congestion control algorithm," *Proc. PFLDnet*. 2006.

[7] A. Baiocchi, A. P. Castellani, and F. Vacirca. "YeAH-TCP: yet another highspeed TCP," *Proc. PFLDnet.* Vol. 7. 2007.

[8] P. Orosz and T. Skopko, "Multi-threaded Packet Timestamping for End-to-End QoS Evaluation," 8th International Conference on Systems and Networks Communications, ICSNC 2013, October 27-31, 2013, Venice, Italy, ISBN 978-1-61208-305-6, Proceeding p. 94-99.

[9] P. Orosz, J. Sztrik, and C. Kim, "Dynamics and Congestion Control of Alternative TCP Variants on Asymmetric Lines," *Communications and Networking*.

[10] T. Uchida, "Hardware-based TCP processor for gigabit ethernet," *Nuclear Science, IEEE Transactions on* 55.3 (2008): 1631-1637.

[11] T. Bova, and T. Krivoruchk, "Reliable UDP protocol, Internet Draft, Network Working Group, draft-ietf-sigtran-reliable-udp-00.txt," February, 1999.

[12] Y. Gu, and R. L. Grossman. "UDT: UDP-based data transfer for high-speed wide area networks," *Computer Networks* 51.7 (2007): 1777-1799.

[13] E. Kohler, M. Handley, and S. Floyd. "Designing DCCP: Congestion control without reliability," ACM SIGCOMM Computer Communication Review. Vol. 36. No. 4. ACM, 2006.

[14] R. Stewart, and C. Metz. "SCTP: new transport protocol for TCP/IP," Internet Computing, IEEE 5.6 (2001): 64-69.

[15] R. Pandit, "Reliable Datagram Sockets (RDS)," OpenIB Developers Workshop, Feb 2006. https://www.openfabrics.org/index.php/ofa-documents/presentations/doc_download/169-ranjit-pandit-silverstorm-reliable-datagram-sockets-rds.html, [retrieved: June, 2014]

[16] Y.J. Kim, V. Kolesnikov, H. Kim, and M. Thottan, "SSTP: a scalable and secure transport protocol for smart grid data collection," Smart Grid Communications (SmartGridComm), IEEE International Conference, 2011.

[17] Y.J. Kim, and M. Thottan, 'SGTP: Smart grid transport protocol for secure reliable delivery of periodic real time data," Bell Labs Technical Journal 16.3 (2011): 83-99.

[18] C. Benvenuti, "Understanding Linux network internals,“ O'Reilly Media, Inc., 2006.

Péter Orosz received his MSc. in Computer Science from the University of Debrecen, Hungary in 2003 and a PhD degree on traffic analysis of high performance networks in 2010. As research associate he worked in the Network Research Laboratory at Queen Mary, University of London where he was engaged in the performance analysis of network monitoring architectures. Since 2009, he has been working at the Faculty of Informatics, University of Debrecen and currently holds an assistant professor position. In 2010, involving PhD students he established a research group called ATMA focusing on IP network management including high performance network monitoring and traffic analysis using reconfigurable architectures. Currently, he is the leader of the ATMA research group and is participating in several research projects.

Tamás Skopkó received his MSc. in Computer Science from the University of Debrecen, Hungary in 2002. He started his PhD studies in 2009 and from 2010 is a member of the ATMA research group from its foundation. His research focus is at network monitoring and media quality evaluation. He is currently a PhD candidate and was participating in several research projects.

Máté Varga received his BSc. in system engineering from the University of Debrecen, Hungary, where he is currently a master student. Since 2012 he has been working with the ATMA group and participating in a research project focusing on reconfigurable architectures and high performance network measurement.

Internet of Things: Application Areas and Research Results of the FIRST Project

Zoltán Gál, Béla Almási, Tamás Dabóczi, Rolland Vida, Stefan Oniga, Sándor Baran, István Farkas

***Abstract* - The FIRST/IoT project coordinated by the Faculty of Informatics, University of Debrecen, Hungary has important impact on the R&D work in this field. Six activity areas have been covered in the twenty-seven months long project. More than thirty researchers from half dozen Hungarian and other universities and research institutes have been involved in this activity. The results of this work are planned to be used for other international IoT projects in the following time period. Other institutes and individual researchers from abroad are invited to join to this open initiative and become partner. In the paper are presented the results and the most exciting aspects of the research activity.**

***Index Terms* - IoT, MPT, Sensor/Actuator, Big Data, Data Clusterization, Cyber-Physical Space, Bloom-filter, E-health, Ensemble Forecasting, Virtual Organization.**

1. Introduction

Several universities and academic research institutes in Hungary working together with over forty professors and researchers from the United Kingdom, Russia and Romania are involved in effective research activity in the topics of IoT. Faculty of Informatics of the University of Debrecen in Hungary has leader role in the IoT research based on consortium project financed by the EU structural fund and the government of Hungary.

The R&D activity includes six topics: i) Integration of the IoT into the IPv4/IPv6 systems (development and analysis of multipath protocol stack networks; evaluation of L1/L2 transmission mechanisms of the sensor networks; energy usage efficiency of WSNs; analysis of the random fields defined on space-time domain to model the transmission events of radio channels - kriging; cluster analysis of sensor variables; surprise event detection at CEP - Complex Event Processing and ESP - Event Stream Processing supported services; bilateral teleoperation over wireless networks). ii) Cyber physical systems (embedded digital systems and integration of the network technologies; analysis of the complex, real time, dynamic reconfigurable systems; network security and intrusion detection in sensor network critical infrastructures). iii) Self-optimizing and self-managing communication mechanisms of the IoT systems (context dependent addressing for IPv4/IPv6 and 6LoWPAN systems; context dependent clustering, routing and multicast on the IoT; opportunistic networking; context-aware communication for the IoT). iv) E-health powered by IoT (development of intelligent home and vital technologies; real time human activity monitoring; remote supervision; elder people activity recognition; life quality enhancing services; indoor localization techniques using wireless sensor network). v) Weather prediction network tool development and analysis (statistical calibration by BMA and EMOS methods of the temperature and wind velocity ensemble prognosis; analysis of the cosmic background relay with the spectrum of random fields defined on the sphere). vi) Development of testbeds and virtual service platforms (authentication method with two factors and increased security level). In the following chapters the subjects listed above are presented.

2. Integration of the IoT with the IPv4/IPv6 systems

In this topic two R&D fields were included. The importance of the multipath transmission (MPT) of the packet switched technology on network and transport layers was analysed. The effect of the MPT to the IPv4/IPv6 protocol stack was demonstrated by an own developed software library. The other group of tasks was oriented to the statistical analysis of the multicast traffic, to the cluster analysis of the data coming from network with high number of variables and to the frequency resource usage of a supercomputer system.

2.1 The MPT software library

The integration of the IoT with the IPv4/v6 systems opens questions on the efficient bandwidth usage of the available multiple interfaces (e.g. RJ-45, WiFi, 3G, Bluetooth) of the hosts (especially of mobile hosts) especially in the transition process from IPv4 to IPv6. The traditional IP communication infrastructure is restricted to a single IP address (and single interface) usage on the communication endpoints. The IP address is used not only to identify the interface of

[1] Manuscript received June 16, 2014, revised September 8, 2014.

Zoltán Gál, Béla Almási, Stefan Oniga and Sándor Baran are with the Faculty of Informatics, University of Debrecen, Hungary.

Tamás Dabóczi and Rolland Vida are with Inter-University Cooperative Research Centre, Hungary.

István Farkas is with National Information Infrastructure Development Institute, Hungary.

the node, but it is also used to identify the communication session (i.e. socket id). Distributing a communication session between different paths is an interesting question, and it is a focused research area today. Easy to see, that the usage of multiple interfaces and paths will increase the throughput of the communication (see e.g. [1]). If the communication session is terminated on a moving node (e.g. computer located on a moving car) the request of changing the IP address inside a communication session may appear. The traditional L3 roaming solution suffers from the efficiency problem of "triangular inequality". Opening the possibility of changing the IP address of the end node (with the assumption, that the communication session must continue the work), could open a quite new solution area for these situations: The moving computer could easily change its IP address without losing the communication session's state, and this solution could eliminate the triangular inequality problem.

At the Faculty of Informatics, University of Debrecen a software library was created (named as "MPT software library), which opens the possibility of using multiple interfaces (and multiple paths) inside a communication session between the endpoints. The individual paths can be turned off and on without losing the connection. The MPT introduces a new conceptual working mechanism, which differentiates the identification of the communication session (i.e. the socket id) and the identification of the physical interfaces. The solution is based on creating a logical (tunnel) interface on the endpoint. The logical interface is used to identify the node's communication sessions, and it is independent of the physical interfaces. The MPT software library maps the logical interface to multiple physical interfaces dynamically, so offering a L3 multipath working environment. Measurement results show, that the MPT library is able to aggregate the throughput of independent paths very efficiently (see [1], [2]). As the logical interface and the physical interfaces are handled independently, it is also possible to use different IP versions on the logical interface (i.e. by the communication software) and in the physical network environment (see [2]), so the MPT library also offers a seamless IP version changing solution. The detailed description on the MPT library can be found in [3].

2.2. Analysis of the IPv4/IPv6 data traffic and control signals transmitted through the sensor networks

The service effect of the new virtual interfaces based on the new IEEE 1905.1 technology was analysed in PAN/SOHO environment. The current smart devices (tablets, phones, etc.) have multiple physical interfaces with different communication technologies (i.e. Bluetooth, NFC, WiFi, USB) able to communicate concurrently. In the classical protocol stack architecture each interface should have own logical address to communicate simultaneously. A given logical address is mapped to the unique physical address of the interface and each logical address should be placed in separated logical network. Introducing a virtual interface function between the LLC and MAC sublayers, the smart device becomes a switch in the OSI layer L1.75 with only one logical address in the network layer. All the physical interfaces remain active with the own communication technology and participate in the merged group of layer L1.5 channels.

Nice results were obtained by the analysis of the congestion effect to the streaming transmission in low bandwidth, sensor based network environment. It was found that both, the channel load and the channel intensity need to be considered for proper evaluation of the congestion in homogeneous TCP or heterogeneous aggregated TCP/UDP multimedia traffic. The aggregated traffic of the congested streams has long range memory (LRD) characteristic [4].

The coexistence of different wireless transmission technologies (i.e. IEEE 802.11 and IEEE 802.15.4) on the same physical environment was studied in function of the frame size transmitted [5].

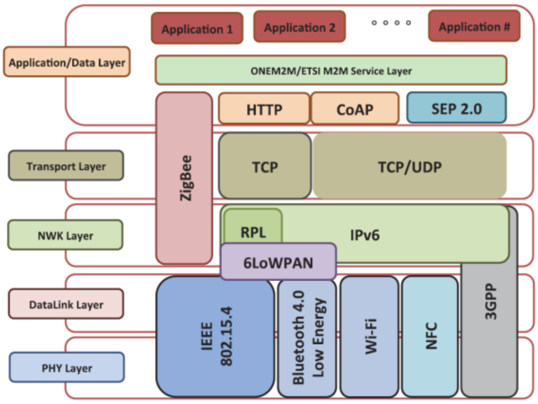


Fig. 1. Relation of the IoT technologies and the OSI model

The radio interference created in the 2.4 GHz radio channels produces three times higher error rate for the IEEE 802.15.4 channel as for the WiFi. Development of stochastic models for systems distributed in space and time and their application in the description of radio channel noise characteristics in WiFi system with high number of base stations serving as sensor nodes [6]. Based on this idea a new kriging method is proposed for continuous extrapolation of the signal field intensity in 4D physical coordinates (space-time domains) not sampled by the discrete sensor nodes [7].

Clustering method was developed and applied to extract information content from sensor network data sets and application of it to characterize the resource usage of a supercomputer system. The method based on artificial neural networks, cluster analysis and wavelets reduces by one order of magnitude the number of variables needed to be sampled to presage surprise events at the CEP (Complex Event Processing) and ESP (Event Stream Processing) supported services based on huge number of logical and physical sensor nodes [8].

3. Cyber-Physical Systems

A Cyber-Physical System is a special case of the Internet of Things. It is characterised by a very intense interaction with the physical processes, and usually cooperating nodes solve a common task. Within the frame of this project we aimed at combining the advantageous behaviour of embedded- and IT systems. We are going to extend the possibilities of embedded systems through utilization of high performance IT solutions and through the possibility of strong cooperation of separate nodes by means of interconnections through the high speed internet. However, in our view, the interconnection of large set of embedded systems serve as general purpose cyber-physical resources, rather than resources for dedicated purposes.

We envision a farm of embedded systems, with a large set of sensors and actuators as a universal infrastructure for gathering information from the physical world, for interacting with it through actuators, and also as a universal computation resource [9]. A user utilising this infrastructure can develop a new application, based on the available new and historic sensor information and can influence the environment (in a controlled way).

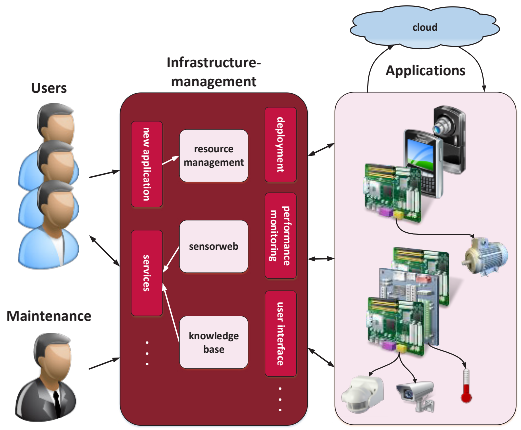


Fig. 2. Architecture of the Cyber-Physical infrastructure

New applications are automatically and dynamically allocated to embedded computing devices. Based on the measured resource utilization, the tasks are reallocated among devices in run-time by means of *Design-Space Exploration*. The above concept requires high level virtualization [10]. We use sensor virtualization (common interface, description structure and database) to access the information about physical processes from any embedded nodes. The possibility of reallocation of tasks also requires a certain level of virtualization, which might range from process *virtualization* to full platform virtualization. The computing nodes are strong enough to host virtual machines, guest operation systems and several applications at a time. However, if the application is very compute intensive, we can delegate certain parts of the calculations to the cloud (*cloud computing*) [11]. Sensor information collected by embedded nodes are accessible through an *ontology*, which allows the users to search for special types of sensors, or based on location, availability, accuracy etc.

4. Context-aware communication in the IoT

The 128-bit IPv6 addresses provide an unthinkably large address space, making it possible to assign trillions of addresses to each square centimetre on the surface of the Earth, so it is hard to envision any future scenario, including the "wildest" IoT-related predictions, whose needs would not be satisfied. However, in certain cases, size does not matter, or at least it is not the only thing that matters. The more relevant question is how can be those addresses used, how large will grow the routing tables, or how fast and how efficient can be the subsequent routing protocols and communication schemes. In the IoT we will probably very rarely use individual IPv6 addresses as is, we will not address a given sensor individually, but rather a group of smart "things" having in common some context-related characteristics. Therefore we propose to use a context-aware addressing and routing scheme, in which the network routes the queries to the proper place(s) based on a set of context parameters, but without knowing the IP addresses of the concerned objects.

We propose to encode context parameters in Bloom filters, which are considered a very resource-efficient and easy-to-process solution to handle set operations. IoT nodes will probably be grouped together in smaller areas behind several edge nodes connecting them to the traditional Internet architecture. The devices behind a specific edge will build and maintain a multi-hop tree over which context information in Bloom-filters can be easily exchanged and aggregated. When a context-based query is initiated, it will be rapidly routed to areas where IoT nodes exist, conforming to the requested context. The basic idea of this context-aware addressing solution was described in [12]. Currently we are working on implementing this approach in an IoT simulator and analysing its efficiency in different setups.

However, context-information can be very complex, involving several temporal and spatial correlations between the different context parameters. Capturing the evolution of most of these parameters is important, but usually only a very reduced set of these parameters affect effectively the behaviour of a given device, application or person. Another aspect of our research was therefore to provide a solution for filtering out these parameters based on the Hierarchical Temporal Memory approach (HTM), as described in [13].

5. E-health powered by IoT

In the present world, millions of people die every year due to lack of information about their health. Increased costs in the healthcare system could be reduced, if it would give more attention to disease prevention through regular assessment of health status and their treatment in the early stages. Our research is oriented to develop technologies for independent daily life assistance of elderly persons or others with disabilities and to improve the quality of human life using Internet of things (IoT) techniques.

Our scope is to bring together latest achievements in the domains of IoT and of assistive technologies in order to develop a complex assistive system with adaptive capability. Learning behaviour that allows living for as long as possible in familiar environment is also in focus of our research work. We use IoT technologies to monitor in real time the state of a patient or to get sensitive data in order to subsequently analyse and to enhance the medical diagnosis.

We have developed an assistive assembly consisting of a smart and assistive environment. This equipment allows also indoor localization based on wireless sensor network and Wi-Fi infrastructure [14]. It was developed a human activity and health monitoring system [15], an assistive and telepresence robot, together with the related components and cloud services. For activity and health pattern recognition we developed a hardware module for vital parameters monitoring (temperature, heart rate, acceleration). The acquired data is used to train neural network that allows recognition of activity or health status of the patient and trigger alert signal in case of unusual state detection. We implemented and tested a recognition system for arm posture, body postures and simple activities like standing, sitting, walking, running, etc., see Fig. 3. These states and movement forms were correlated with the data acquired from a heart rate sensor. The recognition rate of the body postures was over 99 % on the data sets used for training.

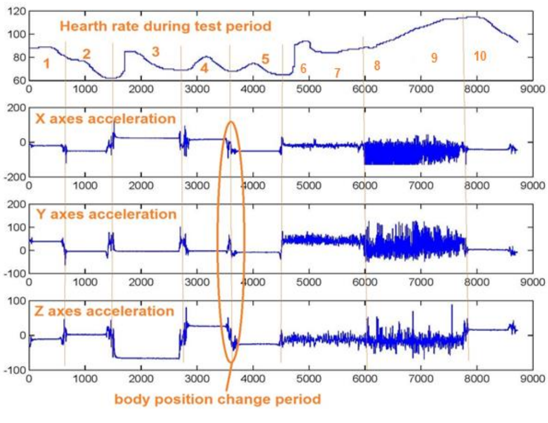


Fig. 3. Data acquired for 10 activities to be recognized.

On Fig. 4 can be seen that by setting the right threshold (red line) for the standard deviation, the static – dynamic postures discrimination can be easily differentiated. For recognizing the walking and running activities, we have extracted further relevant features from raw data set. The FFT transform was used to determine the stepping rate of a person as the most dominating frequency in the acceleration signal's spectrum.

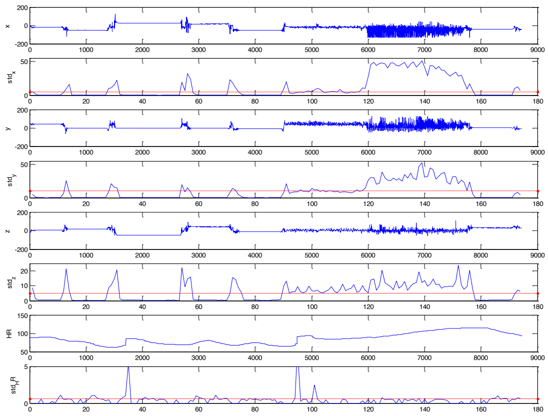


Fig. 4. Differentiating dynamic activity from static.

It was developed an own simulator application based on Qt application framework for feed forward neural networks [16]. Based on neural networks simulated in Matlab environment FPGA circuit was developed.

In our future development this monitoring module will be extended by new sensors (ECG, EMG, breath, etc.) and rules for sensor data fusion and fuzzy logic will be applied to enhance the body activity recognition. Further research based on different types of ANN is needed to simulate other activity/health status recognition. The best performing ANN type will be used to implement new recognition module based on FPGAs.

6. Weather prediction systems and analysis

The main area of our research is statistical post-processing of ensemble forecasts which is a pioneering work in this direction in Hungary. A forecast ensemble is obtained from several runs of a numerical weather prediction model with different initial conditions and makes possible the estimation of the probability distribution of future weather variables. This allows probabilistic weather forecasting, where not only the future atmospheric states are predicted, but also the related uncertainty information [20]. Recently several meteorological services provide ensemble forecasts, the leading organization is the ECMWF, while the Hungarian Meteorological Service (HMS) operates the ALADIN-HUNEPS ensemble prediction system. However, the spread of these forecast ensembles is often too small, they are uncalibrated and statistical methods are needed to account for this deficiency. The most popular tools of post-processing are the Bayesian Model Averaging (BMA) [22] and the Ensemble

Model Output Statistics (EMOS) [21]. Both approaches provide estimates of the densities of the predictable weather quantities and once a predictive density is given, a point forecast can be easily determined (e.g., mean or median value). As a first step we tested the existing BMA models implemented in the R package ensembleBMA on ALADIN-HUNEPS ensemble forecasts of wind speed [18] and temperature [19]. We found that statistical post-processing significantly improves the calibration of probabilistic and accuracy of point forecasts. We also developed a new univariate BMA model for wind speed prediction [17] and a bivariate BMA model for joint calibration of ensemble forecasts of wind speed and temperature. Both methods were successfully tested on ALADIN-HUNEPS ensemble forecasts and on forecasts of the University of Washington Mesoscale Ensemble and the results were compared to the predictive performances of the existing methods. We also performed a detailed comparison of BMA and EMOS calibration of ALADIN-HUNEPS temperature and wind speed forecasts and recently we are working on a new EMOS model for wind speed prediction. The predictive performance of this new model has already been tested on forecasts of wind speed of the UWME and of the ECMWF and ALADIN-HUNEPS ensemble prediction systems.

All new models are implemented in R and compatible with the existing ensembleBMA and ensembleMOS packages. The final goal of our research is the operation application of some statistical post-processing methods at the HMS.

7. Virtual service platforms and testbeds

More than 25 years of continuous development in the research networking area and later in the areas of those higher level e-Infrastructure services as grids, clouds, HPC, storage, collaboration and data infrastructures, have resulted in a leading edge e-Infrastructure system in Hungary that offers the provision of national and international services for the entire Hungarian research and education as well as public collection communities. The service portfolio includes, among others, communication, information access, and collaboration tools and platforms (e.g. remote co-operation and virtual community environments). The country-wide Hungarian e-Infrastructure is connected into the European and global e-Infrastructures via GÉANT, the European backbone of the research and education community. The services, having been developed and being operated by the NIIF Institute, are available also for the Future Internet research communities, and are extended to novel opportunities such as providing Virtual Research Environment (VRE) platforms and supporting Virtual Research Organisations (VRO) by making applications VO ready. An important special example of the major activities related to the e-Infrastructure is the development of a Shibboleth 2.x IdP X.509/LDAP authentication module. The basic motivation is to provide the opportunity of using hardware tokens as authentication source. SPs can decide if they want to force the X.509 authentication, or intend to simply keep a password based solution. Besides Shibboleth X.509 authentication (with or without PKI), also X.509 + LDAP certificate authentication and combining X.509 with username/password authentication are also possible options.

Based on GÉANT, also a specific, reconfigurable testbed operating in a federated virtual networking environment is provided by NIIFI, and its European partners, to the R&D community. The Hungarian segment of the testbed infrastructure is built on the high speed network of NIIFI and, together with its international connections, it is also available for supporting Future Internet research activities. Application of a two-factor authentication module for simpleSAMLphp in the federated virtual networking environment and in the testbed system has been developed, in order to achieve increased security by pairing a time-based token with other credentials, such as a username and a password. SimpleSAMLphp is used as a SAML2 Single-Sign-on solution based on php. Google Authenticator implements time-based one-time password (TOTP) security tokens from RFC6238 in mobile apps made by Google. The Authenticator provides a six digit one-time password users must provide in addition to their username and password to log into Google services. The Authenticator can also generate codes for third party applications, such as password managers or file hosting services.

8. Summary

The FIRST/IoT R&D project executed with the collaboration of several universities and institutes from Hungary, United Kingdom, Romania, Ukraine and Serbia has considerable effect on the Internet of Things topic. More than thirty journal and conference proceeding papers were published based on the theoretical and practical research work during the last two years. The results obtained in this way are considered promising basics for the continuation of the IoT field by next international joint projects.

Acknowledgement

The publication/poster was supported by the TÁMOP-4.2.2.C-11/1/KONV-2012-0001 project. The project has been supported by the European Union, co-financed by the European Social Fund. Tamás Dabóczi also acknowledges the support of ARTEMIS JU and the Hungarian Ministry of National Development (NFM) in frame of the R5-COP (Reconfigurable ROS-based Resilient Reasoning Robotic Cooperating Systems) project.

References

[1] Almási, B., Szilágyi, Sz., "Throughput Performance Analysis of the Multipath Communication Library MPT", Proceedings of the 36th International Conference on Telecommunications and Signal Processing (TSP 2013, ISBN:978-1-4799-0403-7), pp. 86-90, Rome, Italy, 2013

[2] Lencse, G., Kovács, Á., "Testing the Channel Aggregation Capability of the MPT Multipath Communication Library", Proceedings of International Conference on Internet Technologies and Applications 2014 (ICITA 2014) (part of World Symposium on Computer Networks and Information Security 2014, WSCNIS 2014), Hammamet, Tunisia, Accepted to appear, 2014.

[3] Almási, B., "MPT Userguide", availale at http://irh.inf.unideb.hu/user/almasi/mpt

[4] T. Balla, Z. Gal, A. Sz. Karsai, "Resource allocation for real time application in low bandwidth network environment", Advances on Wireless Sensor Networks (AWSN 2013) Conference Proceedings, Debrecen, Hungary (ISBN: 978-963-318-356-4), pp 21-25, 2013.

[5] T. Balla, Z. Gal: "Overview of the Internet of Things Short Range Communication Technologies", Proceedings of the 9th International Conference on Applied Informatics (ICAI 2014), Eger, Hungary, January 29-February 1, 2014 (to appear)

[6] Z. Gal, T. Balla, A. Sz. Karsai, "On the WiFi interference analysis based on sensor network measurements", IEEE 11th International Symposium on Intelligent Systems and Informatics (SISY 2013) Conference Proceedings, September 26-28, 2013, Subotica, Serbia (ISBN: 978-1-4799-0303-0), pp 215-220, 2013.

[7] T. Subba Rao, Gy. Terdik: "A space-time covariance function for spatio-temporal random processes and spatio-temporal prediction (kriging)", arXiv:1311.1981 [math.ST], 2013.

[8] Z. Gal, Tibor Tajti, "Complex Event Processing in Supercomputer Environment: Sensor and Neural Network Based Analysis", Proceedings of IEEE 4th International Conference on Cognitive Infocommunications (CogInfoCom 2013), Budapest, Hungary (ISBN: 978-1-4799-1544-6), December 2-5, pp. 735-740, 2013.

[9] Dabóczi, T., Pataricza, A., Strausz, Gy., and Varró, D., "Next Generation Embedded Systems," Advances on Wireless Sensor Networks (AWSN 2013) Conference Proceedings, Debrecen, Hungary (ISBN: 978-963-318-356-4), pp. 47-52.

[10] Kovácsházy, T., Wacha, G., Dabóczi, T., Erdős, Cs., and Szarvas, A., "System Architecture for Internet of Things with the Extensive Use of Embedded Virtualization,", *4th IEEE International Conference on Cognitive Infocommunications - CogInfoCom 2013*. Budapest, Hungary (ISBN: 978-1-4799-1-1543-9), Dec. 2-6, 2013. IEEE Press, pp. 549-554.

[11] Kocsis, I., Toth, Á., Szatmari, Z., Daboczi, T., Pataricza, A., and Guta, G., "Towards Cyber-Physical System Technologies over Apache VCL," *2nd International IBM Cloud Academy Conference ICA CON 2014*, May 8-9, 2014, Atlanta, Georgia USA. pp 1-10.

[12] A. Kalmar, R. Vida, M. Maliosz. Context-aware Addressing for Internet of Things based on Bloom Filters, in proc. of IEEE CogInfoCom 2013, Budapest, Hungary, December 2013.

[13] A. Kalmár, R. Vida, „Towards Context-Aware Mobile Services Through the use of Hierarchical Temporal Memory", in Proceedings of SoftCom 2013 – 21st International Conference on Software, Telecommunications and Computer Networks, technically co-sponsored by IEEE Communications Society, Split-Primosten, Croatia, September 2013.

[14] A. Buchman, C. Lung, "On the relationship between received signal strength and received signal strength index of IEEE 802.11 compatible radio transceivers", Carpathian Journal of Electronic and Computer Engineering, Volume 6, no. 2, 2013, pp 15-20.

[15] S. Oniga, J. Sütő, "Human activity recognition using neural networks", IEEE 15th International Carpathian Control Conference - ICCC 2014, Velke Karlovice, l, Czech Republic, 28-30 May 2014, pp 403-406.

[16] J. Sütő, S. Oniga, "A new C++ implemented feed forward artificial neural network simulator", Carpathian Journal of Electronic and Computer Engineering, Volume 6, Number 2, 2013, pp 3-6.

[17] Baran, S., "Probabilistic wind speed forecasting using Bayesian model averaging with truncated normal components," Computational Statistics and Data Analysis, vol. 75, pp. 227-238, 2014.

[18] Baran, S., Horányi, A., and Nemoda, D., "Statistical post-processing of probabilistic wind speed forecasting in Hungary," Meteorologische Zeitschrift, vol. 22, pp. 273-282, 2013.

[19] Baran, S., Horányi, A., and Nemoda D., "Probabilistic temperature forecasting with statistical calibration in Hungary," Meteorology and Atmospheric Physics, vol. 124, pp. 129-142, 2014.

[20] Gneiting, T. and Raftery, A. E., "Weather forecasting with ensemble methods", Science, vol. 310, pp. 248-249, 2005.

[21] Gneiting, T., Raftery, A. E., Westveld, A. H., and Goldman, T., "Calibrated probabilistic forecasting using ensemble model output statistics and minimum CRPS estimation", Monthly Weather Review, vol. 133, pp. 1098-1118, 2005.

[22] Raftery, A. E., Gneiting,T., Balabdaoui, F., and Polakowski, M., “Using Bayesian model averaging to calibrate forecast ensembles,” Monthly Weather Review, vol. 133, pp. 1155-1174, 2005.

Zoltán Gál (ZG’66) received MSc degree in electrical engineering and informatics from Politehnica University of Timisoara in 1990 and PhD degree in informatics from the University of Debrecen (UD) in 2010, respectively. Since 1991 he is lecturer at the Informatics Institute of Lajos Kossuth University of Debrecen and Department of Informatics of the UD. In 1999 he joined Cisco Network Academy Program where he is CCNA and CCNP registered instructor. He was chief information officer (CIO) of the UD and of the Center of Arts Humanities and Sciences of the UD for 2001-2006 and 2007-2013, respectively. He was presidency member of Hungarnet Association for ten years starting in 2005. As CIO he managed several infrastructure development projects at national and international levels including LAN, MAN, WAN and supercomputer installations. His research interest includes computer and network architectures, high speed and multimedia networks. He is member of IEEE Communications Society since 1997 and he is currently head of the Internet of Things subproject of the Hungarian national project named Future Internet Research Services and Technologies (FIRST). Dr. Gál is author of 50 papers and 75 conference publications and he wrote two books in computer architectures and cloud computing, respectively.

Béla Almási is an associate professor at the Department of Informatics Systems and Networks at the Faculty of Informatics, University of Debrecen, Hungary. He received his Ph.D. in 1998. at the University of Debrecen, Hungary and Habilitation from University of Debrecen in 2006. His primary research interests are network systems, multipath communication and performance analysis of queueing systems and their application.

Tamás Dabóczi received the M.Sc. and Ph.D. degrees in electrical engineering from Budapest University of Technology, Budapest, Hungary, in 1990 and 1994, respectively. In 1990 he joined the same university. He spent several months at the Swiss Federal Institute of Technology (ETH), Zürich, Switzerland as academic guest in 1992; at the Technical University of Karlsruhe, Germany as visiting researcher in 1994. In 1995 he was invited to National Institute of Standards and Technology, NIST, USA, Gaithersburg, MD as guest researcher, where he developed and implemented inverse filtering algorithm for calibration of ultra-high-speed digital sampling oscilloscopes. Currently he is an associate professor and deputy head of Department of Measurement and Information Systems, Budapest University of Technology and Economics, Budapest, Hungary. His research area includes embedded systems, cyber-physical systems, digital signal processing, especially inverse filtering. He is senior member of IEEE.

Rolland Vida, born in August 1974, obtained his BSc and MSc degrees in Computer Science, as valedictorian, from the Babes-Bolyai University Cluj-Napoca, Romania, in 1996 and 1997 respectively. He then pursued his PhD studies in Computer Networking at Universite Pierre et Marie Curie Paris VI, France, and obtained his PhD degree in 2002. In 2003 he joined Budapest University of Technology and Economics, Department of Telecommunications and Media Informatics, where he is currently an associate professor. Between 2007 and 2010 he was a Bolyai Janos Postdoctoral Research Fellow of the Hungarian Academy of Sciences. His research interests are mostly related to wireless sensor networks, internet of things, vehicular networking, peer-to-peer communications and multicast technologies. He is member of IEEE, and is involved in the organization of many international conferences, either as a member of the TPC or the Organizing Committee. He is currently the Operations Manager of the GLOBECOM/ICC Management & Strategy Committee of the IEEE Communications Society. Dr. Vida was involved in several national and international research projects, participating in numerous FP5, FP6 or FP7 consortia. He has published more than 65 research papers, which generated more than 600 independent citations.

Stefan (István) ONIGA was born in Baia Mare, Romania, on Apr. 20, 1960. He received his M.Sc. and Ph.D. degrees from PUT (“Politehnica” University of Timisoara) in 1985 and 2005, respectively. Since 1992 he has been working at Faculty of Engineering, North University of Baia Mare (now North University Center of Baia Mare, Technical University of Cluj-Napoca), where he holds a position of associate professor in Digital Design, Microcontrollers and Embedded Systems. Between 2008 and 2010 was director of Electronic and Computer Engineering Department. Since 2010 he is associate professor at Faculty of Informatics, University of Debrecen where he teaches Digital Design with Programmable Logic, Digital Technologies, and Reconfigurable Embedded Systems. He is the author and co-author of more than 80 journal and conference papers, 20 research projects and 6 books and books chapters. His major areas of interest are embedded systems design and applications, implementation of artificial neural networks in field-programmable gate arrays, digital neural network based smart devices, e-Health systems and ambient assisted living and assistive robots. Dr. Oniga is chief editor of the peer reviewed journal Carpathian Journal of Electronics and Computer Engineering. He is a member of IEEE, Industrial electronics society since 2006.

Sándor Baran (SB'73) graduated as a mathematician in 1995 and as a teacher of mathematics and English-Hungarian technical translator of mathematics in 1996 at the Kossuth Lajos University in Debrecen. He received PhD degree in Mathematics and Computer Science at the University of Debrecen (UD) in 2001 and in 2006 he defended his habilitation thesis at the same institution. From 1999 instructor, from 2001 assistant professor and from 2006 associate professor at the Department of Applied Mathematics and Probability Theory of the Faculty of Informatics of the UD. Between March and September 2013 he worked as a visiting professor at the University of Heidelberg. For his research work in statistics and probability theory he was twice awarded by the Bolyai János Mathematical Society, between 2001 and 2004 he was a grant holder of the Bolyai Grant of the Hungarian Academy of Sciences (HAS) and in 2012 he was awarded by the Gyires Béla Price of the HAS. From 2012 he is a member of the Regional Committee of the Bernoulli Society. Sándor Baran's research topics are measurement error regression models, parameter estimation of random fields, autoregressive and other time-series models, spatial autoregression, stochastic optimization (simulated annealing), probabilistic weather forecasting, applied statistics (biological, geological and medical applications). He has published 33 papers is refereed journals, 6 papers in conference proceedings, co-author of a university lecture notes and has given 40 talks and 2 poster presentations on international conferences.

István Farkas (1976–) is a network systems engineer at the National Information Infrastructure Development Institute (NIIFI) since 2006. He is head of the Network Operation Department at this institute. Prior to this he was research fellow at the Hungarian Academy of Sciences Institute for Computer Science and Control (MTA SZTAKI). He worked together with the NIIFI team since 1999, where his main activities included operation and maintenance of the Hungarian research network (HBONE) connection to the European TEN-34/TEN-155 network. He was technical representative of the Hungarian research backbone network (HBONE) in the European research network project, GÉANT/GÉANT2 since 2000 and 2004, respectively. He is member of the HBONE project leaders since 2004.

Investigation of spreading phenomena on social networks

Gergely Kocsis and Imre Varga

***Abstract* — In this paper the results of our investigations related to social spreading are summed up and concluded. In our work we studied information spreading on different network topologies. Based on a novel complex network generating method we managed to generate several test cases for social simulations focusing mainly on the case of declining social networks. We ran simulations using a previously presented model of information spreading. As a result we showed how the effectiveness of the spreading depends on the way and the intensity of declining. Later, using a modified version of the model we examined the effect of dynamically active agents in the system. As the most important result of this study we showed that increasing the activity of central nodes of a social network alone does not make the spreading significantly more effective.**

***Index Terms* — information spreading, declining social networks, complex networks, cellular automata simulation**

I. INTRODUCTION

THE investigation of social spreading phenomena has been in the focus of research for a couple of decades. However in the last 15 years with the appearance of online communities its importance has become much greater as it turned out that the models used to model classical societies based on personal contacts are also applicable for these social structures [1-3]. While the classical topics of the field were in most of the cases related to disease spreading, rumor spreading, opinion spreading, etc. [1,2], today – reflecting to the questions of the informational society – one of the most important questions is information spreading. In our work we study the spreading on the most widely known group of online societies, on social networks. In this very work we focus on to major topics: *(i.)* declining social networks and *(ii.)* networks of dynamically active agents. In the first part we try to understand how the dynamics of information spreading change on social networks that have passed their golden age, and start to decline. Based on this result we give a hint, where is the point when it does not worth anymore to spread information (e.g. advertise) on these networks. In the second part of our work we examine the change of the spreading when we assume that agents may not be always active and ready to spread or receive information. More practically our goal is to find out that in such a network what can one do to speed up the spreading only by manipulating the activity of agents. The results of our work may be used later in planning advertising campaign strategies, or anti-spam actions.

Manuscript submitted June 14, 2014, revised September 10, 2014.

Kocsis, G. Is assistant professor at the University of Debrecen, Faculty of Informatics, Department of Informatics Systems and Networks (e-mail: kocsis.gergely@inf.unideb.hu).

Varga, I. Is assistant professor at the University of Debrecen, Faculty of Informatics, Department of Informatics Systems and Networks (e-mail: varga.imre@inf.unideb.hu).

II. SPREADING ON DECLINING SOCIAL NETWORKS

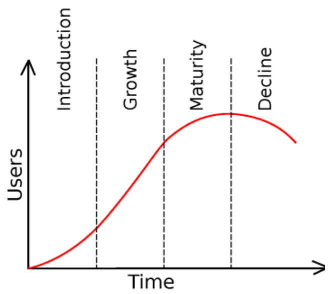


Fig. 1. Life cycle of a social network. First, when the network is introduced more and more users come and it keeps growing till the maturity. The last phase however is always declining.

As history shows, online social networks follow the universal rules of diffusion of new technologies [1] during their life cycle and go through four different stages [4] (see Fig. 1.). At the very beginning, when they are just introduced, the number of users starts to increase slowly. Later more and more users join and the network grows much faster. After a while however the system reaches its possible maximal size, and enters to its maturity phase. The aim of the owners of all social networks is of course to stay as long as possible in these latter two stages. Several examples show however (MySpace, Orkut, iwiw, etc.) that after a while social networks start to get out of fashion. This means that users leave them, they arrive to their fourth stage and the declining starts. In our work presented in [5] we focus on information spreading on social networks of this fourth kind. As one possible result we wanted to get an idea if it is worth to advertise such networks or not.

In the next three sections the two steps of this work is presented. First we model the network structure and then we model the behavior of the nodes of the network.

A. Reproducing the topology of declining social networks

To be able to run our simulations the first step was to find a way to generate topologies similar to the ones that online social networks have. Finally we worked out a two step method for this [6,7]. First we generate a network that is topologically more or less similar to a real social network in its mature phase. And then we attack it to catch the declining.

As it is widely known from the literature real social networks have a so called scale-free topology meaning that there are a huge number of actors with a low number of connections and only some that have a high number of neighbors [8-10]. More precisely this means that the degree distributions of these networks follow a power law form

($P_k \sim k^{-\gamma}$), where k is the number of neighbors and the exponent γ can be used to identify different types of networks. By comparing a the properties of a real social network sample [11] (with the degree distribution in the first place) it can be seen that unfortunately the most widely used Barabási-Albert (BA) model [12] does not provide us a similar network [13]. Namely there is a clear difference between the exponent of the degree distribution of the sample ($\gamma \approx 2.5$) and of the BA generated network ($\gamma = 3$). We found a mixed model of generating scale-free networks to be functional for our needs. Namely we used the model of Lee et. al. [14] which is a mixture of the simplest Barabási-Albert (BA) model and the model of Bianconi et. al. [15]. In both models the generation of the network starts the same way: we start from a fully connected network of m_0 nodes and then we add new nodes one-by-one to the network. Each new node is connected to the network by m edges. The difference between the classical BA model and the model of Bianconi et. al. is the way how we choose the other endpoints of these new edges. The BA model is often referred as a so called popularity driven model, since the endpoints of the new edges are selected based on the existing number of links of the nodes, i.e. a node with a bigger number of edges has a bigger probability to be the end of the new edge as well. On the other hand the model of Bianconi et. al. uses a so called fitness driven algorithm i.e. we define a random real fitness value between 0 and 1 for all nodes, and the probability to connect the new edge to a node is proportional to the product of its fitness and the number of their existing connections.

In our case the two models above were mixed in the following way [5]. We also started from the fully connected core of m_0 nodes and we also added $m = m_0 - 2$ new edges with each new node. However when we selected the endpoints of the new edges we used the classical BA algorithm with probability p, and the fitness driven algorithm with probability $1 - p$. We iterated this process till the total number of nodes in the system N_0 reached a desired value. Using this mixed model we were able to generate networks that are fairly similar to a sample network that we had containing data of almost 60 million Facebook users. Or more precisely the exponent of the degree distribution matched to the respective exponent of the sample [6].

After we became able to generate a topology that is similar to a real social network, we tried to find a way to model the declining itself. However unfortunately we have not found any published results in the literature how this declining goes. One obvious solution would have been for this problem to sample a social network through its whole life cycle (or at least at the end). However with respect to the time cost of this process we found that this is out of the scope of our current work. Instead we worked out different node-removing processes for three different possible declining scenarios. Following the terminology of the literature [16], hereafter we refer to this node-removing as *attack*. Namely we used central, peripheral and general attack to model the situations where the most popular, the least popular or random actors leave the network more likely respectively. From the algorithmic point of view this means that in all three cases we started from the above generated network of N_0 nodes and $\sim N_0 m$ edges. After this we removed $N_0 \eta$ nodes with their edges (where η is the strength of the attack). The difference between these options is that in the central case the probability of a node to be removed is linearly proportional to its degree. In the case of peripheral attack the probability of removing a node is inverse linearly proportional to the degree, while in the case of general attack all nodes are removed with the same probability.

Not surprisingly the topological examination of these attacked networks showed that the most dramatic changes appear in the case of central attack [16]. The general attack has only a minor effect on the topology, while the peripheral attack almost leads back to a previous state of the same network. (Since removing the lowest degree nodes means removing the nodes that were added at the end of the generation.)

With the use of this two step grow-and-destroy model we became able to produce networks with the same properties as declining social graphs (assuming that the declining of real social networks follows one of the three scenarios presented above). As a next step of our research we focused on the spreading of information on these topologies.

B. *Information spreading*

To model the spreading of information in social systems we used a simplified version of a previous model introduced by Kun et. al. [17]. In this cellular automata model the actors of the system get information from two different channels: *(i.)* there is an external source that provides a constant amount of information for all actors and (ii.) actors can also get information from each other. As an example if we look at the case of online advertising the external information channel may model public advertisements placed on web pages, while the inner channel represents discussions of users about the advertised items. In the model the actors are represented by interacting agents sitting on the nodes of an underlying topology. The agents can be in two different states: uninformed σ_0 and informed σ_1. We use the variable S_i to denote whether an agent i is in an informed state or not. If agent i is in state σ_0, $S_i = 1$, and if it is in state σ_1, $S_i = 0$. We introduce two parameters of the system to describe the sensitivity of agents to the information channels. The parameter α tells how sensitive are the agents for the inner information channel, while β describes the sensitivity for the external information channel. Originally α and β were both random real numbers for each agents, however based on the results of the original study of the model we set both of them to be the same for all agents.

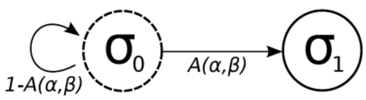


Fig. 2. The states and the possible state changes of the basic information spreading model. Uninformed agents σ_0 can get informed σ_1 or stay uninformed stochastically based on the amount of the received information. Informed agents do not forget so there is no way back from σ_1 to σ_0.

The model evolves in discrete time-steps using synchronous update rules. i.e. the state of agent i at time t depends on the states of itself and its neighbors at time $t-1$. In each time-step agents get stochastically informed based on the amount of the received information at the given time-step (see Fig. 2.). The probability of changing to informed is

$$A(\alpha,\beta) = 1 - \exp\left(-\left(\alpha S_i \sum_{j=1}^{n_i}(1-S_j) + \beta S_i\right)\right), \qquad (1)$$

where in the *exp()* function the amount of information received by agent i is caught. Namely in the sum we add all the information coming from the informed interacting partners and multiply the result with the sensitivity to the inner channel. Then we add the amount of information from the outside multiplied by the respective sensitivity. Note that for simplicity both the amount of information coming from the

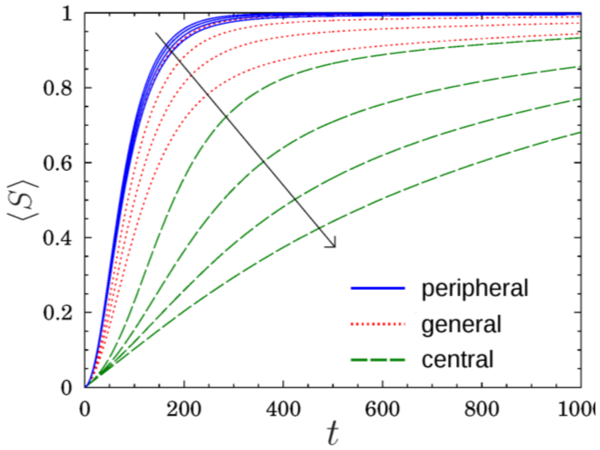


Fig. 3. The effect of the attack on the spreading. On the figure the ratio of informed agents in the system $\langle S\rangle$ is presented as a function of time t Different colors and styles are for different ways of attack. Note that with the increase of the amount of removed nodes (marked by the arrow) the spreading slows down in all cases but in the central attack case especially.

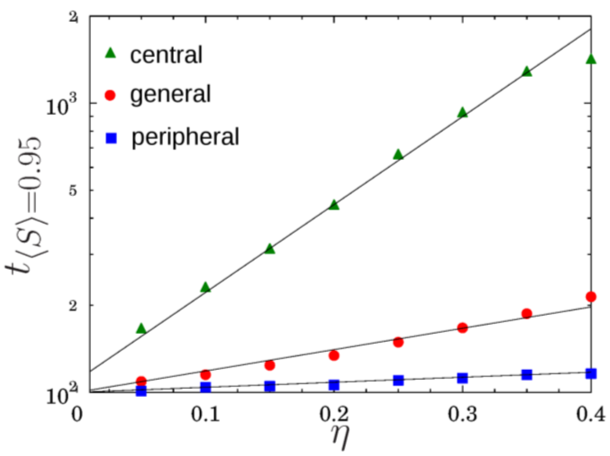


Fig. 4. The time needed to reach an almost homogeneous informed state in the system depends on the attack strength exponentially. Note that this exponential form is independent of the exact type of the attack, however different attack types result in different levels of dependence.

external source in one time-step and the amount of information coming from an informed neighbor are set to 1 "unit".

Since our main goal here was to study the effect of the declining of the network on the spreading, and not the spreading model itself, in all our simulations we used the same parameter set that proven to be quite good in representing real life scenarios [17]. Namely we set the sensitivity parameters to α=0.01 and β=0.001.

C. *The effect of declining*

To study the effect of declining on the spreading of information, we ran computer simulations on various network topologies attacked in different ways. As a key property of the spreading, we focused on the ratio of informed agents in the system $\langle S\rangle$ as a function of time t. Our results showed that the spreading process is qualitatively independent of the system size, but it highly depends both on the way how we attack the system, and the strength of the attack. It can be seen in Fig. 3. that while peripheral attack hardly affects the spreading even in the case of removing 40% of the agents, general attack makes the spreading increasingly slower as more and more nodes are removed from the network. In the third case, when we applied central attack the spreading dramatically changed. Even for a small number of removed agents the spreading gets much slower than in the case without attack. Of course this effect again increases with the increase of the amount of removed nodes.

Another interesting result of our work was found when we plotted the needed time for the system to reach an almost homogeneous informed state $t_{\langle s\rangle=0.95}$ (i.e. where 95% of the agents are informed) as a function of the strength of the attack η. We had to use this almost homogeneous state instead of totally homogeneous one because close to saturation the evolution of the system slows down dramatically implying that the respective simulation time increases in this sense as well. However this almost homogeneous informed state also fits our needs to describe consensus. The results showed that the needed time to reach this state depends exponentially on the strength of the attack (see Fig. 4.). In a practical sense this result means e.g. that if a social network is losing its users linearly the effectiveness of information spreading (e.g. advertising) is decreasing much faster.

III. THE EFFECT OF THE INACTIVITY OF AGENTS

In the first part of our research we focused on systems, where the agents of network were handled to be always active. Most social networks however do not share this property [18,19] so we found it reasonable to examine how the activity of agents affects the spreading. Even though the topic is marginally present in some related works [20,21], we found that the exact question have not been answered yet. What we did to get a better insight, is that we modified the previous model and examined the behavior of it in networks from simple square lattices to complex topologies [22]. Note that in this case instead of reactions for the underlying topologies we focused on the spreading itself.

A. *The modified model*

To make our model able to describe activity of the agents we added a new parameter of the agents $\gamma_i \in [0;1] \in \mathbb{R}$ to describe how often an agent goes inactive. This means that for a value of γ_i close to 1 the agent is almost always inactive while in the opposite case when γ_i is close to 0 it is active in most of the time. Note that in contrast to α and β, γ_i is not a system wide parameter but it can be different for different agents. However at the beginning of our investigations to make our results more clear we set $\gamma_i = \gamma_j$ for all $i, j \leq N$ where N is the number of nodes. The property that agent i is active or not is caught by the variable R_i, where $R_i = 0$ if the agent is inactive, and $R_i = 1$ if it is active. By introducing these new parameters of the agents of course the number of possible states is increased and the state-change rules themselves have changed as well.

First of all, based on the informed/uninformed property S_i and the activity R_i now agent i can be in four different states $\{\sigma_0, \sigma_1, \sigma_2, \sigma_3\}$. These possible states are presented on Table 1. The basic state change rule related to the activity of agents is the following: In each time-step agent i stays or changes to inactive with probability γ_i and stays or changes to active with probability $(1-\gamma_i)$. The most crucial state-change of course in this case again is the change from the active uninformed state σ_0 to active informed state σ_1 (Inactive agents can not get informed or spread information, and informed agents do not go uninformed regardless of their activity). This state-change rule again depends on the received information, however to catch the effect of our new parameters we had to alter eq. (1) a little bit. In this modified model the probability of agent i to get informed is:

$$A'(\alpha,\beta) = 1 - \exp\left(-\left(\alpha S_i R_i \sum_{j=1}^{n_i}\left((1-S_j)R_j\right) + \beta S_i R_i\right)\right). \quad (2)$$

Note that eq. (2) differs from eq. (1) only through the parameters R_i and R_j, where these parameters describe that inactive agents can not receive or spread information. Based on this modified equation and the parameter γ_i. The state-change rules are presented on Fig. 5.

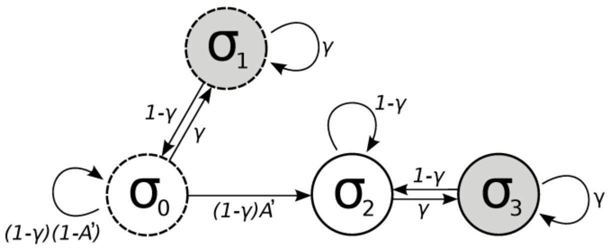


Fig. 5. The states and the possible state changes of the modified information spreading model. Agents can stay or change to inactive (*grey filled*) with probability γ and stay change to active with probability $(1-\gamma)$. The state-change from uninformed to informed is led by the modified probability A'.

In order to study the evolution of our modified model we ran simulations on different network topologies. As a first try to get a qualitative insight of the evolution we used a small regular square lattice. This simple topology made it very easy to visualize what is happening, and compare the original and

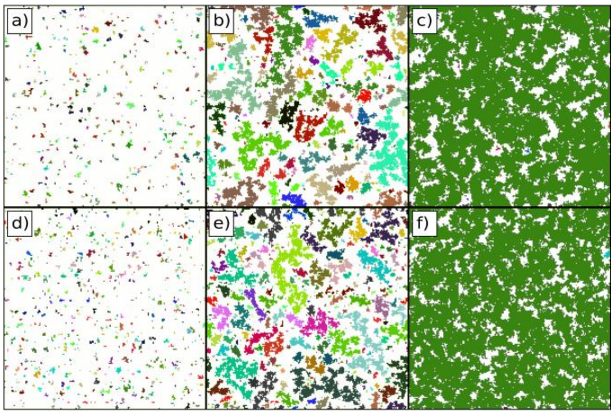


Fig. 6. Snapshots of the evolution of the model on a square lattice of $N = 100x100$ agents. *a, b* and *c* are for the original model while *e, f* and *g* are for the modified model including dynamic activity. Note that the respective pictures in the two rows show the same structure, however in the second row a longer interval of time is covered. (The snapshots were made respectively for *a, b, c, d, e, f* and *g* at $t = 100, 250, 375$ and $t = 350, 800, 1250$. Colors identify separate clusters of informed agents.)

the modified model. Snapshots of a system containing $N = 100x100$ agents at different time-steps are presented on Fig. 6. In the first row the pictures are from the original model while the second row is for the modified model ($\gamma_i = \gamma = 0.5$). Our results show that the evolution of the system does not change qualitatively however it takes noticeably more time for the modified model to reach an almost similar state as the original model. This means that the way of the spreading is

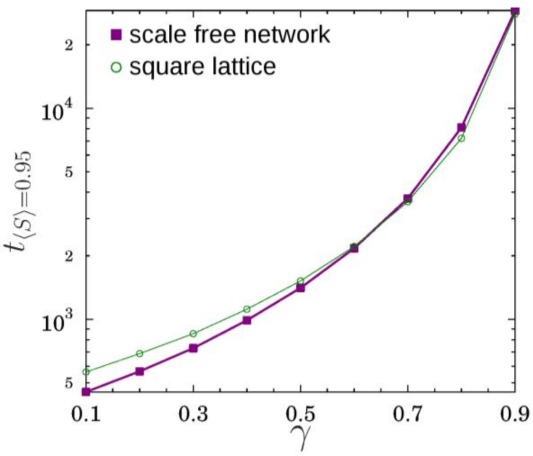


Fig. 7. The time needed for the system to reach an almost heterogeneous informed state in the case of square lattice $N = 10^6$, and on a scale free network similar to real social network topologies. The dependence in both cases are faster than exponential.

TABLE I
POSSIBLE STATES OF AGENTS

	$S = 1$	$S = 0$
$R = 1$	σ_0: active, uninformed	σ_2: active, informed
$R = 0$	σ_1: inactive, uninformed	σ_3: inactive, informed

Depending on the values of S and R agents can be in four different states. S describes whether an agent is informed (0) or not (1), while R tells whether the agent is active (1) or not (0). The values of S and R have been chosen so that the form of the later state-change rules stay reasonably easy.

not changed, but it is much slower. Of course this slowing is not similar for different values of the probability of being inactive γ. In order to find out how the spreading depends on the probability of being inactive we plotted the time needed to reach an almost homogeneous state as a function of the parameter γ. The results are presented on Fig. 7. Note that we again used the 95% informed state of the system ($\langle S \rangle = 0.95$) because of the same reason as in the case of the examination of declining social networks. On the plot we used a semi-log scale in order to show that the dependence of $t_{\langle S \rangle = 0.95}$ is faster than exponential.

Our observations on the square lattice gave us a first insight of the effect of the inactivity of agents, however since we wanted to get results that are closer to real social systems we had to apply complex network topologies again. So as a next step we ran the modified model on a scale-free network. Based on a Facebook data sample we also used previously [11], this network was generated so that it has similar topological properties as online social networks, however to make our simulations faster we used a smaller number of nodes. What we have found in this case was very similar to the square lattice case despite of the obvious differences in the topology. On Fig. 7. we also plotted the time to reach the almost homogeneous state as a function of time in the case of our scale-free generated network. Note that not surprisingly, especially at small values of γ, spreading on the scale-free network is a bit faster. However by observing the whole picture the same sentence becomes true again: As the probability of being inactive increases linearly, the time needed to reach the almost homogeneous state $t_{\langle S \rangle = 0.95}$ increases faster than exponentially.

B. Heterogeneous activity

In the previous cases we always assumed that all agents of the system have the same probability of being inactive i.e. $\gamma_i = \gamma$ for all $i \leq N$. However this is not a realistic assumption in the case of social networks. Thus the next step of our research was to examine the effect of heterogeneous activity in the system.

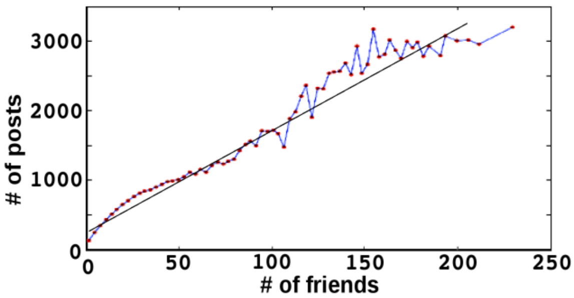


Fig. 8. The activity of users in a social network is linearly dependent on their number of friends. Here we defined activity as the number of posts a user posts in his/her timeline. Figure based on [23].

In order to get realistic results we studied the work of Huberman et. al. [23] and found that in real social networks the activity of users is in linear connection with the amount of friends (see Fig. 8.). In our case this means that if we would like to make our model more realistic, we have to make γ_i a function of the degree of agent i. To do this we chose the following form:

$$\gamma_i = \gamma \left(1 - \frac{n_i}{n_{max}}\right), \tag{3}$$

where n_i is the number of connections of agent i and n_{max} is the highest degree in the system. Eq. (3) means that the lowest degree nodes keep their level of activity, while agents with a lot of connections become more active. In the case of scale-free networks this simply means that we make the central nodes more active.

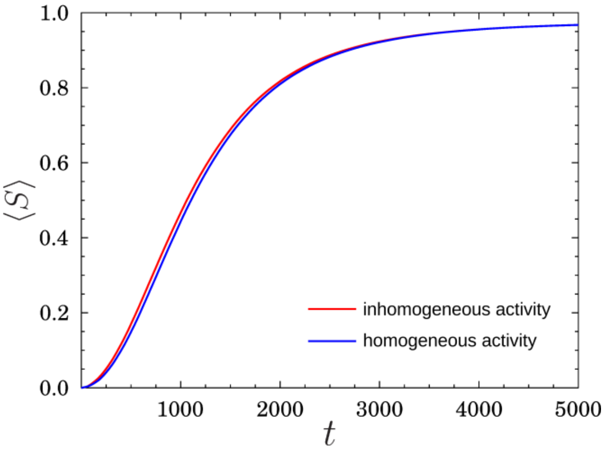


Fig. 9. The average amount of informed agents in the system as a function of time t in the case of homogeneous and inhomogeneous activity. Note that the increased activity of high degree nodes does have as much impact on the spreading as one would expect.

As a result of our simulations surprisingly we found that despite of the inhomogeneous probability of being inactive the spreading process shows hardly any noticeable changes. On Fig. 9. we plotted the average amount of informed agents in the system $\langle S \rangle$ as a function of time t for both homogeneous and inhomogeneous inactivity. It can be clearly seen that making the central nodes of the network more active does not have a major effect on the spreading. In a practical sense this would result for example that in the case when one wants to improve the efficiency of an online advertising campaign, making the most active user more active alone does not have the required effect. The reason of this is that however these central nodes became more active and ready to spread information more likely, since the contacts of them are still inactive with the same probability there is no one to interact with. A possible solution would be of course to focus instead on the increase of the activity of all the agents of the system (i.e. decrease the value of γ, see Fig. 7.).

IV. DISCUSSION

In this paper the results of our investigations related to spreading phenomena on social networks have been presented. Our work built up from three major parts. At first, starting from a network sample we developed a way to generate network topologies with similar key properties as real social networks. To do this we examined the topological properties of the real and the generated networks. With the use of an information spreading model we also studied how spreading behaves on these networks. As a second part of our research we investigated spreading on declining social networks.

Namely we studied the effect of different types of attack on the spreading, using the same information spreading model as above. We showed that independently of the type of the attack the time to reach an almost fully informed state of the system depends exponentially on the strength of the attack (from 0% to 40% nodes removed). In the third part of the work we examined how the presence of dynamically active/inactive agents effects the spreading. For this we altered the original model a bit, and introduced some new states, and state-change rules. We found here that even though the dynamic activity does not change the spreading qualitatively, it slows it down. Namely the needed time for the domination of informed agents shows faster than exponential dependence on the probability of being inactive. Finally we studied the effect of inhomogeneous activity and found that increasing only the activity of high degree nodes in social networks does not have the expected result. The activation of low degree nodes is also needed to make the spreading significantly faster.

It is clear that additional investigations are required to make our findings more precise. However these results may be found useful in the future when planning online advertisement campaigns or anti-spam actions. As a further step we would like to examine spreading phenomena on spatially and temporally dynamic networks in order to take one more step to bring our results closer to reality. It is also known from the literature that beside the degree distribution the structure of the community also plays an important role in spreading phenomena [24]. Based on this idea it would be also promising to examine the spreading on networks with different community structures.

V. ACKNOWLEDGEMENT

The publication was supported by the TÁMOP-4.2.2.C-11/1/KONV-2012-0001 project. The project has been supported by the European Union, co-financed by the European Social Fund.

The computations were performed on the HPC centre provided by the Hungarian National Information Infrastructure Development Institute at the University of Debrecen.

REFERENCES

[1] E. M. Rogers, "Diffusion of Innovations", Fifth edition, New York, Free Press, 2003.

[2] W. Weidlich, "Sociodynamics: A systematic approach to mathematical modelling in the social sciences", Dover Publications, Mineola, NY, 2000, pp. 113–187.

[3] M. E. J. Newman, "Networks - An introduction", Oxford University Press, 2010, pp. 627–675.

[4] B. M. Varghese, „The Life Cycle of a Social Network", Techpedia, [Online] Available: http://www.techipedia.com/2011/socialnetwork-life-cycle, Sept. 2011.

[5] G. Kocsis and I. Varga, "Investigating the effectiveness of advertising on declining social networks", accepted to Creat. Math. and Info., 2014.

[6] I. Varga, A. Németh and G. Kocsis, "A novel method of generating tunable underlying network topologies for social simulation", in Proc. of 4th IEEE Int. Conf. on CogInfoComm., Budapest, 2013, pp. 71–74.

[7] I. Varga and G. Kocsis, "Scale-free network topologies with clustering similar to online social networks", submitted to SMSEC, Kobe, 2014.

[8] R. Albert and A.-L. Barabási, „Statistical mechanics of complex networks", Rev. Mod. Phys. vol. 74, pp. 47-97 Jan. 2002.

[9] T. Vicsek, „Fractal Growth Phenomena", World Scientific, Budapest 1989, pp. 103-112.

[10] F. Papadopoulos, M. Kitsak, M. A. Serrano, M. Boguna, and D. Kriukov, "Popularity versus similarity in growing networks", Nature, vol. 489, pp. 537–540, Sep. 2012.

[11] M. Gjoka, M. Kurant, C. T. Butts and A. Markopoulou, „Walking in Facebook: A Case Study of Unbiased Sampling of OSNs", in Proc. of IEEE INFOCOM '10, San Diego, CA, 2010, pp. 2498-2506

[12] R. Albert and A.-L. Barabási, "Emergence of Scaling in Random Networks", Science, vol. 286, no. 5439, pp. 509-512, Oct. 1999.

[13] G. Caldarelli and A. Capocci and P. De Los Rios and M. A. Muoz, "Scale-Free Networks from Varying Vertex Intrinsic Fitness", Phys. Rev. Lett., vol. 89, p. 258702, Dec. 2002

[14] H. Y. Lee, H. Y. Chan, and P. M. Hui, „SF networks with tunable degree distribution exponents", Phys. Rev. E, vol. 69, p. 067102, Jun. 2004.

[15] G. Bianconi and A.-L. Barabási, „Competition and multiscaling in evolving networks", Europhys. Lett. vol. 54, pp. 436-442, May 2001.

[16] R. Albert, H. Jeong and A. L. Barabási, "The Internet's Achilles' Heel: Error and attack tolerance of complex networks", Nature, vol. 406, pp. 378-382, Jun. 2000.

[17] G. Kocsis and F. Kun, „Competition of information channels in the spreading of innovations", Phys. Rev. E, vol. 84, p. 026111, Aug. 2011.

[18] M. E. J. Newman and J. Park, „Why social networks are different from other types of networks", Phys. Rev. E, vol. 68, p. 036122, Sep. 2003.

[19] E. Cho, S. A. Myers and J. Leskovec, „Friendship and mobility: User movement in location-based social networks", in Proc. Int. Conf. on Knowledge Discovery and Data Mining, 2011, pp. 1082-1090.

[20] K. Kosmidis, A. Bunde, "Propagation of confidential information on SF networks", Phys. A, vol. 376, pp. 699-707, March. 2007.

[21] X.G. Tang, E.W.M. Wong, "Information traffic in scale-free networks with fluctuations in packet generation rate", Phys. A, vol. 388, pp. 4797-4802, Nov. 2009.

[22] G. Kocsis and I. Varga, "Agents based simulation of spreading in social-systems of temporarily active actors", accepted to Cell. Autom. for Research and Industry, 2014

[23] B. A. Huberman, D. M. Romero and F. Wu, "Social networks that matter: Twitter under the microscope", 1st Mon., vol. 14, p. 1, 01. 2009.

[24] L. Weng, F. Menczer and Y. Ahn, "Virality Prediction and Community Structure in Social Networks", Sci. Rep., vol. 3, 2522. 2013.

Gergely Kocsis was born in Hatvan, Hungary, in 1983. He received the M.S. degrees in computer science and Hungarian-English special translation from the University of Debrecen, Hungary, in 2007 and the Ph.D. degree in computer science also from University of Debrecen, in 2013.

From 2010, he was an instructor and from 2013 an assistant professor at the University of Debrecen, Department of Informatics Systems and Networks. His current research interests are mostly focused on social spreading phenomena, complex networks and dynamic network topologies.

Imre Varga was born in Hajdúböszörmény, Hungary, in 1979. He received the M.S. degrees in physics and informatics teaching from the University of Debrecen, Hungary, in 2002 and the Ph.D. degree in physics also from the University of Debrecen, in 2008.

From 2002, he was a computer specialist and part-time teacher and then he is now the assistant professor of University of Debrecen, Department of Informatics Systems and Networks. He is the author of a book and several articles. His research interests include computer simulation of structure formation in dipolar monolayers and study of complex networks and information spreading on them.

Multimedia Communications: Technologies, Services, Perspectives
Part II. Applications, Services and Future Directions

Leonardo Chiariglione and Csaba A. Szabó

Abstract—**This survey/position paper gives an overview of the state-of-the art multimedia communications technologies and services, analyses their present significance and expected future role, and attempts to identify development trends. The paper analyses the evolution of networking infrastructure and multimedia services over the last decade and identifies future directions. It consists of two parts. Part I, published in the preceding issue of this journal, dealt with the technologies and systems for multimedia delivery, and covered the dedicated networks such as digital broadcasting systems and IPTV and the technologies of Internet based multimedia delivery. The present paper, Part II, addresses applications, services and future directions.**

Index Terms—**Multimedia communication, IP networks, Internet, mobile communications.**

I. INTRODUCTION

The survey paper, written a decade ago by Stephen Weinstein and Alexander Gelman [1], was cited in detail in Part I [2] since one of the objectives of the paper was to analyse the state-of-the-art and to see what trends could be observed, after ten years since the paper was published. It was shown how the networking infrastructures and services have developed, and now in this paper we want to show whether the forecasted applications have gained wide acceptance and implementations and what new trends can be identified that were not foreseen that time by Weinstein and Gelman.

The paper is organized as follows.

In Sections 2.1 and 2.2, we come back to digital broadcasting and IPTV, with short comments on the service aspects.

In Section 2.3, the technology as well as service-related issues around the emerging Internet TV and OTT – Over-The-Top content services will be addressed.

The social element is gaining increasing role and importance in media consumption. In Section 3, dealing with “social media” and “social TV”, we look into issues around these terms and point out to the importance of social network based interactions among users of multimedia services.

Submitted on May 14, 2014, revised September 7, 2014.

L. Chiariglione is with CEDEO.net, Via Borgionera, 103, 10040 Villar Dora (TO) Italy (e-mail: leonardo@leonardo.net).

Cs. A. Szabo is with the Department of Networked Systems and Services, Budapest University of Technology and Economics, Magyar Tudósok krt. 2, Budapest, 1117, Hungary (e-mail: szabo@hit.bme.hu).

Section 4 deals with key application areas such as entertainment, e-health and telemedicine, visual collaboration and e-learning, and smart city applications and services.

Our concluding Section 5 outlines some promising directions including multi-screen TV, Free-viewpoint TV and the trend of moving from traditional broadcasting platforms to wireless broadband Internet.)

II. TELEVISION SERVICES PROVIDED OVER DEDICATED NETWORKS

A. Digital television broadcasting

In Part I, we discussed the technologies of digital television systems. While the technical solutions are interesting for engineers, the question is what do they mean for consumers. The short answer is: improved picture quality including HD. This is not too much, given the expenses the customers need to bear by buying a digital TV set or a set-top box that allows for using the existing analogue receiver and a new rooftop antenna. Additional benefit is a greater choice of channels. A digital multiplex as it is called, a package in which TV channels are grouped, usually contains up to 8 TV channels and several radio channels, and is being provided for free. Additional multiplexes are included in a paid monthly subscription.

We should mention the importance of digital terrestrial broadcasting for low-income population groups, living in rural areas, in less developed countries. For them, the digital switchover means that the terrestrial broadcasting will survive in its new form, and thus their only way of accessing news, entertainment programs and taking part in educational programs will remain available in the foreseeable future.

Interestingly enough, the transition from analogue to digital broadcasting paves the way towards replacing this relatively new (since it is digital) but at the same time old (since it uses dedicated broadcasting systems) method by a new one: distributing and consuming TV programs using the public Internet. This is because, as a result of the analogue-to-digital switchover, a considerable amount of bandwidth within the so-called digital dividend has been/will be freed up and will be available for other purposes. There are several applicants for these frequencies, and most likely a great deal of it will be allocated to mobile cellular service providers, which then will use it for expanding their new generation services, and first of all mobile broadband Internet access. This process could eventually lead to the decline of digital broadcasting, although it may be too early to say that. We shall come back to this issue later in this paper.

B. IPTV

In Part I, the technology and architecture of IPTV systems were briefly presented. Although the term IPTV itself denotes TV delivery over any IP network, the most common form is the existing last mile network, the subscriber loop, of the telecom service providers. Thus it is obvious that IPTV is an excellent opportunity for telecom operators to enter into broadcasting business without having to build a new network infrastructure for it. In addition to the original telephone service, and the already provided Internet access over xDSL, TV distribution becomes the third component of their „triple play“ – voice + Internet + TV – offering.

C. Internet TV and OTT

Currently we use two distinct ways of media consumption. On the one hand, we have the linear media consumption which is the traditional TV broadcasting, with its edited programmes, that are broadcasted and distributed based on a schedule known in advance. On the other hand, non-linear media consumption has been gaining increasing importance during the last years. Consumers use their desktops or laptops with Internet access to access and download media they choose to watch/listen, whenever they want, repeatedly or in parts if they wish so, very similarly to the old VHS or the recent DVD players. Non-linear way of accessing content is preferred by young generations, although the principle itself should not be unfamiliar also to non-Internet generations: after all books are a non-linear content, too.

OTT or Over-the-Top Content can be seen as an effort to bridge this gap. Rigid boundaries seem to be ceasing anyway. Non-linear content coming from home media players are being consumed on TV sets. Linear content can be viewed also on desktops, laptops or smartphones. Smart TVs are penetrating, they seem to be the device to combine these two media consumption ways. Smart TVs or connected TVs or hybrid TVs integrate Internet and Web2.0 capabilities into TV sets and set-top boxes. Currently, about 10 percent of the total Internet video traffic is being delivered to TV sets and this number will grow to 14 percent by 2017, according to [3].

What does the term OTT cover? Technically, it is content distribution over the unmanaged public Internet. From the business and service model point of view it is about separating the content provider and distributor/network service provider roles (the latter are the ISPs). This model is fundamentally different from that of the traditional broadcasting companies or of the majority of IPTV providers (who can be called “vertical” service providers). OTT could be called Internet TV as opposed to IPTV that uses a dedicated, managed IP network. On the other hand, it is more adequate not to call it Internet TV, since this term is too general and only refers to the technical side, while OTT is also a service/business model. Table 1 compares OTT and IPTV.

	OTT	**IPTV**
Delivery method	Over the open Internet	Using a proprietary network
Network ownership by the content provider/ aggregator	No, network provider and content provider roles are separated	Usually the content distributor owns the delivery network
Quality of Service	In general, best effort provided by an ISP, a CDN may be used to improve it	QoS can be guaranteed
Protocol for media transport	HTTP/TCP, adaptive streaming like HLS is emerging	Transport stream over RTP/UDP
Routing topology	Unicast	Multicast

Table 1 Comparison of OTT and IPTV

The components of the OTT “ecosystem” are the following entities:

- Content providers. Example: BBC (UK).
- Content aggregators and distributors. Example: Netflix (USA, UK).
- ISP – Internet Service Providers. Examples: Verizon (USA), T-Home (in many European countries).
- CDN – Content Delivery/Distribution network provider. Examples: Akamai (USA), StreamZilla (The Netherlands).
- Access/core network provider. Examples: most traditional telecom service providers and mobile operators.

Most well known OTT service providers are Amazon, Apple, Hulu, Netflix [4], [5], [6], [7].

How can the service provider ensure QoS without owning the distribution network or having access to it (because the distribution network is the Internet)? Only monitoring and collecting information at the receiver side is possible, without intervention. ISPs cannot monitor the content of the IP packets (otherwise they would violate the “network neutrality” principle). Fortunately, broadband Internet is penetrating, with increasing quality and reliability. In EU countries the goal is to grant at least 100 Mbps access data rate for the whole population by 2020. Also the inclusion of a CDN (Content Delivery Network) provider into the delivery process supports quality of service.

Can we combine IPTV, the TV broadcasting technology over dedicated, managed IP networks and the OTT? It seems yes, the term Hybrid TV refers to a consumer TV set where TV content is delivered by digital television technology (DVB/T/S/C), or, by an IPTV service, and, since the set is connected to the public Internet, all kind of multimedia content, including TV programs, can be accessed. Current hybrid TV platforms of leading vendors (Samsung, LG, Sony) are proprietary ones, thus standardization is a prerequisite for penetration.

HbbTV (Hybrid Broadcast Broadband TV) is a new standard providing an open and neutral technology platform that seamlessly combines TV services delivered via broadcast with services delivered via broadband and also enables access to Internet only services for consumers using connected TVs and set-top boxes. Founders are European telecom and broadcasting companies, TV equipment vendors and other organizations [8]. It is an ETSI standard [9]. As for the current status, HbbTV has been deployed, among others, in Germany, Austria, Switzerland, France, and is being tested in many other countries in Europe. Analysts predict a distribution of 23 million HbbTV-capable devices by 2014 [10]. Nevertheless its future depends on whether it remains a pure European initiative or it will gain acceptance by and support of the US entertainment industry to ensure worldwide penetration.

III. THE SOCIAL ELEMENT IN NETWORKED MULTIMEDIA APPLICATIONS

A. *Social media*

Social media is exactly what the two words suggest: it is an intersection of social networks and personal media. On the one hand, social networks have become an integral part of our everyday life, we exchange information and communicate to a large extent through FaceBook and similar networks. On the other hand, personal media is becoming more and more social. Let us consider photos as an example. Old ways of collecting and storing of personal photos about ourselves, family events, travels and presenting them to family members and friends are being replaced by doing the same via photo and video sharing sites on the Web. Replacing our paper boxes at home where we used to store photos by „electronic boxes“ like the hard drive of our laptop or external storage devices is just the first step. We use new ways of „showing“ them to all people who might be interested, and given the enormous amount of photos we take, due to the ubiquitous use of digital cameras, smartphones and tablets equipped with cameras, we really need new ways for the storage and presentation. We can easily select photos we would like to share and upload them to some sharing site, and we can select photos that we want to see from those available on these sites. Moreover, we can use social networks to disseminate information about our personal media, and to search for the desired content on these sites. So, social media is also a new approach to multimedia content search that uses collaborative annotations of the members of social networks.

Ramesh Jain, Univ. of California at Irvine, in his lecture at a meeting of the European project „nextMEDIA” defined the characteristics of social media as follows [11]:

- it is produced and consumed by many,
- production and consumption is democratized,
- sharing plays an essential role,
- media is multimedia content (today: picture, video, voice, data, in the future: smell, tactile info etc.).

We started to talk about photos as an example of personal media, which is just one kind of UGC – User Generated Content. Another example is digital storytelling, which is a combination of narrative with digital content, including images, sound, and video, to produce a short movie, typically with a strong personal character and usually with emotional component. Who is doing digital storytelling? BBC was pioneering in helping digital storytelling to penetrate, by encouraging and sponsoring people all around the UK to capture videos on local histories and cultures. Another early example is the San Francisco broadcasting station KQED which solicited high school students to shoot stories about how they are living in California. Digital storytelling has a great significance as an educational tool, as well.

It is very little what people need to produce a digital storytelling video: a script and some hardware and software, including a capturing device (video camera, smartphone, iPad etc.) and an editing environment (a desktop or laptop with the necessary inputs and with an editing software, e.g. Apple iMovie, Microsoft MovieMaker).

Since it is so easy to produce some video clips or sound recordings, and, even more importantly, there are always people around when something interesting and surprising happens, UGC is a part of professional media production these days. Mobile users frequently generate content for broadcasters and news portals. Watching a TV news channel, we often see inserts produced by people who were just there where something happened before the TV team could have arrived at the spot.

B. *Social television*

Today's media consumption is characterized by three independently developed and operated systems/environments/ opportunities [12]:

- *Home living room environment.* It is characterized by large plasma, LCD or LED screens, HD and 3D presentation, with high-fidelity sound system are the main components. Content comes (i) from the Internet and presented on living room equipment (supported by proprietary „media center“ type solutions), and (ii) from air or cable as before, and more recently, via IPTV. This set-up can be called a „linear“ media consumption environment, with limited interaction possibilities.

- *Desktop environment.* Since people spend long hours in front of the screen of their PCs or notebooks, both in the office and at home, no wonder that this environment is used more and more also for media consumption. Numerous TV and radio broadcasting programs are directly available on the Internet, but content can be acquired from the social sites (Facebook, YouTube) where content consumption is often accompanied by annotation and recommendation services, thus bringing more interactivity than can be achieved in the living room environment.

- *Mobile media consumption.* As we discussed earlier, people use their mobile and portable devices (smartphones, tablets) for media consumption, production and interactions within their communities. Radio broadcasting receivers are often built in smartphones. TV channels are accessible via Internet, alternatively DVB-H broadcasts can be received where available. Internet connectivity is available almost everywhere at high speed via Wi-Fi and 3G/4G mobile. Interactivity is supported by the inherent communication capabilities: voice calls, SMS, MMS, video calls. Users generate personal content and share it via social networks.

If we combine the three settings and use the social networks for tagging, recommendations etc. we can create community interactions in the context of viewing TV programs. For example, users can collect multimedia information related to the piece of media just being watched/listened to. Another example is collecting and presenting multimedia information related to the geographical location the user is currently visiting. In general, we can make decisions on what to watch based on peers; share programs or edited versions, directly communicate via chat, audio, or video with other peers and comment about a television program within a (possibly) large community, and make available to others what we are watching.

In the yearly publication of MIT, social TV was listed among the „10 most important emerging technologies“ in 2010 [13], see also [14] . The prototype developed at MIT was built around a central database that aggregated video from online sources like YouTube, shares user-specified data with social networks, delivers video to the user's TV, lets users and the people in their networks send comments and ratings via an iPhone app. On social TV, see also [15] and [16].

IV. Main application areas of multimedia communications

A. Entertainment

By 2017, the total spending on online games will reach the amount spent on offline console games [17]. The latter is a huge industry, for global data see [18].

According to Spilgames, a Dutch-based global provider of online gaming platforms, the percentage of online population who plays online games is 44%. [19]. Out of roughly 1.2 billion gaming users 700 million was playing online games in 2013. This activity is on the 2nd place among online activities after watching videos, and before watching TV and radio programs and movies. The share of gaming traffic on the Internet is even more significant if we take into account the time spent on gaming compared with other activities. According to the aforementioned report, the average session duration is 40 min for games, 15 min for YouTube and 5 min for news portals. Cisco forecasts that in 2017, the amount of gaming traffic is expected to be 59 petabytes, compared to around 53 pB of video traffic, around 15 pB of web, e-mail and data traffic and 9 pB of file transfer traffic (data from [3], rounded figures).

Online games on mobile devices is a new but potentially very large market and a fast growth is expected in terms of time subscribers spent on games and the traffic generated. There are technical challenges such as ensuring adequate QoE.

The Entertainment Software Association, ESA, provides data on what people play. About one-third, 34%, is the group containing puzzle, board game, game show, card games, trivia. 26% of games played fall into the category of action, sports, strategy and role-playing. Social games is an increasing group, currently (2013) it constitutes 19%.

The massively multiplayer online role-playing game (MMORPG) is a specific genre characterized by a very large number of players, usually in the order of several thousands in one server. The most popular ones are World of Warcraft, that has close to ten million subscribers or Star Wars: The Old Republic, which, after its release in 2011, became the world's fastest-growing MMO after gaining 1 million subscribers within the first three days of its launch. Guild Wars is also among the most popular ones.

The business models of MMORPGs are based on monthly subscription fee, or one-time fee, and users have to purchase the client software as well. An increasing trend is to use web browsers as clients. There are different additional fees to be paid for access to extended areas of the game, for example. There are games available free of charge too.

The share of the game traffic on the Internet, although growing, is still and for some time remains to be a small fraction of the total Internet traffic, due to the nature of data transmitted while gaming (as compared with other web applications, file sharing, voice and video applications). Gaming data include the inputs of the players, chats among players, virtual world updates etc. The reason is that most popular games were developed long ago when bandwidth was scarce and access was mostly dial up at low speed. The second generation of game software will make use of the since then dramatically improved access bandwidth and will possibly generate significantly larger data traffic.

As for the technology aspects, most MMORPGs, as other kinds of online games, are based on a client–server system architecture. The server software generates the environment, the virtual world, and players connect to it via client software. A game is usually hosted by several servers in parallel so that the number of players does not exceeds a few thousand, although there are exceptions when a server hosts several tens of thousand users. P2P architecture plays a negligible role and is currently a research issue. Game development technology is about 3D modeling, graphics, animation and games programming.

Finally, let us mention a related area, serious games which have been used for quite a long time as educational tool sin different areas. Currently the majority of serious games applications are being used off-line but the on-line share is growing due to the high demand for flexible forms of education.

B. Health care

The use of information technology in health care, briefly called e-health is another rapidly growing application area these days. It includes telemedicine and home health care services. Telemedicine is about providing health care services to distant locations. According to ATA (American Telemedicine Association) telemedicine is "exchange of medical information by telecommunications means to improve treatment". The first documented telemedicine service dates back to late '60s when a microwave link between the Logan airport in Boston and the Massachusetts General Hospital was installed for transmission of diagnostic data from the emergency room of the airport to the hospital for remote consultation. Already this set-up consisted of the two main components of a telemedicine system: a store-and-forward transmission of diagnostic picture and video and live audio/video communications for medical consultation.

M-health or mobile health solutions are of great importance where wireline communications is missing or of bad quality (there are many such areas and not only in developing countries), and are useful for patient monitoring at home and while the patient is moving (e.g., permanent remote ECG monitoring). Portable screening stations make mass screening possible in remote, sparsely populated areas. Medical services can be provided to disaster areas where the wireline infrastructure is usually damaged.

An ongoing European m-health project addresses scenarios like the following [20]. A young man is jogging in a park. Suddenly he falls down and remains laying, apparently he cannot move, possibly also collapsed. His cellphone (a smartphone) detects the event, searches for mobiles in the vicinity, founds one, sends a message to it, and its owner, a lady runs to the young man. She recognises that the young man collapsed, alarms the emergency service. Contacts the emergency center and using her smartphone, transmits pictures and videos of the patient, and also describes verbally what she sees. Based on this preliminary information, the emergency center decides which hospital the young man shall be delivered. In the emergency car, hearth ultrasound pictures are taken and transmitted to the hospital using a wireless link between the car and the hospital. Then the doctors in the hospital can plan what the best treatment shall be.

An important area of telemedicine where networked multimedia plays a key role is telesurgery. There have been telesurgery experiments for more than a decade. Since the first experiments, the robotic surgery has developed significantly and the Da Vinci robots [21] have become standard technology in developed countries. Telesurgery is just one step from the Da Vinci robotic surgery room: the workstation has to be displaced from the operation room to a large distance. Of course this simple step poses serious challenges in the design and implementation of the communication system: to ensure the required quality of service for video transmission over various telecommunications links, including wireless and mobile systems and services. Space surgery presents additional challenges because of the large propagation times, smaller available bandwidth, and lower reliability.

Some methods, solutions and systems developed for telemedicine purposes have been implemented in a slightly different and highly significant area, namely in providing health care and in general, providing ambient assisted living (AAL) services to elderly people in their homes. According to a EU statistics, the percentage of citizen over 65 years in the 27 European Union member countries is expected to become 25% by 2020 [22]. The majority of elderly live alone at home, thus the various means helping them are of paramount importance. Solutions include regular or permanent monitoring of important vital parameters and transmission of these data to a health care provider site, or monitoring the everyday activity within the homes and outside the homes to detect unusual and dangerous situations such as a sudden fall. Since home health care and AAL are highly important for society, governments and international organizations allocate significant funds for the research, development, implementation and operation of such services. An example is the Ambient Assisted Living Joint Programme of the European Union [23].

C. Audio-visual collaboration and e-learning

Interactive audio-visual communications is a basis for many important applications including virtual meetings, virtual conferences, and e-learning. The technology is based on the call control protocols. Currently two call/session control protocols co-exist: the H.323 ("umbrella standard"), developed in the telecom world and standardized by ITU [24] and SIP – Session Initiation Protocol, developed almost in parallel in the Internet world and standardized by IETF [25]. (SIP was dealt with in Sub-section 3.2.3 within the context of IMS.) Based on these protocols as well as on sophisticated audio/video compression methods, a large variety of audio-visual collaboration systems (also called video conferencing systems) have been developed, ranging from soft clients for personal computers through the different kinds of conference room systems to multi-screen "video walls". Example is the widely used is TelePresence systems by Cisco [26].

Videoconference systems are the technical means for "virtual classroom"-type distance learning, one of the two main methods used for distance learning. As the name suggests, it is about delivering a lecture to a remote lecture hall or halls. Thus the goal is to bridge large geographical distances. A related application is the "extended" lecture hall, meaning the extension of the real lecture hall, where the presenter gives a live lecture to the participants, to additional (remote) rooms. Examples are conferences with a large number of participants or inter-university lectures.

The other main method, gaining increasing importance, is web-based learning, or e-learning, which is based on hypermedia learning materials accessed over the public Internet or on the intranet of a company or institution. In addition to complete on-line courses, shorter forms such as webinars and webcasts can be used in web-based learning. The advantages are obvious: it provides personalized learning from any place and time. Instructors still play an important role, their active help (e.g., tutoring via e-mail, phone, Skype, video conference) is an integral part of the learning process.

E-learning is a fast growing Internet application area due to strong interests both from the learners as well as from the educational institutions. Learners are interested in flexible forms of learning. They want to obtain additional degrees while working which often requires enrollments in programs of distance universities. Educational institutions can make their programs accessible for a wider population at lower costs.

Online courses are provided either by universities themselves (a well known example is MIT - Massachussetts Institute of Technology - with its OpenCourseWare program) or by independent organizations that are still connected to and based on the offering of a university or several universities. edX is a non-profit organization established by the Gates Foundation and is based on a cooperation with leading US universities such as MIT, Harward, University of California at Berkeley and University of Texas [27]. Another major player in the so-called MOOC - Massive Open Online Courses – business is Udacity, founded by Harvard University [28]. The largest is perhaps Coursera, established by two Stanford professors, with its 22 million enrollments from over 190 countries across 571 courses [29].

Discussing the issues of educational methodology related to open courses is outside of our scope. Let us only mention that providing course materials in the form of video lectures (which is currently the most common format, and, from educational methodology point of view, does not really differ from classroom lectures) is just one component of the learning process. One needs to add tools and create environment for the whole distance or on-line learning process, e.g. consultations with the professors, tutoring, certification of courses and programs, examination, etc. There are also organizational and legal issues such as how the participation in open lectures can lead to degrees, and also what are the suitable business models. Most of these and the like are unanswered questions at the time of writing of this paper.

D. Smart environments

Networked multimedia plays a central role in another emerging application area which can be broadly denoted by the term „smart environments“. Related terms are, in the context of cities: „digital cities“, „digital communities“, „smart cities“ [30], in the context of the citizens‘ immediate environments: „smart homes“ or „intelligent“ homes or „ambient assisted living“, in the context of exhibitions and other mass events: „smart spaces“. In general, we are talking about new applications and services for the benefit of citizens, organizations and businesses. For example, in the context of cities, let us have a brief look at the most important smart city application areas.

Smart metering and control for utility services.

Smart metering or automatic meter reading (AMR) is the technology of automatically collecting consumption, diagnostic, and status data from different utility metering devices such as water, gas, electricity and heat meters and transferring that data to a central database for billing, analysing and troubleshooting. Using an AMR system, failures or misuses can be detected immediately, making possible instant intervention. Billing can be based on near real-time consumption to better control the production and consumption of public utility services.

Parking is a big problem in most modern cities in particular in downtown areas, and many cars are looking for parking lots in every moment, resulting in waste of time, fuel, increasing air pollution and noise load. Assisting drivers in finding parking places is similarly important in parking garages. An intelligent parking assistance system navigates cars to the appropriate available parking lots, navigates drivers back to their cars, and, among other related applications, supports billing.

Let us only list some otherequally important application areas:

- *Smart city transportation*
- *Ubiquitous access to government and community portals*
- *Smart applications for tourists*
- *Meeting the needs of elderly people*
- *Urban safety*
- *Public health applications*

Key technologies for smart cities are (i) a networking infrastructure, mostly wireless, and (ii) sensors installed everywhere, in parking lots (infrared sensing of occupancy), on buildings (e.g. surveillance cameras), in citizens homes (metering devices), just to mention a few examples. Multimedia communications plays a key role on most of the aforementioned applications.

IoT – Internet of Things is a frequently used term to denote the system of networked sensors embedded in physical objects and machines. „Internet“ because this network operates using the same TCP/IP protocols stack as the public Internet, and uses Internet infrastructure as well.

V. Outlook and summary

In this paper, we gave an overview of the state-of-the art networked multimedia technologies and services.

As for the future, it is impossible to identify a single unified direction where multimedia services are developing. Instead, in this concluding section, we briefly address some interesting and promising approaches, including solving challenges in providing broadband access, novel multimedia technologies, and new business models for providing multimedia services.

A. Multi-screen TV

While until recently the term multi-screen denoted different kinds of video walls with several screens, for example displaying the participants of a video conference plus presentations and videos delivered by participants, different stakeholders of the TV and mobile industry gave it a different meaning. For TV service providers, it means extending the viewing experience of customers from smart TV sets to a second device such as a table PC or a smartphone, basically adding tablet and smartphone participation to classical TV viewing. The terms “Companion TV“, „second screen TV“ refer to the same thing.

In general, multi-screen TV means enabling the service providers to deliver any content to any screen via any network within the framework of a single integrated system (“TV anywhere“). It seems that while smart TV offers a combination of traditional broadcasting content and interactive content accessible from the Internet, there is a significant demand for an application where second-screen or multiple-screen viewing can be delivered and customized via a mobile handheld device, as a flexible extension to the primary large-screen TV set.

As an example, WatchON™ of Samsung [31] allows for accessing content across different Samsung devices. With WatchON, a mobile smartphone or tablet PC can be paired with the main smart TV, and the selection of content and device is controlled via a smartphone application. In addition to multi-screen viewing, this application provides intelligent program search, personalized recommendations, complementary content and social network sharing.

Ericsson claims that "TV consumers today demand a seamless way to consume any content, on any screen, anywhere. They want to have a social experience while watching TV. They want personalized and rich experience across devices." [32] Ericsson's „Multiscreen TV Solution" is a complete system consisting of Multi-screen clients and Multi-screen Core Platforms. The core functionality includes content management, middleware, VOD backoffice, digital rights management among others. The client functionality includes common client framework, SDK and portals.

B. Future 3D Media Internet

As it was discussed in Part I, there have been important research initiatives on both sides of the Atlantic to support research of Future Internet network architectures. Several directions have been listed in which new solutions were required to meet the new demands. Perhaps the most important challenge is to support 3D media delivery on the Future Internet, and the solutions embedded in the Future Internet to meet this challenge are denoted by the term „Future 3D Media Internet" [33].

Requirements toward the Future Media Internet, listed and analysed in [34], [35], and [36] include coding, transmission and presentation of 3D audio/video + additional sensorial information (pressure, vibration, smells etc.), interactivity with 3D media from any user devices, inclusion of 3D handling capability in browsers.

The main challenge is transmission and delivery of 3D information while meeting the desired QoS and QoE requirements. There are several directions in which the current Internet should evolve to become Future Media Internet. Just to mention one: 3D media transport. The current protocol stack of RTP/UDP/IP will most likely be replaced by RTP/DCCP/IP. TCP is usually not an option in most media applications (streaming, interactive media, online games) because fast delivery is needed and the reliable transport mechanism of TCP causes unaffordable delays. On the other hand, UDP lacks a congestion control feature which is important for the transmission of large amounts of multiview 3D video. DCCP or Datagram Congestion Control Protocol seems to be a solution. DCCP provides congestion control for datagrams in bidirectional unicast connections. The transmission of datagrams occurs unreliably just as with UDP.

C. Free viewpoint 3D TV

As it was noted earlier in this paper, current 3D stereoscopic technologies, used in cinemas and blue-ray players, are not attractive enough for 3D television due to the reasons mentioned above. Besides some limitations of the current technologies, such as specific glasses or the limitations of the current displays, it seems that viewers are not enthusiastic enough just about some depth experience, rather many prefer 2D viewing but with increased resolution. The main reason maybe is that these technologies provide only a single 3D view and the viewers cannot control their viewpoint.

A breakthrough technology could be free-viewpoint TV, a multiview system where the user can freely choose the viewpoint he/she wants to watch a scene from. The basic technology is the free viewpoint video, or FVV, which has been around for several years as a novel production and reproduction technology. In FVV, the scene is recorded simultaneously by several cameras. The recording is controlled by the master camera and the operator, who selects a part of scene to focus on and all other cameras are directed to the same part of the scene, but each from different location/angle, in a computer controlled way. The resulting multiple video stream is processed by appropriate coding technologies, transmitted to the receiver side, where it is decoded and rendered to a special display. So far this technology was commercially utilized mainly on the production side, examples being the stop-motion animation in the movie "Matrix" or in the "Eyevision" system for sport effects [37].

According to [38], the world's first real-time free-viewpoint TV system, including the complete chain of operation from image capture to display, was constructed by Masayuki Tanimoto and his team.

Part of this complicated chain, namely, the multiview coding, has been already standardized. The Multi-view video coding (MVC) standard is an extension of H.264/AVC, and is based on exploiting the redundancies not only between frames of a given view but also between frames of different views [39]. MPEG is now (in 2013) engaged in the third phase of FTV standardization, with the objective of establishing a new FTV framework. The first document [40] deals with use cases and requirements. There are important developments in the other elements of the FTV chain, e.g. in multiview 3D displays, so one can expect the appearance of the first commercial TV applications in the near future.

D. OTT vs digital broadcasting

We noted earlier that the transition from analogue to digital terrestrial broadcasting has been recently completed in most countries worldwide. Analogue TV distribution in CATV networks will be terminated also soon. But what will be the future of digital TV broadcasting? As we also mentioned, the bandwidth savings due to digital switchover will help the penetration of wireless broadband Internet access which is going to be increasingly used by customers for enjoying TV content as opposed to using traditional digital broadcasting and distribution services. Many think – and the authors belong to them - that the days of DVB-T are numbered. This service will probably continue to grow in the next few years but then a decrease will likely start. IP-based TV distribution over dedicated networks – IPTV - will remain significant in the next few years. And mobile broadband Internet access will grow due to reuse of the digital dividend and later due to freeing up new bands. Consequently the share of the non-linear media consumption will grow. In the near future, customers will use a multiplicity of platforms, interfaces etc., no universal worldwide standard is expected to be established. As we already noted, the future of HbbTV is unclear.

This trend was recognized by the Communications Committee of the British Parliament. In their recent document [41], it is stated: “We recommend that the Government, Ofcom and the industry begin to consider the desirability of the transfer of terrestrial broadcast content from spectrum to the Internet and the consequent switching off of broadcast transmission over spectrum, and in particular what the consequences of this might be and how we ought to begin to prepare.”

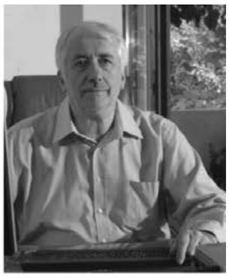
The development in the mobile broadband technologies seems to be supporting this trend. Existing HSDPA and LTE technologies already offer significant download speeds as compared with 3G services. And the move to the next generation is underway. Ericsson, one of the market leaders in mobile system technology, a company investing a lot in related research, is talking about 5G as evolution of existing standards plus complementary new technologies. They say “5G will enable the Networked Society and realize the vision of unlimited access to information for anyone and anything. This vision will be achieved by combining evolved versions of today’s radio-access technologies (RATs), including LTE and HSPA, with complementary RATs for specific use cases, not by replacing existing technologies. Future mobile broadband users will expect “unlimited” performance from the network.” [42]. Moving to new generation mobile broadband services is also in the focus of the recently launched program, 5G PPP, within the new Horizon 2020 framework program of the European Union. According to [43] the 5G PPP will deliver solutions, architectures, technologies and standards for the ubiquitous next generation communication infrastructures of the coming decade. It will provide such advancements as 1000 times increase in wireless capacity serving over 7 billion people (while connecting 7 trillion “things”), saving 90% of energy per service provided, and creating a secure, reliable and dependable Internet with zero perceived downtime for services.

To sum up what was outlined in this concluding section, it is likely that we will be witnessing dramatic changes in networked multimedia to take place in the near future: a move from the present digital broadcasting to mobile broadband Internet, at the same time a transition from linear to non-linear media consumption, and the commercial introduction of novel 3D technologies that will provide unprecedented viewing and listening experience.

REFERENCES

[1] S. Weinstein and A. Gelman, “Networked Multimedia: Issues and Perspectives”, IEEE Communications Magazine, Vol. 41, No. 6, pp. 138-143, June 2003.
[2] L. Chiariglione and Cs. A. Szabo, „Multimedia Communications: Technologies, Services, Perspectives. Part I. Technologies and Delivery Systems”. Infocommunications Journal, Vol. VI, No. 2, pp. 27-39, June 2014.
[3] Cisco Visual Networking Index: Forecast and Methodology, 2012–2017. Cisco White Paper, May 29, 2013. Downloaded from http://www.cisco.com/c/en/us/solutions/collateral/service-provider/ip-ngn-ip-next-generation-network/white_paper_c11-481360.html, on February 24, 2014.
[4] www.amazon.com
[5] http://www.apple.com/hu/appletv/#tv-overview
[6] www.hulu.com
[7] www.netflix.com
[8] www.hbbtv.org
[9] "Hybrid Broadcast Broadband TV” ETSI TS 102 796 V1.1.1 (June 2010).
[10] S. Arbanowski, “Hybrid-, connected-, Smart-, TV? Was steckt hinter der neuen Medienwelt.” Universitaet Potsdam, 2013. http://www.medienengineering.de/Ringvorlesungen/WS20122013/Arbanowski15.1.2013.pdf
[11] Social Media: Creation, Distribution and Access. Ramesh Jain, UCI. Presentation at the 1st Spring School on Social Media Retrieval (S3MR), Interlaken 2010. Accessible on http://videolectures.net/s3mr2010_jain_cda/.
[12] L. Galli, R. Guarneri, and J. Huhtamaki, “VERTIGO: Find, Enjoy and Share Media Trails across Physical and Social Contexts”, Digibiz, London, 2009.
[13] TR10: Social TV. Relying on relationships to rebuild TV audiences. „10 most important emerging technologies“, MIT Technology Review, May/June 2010, downloaded from: http://www.technologyreview.com/printer_friendly_article.aspx?id=25084.
[14] M.-J. Montpetit and M. Medard, „Social Television: Enabling Technologies and Architectures,“ Proc. IEEE, Vol. 100, No 5, May 2012, pp. 1395-1399.
[15] Pablo Cesar, David Geerts, “Understanding Social TV: a survey”. http://nem-summit.eu/wp-content/plugins/alcyonis-event-agenda//files/Understanding-Social-TV-A-Survey.pdf
[16] P. Cesar, D. Geerts and K. Chorianopoulos: Social Interactive Television: Immersive Shared Experiences and Perspectives. IGI Global, 2009.
[17] “Global Online Gaming Market 2014” YStats report, February 2014, downloaded from http://www.ystats.com/uploads/report_abstracts/1064.pdf?PHPSESSID=3656a3c0b0e2afdb46fab3740e414327 on March 2, 2014.
[18] Sales and demographic data, ESA – Entertainment Software Association, 2013.
[19] “State of the online gaming report 2013.” Spilgames, http://www.spilgames.com/press/2013-state-online-gaming-report-released-spil-games/
[20] A. Takács, L. Bokor, A Distributed Dynamic Mobility Architecture with Integral Cross-Layered and Context-Aware Interface for Reliable Provision of High Bitrate mHealth Services., Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering 61: pp. 369-379. 2013.
[21] http://www.davincisurgery.com/da-vinci-surgery/
[22] European Commission “2009 Ageing Report: economic and budgetary projections for the EU 27 Members States”.
[23] http://www.aal-europe.eu

[24] http://www.itu.int/rec/T-REC-H.323-200912-I/en
[25] http://www.ietf.org/rfc/rfc3261.txt
[26] https://learningnetwork.cisco.com/docs/DOC-8096
[27] www.edx.org
[28] www.udacity.com
[29] www.coursera.org
[30] Magnus Boman: Digital Cities. White Paper. EIT ICT Labs Innovation Radar, 31 December 2012, downloadable from: http://www.eitictlabs.eu/fileadmin/files/publications/130426_wp_digital_cities.pdf
[31] http://www.samsung.com/us/watchon/multi.html, downloaded January 23, 2014
[32] http://www.ericsson.com/ourportfolio/media-industry/multiscreen-tv
[33] Future Internet and NGN Design requirements and principles for a Future Media and 3D Internet Created by "Future Media and 3D Internet Task Force". Coordinated and supported by the Networked Media Unit of the DG Information Society and Media of the European Commission. February 2009. Available at ftp://ftp.cordis.europa.eu/pub/fp7/ict/docs/netmedia/20090220-fid-rp-3-dg_en.pdf
[34] P. Daras and F. Alvarez, "A Future Perspective on the 3D Media Internet", in: G. Tselentis et al (Eds.), "Towards the Future Internet". IOS Press, 2009.
[35] Position Paper on Networked 3D Multimedia. Accessible on www.nem-initiative-org.
[36] T. Dagiuklas, "3D Media over Future Internet: Current Status and Future Research Directions", International Journal of Computer Science Issues, Vol. 9, No 2, January 2012.
[37] A. Smolic, "3D video and freeview point video - From capture to display." Pattern Recognition 44 (2011) 1958 -1968.
[38] M. Tanimoto, M. P. Tehrani, T. Fujii, and T. Yendo, "Free vievpoint TV. A review of the ultimate 3DTV and its related technologies." IEEE Signal Processing Magazine, January 2011, pp. 67-76.
[39] ITU-T and ISO/IEC JTC 1, "Advanced video coding for generic audio- visual services", ITU-T Recommendation H.264 and ISO/IEC 14496-10 (MPEG-4 A VC), 2010.
[40] M. P. Tehrani, S. Shimizu, G. Lafruit, T. Senoh, T. Fujii, A. Vetro, M. Tanimoto, "Use Cases and Requirements on Free-viewpoint Television (FTV)". ISO/IEC JTC1/SC29/WG11 MPEG2013/N14104. Geneva, Switzerland. October 2013.
[41] Broadband for all - an alternative vision - Communications Committee of the British Parlament, 2011. Available at: http://www.publications.parliament.uk/pa/ld201213/ldselect/ldcomuni/41/4106.htm#a18
[42] Ericsson Labs Research Topics, downloaded from http://labs.ericsson.com/research-topics/5g, 17 February 2014.
[43] 5G Infrastructure PPP: The next generation of communication networks will be "Made in EU". Downloaded from http://ec.europa.eu/research/press/2013/pdf/ppp/5g_factsheet.pdf, 17 February 2014.



Leonardo Chiariglione obtained his Ph. D. degree from the University of Tokyo in 1973. During his career he launched or participated in several initiatives. Among these are MPEG and the Digital Media Project (DMP). He is currently CEO of CEDEO.net a company developing its own portfolio of technologies, products and services in the area of digital media, such as WimTV (http://wim.tv/). Leonardo is the recipient of several awards: among these the Eduard Rhein Foundation Award, the IBC John Tucker Award, the IEEE Masaru Ibuka Consumer Electronics Award and the Kilby Foundation Award.

Csaba A. Szabó obtained his MSc from St. Petersburg State University of Telecommunications, Dr. Univ., Ph. D. and Dr. Habil. from the Budapest University of Technology and Economics (BME) and the title Doctor of Technical Science from Hungarian Academy of Sciences. He is a Professor and Head of Laboratory of Multimedia Networks and Services at the Dept. of Networked Systems and Services of BME. Prof. Szabo has long-term experience in academia, R&D and telecommunication business. He co-founded a leading network system integrator and service provider company that was subsequently integrated into Hungarian Telecom. Prof. Szabo has been a member of editorial boards and EiC of several international journals. He is Senior Member of IEEE.

Guidelines for our Authors

Format of the manuscripts

Original manuscripts and final versions of papers should be submitted in IEEE format according to the formatting instructions available on

http://www.ieee.org/publications_standards/publications/authors/authors_journals.html#sect2,

"Template and Instructions on How to Create Your Paper".

Length of the manuscripts

The length of papers in the aforementioned format should be 6-8 journal pages.

Wherever appropriate, include 1-2 figures or tables per journal page.

Paper structure

Papers should follow the standard structure, consisting of *Introduction* (the part of paper numbered by "1"), and *Conclusion* (the last numbered part) and several *Sections* in between.

The Introduction should introduce the topic, tell why the subject of the paper is important, summarize the state of the art with references to existing works and underline the main innovative results of the paper. The Introduction should conclude with outlining the structure of the paper.

Accompanying parts

Papers should be accompanied by an *Abstract* and a few *index terms (Keywords)*. For the final version of accepted papers, please send the *short cvs* and *photos* of the authors as well.

Authors

In the title of the paper, authors are listed in the order given in the submitted manuscript. Their full affiliations and e-mail addresses will be given in a footnote on the first page as shown in the template. No degrees or other titles of the authors are given. Memberships of IEEE, HTE and other professional societies will be indicated so please supply this information. When submitting the manuscript, one of the authors should be indicated as corresponding author providing his/her postal address, fax number and telephone number for eventual correspondence and communication with the Editorial Board.

References

References should be listed at the end of the paper in the IEEE format, see below:

a) Last name of author or authors and first name or initials, or name of organization
b) Title of article in quotation marks
c) Title of periodical in full and set in italics
d) Volume, number, and, if available, part
e) First and last pages of article
f) Date of issue

[11] Boggs, S.A. and Fujimoto, N., "Techniques and instrumentation for measurement of transients in gas-insulated switchgear," IEEE Transactions on Electrical Installation, vol. ET-19, no. 2, pp.87–92, April 1984.

Format of a book reference:

[26] Peck, R.B., Hanson, W.E., and Thornburn, T.H., Foundation Engineering, 2nd ed. New York: McGraw-Hill, 1972, pp.230–292.

All references should be referred by the corresponding numbers in the text.

Figures

Figures should be black-and-white, clear, and drawn by the authors. Do not use figures or pictures downloaded from the Internet. Figures and pictures should be submitted also as separate files. Captions are obligatory. Within the text, references should be made by figure numbers, e.g. "see Fig. 2."

When using figures from other printed materials, exact references and note on copyright should be included. Obtaining the copyright is the responsibility of authors.

Contact address

Authors are requested to send their manuscripts via electronic mail or on an electronic medium such as a CD by mail to the Editor-in-Chief:

Csaba A. Szabo
Department of Networked Systems and Services
Budapest University of Technology and Economics
2 Magyar Tudosok krt.
Budapest, 1117 Hungary
szabo@hit.bme.hu

WCNC 2015

IEEE WIRELESS COMMUNICATIONS AND NETWORKING CONFERENCE

NEW ORLEANS, LA, USA • 9-12 March, 2015

IEEE | IEEE COMMUNICATIONS SOCIETY

www.ieee-wcnc.org

IEEE WCNC is the premier event for wireless communications researchers, industry professionals, and academics interested in the latest development and design of wireless systems and networks. Sponsored by the IEEE Communications Society, IEEE WCNC has a long history of bringing together industry, academia, and regulatory bodies. In 2015, New Orleans will become the wireless capital by hosting IEEE WCNC 2015. The conference will include technical sessions, tutorials, workshops, and technology and business panels. You are invited to submit papers in all areas of wireless communications, networks, services, and applications. The instructions for authors will be posted on the conference website www.ieee-wcnc.org/2015. Potential topics include, but are not limited to:

CALL FOR PAPERS

Track 1: PHY and Fundamentals

- Interference characterization
- Cognitive radio, ultra-wideband
- Multihop and cooperative communications
- Modulation, coding, diversity
- Equalization, synchronization, channel estimation
- Space-time coding, MIMO, adaptive antennas
- OFDM, CDMA, spread spectrum
- Channel modeling and characterization
- Interference cancellation and multiuser detection
- Iterative techniques
- Information-theoretic aspects of wireless communications
- Signal processing for wireless communications
- Propagation models for high frequency channels

Track 2: MAC and Cross-Layer Design

- Multiple access techniques
- Cognitive and cooperative MAC
- Collaborative algorithms
- MAC for mesh, ad hoc, relay, and sensor networks
- Network information theory
- Radio resource management and allocation, scheduling
- Cross-layer design, cross-layer security
- Software defined radio, RFID
- Adaptability and reconfigurability
- Wireless MAC protocols: design and analysis
- B3G/4G Systems, WiMAX, WLAN, WPAN
- QoS provisioning in MAC

Track 3: Mobile and Wireless Networks

- Localization for wireless networks
- Network estimation and processing techniques
- Mesh, relay, sensor and ad hoc networks
- Mobility, location, and handoff management
- Mobile and wireless IP
- Wireless multicasting, routing
- Robust routing
- Multimedia QoS and traffic management
- Wireless broadcast, multicast and streaming
- Congestion and admission control
- Proxies and middleware for wireless networks
- Wireless network security and privacy
- Performance of E2E protocols over wireless networks
- Interworking heterogeneous wireless/wireline networks
- Capacity, throughput, outage, coverage

Track 4: Services, Applications, and Business

- Emerging wireless/mobile applications
- Context and location-aware wireless services & applications
- Wireless telemedicine and e-health services
- Intelligent transportation systems
- Cognitive radio and sensor-based applications
- Content distribution in wireless home environment
- Wireless emergency and security systems
- Service oriented architectures, service portability
- SIP based services, multimedia, QoS support, middleware
- Innovative user interfaces, P2P services for multimedia
- Dynamic services, autonomic services
- Regulations, standards, spectrum management
- Test-bed and prototype implementation of wireless services
- Personalization, service discovery, profiles and profiling

CALL FOR TUTORIALS AND WORKSHOPS

Proposals for tutorials and workshops are solicited on hot topics for future wireless communications systems and applications.

CALL FOR PANELS

Panel proposals are also solicited on technical, business and policy-related issues and opportunities for the wireless communications industry.

PLEASE NOTE: To be published in the IEEE WCNC 2015 Conference Proceedings and to be eligible for publication in IEEE Xplore®, an author of an accepted paper is required to register for the conference at the full or limited (member or non-member) rate and the paper must be presented by an author of that paper at the conference unless the TPC Chair grants permission for a substitute presenter arranged in advance of the event and who is qualified both to present and answer questions. Non-refundable registration fees must be paid prior to uploading the final IEEE formatted, publication-ready version of the paper. For authors with multiple accepted papers, one full or limited registration is valid for up to 3 papers. Accepted and presented papers will be published in the IEEE IEEE WCNC 2015 Conference Proceedings and submitted to IEEE Xplore®.

A portion of the accepted papers will be presented as posters. The IEEE WCNC 2015 Technical Program Committee will decide which papers will be presented in oral (lecture type) sessions and which papers as posters, and the decisions will be announced to the authors. Choice between oral and poster presentations will be totally independent of the review scores and of the paper quality.

IMPORTANT DATES

Paper submission deadline:	1 August 2014	Tutorial proposals:	1 August 2014
Notification of acceptance:	15 November 2014	Workshop proposals:	1 June 2014
Final Camera-ready papers due:	15 December 2014	Panel proposals:	1 August 2014

General Chair:
José Roberto B. de Marca,
Pontifical Catholic University of Rio de Janeiro, Brazil

Executive Chair:
Richard Miller,
AT&T (retired), USA

Technical Program Chair:
Nirwan Ansari,
New Jersey Institute of Technology, USA

Technical Program Vice-Chair:
Edit Kaminsky Bourgeois,
University of New Orleans, USA

WCNC Steering Committee Chair:
Khaled Letaief,
Hong Kong University of Science and Technology, Hong Kong

WCNC Advisor:
Sherman Shen,
University of Waterloo, Canada

SCIENTIFIC ASSOCIATION FOR INFOCOMMUNICATIONS

SCIENTIFIC
ASSOCIATION FOR
INFOCOMMUNICATIONS
HUNGARY

FOUNDED IN 1949

Who we are

Founded in 1949, the Scientific Association for Infocommunications (formerly known as Scientific Society for Telecommunications) is a voluntary and autonomous professional society of engineers and economists, researchers and businessmen, managers and educational, regulatory and other professionals working in the fields of telecommunications, broadcasting, electronics, information and media technologies in Hungary.

Besides its more than 1300 individual members, the Scientific Association for Infocommunications (in Hungarian: Hírközlési és Informatikai Tudományos Egyesület, HTE) has more than 60 corporate members as well. Among them there are large companies and small-and-medium enterprises with industrial, trade, service-providing, research and development activities, as well as educational institutions and research centers.

HTE is a Sister Society of the Institute of Electrical and Electronics Engineers, Inc. (IEEE) and the IEEE Communications Society. HTE is corporate member of International Telecommunications Society (ITS).

What we do

HTE has a broad range of activities that aim to promote the convergence of information and communication technologies and the deployment of synergic applications and services, to broaden the knowledge and skills of our members, to facilitate the exchange of ideas and experiences, as well as to integrate and harmonize the professional opinions and standpoints derived from various group interests and market dynamics.

To achieve these goals, we…

- contribute to the analysis of technical, economic, and social questions related to our field of competence, and forward the synthesized opinion of our experts to scientific, legislative, industrial and educational organizations and institutions;
- follow the national and international trends and results related to our field of competence, foster the professional and business relations between foreign and Hungarian companies and institutes;
- organize an extensive range of lectures, seminars, debates, conferences, exhibitions, company presentations, and club events in order to transfer and deploy scientific, technical and economic knowledge and skills;
- promote professional secondary and higher education and take active part in the development of professional education, teaching and training;
- establish and maintain relations with other domestic and foreign fellow associations, IEEE sister societies;
- award prizes for outstanding scientific, educational, managerial, commercial and/or societal activities and achievements in the fields of infocommunication.

Contact information

President: **DR. GÁBOR HUSZTY** • *ghuszty@entel.hu*

Secretary-General: **DR. ISTVÁN BARTOLITS** • *bartolits@nmhh.hu*

Managing Director, Deputy Secretary-General: **PÉTER NAGY** • *nagy.peter@hte.hu*

International Affairs: **ROLLAND VIDA, PhD** • *vida@tmit.bme.hu*

Addresses

Office: H-1051 Budapest, Bajcsy-Zsilinszky str. 12, Room: 502

Phone: +36 1 353 1027, Fax: +36 1 353 0451

E-mail: *info@hte.hu*, Web: *www.hte.hu*